ROCKWOOL FOUNDATION
BERLIN
INSTITUTE FOR THE ECONOMY AND THE FUTURE OF WORK

CReAM Centre for Research & Analysis of Migration

DISCUSSION PAPER SERIES 21/25

People, Practices, and Productivity: A Review of New Advances in Personnel Economics

Mitchell Hoffman, Christopher T. Stanton

JUNE 2025

ROCKWOOL Foundation Berlin
www.rfberlin.com

Centre for Research & Analysis of Migration
www.cream-migration.org

People, Practices, and Productivity: A Review of New Advances in Personnel Economics*

Mitchell Hoffman† Christopher T. Stanton‡

May 2025

Abstract

This chapter surveys recent advances in personnel economics. We discuss new research on incentives and compensation; hiring practices; the influence of managers and peers; and time use, technology, and training. Two main themes emerge from this survey. First, we illustrate the interplay between these topics and productivity differences between people and work units. We discuss evidence showing substantial and persistent productivity variation among workers in the same roles, and we examine the extent to which personnel economics research can explain this variation. Second, personnel economics has benefited from exploration – the willingness to use new data and methods to shed light on existing questions and raise new ones. Since the last handbook chapter, personnel economics has evolved from focusing primarily on compensation and incentives to embracing a broader research agenda that examines various HR practices and their impact on worker and firm outcomes. As many personnel studies use data from individual firms, we discuss external validity and provide concrete guidance on improving discussions of generalizability from specific contexts.

JEL Classifications: M50

Keywords: Incentives, hiring, managers, peer effects, time use, technology at work, training

*__Chapter for the *Handbook of Labor Economics*.__ We are extremely grateful to Kathryn Shaw for guidance, insights, and immense contributions to this handbook chapter, as well as her mentorship to us and many personnel economists. We are also incredibly grateful to Eddie Lazear for his mentorship, insights, and shaping of the field. We also thank Alan Benson, Kevin Bryan, Christian Dustmann, Guido Friebel, Peter Kuhn, Fabian Lange, Thomas Lemieux, and Paul Oyer for helpful comments. We thank Cameron Greene, Shira Aronson, Jessica Arp, and Kazuma Wells for excellent research assistance. We thank the ROCKWOOL Foundation Berlin (RFBerlin) for helping fund the conference for the Handbook.

†UC Santa Barbara and NBER

‡Harvard Business School and NBER

1 Introduction

Personnel economics is the study of labor economics issues inside firms and organizations, with applications to human resources (HR). For many years, HR issues were studied primarily by non-economists, including psychologists, management scholars, and sociologists. However, starting with the pathbreaking work of Eddie Lazear, personnel economics is now researched and taught around the world. Personnel economists study hiring, incentives, managers, peers, technology at work, communication, monitoring, and many other topics. Broadly, personnel economics focuses on the firm side of labor market topics.

What distinguishes personnel economics from the broader field of labor economics? A core differentiator is that, in contrast to labor economists, personnel economists generally take the perspective of the firm, aiming to understand the returns to firms' choices and the benefits and costs of organizational practices. While workers' responses are a core input into firm policies, the outcomes of interest for personnel economists are typically the tradeoffs associated with what firms do. For example, a labor economist studying parental leave policies might focus on how these policies affect children's outcomes and inequality. In contrast, the personnel economist would likely focus on the optimal design of such policies from the firm's perspective, examining how leave substitutes for other compensation, affects retention, or alters worker sorting.[1]

Personnel economics has also developed distinct methodological approaches. The field pioneered "insider econometrics" (Shaw, 2009; Ichniowski & Shaw, 2012), where researchers collect data in close collaboration with companies, often working with proprietary records or designing field experiments within organizations. This approach yields granular insights about worker productivity, incentive responses, and firm practices that would be difficult to obtain from the large administrative or survey datasets more commonly used in labor economics.

Lazear's research agenda in personnel economics grew out of the Chicago tradition, pioneered by Becker (1993) and others, of applying economics principles and methods to a range of social scientific topics. Lazear (2000b) refers to this process as the tendency of economics to expand its "scope of inquiry" and "sphere of influence." In the spirit of Lazear (2000b), we believe that

[1]By focusing on firms' interests, personnel economists are not taking a stand on firms' purpose or social impact (Hart & Zingales, 2017), but rather aim to understand the policy and practice tradeoffs of firms' interactions in the labor market.

personnel economics will benefit by continuing this exploratory process. A variety of developments permit personnel economists to explore further in all aspects of the research process, including data (if the research is empirical), methods, and questions.[2]

On data, firms collect and store huge troves of data about their workforces (McAfee & Brynjolfsson, 2012), and firms increasingly share this data with researchers. There are also many third parties collecting data, from companies specializing in salesforce management to those focused on job testing. Big data enable researchers to perform analyses with statistical precision, to exploit exogenous shocks, and to measure variables that were previously unobserved.[3] Together, these features allow researchers to address new questions and to revisit old ones. For example, big data may help researchers address key questions about heterogeneity, such as how different types of people respond to incentives or value amenities. Sensor technologies and data in digital systems may help to map collaboration patterns or estimate how information or ways of working diffuse through an organization.

On methods, economists increasingly collect their own data using randomized controlled trials (RCTs) (Duflo, 2020) and surveys. In personnel economics, this often takes the form of individual firms collaborating with researchers to run intra-firm RCTs, though there are also RCTs run across firms, such as those done with workforce platforms. The growth of machine learning and artificial intelligence offers tremendous opportunities for researchers, especially when combined with new sources of data. Machine learning enables personnel researchers to create algorithms, to classify large amounts of text into low-dimensional types, and to analyze outcomes like productivity and turnover using high-dimensional control variables.[4]

On questions, Lazear argues that personnel economics is primarily positive rather than normative: it seeks to explain why firms and workers behave as they do (Lazear & Oyer, 2012). We believe this is still true, and that there is tremendous value in understanding the drivers of behavior, especially in terms of decision-making. However, research—especially through RCTs—increasingly examines

[2]The engagement of personnel economics with different sources of data, methods, and questions is not new, as Lazear's own research agenda touched on many new areas (Lazear, 1999, 2006, 2012).

[3]For a discussion of different types of data that are being used inside companies for people analytics, including the uses of digital trace data, see Polzer (2022). For example, Bernstein & Turban (2018) use data from sensors to study face-to-face interactions and the effects on collaboration of a company's move to an open office plan, while Cullen & Perez-Truglia (2023a) use card swipe data on office access to infer when workers take smoking breaks together.

[4]Machine learning is also affecting the behavior of HR practitioners, creating opportunities for research about how access to analytics and data change HR decisions.

the return to different practices, thus providing potential normative guidance to firms. The ability to provide normative guidance can also generate a positive feedback loop, enabling cooperation with firms that allows researchers to access new settings and test new hypotheses. New questions also emerge naturally from new sources of data, e.g., surveys can help build understanding of why employees make certain decisions or have certain beliefs.[5]

A theme throughout the chapter relates to how new data and methods have allowed economists to understand the sources and implications of worker heterogeneity, particularly in productivity or output. Studies of frontline workers, managers, and teams all show that production varies tremendously across people or groups doing the same job or jobs within a firm. Understanding how firm policy and practices contribute to these differences (and overall productivity levels) will likely be a lively direction for future research in personnel.

Since the last handbook chapter by Oyer & Schaefer (2011), much has been learned in personnel economics on a wide range of topics. To summarize, we highlight several important new themes:

1. **Expanded focus beyond incentives:** While the previous chapter emphasized incentives and compensation as the field's primary focus, our survey reflects how the field has broadened its scope in several important ways. While incentives remain important, there has been substantial growth in research on hiring practices, the role of managers (especially those below the C-suite) and peers, and the effects of management, technology investments, and organizational practices.

2. **Productivity variation and benchmarking:** New research has systematically demonstrated substantial and persistent productivity differences across workers doing identical jobs, with differences often exceeding wage variation within a job or occupation.

3. **Methodological innovations:** The field has embraced randomized controlled trials (RCTs) within firms, machine learning techniques, and the analysis of digital trace data that captures worker activities with unprecedented granularity.

4. **New data sources:** Researchers increasingly leverage detailed HR records, online labor markets, and real-time performance metrics to study previously unobservable workplace

[5] E.g., recent research has tackled how sleep (Bessone *et al.*, 2021), noise (Dean, 2022), and financial stress (Kaur *et al.*, 2021) affect worker productivity.

dynamics. The growth of gig and non-standard work arrangements provide new data for researchers to analyze.

Our chapter proceeds as follows. Section 2 walks readers through stylized facts showing that there are tremendous differences in productivity or output between workers doing the same job. This productivity dispersion appears to be ubiquitous across different types of jobs, ranging from entry-level to high-skilled/high-earning occupations. This background motivates questions about the sources of performance differences across people and the role of firms' incentive-setting, hiring, and management practices in either closing or magnifying these gaps.

Section 3 discusses new work on incentives and compensation. This topic has historically been the "bread and butter" of personnel economics, and new research has made great progress in understanding how different incentive schemes influence workers' sorting between firms or jobs. Additionally, while many early studies examined incentive effects in relatively routine tasks, more recent work has shown that behavioral responses to performance pay are present in non-routine occupations including teaching, medicine, and law. We then review several streams of work on topics related to firms' incentive provision, including evidence on behavioral models, motivation for teams and innovators, career and tournament incentives, and tradeoffs when using subjective evaluations. We tie our discussion of incentives to the literature on monopsony in the labor market, where firms set wages strategically based on market conditions. We discuss how market structure may influence incentive provision and suggest that firms' incentive policies under monopsonistic competition likely depend on workers' heterogeneity in productivity.

Section 4 discusses research on hiring, an area which has seen substantial growth since the last handbook chapter. We discuss the role of networks and employee referrals, the role of technology and other procedures in hiring, how hiring practices affect disadvantaged workers, and how workers make decisions about jobs. Research has shown that different hiring methods like employee referrals and job-related tests can provide relevant information about the quality of candidates. While these and other hiring methods have been shown to matter substantially for worker quality and firm outcomes, less is known about how hiring decisions are made. Research also suggests that firm-side policies could be important for improving hiring outcomes for disadvantaged workers.

Section 5 surveys work on managers and peers in the workplace. Work on managers, especially those below the C-suite, represents an emerging focus area for personnel economists. Studies that seek to understand the effects of managers on subordinates' productivity and retention, how to select good managers, and the roles that managers play have proven to be quite influential. Many studies now indicate that individual managers strongly influence their subordinates' productivity and that managers' social and people skills are important inputs. Yet evidence about whether the same skills are valued across all settings and the processes by which firms select good managers remains relatively sparse. The literature on peer effects has also burgeoned, focusing on two mechanisms: peer pressure and knowledge spillovers; we discuss evidence on the magnitude of each. While peer effects can be substantial, frictions appear to prevent peer spillovers in some settings. Some impediments, like the reluctance to ask others for help, may come from workers' preferences or beliefs about how to act among coworkers. Others, like disincentives to share information, may be magnified by incentive contracts chosen by firms. Our read of the literature suggests that firms' practices can influence the extent of peer spillovers at work, even for workers doing autonomous tasks. In fact, for measurement reasons, the literature on peer effects largely focuses on settings where workers produce individually, and important questions remain about how team management and composition affect performance. We conclude this section by mentioning some of this work and calling for additional studies on team production and composition.

Section 6 describes work on important topics that have received less overall attention from the field: time use, technology at work, and training. Recent data advances have made it possible to observe how individuals spend their time at work, including how they sequence their activities. Another core issue is how new technologies will interact with productivity heterogeneity and the boundaries of the firm. A related open question is how tools that wrangle and leverage digital trace data, combined with generative AI, will change the role of managers. Will managers play a greater role in coaching workers, or will AI take on that function? By enabling remote access to workers outside local labor markets, technology may also change the tasks that happen inside or outside of firms. We discuss several papers about how firms use online labor markets to find and manage workers, emphasizing that market design choices by platform operators often have profound effects on matching. Finally, the literature on training has given limited attention to firms. Despite the presence of learning and development functions in most HR departments and the estimated $100 billion that US firms spent on training expenditures in 2022, firms' decision-making concerning investments in workers remains abstract. We mention several recent studies that touch on various aspects of firm training.

Section 7 discusses challenges for personnel economics research related to external validity: how can we generalize from one firm or industry to the labor market overall? We discuss the issue and what personnel economists can do. We also discuss issues related to scaling treatments and general equilibrium effects. Overall, we suggest that personnel economics can help explain aggregate labor market trends, such as the polarization between superstar firms and others, by providing a better understanding of the practices of leading firms and organizations.

Where can personnel researchers continue to advance? In our view, important topics for future work include:

- Analyses of *heterogeneity*, i.e., why do workers perform differently on the job? How might a particular practice affect different types of workers and firms? For example, what types of workers respond best to performance pay.

- Analyses of *worker and firm beliefs* about the returns to different practices. For example, how much value do workers place on good managers? What do workers believe makes a good manager? Do workers and firm leaders have consistent beliefs about the efficacy of different HR practices?

- Research that elucidates how workers and managers *make decisions* about organizational issues. For example, how do managers and organizations vary in making hiring decisions?

While research in personnel economics is extensive, our focus prioritizes novel developments in the field while considering overlap with other chapters of this handbook. The biggest omission of this chapter is work related to the nexus between personnel economics, discrimination, and diversity, as these topics are covered in other chapters.

This chapter can be used in several ways by different audiences:

- For PhD students and researchers: The full chapter provides a thorough review of current personnel economics research, highlighting open questions and future directions.

- For teaching in a full-semester personnel economics course: The entire chapter can be used to structure a curriculum, with each section corresponding to a major topic module.

- For general labor economics or organizational economics courses: Instructors in these courses might focus on Section 2 (productivity variation) and Section 3 (incentives) for a historical view of the field's foundations. Sections 2, 4, and 5 (productivity variation, hiring, managers and peers) could be briefly discussed to provide coverage of newer directions, while the introduction and conclusion may be used as a high-level overview of personnel and adjacent topics.

2 Background and Stylized Facts

Although pay is often used as a proxy for human capital or a worker's marginal product of labor, personnel economists go beyond studying pay to examine output and productivity directly. Empirical evidence shows that there is substantial and persistent productivity variation between firms (Gibbons & Henderson, 2012). For example, Syverson (2004) documents that the average total factor productivity difference between firms at the 75th and 25th percentile of a 4-digit industry code is about 45 log points. Dispersion is similarly substantial even in industries with limited product differentiation. Starting with Bloom & Van Reenen (2007), the research on how different management practices correlate with and cause productivity differences at the firm level has proven to be some of the most important and influential work in economics over the past decade. Evidence from careful observational work, RCTs, and benchmarking shows that the average effect of instituting better management practices is often positive, even for firms that are induced to make changes by external stimuli (Bloom *et al.*, 2013). A core set of managerial "best practices" appears to exist (Bloom *et al.*, 2016), and significant effort has gone into explaining why some firms do not adopt them (Bloom *et al.*, 2014, 2019).

The management best practices identified in the literature encompass both process and incentive dimensions. Many of the studies discussed in this chapter concern worker-level responses to similar practices. Identifying which workers respond to incentive and process changes is critical for understanding within- and across-firm productivity distributions.

To begin this discussion, we focus on one core dimension of worker heterogeneity: differences in baseline output or productivity. Numerous studies document that substantial productivity differences exist at the individual level for workers doing the exact same job in the exact same setting. These productivity differences are comparable in magnitude to those across firms or plants, and often dwarf pay differences across workers, at least in the contexts studied.[6]

Table 1 displays estimates of productivity dispersion at the individual level. Estimates are sorted in ascending order by the approximate skill level required by the job. The "Setting" column describes the context and productivity measure for each paper. The third column reports the "Dispersion Statistic" used to measure productivity in the study. The final column reports the dispersion estimate. An asterisk in the final column signifies that the estimates have been adjusted using a shrinkage or regularization technique, accounting for sampling variation that might load on person-level estimates of heterogeneity.

The first entry, from Mas & Moretti (2009), describes a well-known study of unionized grocery cashiers who, due to union rules, are responsible solely for scanning items and processing customer payments. There is no performance-related pay. The productivity measure used in the study is the number of items scanned per second worked. Mas and Moretti estimate worker fixed effects and find that the difference between workers at the 90th and 10th percentiles of the productivity distribution corresponds to a 21% difference in output. Similarly, Soetevent & Romensen (2024) study bus driver productivity in the Netherlands and report coefficients of variation (standard deviations over means) for fuel economy, acceleration, braking, and cornering that range from 0.05 to 0.86.

Moving beyond relatively routine or manual occupations, Lazear *et al.* (2015) benchmark worker and manager fixed effects in a technology-based services firm with relatively little incentive pay. Using a manager rotation design, they find that the standard deviation of worker fixed effects is approximately 13% of the overall average output per hour. Assuming normality, these results imply a roughly 20% difference in output between frontline workers at the 75th and 25th percentiles. The standard deviation of manager fixed effects on output, when scaled to the entire team, is about 3 times as large as the variation in worker fixed effects.

Sandvik *et al.* (2020) study productivity in a sales call center where workers receive significant incentive pay in addition to an hourly wage. They use the fact that incoming sales calls are randomly allocated to available agents within a division to benchmark baseline productivity differences that existed prior to an experiment. Workers at the 75th percentile of the distribution sell about 48%

[6]As we will do throughout the chapter, we remind readers that these studies come from settings where output can be attributed to individuals; these settings may not be representative of the entire labor market. As such, we urge readers to consider the extent to which findings generalize. We provide some guidance for thinking about external validity and extrapolation to other contexts in the final section.

more on a given call than workers at the 25th percentile. In this setting, the firm's potential profit gain from raising the tail of lower-productivity workers appears substantial, motivating questions about whether such large variation in productivity is optimal and how firms might respond. Can hiring or training practices be adjusted at a low enough cost to make addressing productivity dispersion worthwhile?

Table 1 Comparison of Productivity Differences Across Workers Doing the Same Tasks in the Same Settings

Paper	**Setting**	**Dispersion Statistic**	**Difference**
Mas & Moretti (2009)	Grocery Cashier Checkout Productivity	P90 - P10	21%
Soetevent & Romensen (2024)	Bus Driver Productivity	C.V.	0.05 - 0.86*
Lazear *et al.* (2015)	Technology Service Agents	C.V.	0.13*
Sandvik *et al.* (2020)	Revenue on Random Sales Calls	P75 - P25	48%*
Staiger & Rockoff (2010)	Teacher Value-Added	Std. Dev. of Student Achievement in Math and Language	0.15*, 0.12*
Chan & Chen (2022)	Total Cost of ED Visit Across Providers (separate statistics for NPs and MDs)	Std. Dev, P75 - P25	21%, $650k*
Chan *et al.* (2022)	Physicians' Diagnostic Accuracy on Chest X-Rays	P90 - P10	22%*
Dahlstrand (2025)	Primary Care Outcomes	Std. Dev.	31%*

As discussed further in Section 4, evidence varies on whether firms can adjust hiring processes to accurately identify relative performance differences between workers. Some firms' hiring assessments appear predictive; Sandvik *et al.* (Forthcoming) show that interviewers' pre-hire assessment scores correlate with workers' on-the-job sales, while Hoffman *et al.* (2018) show that pre-hire assessments predict job tenure, a proxy for match quality or job suitability. However, in both cases, firms still hire some workers whose assessments indicate that they are likely poor matches for the job. Why? Three possibilities are: a) the accuracy of assessments in predicting future productivity is not well known to recruiters, b) there are agency conflicts or biases that incentivize recruiters to fill seats regardless of candidates' quality, or c) firms' external labor constraints may make it worthwhile to accept a lower performer rather than raising wages to improve the recruiting pool or letting a slot go unfilled.

Identifying high performers is more difficult in some jobs than others. For example, Staiger & Rockoff (2010) find that a standard deviation increase in a New York teacher's persistent value-added score translates to gains of 0.15 and 0.12 standard deviations in student math and reading scores, respectively.[7] The link between student test scores and future earnings suggests these differences in teacher effectiveness are worth hundreds of thousands of dollars to society. Yet predicting teacher effectiveness at the hiring stage has proven to be enormously difficult. The strongest predictor appears to be prior teaching experience, which offers little help in screening new entrants to the profession.

Shifting to higher-earning occupations, Chan & Chen (2022) show that variation in the total cost of emergency department visits across Veterans Affairs (VA) hospital providers is substantially greater than the variation in their salaries. Because VA hospitals are captive health organizations, cost differences matter for the efficiency of service delivery. Chan and Chen benchmark differences in costs based on quasi-random patient assignment. For both medical doctors (MDs) and nurse practitioners (NPs), the annual spending difference between the 75th and 25th percentiles of their respective distributions is approximately $650,000. Additionally, the average MD (who undergoes more intensive training but also has a higher salary) is more cost effective than the average NP. On average, the higher pay for doctors is offset by the lower costs they generate for the organization. Of course, appropriate case assignment and job design can leave a substantial role for NPs, but this study illustrates that hard-to-measure differences in productivity can swamp differences in pay across groups with different credentials and levels of training. Even more stark, the variation within each professional group exceeds the average difference between MDs and NPs. This variation among practitioners remains present even when focusing only on commonly-performed tasks. For example, Chan *et al.* (2022) find significant variation in the diagnostic accuracy of chest X-ray interpretations between the 90th and 10th percentiles. They attribute this disparity to differences in skill and conclude that enhancing practitioners' skills is a more effective policy than imposing treatment or process guidelines.

Dahlstrand (2025) studies the productivity of primary care doctors on three different outcomes, in a setting with no incentive pay for individual workers. For subsequent Emergency Room visits,

[7] Value-added scores capture the portion of changes in student test scores that can be attributed to an individual teacher.

proxying for doctors' case resolution, a one standard deviation change in doctor performance translates to a 31% difference in ER visits. These are costly to the healthcare payer and suggest primary care has not been successful. An outcome that can be even more closely controlled by the doctor, counter-guideline antibiotics prescriptions, a 1 standard deviation change translates to a 75% change in prescriptions against the national guidelines. Dahlstrand also finds large aggregate gains from matching patients and doctors according to patient needs and doctor productivity in specific tasks, suggesting a multi-dimensional view of productivity may enable gains from trade by comparative advantage.

In the rest of the chapter, we discuss several productivity-related questions that have been studied by personnel economists. To organize our discussion, we refer to an illustrative production function where worker i's individual output in setting j is

$$y_i = T_j \times h_{ij} \times e_i. \tag{1}$$

We let T_j be a technology that may include factors like IT, management practices, managers themselves, or peers on a team. The variable h_{ij} is a workers' human capital or ability in setting j, and e_i is effort. Notice that there is no explicit role for capital in this production function, as many of the studies that we examine hold fixed capital per worker by examining differences across people in the same setting. This framework serves as a starting point, and we will note throughout the text when it requires expansion, such as in cases where output is attributable to a group rather than an individual.

Workers' choices of effort given employment depend on the firm's compensation policy function, $w(y)$, and their individual cost of effort function, $c_i(e)$.[8] Compensation policies may take many forms, including fixed wages, individual performance pay, group incentives, or tournaments, meaning the argument entering the function $w(.)$ can be quite general. Both effort and employment choices depend on $w(.)$.

A substantial literature in contract theory has focused on a problem where firms set $w(.)$ to

[8] While worker heterogeneity is usually modeled as differences in either skill or the cost of effort, some settings allow for distinguishing between the two. For example, a highly skilled plumber with decades of experience may be capable of handling any job, but increasing effort could be more challenging compared to a newcomer who has less accumulated wear and tear.

maximize some overall production function $f(.)$ less wages,

$$f(\sum_i y_i) - \sum_i w(y_i). \tag{2}$$

Most of the focus has been on how changes in $w(.)$ affect e (i.e., incentive compatibility) and how they may alter the composition of workers at the firm (i.e., individual rationality). Until recently, there had been much less focus on how the hiring, selection, and training of managers and workers influence h and T. These factors are especially important in shaping how the individual rationality constraint functions in real-world contracting scenarios.

Practices that influence the individual rationality constraint and worker selection are likely highly important, as variation across the broader labor market is almost certainly greater than the within-firm differences we have discussed. For example, when comparing knowledge workers in research, Levin & Stephan (1991) document that in most fields, the standard deviation of research productivity is larger than the mean, with a distribution that is significantly skewed. Hiring the most productive workers, promoting the right managers, and getting teams to gel are consequential, practical problems for firms. Yet these problems are difficult to navigate, both for firms and researchers.

For researchers, it is rare to observe the same individuals' productivity or output across different settings - an important factor in assessing the benefits of external hiring or determining whether differences across firms stem from the firms themselves, the workers, or match effects. There are, however, important exceptions. For example, Groysberg (2010) measures the accuracy of security analysts' forecasts as they move between firms. His work shows that firm-level factors influence the productivity of star analysts, raising questions about the portability of human capital. Lerner *et al.* (2024) use a related related design in their study of university researchers, finding that institutional and geographic factors (firm effects) account for a significant share the variation in commercialization outcomes. Similar studies on other professions yield mixed results; doctors only partially adapt their practice styles after moving to a new region (Molitor, 2018), while teachers' value-added appears to translate across districts (Biasi, 2021).

For firms, implementing the optimal hiring and promotion practices requires matching the appropriate evidence to their particular setting, understanding contingencies, and testing out different policies and practices. Although our discussion begins with recent work on incentives, a core message from our chapter is that other channels beyond pay are important to understand how firms make decisions about managing people.

3 Incentives, Compensation, and Labor Markets

Personnel economics research has demonstrated that firms' incentive and compensation policies can affect output. This research is often grounded in agency theory. A key implication of the basic agency model is the tradeoff between risk and incentives. Input-based contracts (i.e., fixed wages) may offer only weak motivation for workers to exert effort, while output-based contracts (i.e., pay-for-performance) expose workers to risk from factors beyond their control (Lazear, 2018). Due to measurement challenges arising from unmodeled production factors (e.g., agents' local information advantages (Prendergast, 2002)), Oyer & Schaefer (2011) argue that personnel economists should extend their focus beyond risk/incentive tradeoffs and optimal incentive contracts. Since Oyer & Schaefer (2011), a rich body of research has explored the various ways in which incentives influence worker behavior. Compared to earlier work that often focused on theory and contract design, this research has taken a more empirically-driven approach, emphasizing estimates of responses to incentives rather than the derivation of optimal contracts. The literature has also explored new approaches and settings, such as behavioral models, compensation for innovators, group-based incentives, and subjective evaluations. Given the breadth of this literature, our discussion will necessarily omit some core contributions. In particular, due to excellent recent surveys on executive compensation by Frydman & Jenter (2010) and Edmans *et al.* (2017), we focus instead on incentives for lower-level employees. While much of our review covers empirical applications, we hope that readers will be able to translate from estimates in different contexts to inform the theoretical models that originally launched the field.

Conceptually, incentives can affect productivity through three channels. First, incentives can affect effort. In the baseline model, performance-based pay (piece rates, tournaments, or subjective bonuses) encourages greater effort compared to a fixed salary or hourly pay since the agent's compensation is directly linked to their output. Second, incentives can influence the composition of a firm's workforce through sorting. Under performance-based pay, for example, less able workers may leave the firm, while more able workers are often thought to be attracted to join (Lazear, 2000a). A third possibility is that incentive pay may alter the extent to which workers accumulate human capital on the job. For example, the introduction of high-powered individual incentives or career-incentives may cause workers to learn more rapidly.[9]

We first examine the literature on performance pay. We discuss recent theory and evidence on sorting and effort responses, tests of behavioral models that depart from the standard agency-theoretic paradigm, and incentives in "non-standard" production environments, particularly those involving teams and innovators. At the time of the previous handbook chapter, substantial research showed that workers in routine jobs (e.g., fruit pickers, tree planters) respond strongly to performance pay. Research over the last 15 years has extended this finding to show that workers in non-routine jobs, and especially high-performers, similarly respond strongly to performance pay. But this does not mean firms should necessarily use pay that is strongly based on output, as there will be tradeoffs between what is measured and not measured – leading to multi-task concerns (Holmstrom & Milgrom, 1991). In fact, there are many settings in which firms do not utilize output-based contracts. In the second part of this section, we examine research on alternative incentive structures, including pay levels, subjective performance evaluation, career incentives, and how firms establish standards or requirements to encourage worker effort.

We would encourage the next generation of personnel economics researchers to push further linking the work on incentive provision to the external labor market context where firms operate. Standard models typically assume that workers' options outside the firm are fixed. However, the literature on monopsony (reviewed elsewhere in this handbook) suggests that understanding how incentive provision interacts with external labor market conditions is a key area for future research. Beyond the focus on labor market competition in the monopsony literature, demographic factors, such as aging societies, may also influence whether firms increase reliance on career incentives or short-term contracts.

[9]This third aspect of how incentives affect on-the-job learning has not received as much attention. The closest literature tends to look at signaling and rat race equilibrium through the lens of work hours (Holmström, 1999; Landers *et al.*, 1996). We would welcome additional research on how high-powered incentives influence workers' human capital acquisition.

3.1 Performance pay

3.1.1 Sorting and effort responses

Effects of the introduction of performance pay As Lazear (2018) observes, "the literature is full of examples where manipulating the pay structure alters worker behavior, by affecting either hours of work or output associated with it." The canonical example is a study of Safelite Auto Glass installers, where a shift from hourly wages to piece-rate pay per successfully installed windshield led to a 44% increase in productivity, largely driven by sorting (Lazear, 2000a).

Recent studies examine workers' responses to performance pay in alternative settings. For example, theory recognizes that workers who do a variety of different tasks face multi-task incentives, meaning that the incentive structures for doctors, lawyers, or teachers may differ from those in roles focused on a single output dimension. Additionally, workers in different jobs may vary in cognitive skills and their attention to financial details – artists, for instance, may be less responsive to monetary incentives, while other workers may struggle to interpret complex or opaque links between performance and pay.

To illustrate how the literature has evolved, we begin with results on performance pay for teachers. Studies of teachers' incentive responses should interest personnel economists for a number of reasons. First, teacher output measures – value-added scores – can be standardized across schools or districts, allowing economists to analyze not only how teachers respond to incentive changes in one establishment or school, but also how those who move between schools or districts adjust their behavior or performance after relocating. Second, even though debate continues over whether teacher value-added is the right indicator of teachers' performance, this output measure is one that economists understand.[10] Third, the bureaucratic structure of most school systems ensures that complementary inputs remain relatively consistent across settings within the same local area or labor market.[11] However, as with many personnel studies, questions remain about whether findings from regions with particular cultural institutions or behavioral norms generalize to other contexts. Finally, and perhaps most importantly, teaching is a non-routine job that likely attracts workers

[10] For readers interested in best practices for addressing gaming – along with a discussion of commonly-used contract structures, such as pay-for-percentile – see Neal's (2011) excellent handbook chapter on teacher incentives.

[11] The primary concern when comparing settings is differences in student or parent quality. While value-added measures focus on achievement gains rather than levels, sorting based on match effects may still occur if some teachers have a comparative advantage in working with specific segments of the achievement distribution.

with intrinsic or pro-social motivations – two factors that studies have scrutinized when questioning whether findings on performance pay in simpler, routine settings apply more broadly.

We build our discussion around Brown & Andrabi (2023), who develop a simple model of sorting and effort effects that they test empirically with data from a two-part experiment. In their model, there are two contracts: one paying a fixed wage (contract j_F), and one that pays based on performance (contract j_P). Teachers' utility from choosing each contract is given by

$$u = \begin{cases} w_0 + \epsilon_{iF} & \text{if } j = j_F \\ p(\hat{\theta}_i + \hat{\beta}_i) - 0.5p\hat{\beta}_i + \epsilon_{iP} & \text{if } j = j_p \end{cases} \tag{3}$$

where p is performance pay and $\hat{\theta}_i + \hat{\beta}_i$ is output. Output is determined by baseline ability, θ_i, and the individual's effort responsiveness to incentives, β_i. In the language of our framework in Section 2, θ_i maps to h_{ij}, although it does not have the firm-specific match component (j in our framework), and β_i roughly maps to $c_i(e)$ when the cost of effort function is quadratic above some normal level. The hat notation indicates that teachers have prior beliefs with some uncertainty around the true values. The ϵ terms represent the non-wage amenities for each job.

Differences in output between fixed and performance-based pay can be broken down into three components: sorting on ability, sorting on the individual's effort response to performance pay, and the average effort increase under performance pay. Brown & Andrabi (2023) estimate these effects using a creative 2-part experiment in a chain of private schools in Pakistan. Prior to the experiment, the standard deviation in teachers' value-added – which they use to measure productivity dispersion in the absence of incentive pay – was 0.15; this estimate is remarkably similar to the one for New York City teachers in Table 1. Teachers in their sample earn about $4,000 per year.

In the first phase of the experiment, researchers elicited how individual teachers would like their annual pay raise allocated between a performance-based component and a flat increase.[12] Teachers who wanted more than 50% of their raise to be performance-linked had value-added scores that were 0.05 standard deviations higher than teachers who preferred a larger fixed raise – about one-third of a standard deviation in baseline productivity.

[12]To make the exercise incentive compatibility, teachers' stated contract preferences had a roughly 1-in-3 chance of being implemented.

In the second phase, randomization was conducted at the school level. Some schools implemented teachers' individually-chosen pay structures, while others were assigned to either a fixed 5% base salary increase or a raise tied to performance. Under the performance-based wage increase, raises were determined using a pay-for-percentile scheme within each school.[13]

This second phase has two attractive features for inference. First, there is mobility across schools in the network, allowing the authors to analyze how teachers sort to schools with different pay structures. Teachers who moved from flat-pay to performance-pay schools after the second phase was implemented had value-added scores that were 0.064 standard deviations above the mean. However, not all teachers relocated. Overall, teacher re-sorting accounted for a 0.022 standard deviation difference in baseline value-added between performance-pay and flat-pay schools. These effects were observed over a one-year period and are averaged across both movers and stayers, suggesting that long-run effects may be even larger.[14] Second, the design enables Brown and Andrabi to compare performance changes between teachers who receive their preferred contract versus those who do not. For teachers who preferred and were assigned to the performance-based raise contract, value-added increased by 0.09 standard deviations. In contrast, for teachers who favored a fixed raise but were assigned performance pay, value-added increased by only 0.01 standard deviations. The authors interpret this difference as evidence of heterogeneity in effort responses because they can hold fixed value-added scores prior to the contract assignment, capturing ability heterogeneity.[15]

[13]The assessment for the percentile-ranking was based either on test scores or on administrators' subjective evaluation, and was cross-randomized across schools. Teachers in the 90th percentile or above received a 10% raise, those between the 61st and 90th percentiles received a 7% raise, those between the 16th and 60th percentiles received a 5% raise, and those below the 15th percentile received either a 2% or 0% raise.

[14]Some other studies of market-level introductions of incentive pay also document positive sorting. Biasi (2021) studies a policy change in Wisconsin that allowed school districts to implement flexible pay structures, moving away from traditional, seniority-based compensation models. After the reform, high-quality teachers increasingly sorted into districts with flexible pay, raising student achievement. Here, the sorting channel accounts for approximately two-thirds of the relative achievement gains in districts that adopted flexible pay. Leaver *et al.* (2021) use a similar design to study the effects of pay-for-performance for teachers in Rwanda. The first stage of their experiment involved randomly assigning districts to advertise contracts as either pay-for-performance or fixed-wage positions. They then re-randomized contracts after selection (and offered a transfer to make sure no teacher was worse off) to study performance effects. They find that effort increases outweigh sorting on ability; teachers who opt into performance-based contract tend to have lower intrinsic motivation, but compensate by increasing their on-the-job effort. Compared to other designs, the lack of contract variation within a district labor market may limit sorting effects because of the search, moving, or commute costs that teachers might incur by changing jobs.

[15]While performance pay for teachers in developing countries has been shown to lead to substantial performance gains, the evidence on pay-for-performance in advanced economies is much less clear. For example, Muralidharan & Sundararaman (2011) find that introducing a student achievement bonus for teachers equal to 3% of their salary lifts performance in rural Indian schools by 0.27 and 0.17 standard deviations in math and language tests, respectively. These gains appear to stem from increased teacher effort, and are larger when incentives are tied to individual teachers' performance rather than awarded at the school level. In contrast, Fryer *et al.* (2022) argue that most teacher incentive pilots in the US do not produce comparable achievement gains. Several factors may explain these differences, including

Given concerns about multitask incentives and the potential crowding out of prosocial behavior, the authors test for cheating, changes in classroom environments, and reduced altruism. They detect some deterioration in the classroom environment under performance pay (e.g., higher pressure put on students or more yelling), but these effects appear concentrated among teachers who did not want the performance pay contract. Allowing them to sort out of the contract would have likely improved aggregate outcomes. Overall, the paper's creative design shows heterogeneity in both preferences for and responses to performance-pay contracts. We are aware of few other studies that combine pre-treatment measures of preferences with actual sorting behavior to uncover heterogeneous incentive effects. Methodologically, this innovative experimental design offers valuable insights for other researchers seeking to advance the study of incentive responses while accounting for variation in how different workers react. Understanding these response margins is likely useful for understanding why firms might tailor their incentive offers to attract and retain distinct groups of workers.

Other recent work also estimates workers' incentive responsiveness in a variety of non-routine jobs. Clemens & Gottlieb (2014) examine a 1997 change in how Medicare – a nearly universal health program for US residents over 65 – reimbursed doctors. Before 1997, reimbursement rates for a given procedure varied across 210 payment regions using a regional multiplier. In 1997, Medicare reduced the number of payment regions to 89, causing procedure prices to increase in some areas and decrease in others. Although this case does not offer an exact principal-agent analogy, about 60 percent of doctors were self-employed at the time of the change, and approximately 85 percent of those in group practices had compensation tied to patient care revenue. Leveraging this variation in prices, Clemens and Gottlieb find a long-run elasticity of care supply with respect to prices of 1.5, as doctors change elective and in-office procedures.[16] Across a wide range of settings, like publication incentives in academia (Checchi *et al.*, 2021), to incentives for lawyers to settle assigned criminal defense cases quickly for indigent defendants (Agan *et al.*, 2021a), to doctors who respond to higher

implementation challenges, designs that limit sorting effects, and baseline conditions where higher levels of monitoring and established norms for basic behaviors, such as attendance, are already in place.

[16]In some settings, like medicine, higher care utilization is not always better, leading to concerns that high-powered incentives come at the expense of patients' best interest. Johnson & Rehavi (2016) suggest there is some merit to these concerns. They study C-sections when the patient is herself a physician compared to when one parent is highly educated but neither is a physician. Physician patients are 7-8% less likely to have a C-section compared to other highly educated patients, and the gap between physician and non-physician patients widens in hospitals where there are financial incentives for doctors to perform C-sections.

reimbursement rates (Clemens & Gottlieb, 2014), high-skilled workers in non-routine tasks respond to incentive pay.

Yet this strong incentive response is not always useful for organizations. In a leading international law firm that reduced the strength of incentives for lawyers, Bartel *et al.* (2017) reveal a common issue with output-based incentives: they can disproportionately focus effort on measurable tasks, such as billable hours, while discouraging contributions on less quantifiable but valuable activities, like mentoring or fostering firm culture. This reflects a broader challenge in incentive design—when compensation is tied to specific, easily measured outputs, workers may shift their efforts toward those metrics at the expense of other important but harder-to-measure responsibilities.

Another concern is agents gaming incentive schemes, which may muddle inference about incentives in experimental tests. For example, Alexander (2020) makes the point that tests of incentives to reduce medical costs are often done in a way that allow health providers to sort patients to the test scheme or out of it. Alexander studies a New Jersey Gainsharing Demonstration, a pilot experiment where hospitals paid doctors bonuses for reducing total costs for Medicare patients. A maximum bonus was assigned for each patient type, with doctors getting some fraction up to the maximum bonus, depending on treatment costs for the patient. However, doctors often have privileges to admit across different hospitals, some of which are in the scheme and others that are not. Alexander finds that doctors "responded to the bonus by reallocating admission across patient – both by changing admission thresholds and diverting healthier patients into participating hospitals." There were no changes in actual health spending or healthcare utilization when holding fixed patient characteristics, but doctors' gaming behavior could have masked the lack of changes in actual patient service provision. This point makes clear that gaming is likely of considerable importance, and gaming can muddle inference in experiments done at small scale. We return to this point later in the chapter when discussing external validity.

Effects of performance pay on the intensive margin Much of the work on incentive pay has focused on the extensive margin shift from fixed to incentive-based pay. However, the magnitude of incentive pay also likely matters. A common theme across studies looking at variation in the strength of incentive pay on the intensive margin is that increasing the power of incentives often helps with the attraction and retention of more productive workers.

Sandvik *et al.* (2021) examine the effects of incentive pay changes on the intensive margin in a quasi-experiment in which a sales call center reduced worker commissions in one of its six divisions. Average pay in the affected division was predicted to drop by about 7%. As a result, turnover increased significantly among the most productive workers. However, retention rates for lower-productivity workers, sales levels, and reported effort showed little change. The authors suggest that workers may have increased effort to maintain earnings (an income effect channel), but at the same time, lower commissions reduced the incentive to exert effort (a price effect channel), leading to a negligible change in net effort. Ultimately, the firm lost more revenue due to the departure of high-performing employees than it saved in reduced compensation.

In other settings, the income and price effects of compensation changes may not cancel out. For example, Krueger & Friebel (2022) study the equalization of pay across divisions in a consulting firm, where some divisions initially offered higher base salaries with smaller bonuses, while others relied more heavily on bonus-based compensation. Under the new policy, divisions that initially had higher bonus pay saw a decrease in their bonuses and an increase in their base salaries. In contrast to the pure reduction in commissions in Sandvik *et al.* (2021) that reduced overall earnings, cutting variable pay but increasing salaries may have muted the income effect channel, leading to a significant drop in effort and an increase in departures among highly productive workers. Less productive workers and new entrants to the firm had limited responses. A different possibility is that sales workers in a call center have less discretion than those in consulting, giving consulting employees greater latitude to adjust discretionary effort.[17]

Extrapolating from these studies, it is likely inappropriate to conclude that increasing the strength of intensive margin incentives will always improve firms' recruitment of high performers. Castro-Pires & Georgiadis (2025) demonstrate this result by developing a moral-hazard model that accounts for workersâ outside options, showing that steeper incentives can backfire by disproportionately attracting low-ability applicants if their change in expected utility is greater than the change for high-ability ones. Beyond the studies of teachers cited previously, empirical work that builds in workers' outside options is likely an opportunity to test how changes in incentive power alters worker

[17] Coviello *et al.* (2022a) suggest that looking just at retention and effort proxies may miss other ways that workers might respond. They study an incentive change in a retail call center where commission thresholds increased, forcing workers to exert more effort to maintain their initial income. In this setting, workers could misallocate customers to certain products, leading to higher return rates that harmed the firm. In response to the adverse incentive change, some workers appear to signal displeasure by increasing the return rate. The firm subsequently responded by more heavily penalizing refunds, but the short-term effect of the change reduced profitability for the firm by more than the compensation savings.

flows between firms. We discuss this more in Section 3.3 on external labor market conditions.

3.1.2 Behavioral models and non-monetary motivation

Machiavelli's writing about mercenaries suggests paying for output is a fool's errand because it reduces allegiance or intrinsic motivation, saying "The fact is, they have no other attraction or reason for keeping the field than a trifle of stipend, which is not sufficient to make them willing to die for you." A growing behavioral literature has explored incentive-related questions using models beyond the agency-theoretic framework and interacts with non-financial aspects of motivation. Some findings suggest that factors such as inequality aversion, social preferences, and stress can weaken the effectiveness of high-powered incentives. Other factors, like loss aversion or limited cognition about complexity, may be leveraged by firms to magnify the power of incentives. These behavioral biases likely vary across individuals, raising the possibility that, in the long run, firms may design different incentive structures that allow workers to sort on behavioral factors in addition to ability and effort costs. A potential area for future research might involve quantifying the importance of behavioral factors in how workers sort into and out of particular incentive schemes. To illustrate how these studies might be conducted, we briefly walk through a baseline model and some key tests from DellaVigna & Pope (2018), which examine behavioral incentive effects in a large-scale experiment run on Amazon's Mechanical Turk. Our goal is not to conduct an exhaustive review, but instead to provide background to readers about strategies for estimating a particular subject's behavioral biases. We then review findings from studies that examine behavioral incentive concepts in real firms.

DellaVigna & Pope (2018) evaluate models where a worker chooses effort, e, to solve

$$\max_{e \geqslant 0} \, (s+p)e - c(e). \tag{4}$$

In the objective function, s captures a norm to supply effort in an employment relationship (possibly due to gift exchange or gratitude), and p denotes a piece rate. The function $c(e)$ captures the cost of effort and is assumed to be convex. Optimal effort is $e^* = c'^{-1}(s+p)$. The authors then extend this baseline model to account for how various treatments induce different behavioral responses. Here, we describe some of the extended models:

- Crowd Out: The concept of crowd out suggests that small monetary incentives interfere with intrinsic motivation. This idea is captured by the first order condition for effort $e^* = c'^{-1}(s + \Delta s_{CO} + p)$. Here, the parameter Δs_{CO} captures any change in effort that occurs when shifting from no piece rate to a very low one, which has limited direct incentive effects. Since the monetary reward is too small to meaningfully impact effort through standard incentives, any observed change in effort suggests that introducing financial compensation has displaced intrinsic motivation.

- Altruism and Warm Glow: These concepts suggest that firms might improve the power of incentives by tying workers effort to some socially desirable outcome. The first order condition is $e^* = c'^{-1}(s + \alpha p_{CH} + a \times .01)$, where α is an altruism parameter that multiplies a piece rate that is paid to a charity instead of to the individual. An alternative model, warm-glow, is meant to capture that any small amount of money given to the charity (in this case $0.01) may alter the utility of effort.

- Gift Exchange and Psychological Treatments: Gift exchange models suggest that tying payment to output may be unnecessary, as workers will reciprocate positive upfront payments. The first order condition $e^* = c'^{-1}(s + \Delta s_{GE})$ captures a gift exchange treatment where a bonus is paid up-front, independent of the workers' amount of output. There is no piece rate.

- Psychological Manipulations: These models suggest that firms can alter the information environment to manipulate productivity with limited need to pay high monetary costs. The first order condition $e^* = c'^{-1}(s + \Delta s_{Psych})$ captures treatments with different psychological manipulations. In one psychological treatment on social comparison, subjects were told that many prior participants exceeded a certain score. In another, subjects were told that they would learn where they stand relative to other participants. A final treatment targeted meaning at work, in which subjects were informed that the task is significant to the researchers.

- Reference Dependence: Reference-dependent models suggest that firms can frame incentive schemes in different ways to alter effort for the same amount of expenditure. The objective function with bonus G in a gain-framed treatment is $max_{e \geqslant 0} se + 1_{e \geqslant T} G + \eta(1_{e \geqslant T} G - 0) - c(e)$, where workers get a bonus of G if they achieve a threshold T and get zero otherwise. The utility function features $1_{e \geqslant T} G$ as part of consumption utility, while the parameter η captures additional utility from hitting the bonus relative to the reference point of zero. Under loss framing, the utility becomes $max_{e \geqslant 0} se + 1_{e \geqslant T} G + \eta\lambda(0 - 1_{e<T} G) - c(e)$ in a loss framed treatment. The utility to hit the threshold in the gain framed treatment is thus $(1+\eta)G$ and is $(1+\lambda\eta)G$ in the loss-framed treatment. When $\lambda > 0$ there is loss aversion, which predicts higher effort in the loss-framed treatment than the gain-framed treatment for the same level of bonus.

Results from DellaVigna and Pope's real effort experiment lend support for both classical incentive effects and behavioral factors. Strikingly, their results reveal that even small piece rates generate substantial effort responses, exceeding the average effects of gift exchange, meaningful work, and social comparisons. The authors find no evidence that small incentives crowd out intrinsic motivation. In their setting, even minimal piece-rate payments lead to significant effort increases. However, this finding may potentially be a result of the task they utilized (we wouldn't expect subjects to be intrinsically motivated to repeatedly pressing keys).[18] Finally, the loss-framed treatment in their experiment was slightly more effective than the gain-framed one. However, the primary effect of both treatments, relative to controls, seems to stem from the introduction of a clear threshold, motivating workers to adjust their effort levels to ensure they qualify for (or avoid losing) the bonus.

These findings highlight the importance of setting clear targets, suggesting that goal-setting itself may be a powerful incentive tool. Despite its potential impact, the way firms design and implement goals as part of their incentive structures has received relatively little attention in the personnel literature. An exception is Kuhn & Yu (2024), who draw on the extensive psychology literature on goal setting and apply it to personnel economics. While goal setting has been widely studied in psychology, its role in workplace incentives is harder to isolate because achieving a goal is often tied to a financial reward, making it difficult to separate the motivational effect of the goal itself from the incentive from the payout. To separate these effects, Kuhn & Yu (2024) study small retail teams where the pay schedule becomes kinked (but does not have a discrete jump) at the goal threshold. The kink means that the marginal return to effort is slightly higher just after reaching

[18] More generally, other research has suggested that incentives may undermine an agent's perception of a task, eroding intrinsic motivation in the long-run (Benabou & Tirole, 2003; Gneezy *et al.*, 2011).

the goal than just before it. If the goal is psychologically important, workers should bunch their effort near the achievement level. In contrast, if the incentive effect dominates, workers should continue increasing their effort beyond the kink to take advantage of the higher returns. Kuhn & Yu (2024) find significant bunching, suggesting that teams derive intrinsic rewards – or value some symbolism, praise, or extra positive attention – from achieving goals.

We would be interested in work that explores how firms set goals or standards. Both DellaVigna & Pope (2018) and Kuhn & Yu (2024) find that goal setting can lead workers to supply extra effort to hit quotas or targets. The efficiency wage-like models of Lazear *et al.* (2016) and Coviello *et al.* (2022b) feature minimum performance standards for retention, consistent with goal setting. In practice, firms often use different frameworks to set objectives and specify key performance indicators, and resources targeted toward managers highlight different rationales for using one approach versus others. Work that uncovers how firms set goals jointly with compensation, and whether managers anticipate workers' behavioral responses (e.g., ratchet effects) to different goal setting methods, would be very useful to link theory to practice.

In practice, how can firms leverage other behavioral factors to induce effort? Several studies suggest that understanding behavioral and psychological incentive responses might be fruitful for firms, but we caution that worker responses may be heterogeneous across settings.[19] Several prominent studies have tested behavioral incentive responses involving loss-framing or crowd out. In a study revisiting the limited incentive effects of teacher pay in high income countries, Fryer *et al.* (2022) show that using loss-framed incentives significantly improves teacher value-added compared to gain-framed incentives or no incentives. In their design, teachers received bonuses upfront based on expected achievement gains that were then clawed back if teachers fail to improve test scores. Similarly, Brownback & Sadoff (2020) find that loss-framed incentives for community college instructors are tied to significant improvements in student educational attainment. While instructors tend to dislike the loss-framed bonus treatments initially, the initial preference for gain-framed bonuses diminishes with time and experience. In other cases, loss-framing may be too powerful, leading to multitask concerns. For example, Pierce *et al.* (2020) conduct a field experiment on loss-framed bonuses in a car dealership network. Contrary to their expectations,

[19]For example, DellaVigna & Pope (2018) show that expert forecasts of incentive effects are more variable in treatment arms where the standard deviation of worker output is larger, suggesting that some treatments (especially gift exchange, social comparison, and task significance) will induce heterogeneous worker responses.

the treatment resulted in a 5% reduction in sales. The authors attribute this outcome to gaming behaviors, where workers shifted their focus toward the incentivized task while neglecting other important but uncompensated responsibilities.

In the spirit of connecting to the prior literature on sorting, heterogeneity in behavioral factors may also lead to sorting on different dimensions than those emphasized in traditional models. For example, there is emerging work suggesting that behavioral biases differ across individuals (see, e.g., Chapman *et al.* (2022) on heterogeneity in loss aversion). Larkin & Leider (2012) test this notion in the lab by focusing on overconfidence. They measure how subjects expect themselves to perform on a multiplication test, and offer them the the choice of linear or convex pay schemes. Overconfident workers (as measured by differences between actual and expected performance) are more likely to choose the convex scheme, work hard to fulfill their expectations, and ultimately produce more at lower cost than if they did not have this behavioral bias. Connecting this idea to personnel outcomes, Hoffman & Burks (2020) study overconfidence in the context of long-haul truckers. They show that new workers, even when given strong incentives for accurate predictions, systematically over-predict their productivity and correct these predictions much slower than predicted by Bayes' rule. Over-prediction in this case is consequential because drivers are paid by piece rate (i.e., by the mile). Conditional on actual productivity, drivers with higher productivity beliefs are more likely to stay with the firm, and firm profits appear to benefit substantially from worker overconfidence. Similarly, in a field experiment of managers at a food and beverage store chain with tournament-like pay based on store performance, Huffman *et al.* (2022) find that managers have overly positive memories of their past performance and make overconfident predictions about their future performance. When managers are overconfident and slow to update their beliefs, they tend to continue supplying effort in tournaments even when it may be optimal to give up.[20]

To what extend are framing strategies that firms use to target crowd out, meaning, or intrinsic motivation more effective than pay levels or performance pay?[21] Papers in this area often assess whether awards and recognition, communicating task significance, and financial rewards are substitutes or complementary practices. For example, Ashraf *et al.* (2014) show that a "star"

[20]Overconfidence is a highly robust behavioral phenomenon that is widely documented across many lab and field settings (De Bondt & Thaler, 1995; Bernstein *et al.*, 2023).

[21]Frederiksen *et al.* (2024) point out that framing and negotiations around incentive changes are often important tasks for managers, and likely precede implementation of large-scale changes.

recognition treatment for top-selling retailers boosts condom sales for HIV prevention in Zambia. For most of their sample, recognition has a greater marginal effect than financial incentives,[22] but financial incentives nevertheless boost the performance of the sellers who are likely to have the highest intrinsic motivation. In contrast, Deserranno (2019) finds that higher compensation for a health promoter position appears to signal a lower positive benefit to the community from the role, discouraging those with strong pro-social preferences from applying. Rationalizing the mixed results on crowd out due to financial incentives deserves further attention. We would be curious to know whether differences across studies come from variation on the intensive margin or the extensive margin. It is also possible that differences arise because of variation in the target population, where some workers or groups may be less inclined to have intrinsic motivation at baseline.

More complex incentive schemes can also lead to varied reactions among workers. For example, Abeler *et al.* (2023) find that cognitive ability influences how workers respond to incentives. In a field experiment with warehouse workers, the introduction of a dynamic incentive scheme with ratchet effects led to increased effort, even though the optimal strategy, based on the scheme's design, was to reduce effort. Workers' responses varied as a function of a pre-experimental assessment of cognitive skills. Complex schemes can also lack salience, reducing attention. Prior research has shown that simply providing reminders about a complex scheme can encourage workers to increase effort (Englmaier *et al.*, 2017). These findings highlight complexity as an important but underexplored factor in the design of incentives, suggesting that more field studies are needed to examine its effects.

3.1.3 Incentives for teams and innovators

Team incentives The prevalence of team-based work has grown significantly in recent years, with estimates suggesting that the time spent on collaboration has increased by 50 percent over the past 2 decades (Cross *et al.*, 2016). From 1980 to 2012, jobs requiring high levels of social skills increased by nearly 12 percentage points as a share of the US labor force (Deming, 2017). Evidence from scientific papers and patents shows that teams increasingly dominate knowledge production (Wuchty *et al.*, 2007). As teamwork becomes increasingly important, personnel economists may need

[22]Gallus (2017) also finds that recognition for Wikipedia editors improves their retention at zero financial cost, while the literature on social comparisons yields mixed evidence on these effects (Blanes i Vidal & Nossol, 2011; Reiff *et al.*, 2022).

to consider whether individuals or teams should be the primary unit of analysis in future studies. Forming effective teams—and drawing meaningful inferences about them—is challenging due to the vast number of possible coworker combinations and interactions between people. This section focuses specifically on team-based incentives, while Section 5.2 examines aspects of multi-agent production and team effectiveness—such as peer learning, mutual monitoring, and coordination.

Workers in teams face an inherent tradeoff between spending time on individual work and spending time helping others (Lazear, 2018). A potential way to incentivize collaboration is to tie an agent's pay to some measure of team output. The core issue with doing so is that paying based on group performance reduces incentive power compared to individual incentives due to free-riding—workers receive only a fraction of the returns from their own effort while still bearing the full cost of exerting it (Holmstrom, 1982). This problem worsens as team size increases, since the personal benefit of additional effort declines while the individual still incurs the full cost.

Empirically, however, team incentives have proven to be effective in various field settings. A famous example comes from Knez & Simester (2001), who find that implementing a firm-wide bonus for all 35,000 hourly employees at an airline improved performance. The firm's structure—organizing employees into autonomous work groups—facilitated mutual monitoring, allowing smaller teams to reinforce accountability. This highlights the importance of interactions between social and team incentives in making group-based pay effective.

Bandiera *et al.* (2013) illustrate this connection in their study of team-based rank incentives and tournament pay at a soft fruit producer. The authors design a field experiment based on a model that nests both pure free-riding and pure social preferences, where workers internalize spillovers. Using the notation of our framework, the utility of worker i who forms a team with worker k is

$$U_{ik} = \frac{p(1+\pi_{ik})}{2}(h_i e_i + h_k e_k) - \frac{1}{2}e_i^2 + s_{ik},$$

where $\pi_{ik} \in [0,1]$ is a measure of the weight worker i places on the earnings of worker k in her team, p is the group incentive, h_i and h_k are the respective productivities of each worker, e_i and e_k are workers' efforts, and s_{ik} is a productivity-invariant bonus. For worker i in a given team, the optimal level of effort is given by

$$e_i^* = \frac{p(1+\pi_{ik})}{2}\theta_i.$$

Here, $\pi_{ik} \in [0, 1]$ represents the weight that worker i places on the earnings of worker k in the same team, effectively amplifying the incentive power p as workers internalize the impact of their effort on their teammates' earnings. If $\pi = 0$, free riding is optimal, while if $\pi = 1$, the optimal level of effort is equivalent to that in a single-agent setting.

In their experiment, Bandiera *et al.* organized workers into teams of five. Workers were able to change teams once a week, enabling the authors to study team composition changes under different incentives. In the control group, workers were paid piece rates based on aggregate productivity. In the first treatment, team ranks—each team's absolute productivity level— were displayed on a daily basis. In the second treatment, tournaments were introduced, with a monetary bonus awarded to the most productive team each week.

The introduction of rank displays reduced average team productivity by 14 percent, while the introduction of tournament incentives increased productivity by 24 percent. Both treatments significantly altered team composition compared to the control group, primarily due to sorting on ability. In the rank display treatment, workers sorted more by ability rather than friendship, which weakened social incentives in average-performing teams and led to lower productivity. In the control group, friendship ties helped even mid-level teams internalize free-riding spillovers, maintaining effort levels. However, once rank information was introduced, these ties broke down, reducing cooperation. In contrast, the tournament regime provided incentives strong enough to elicit greater effort, outweighing any increase in free riding within teams and leading to a net productivity gain.

Friebel *et al.* (2017) explore the effectiveness of team-based bonuses when the firm, rather than workers, determines team composition. In a retail bakery chain, each shop employed an average team of seven workers responsible for a variety of interconnected tasks, including handling goods, operating the oven, and serving customers. In control stores, workers received fixed wages based on job tenure. In treatment stores, the firm implemented a shop-based bonus for exceeding predefined sales targets. Team bonuses were found to increase sales through increased effort serving the queue of incoming customers. The impact of the bonus on overall effort diminished as the proportion of non-incentivized team members increased, suggesting that the effect was driven by monetary incentives rather than peer pressure.

What is surprising about this literature is how such small incentives matter in practice. More empirical research is needed to understand how team incentives influence productivity, particularly in service industries and knowledge work, which account for a growing share of economic activity. However, measurement challenges in these fields make drawing clear inferences more difficult, highlighting the need for innovative research designs.

Incentives for innovation Motivating innovation or fostering creativity often requires long-term contracts that avoid reliance on short-term performance metrics. Emerging evidence suggests that innovators benefit from greater freedom to experiment and fail early in the process, as innovation typically involves exploring new approaches with a high likelihood of failure. Intense monitoring and pay linked to short-term progress may hinder exploration and risk taking. On the other hand, a fixed wage may provide little incentive to explore. Manso (2011) proposes that the optimal incentive scheme to motivate innovation should tolerate early failure while rewarding long-term success. Supporting this idea, Ederer & Manso (2012) find in a laboratory experiment that subjects are more likely to innovate under a failure-tolerant incentive scheme compared to fixed-wages or standard pay-for-performance contracts. In their experiment, subjects make decisions about the location of a lemonade stand, product selection, and pricing over 20 periods. Subjects observe profits and customer feedback after each period. In addition, they receive advice on a successful business strategy via a letter from the previous manager. The previous manager's strategy is profitable but not profit-maximizing. Participants thus face the choice between fine-tuning a profitable strategy or innovating. Compared to other treatments, subjects are more likely to discover the profit-maximizing strategy in the failure-tolerant condition, where they are paid for performance only in the last 10 periods of the experiment.

The importance of these contract features extends beyond the lab. Exploiting differences in funding streams in the academic life sciences, Azoulay *et al.* (2011) find that investigators produce high-impact articles at a much higher rate than a control group of similarly-accomplished scientists when the source of funding tolerates early failure and rewards long-term success. This study provides some of the strongest evidence on how incentives shape innovation in high-stakes settings. Understanding how funding agencies—whether through explicit or implicit mechanisms—design incentives for science and innovation is likely to have far-reaching implications.

The time required to innovate in failure-prone environments can also influence sorting into innovative sectors. For example, Venture Capital (VC) investments typically leave founders holding risky equity and earning low salaries in the early stages of a startup, with substantial financial rewards tied to achieving key innovation or product milestones. Ewens *et al.* (2023) exploit the introduction of cloud computing, which accelerated the resolution of uncertainty in some industries but not others. They find that when entrepreneurs can determine more quickly whether an innovation will succeed, highly skilled, highly-paid individuals are more likely to leave traditional jobs and enter entrepreneurship.

Some Intermediate Reflections The literature on performance pay and output-linked incentives is extensive, likely reflecting somewhat diffuse priors about which incentive mechanisms are most effective across different workplace settings. Given the rich space for responses and contracting, a framework for testing the efficacy of different incentives is likely needed, along with disciplining models based on manager's beliefs (an issue we return to at the end of this section). For testing between different contracts, Georgiadis & Powell (2022) consider how a manager can design experiments to improve upon an existing contract. In particular, their paper provides a (relatively technical) characterization of the economic primitives that can be recovered with A/B tests of incentive contracts. We would encourage readers to engage with this work to provide guidance about what experiments can reveal.

3.2 Pay levels, goal setting, career incentives

Firms that do not have objective, contractible performance measures often link pay to performance through subjective evaluation and career incentives. However, even firms that could implement output-based contracts often choose not to, suggesting that pay levels themselves can serve as an incentive device. Important work has been conducted on pay setting, subjective evaluations, and career incentives. In this subsection, we discuss models of how firms set pay and output standards. We then review evidence on how output standards serve as an incentive, and comment on constraints that firms may face when setting pay levels, including those stemming from perceptions of inequity across workers. We then examine subjective performance evaluations, exploring the tradeoffs between incentives and bias, before turning to career incentives, which raise broader questions about how firms assess worker performance and make promotion decisions.

3.2.1 Pay levels

In the prior subsection, the evidence suggests that highly-productive workers are the most sensitive to performance pay. How does this finding extend to workers' reactions to the level of base pay? Lazear *et al.* (2016) examine this issue in the context of fixed hourly pay and find that the least productive workers are the most responsive to changes in pay levels. In other words, raising pay is most likely to improve output and retention among workers who are marginal to the firm. They reach this conclusion by comparing the output of workers performing the same job across states with varying unemployment rates during the Great Recession. When unemployment is higher, workers have fewer outside job opportunities, making their current job relatively more attractive. The authors find that lower-productivity workers respond the most to changes in the local unemployment rate, increasing their effort as unemployment rises. As a result, average output per worker increases, with the largest gains occurring at the lower end of the productivity distribution.

Similarly, Coviello *et al.* (2022b) examine the effect of the minimum wage on worker effort across the distribution of worker productivity. Extending efficiency wage models, their theory allows the firm to offer both fixed and performance-linked pay components. Using a border discontinuity design, they find that increases in the minimum wage raise productivity, especially for lower-performing workers. In their model, increasing the minimum wage "has two opposite effects on incentives: it demotivates effort provision because it flattens the pay schedule (pay-for-performance channel), but it motivates effort provision because of the fear of losing a now higher-paying job (efficiency wage channel)." Workers who are likely bound by the minimum wage increase effort the most, with this efficiency wage effect outweighing the reduced incentive power of performance-based pay. The efficiency wage effect is strongest in settings with a high supervisor-to-worker ratio, suggesting that effective monitoring is needed for this mechanism to operate. However, while minimum wage increases lead to higher worker productivity and retention, the resulting rise in labor costs outweighs these gains to firms, ultimately reducing the firm's profitability. These studies appear to support a channel that is similar to those predicted by efficiency-wage theory, where workers will exert more effort to keep a good job, relative to their best likely outside option.[23]

[23] It is also possible that improving job quality or rewards can reduce stress, which has been shown to increase productivity in some settings. For example, Kaur *et al.* (2021) find that workers' financial concerns significantly impact their productivity.

Given employees' responses to the level of pay, firms may find it profitable to attempt to match incentives to workers who will likely be the most responsive. For example, Parsons & Van Wesep (2013), analyzes workers' self-control problems and suggest that firms can save on wage costs by timing pay to match different workers' consumption. However, firms may be constrained when adjusting pay levels differentially across workers. Dube *et al.* (2019) utilize quasi-experimental variation resulting from formula-based pay raises to investigate the impact of own and peer raises on firm separations. They find that peer wage changes significantly impact separations. These effects are primarily driven by increased turnover among focal workers when a coworker receives a higher pay raise. Similarly, Breza *et al.* (2018) utilize data from a month-long field experiment of manufacturing workers. They show that when productivity differences are difficult to observe, wage inequality between coworkers reduces output by 0.24 standard deviations and reduces attendance by 12%. However, when individual output is easily observable, pay inequality has no discernible effect on turnover or productivity, suggesting that workers are more accepting of within-firm pay disparities when they perceive them as justified by productivity differences.

3.2.2 Subjective evaluation, career concerns, and employer learning

In many settings, individual worker performance is difficult to quantify, requiring firms to use incentives that are not directly tied to measured output. This point is especially pronounced in jobs that involve complex, non-routine tasks and that require a high degree of creativity, interpersonal skills, or judgment (e.g., management, research and development, and professional services).[24] Even in cases where objective measures are available, firms may choose not to use them for incentives because of issues with multi-tasking, sabotage, or culture erosion. In this section, we examine research on alternatives to objective performance measures, highlighting key tradeoffs and emerging findings. We also explore potential directions for future research, particularly as text-as-data techniques become increasingly prevalent.

Subjective performance evaluations The introduction of incentives based on subjective evaluation may yield productivity-enhancing effects when starting from a baseline with limited performance-linked pay.[25] Two opposing forces—the introduction of bias and the ability to provide

[24]The relative prevalence of non-routine jobs is likely to expand in the long-run since tasks with objective output are more prone to automation.

[25]Relative to a baseline with performance-linked pay based on objective measures, however, subjective performance

performance-based incentives—compete to determine the net effects of subjective evaluation. The literature has identified two sources of bias: leniency bias, and differential treatment. With leniency bias, ratings are inflated for an entire group. In these cases, relative performance mechanisms like a forced curve can recover accurate rankings (Frankel, 2014),[26] but comparisons between raters or across different divisions may be challenging. Further, many firms avoid using relative mechanisms to address leniency, indicating that they value the ability to make absolute comparisons across individuals or teams. With differential treatment bias, workers within certain groups are less likely to receive high ratings, even when their performance is identical to that of others.

Frederiksen *et al.* (2020) benchmark the degree of leniency bias in subjective performance evaluations across supervisors. A key empirical challenge is that subjective evaluations are often employed precisely because individual performance measures are lacking, making it difficult to separate the effects of supervisor performance from the influence of leniency bias. To overcome this problem, they use data from a large Scandinavian services firm that includes branch-level performance measures; although objective individual performance measures cannot be observed, they can benchmark supervisors' aggregate subjective leniency relative to overall branch results. They find that a 1 standard deviation change in supervisor leniency leads to a 30% increase in a worker' subjective rating. Mapping ratings to earnings, this finding implies that moving from a supervisor at the 10th to the 90th percentile of the leniency distribution would increase the present discounted value of lifetime earnings at the firm by 6-12%. Differences across supervisors may reflect styles or other factors, but such a large magnitude of bias suggests that firms may want to use at least *some* performance metrics, even if they are imperfect.

Similarly, Li (2017) explores the tradeoff between leniency bias and expertise in NIH grant evaluations. She shows that the presence of intellectual proximity between evaluators and researchers increases the likelihood of being funded by 2.2%. Evaluators can assess intellectually proximate applicants with greater precision due to expertise, but specialization may also bias them against new paths of research, meaning they are less lenient with respect to novel ideas.

Future research will likely explore how text-processing tools and digital trace data can enable

evaluation leads to more compressed pay Macleod (2003).

[26] A tournament in which only some workers are rewarded is consistent with this mechanism. There are other issues with relative performance evaluation, including sabotage, that may prevent these mechanisms from being used in some some settings.

firms to gain deeper insights into worker performance. These advancements open new opportunities to better understand employee activities and to utilize this understanding to mitigate bias in subjective performance evaluations. For example, Impink *et al.* (2024) use email data to recover communication patterns, but more detailed data and topic modeling might allow firms to better observe worker activities when objective measures are sparse. The interaction between these tools, subjective ratings, the rationales for those ratings, and how they map to the activities or actions workers perform will likely be a key area of future research. However, the applicability of these new measures—and the conditions under which they should be implemented—will likely differ across various settings.

Shifting to the second type of bias, differential treatment, Benson *et al.* (2023) use data from a large retail chain to explore how subjective assessments of employee short-term performance and long-term potential contribute to gender gaps in promotion and pay. Despite receiving higher performance evaluations on current job tasks, women were 13% less likely to be promoted than men. Women at the firm received lower assessments of their future potential relative to their male counterparts. These lower potential ratings, however, did not predict future performance after promotion; in fact, the marginal promoted woman outperforms the marginal promoted man. Within the firm studied, the job hierarchy accounts for 70% of the gender wage gap, and bias in assessments of potential for women contributes substantially to this gap and raises important questions about how firms deploy ex-post incentives such as raises or promotions (Holmström, 1999; Gibbons & Waldman, 1999). Bias may undermine incentive effects in some cases, while other forms of bias (like handicapping in tournaments to give some workers an advantage) may induce heterogeneous workers to supply more effort. Bias can also arise via social interactions, for example, if managers are friends with or are dating subordinates (Macdonald *et al.*, 2023).

Tournaments, career incentives, and employer learning The literature on tournaments has been applied to a variety of settings, but one of its original motivations was to describe contests for promotions as a form of career incentive. Waldman (2013) discusses the evidence on promotion tournaments and highlights the perspective that promotions play a market-like role in leveling up wages. Ke *et al.* (2018) explore the constraints firms face when using promotion-based incentives, such as the limited number of available promotion slots. They demonstrate that there is a tradeoff between using promotions to motivate employees and maintaining firm-level efficiency. Other research has examined how variation in tournament composition, such as the presence of strong competitors in the field or in future rounds, affects behavioral responses (Brown, 2011; Brown & Minor, 2014). These studies suggest that the degree of heterogeneity within a workplace may significantly influence the effectiveness of tournament incentives. Further, some workers may be less likely to engage in tournament-competition in the first place, even holding fixed ability, as has been shown in the lab by Niederle & Vesterlund (2007).

Workers' responses to tournament or promotion incentives likely depend on their expectations and knowledge of future rewards. For example, using survey data from workers at an Asian bank, Cullen & Perez-Truglia (2022) show that many workers have limited information about their likely earnings upon promotion and, relatedly, under-estimate the earnings of their managers. When these workers learn that their managers earn more than they thought, their own career incentives appear to become more salient; they adjust their expectations about their own future earnings, and respond by working longer days and generating more revenue. In a related paper, Cullen & Perez-Truglia (2023b) explore why salary information does not diffuse without intervention, and suggest that high earners may fear resentment or increased competition if their earnings are revealed. As a result, lower-earning workers may not know about future earnings growth opportunities.

Perceptions of tournament institutions also likely matter for their efficacy. For example, Deserranno & León-Ciliotta (2025) conduct an RCT with healthcare workers in the Sierra Leone Ministry of Health to investigate whether more meritocratic promotion tournaments encourage greater effort than less meritocratic ones. They find that having a more meritocratic promotion tournament substantially enhances worker performance, and that there is a complementarity between having a more meritocratic tournament with the degree of perceived pay progression. Their results highlight the importance of how workers perceive tournaments, suggesting that these perceptions play a critical role in shaping their effort and performance. Working in a similar context in Sierra Leone, Deserranno *et al.* (2023) conduct an additional RCT, this time focused on the allocation of performance incentives between workers and supervisors. They find that sharing incentives equally between workers and supervisors increases performance more than providing incentives to either workers or supervisors separately.

On the firm's side, the effectiveness of various incentive schemes, subjective evaluations, and promotion structures ultimately hinges on the firm's ability to learn about its workers. This learning process is essential because it determines how well a firm can match incentives to worker types to optimize performance and outcomes, and is a key input into the decision rules about subjective evaluation and promotion (Kahn, 2013). Kahn & Lange (2014) analyze whether learning about workers' types occurs due to (1) employer learning about fixed employee ability over time, or (2) worker productivity that evolves heterogeneously over different growth paths. Using a 20-year panel of repeated pay and performance measures within a large firm, they find that wage dynamics reflect both kinds of employer learning. Specifically, they find that "dispersion of pay increases with experience primarily because productivity differences increase. Imperfect learning, however, means that wages differ significantly from individual productivity all along the life cycle because firms continuously struggle to learn about a moving target in worker productivity." Since employer learning takes time, incumbent firms may have an advantage in identifying worker types, consistent with traditional learning models. Because of heterogeneous growth, promotion rules, and how to use promotion incentives, can become extremely complicated.

Further, even when firms have access to objective performance measures, they may not use these measures in an optimal way. For example, Benson *et al.* (2019) show that, across sales firms, salespeople tend to get promoted based on objective measures of their prior performance. However, prior performance has little predictive value for managerial success. It is unclear whether firms recognize this mismatch and believe that the incentive value of promotion is so significant that they can tolerate the misallocation, or if they are simply making mistakes in their decision-making. We know of few other studies on how firms learn about and utilize information on worker performance for incentive provision and career planning, but we believe that further research in this area would be highly valuable.

3.3 External markets and firms' incentive responses

Given the extensive literature in labor economics on how incentives interact with external labor market conditions, personnel economists' expertise in modeling the effects of incentive contracts can help refine our understanding of monopsony and labor market power. At the same time, incorporating broader labor market conditions into personnel contracting models could enhance their realism and applicability. For example, beyond incentive contracts, incorporating labor market features into research on incentives can shed light on how firms invest in non-wage aspects of employment relationships, such as training, workplace amenities, and other benefits. These investments may be difficult for workers to finance individually but can play a crucial role in attracting and retaining talent.

Imperfectly competitive labor markets are increasingly important for labor economists, and will likely play a growing role in future personnel economics studies. Assuming perfect competition in the labor market implies that the firm faces an infinitely elastic labor supply curve. Firms are therefore unable to deviate from the market wage. In reality, empirical studies suggest that firms do not face perfectly elastic labor supply curves. In workhorse models with homogeneous workers and no pay-incentive effects, this limited elasticity implies that firms mark down wages relative to workers' marginal products of labor.

Manning (2021) identifies two key frictions that give firms labor market power: job search costs and job differentiation. Search costs arise because finding a new job takes time and effort, and temporary unemployment can be costly for workers. Anticipating these barriers, firms take advantage of workers' limited mobility and offer lower wages than in a perfectly competitive market. Job differentiation further strengthens firms' market power, as workers do not view all jobs as identical. When firm-specific attributes, such as location or workplace culture, align with worker preferences, firms can offer lower wages without losing employees, as switching jobs would require workers to sacrifice valued non-wage benefits. Accounting for these frictions and the imperfect competition in labor markets is likely to enhance the depth and predictive power of research in personnel economics. We see three primary areas of analysis where incorporating these factors will likely prove fruitful: firms' labor market power, incentive provision, and performance heterogeneity.

First, studies of labor market power often use models with fixed wages and limited incentive effects. How does building workers' responsiveness to incentives into these models alter conclusions? Emanuel & Harrington (2020) begin to answer this question through the lens of an efficiency wage model. They extend the standard monopsony framework with fixed wage posting to include incentive effects of higher wages. In the standard model, wage setting depends only on the elasticity of labor supply. Equilibrium wages are given by $w = Y'(N)(\frac{2\epsilon_{departures,w}}{1+2\epsilon_{departures,w}})$, where $Y'(N)$ is the marginal product of adding an additional worker and $2\epsilon_{departures,w}$ is the elasticity of labor supply to the firm under the assumption that the recruitment and departure elasticities are symmetric. To introduce incentive effects, Emanuel and Harrington allow worker productivity (either through effort or sorting) to depend on wages, capturing this relationship using the function $e(w)$. When the elasticity of quality-adjusted effort to wages, $\epsilon_{e,w}$, is non-zero, then optimal wage setting (when $\epsilon_{e,w} \leqslant 1$) gives

$$w = Y'(Ne)e(w)(\frac{2\epsilon_{departures,w} + \epsilon_{e,w}}{1 + 2\epsilon_{departures,w}}).$$

This expression is a function of the productivity level for a wage w and the sensitivity of productivity to a change in wages. When $\epsilon_{e,w} > 1$, additional wages exceed the marginal product of the next worker among applicants who could be hired. When the productivity elasticity is sufficiently large, the firm stops hiring and sets wages such that there are no markdowns relative to the last worker's marginal product. By examining this case, it is clear that a higher wage-productivity response will limit firms' markdowns.

Emanuel and Harrington test the theory using data on the implementation of a voluntary minimum wage. Building in a realistic personnel-economics view of the relationship between wages and effort, they show that, based on the elasticity estimates, the optimal wage is very close to having no markdown. That is, forces in personnel economics, like incentive responses, need to be accounted for in inferences around labor market power and firms' wage policies. These issues are not new; Kuhn (2004) states, for example, that "heterogeneity in workers' abilities does not play enough of a role in existing models of search or monopsony." Kuhn argues that attention to workers' heterogeneous abilities, and how these differences in productivity affect the way that workers and firms compete in the market, will likely be crucial to building more realistic models of monopsony and for inferring market power.

Whether firms understand these issues and set pay accordingly is an open question. The closest evidence suggests that firms are especially sensitive to benchmarking information, which may break the link between these personnel economics forces and actual pay setting. Exploiting variation in firms' access to salary benchmarking information, Cullen *et al.* (2022) find large and significant salary compression by skill level in the presence of a benchmark, with the strongest effects among low-skill positions. Specifically, they find a 40% decline in dispersion around the median for low-skill workers due to salary benchmarking, indicating extraordinary responsiveness to providing information about market-level compensation. In practice, firms may also set wages centrally. For example,Hazell *et al.*

(2022) find that 40 to 50 percent of firms set wages nationally, so that wages within a firm for a particular job are fixed across locations.

In alternative models, efficiency wage incentives arise from labor market thinness, meaning that as it becomes harder for workers to find comparable jobs, they have stronger incentives to exert effort to retain their current position. Studies that use concentration or labor market slack to infer wage-setting power thus likely need to grapple with the competing effects of efficiency wage incentives and search frictions, both of which limit the labor supply elasticity but with opposite markdown effects. If firms anticipate efficiency wage responses, it is possible that markdowns are overstated relative to the implied labor supply elasticity facing the firm. Interestingly, however, Emanuel and Harrington show that the way their firm sets wages does not fully account for the efficiency wage channel. More evidence is needed on how firms actually take into account the different forces when setting wages, a topic we return to in Section 3.5.[27]

Second, personnel economists may wish to explore how incentive provision depends on labor market structure. While many compensation contracting models implicitly incorporate some degree of firm market power through a worker's participation constraint, few models or empirical studies explicitly analyze how labor market competition influences the design of equilibrium incentive contracts.[28] An exception is Bénabou & Tirole (2016), whose model predicts that performance pay is more prevalent in competitive labor markets. In their model, workers are classified into two types—high and low productivity—and exert effort across two dimensions: one that is observed and contractible and another that is unobserved and uncompensated. Firms compete for workers by offering two contracts, each designed to attract one of the two worker types. Each contract consists of both fixed and variable pay components. The core prediction for personnel economists is that labor market competition increases the rate of performance pay in both contract types across the ability distribution, as offering higher bonus rates allows competing employers to attract high-productivity workers. By contrast, a monopsonistic firm does not need to compete for high-productivity workers and can instead offer a lower fixed wage to extract rents from low types. The setup with different

[27]Note that the firm's voluntary implementation of a minimum wage may have different implications than the findings on market-level minimum wages that have been studied by Coviello *et al.* (2022b) and Ku (2022).

[28]A related literature examines the effect of product market competition on firm pay structures (Cuñat & Guadalupe, 2005; Karuna, 2007; Raith, 2003). To the extent that product market competition and labor market competition overlap, most of these studies suggest that competition increases the prevalence of performance pay. However, excessive competition may have the opposite effect. Khashabi *et al.* (2021) find that in a bakery chain, performance pay has the strongest effect on sales under moderate competition but becomes detrimental under intense competition.

worker types in Bénabou & Tirole (2016) raises the question of how performance heterogeneity within a particular job affects firms' markdowns and market power. At the time of our writing, we are aware of little empirical work testing these predictions, but we would be excited to see future work on the relationship between market structure and the form of incentive contracts.

A related question is how firms balance external hiring with internal labor markets and how this varies depending on firm's external positioning in the productivity distribution. Friedrich (2016) makes some progress by examining how firms decide whether to fill managerial positions through internal promotions or external recruitment. In his model, a firm's position in the productivity distribution influences its available talent pool, while asymmetric employer learning about internal candidates' ability and firm-specific human capital accumulation makes external hiring riskier. To mitigate this risk, some firms invest in hiring young talent and developing internal promotion pipelines. Using evidence from Denmark, he finds that highly productive firms competing for elite graduates rely more on internal promotion, whereas lower-productivity firms are more likely to engage in market-based hiring for managers.

Finally, demographic issues will likely change how firms interact with external labor markets. A slot-based view, that has become popular among personnel economists, helps to illustrate the impact of economy-wide demographic changes on promotion incentives. Workers and workforces are getting older, potentially leading to congestion in firms. To the extent that firms use promotion incentives, an older workforce that is slow to move or retire may impede younger workers from moving through an organization's ranks. While this has implications for across-cohort human capital and income inequality, slot-based models also have implications for firms' incentive provision. In a study of an Italian pension reform that delayed retirements, Bianchi *et al.* (2023) show that delaying retirements reduces career progressions for young workers. While models with promotion incentives suggest that firms might substitute higher pay as the ability to promote falls, the general equilibrium effect of congestion appears to dominate, lowering young workers' wages. Across countries, relative wages for young workers facing an older age structure have tended to fall (Bianchi & Paradisi, 2022). This evidence points to the importance of slots, promotion, and career progression as spilling over from firms to the general labor market. Training incentives and how firms manage – especially in settings with rigid labor markets – will be an especially fruitful topic as the many economies face older demographic structures.

3.4 Core takeaways

- Performance-linked incentives affect effort in non-routine jobs (e.g., teaching, medicine, law), suggesting that prior work on the effects of performance pay in relatively routine jobs is likely generalizable to other settings.

- High-performing workers appear most sensitive to performance pay on the retention and effort margins.

- Group-based incentive pay often yields effort increases, even though theory predicts that free-riding should weaken its impact. The effectiveness of group incentives may depend on team composition, coworker dynamics, and social connections among employees.

- The level of pay serves as an incentive when firms can effectively monitor performance and set goals.[29] Lower-performing workers appear most responsive to variation in effective pay levels.

- Personnel economics forces – like incentive responses and worker heterogeneity – can alter inferences about firms' labor market power.

3.5 Open questions about incentives

While significant progress has been made since the previous handbook chapter, there remain many promising avenues for future research. We would be very interested in surveys examining how firms set compensation and incentives in practice. While corporate finance has a rich literature on how CFOs make decisions and use models (Graham & Harvey, 2001; Caldwell *et al.*, 2024), similar large-scale investigations into how managers approach incentive provision remain scarce. Indirect evidence, however, suggests that human resource managers set pay using heuristics like salary benchmarks. To our knowledge, few studies incorporate manager beliefs into understanding how firms set incentives. Survey-based studies that explore managers' beliefs about the effects of incentives may provide valuable context for understanding how these incentives are provided. We suggest that researchers can make important leaps forward by understanding what practitioners—and managers on the ground—believe about the role of incentives and the benefits and limitations of using different incentive schemes.

[29]For example, De Ree *et al.* (2018) find that doubling salaries in Indonesian schools had no effect on student achievement, arguably because there was no increase in the expected performance standard.

The question of how firms can personalize incentives—through gamification or individual targets—and whether they should do so has received some attention but deserves more. Further, the literature is only now beginning to recognize the importance of analyzing how firms set and communicate goals related to incentive provision. This area of research will likely benefit from leveraging evolving text-as-data techniques and digital records. These techniques will also likely be valuable in helping future researchers understand how relationships, and relational incentives, work in practice (see Gibbons (1998) for background).

Finally, an important but underexplored question in the study of incentives is how higher-powered incentives affect the pace of workers' human capital investment on the job. Do certain types of incentives—such as relative performance-based pay—hinder learning-by-doing by encouraging behaviors like sabotage? A well-known literature has examined the returns to general experience and firm-specific tenure, but relatively little work has considered how different incentive schemes influence the rate of learning.

4 Hiring

Hiring is the process by which workers and firms match with one another. As Oyer & Schaefer (2011) noted in the previous handbook chapter, hiring had received relatively little attention from personnel economists at the time, particularly in comparison to research on incentives. Since then, substantial progress has been made in the study of hiring within personnel economics. In the spirit of exploration, we believe that there are several angles by which it is useful to better understand hiring. These include:

1. How and why do firms use networks and informal methods in hiring?
2. What recruiting practices and hiring decision rules do firms use? Why? What are the tradeoffs?
3. How do hiring methods affect disadvantaged workers?
4. How do workers decide on firms and what do workers know about firms?

Starting with our initial Equation (1), personnel economists' interest in hiring can be broadly framed as studying the determinants of h_{ij}, and the effects of firm policies designed to increase h_{ij}.

4.1 Networks and information

While workers may find jobs by reading a newspaper or searching the internet (Kuhn & Mansour, 2014; Kuhn & Shen, 2013), a large share of jobs are found informally via referrals. In his pioneering study on job-finding in Massachusetts, Granovetter (1973) documented that over half of jobs are found through referrals. A large literature in sociology and economics has since documented the prevalence of referrals in the job-finding process more broadly.[30] Historically, much of this literature has examined hiring from the perspective of employees. For example, Bayer *et al.* (2008) find that individuals are more likely to find jobs alongside others from their own census block rather than from nearby blocks, which they interpret as evidence of word-of-mouth hiring.

Studies have increasingly examined referrals from firms' perspectives. A central question is whether firms benefit from the use of informal networks in hiring. The majority of these papers work with observational data, either from firms' personnel data (Fernandez & Weinberg, 1997; Fernandez *et al.*, 2000; Castilla, 2005; Burks *et al.*, 2015; Brown *et al.*, 2016) or from administrative datasets (Hensvik & Skans, 2016). Across a number of papers, two consistent findings papers are that referred applicants are substantially more likely to be hired than non-referred applicants, and that referred workers have lower turnover than non-referred workers (Hoffman, 2017).

Whether the observational patterns are causal effects of referrals remains unclear. For example, it is possible that the observed data are the result of firms' setting different hiring standards for non-referrals or tolerating lower match quality for non-referrals because of unobserved availability of referred candidates. As a result, observational studies may fail to provide guidance to firms on how to handle referred candidates. Suppose a firm has one referred and one non-referred candidate for a job, and they appear observationally similar. Is it better to hire the referred candidate?

Pallais & Sands (2016) make progress on this issue by running their own firm and hiring oDesk contractors, allowing them to test the effects of referrals directly. Their methodological contribution is addressing the endogeneity of the hiring process by hiring all referred and non-referred applicants. The authors run two RCTs. In the first, they vary whether referrers receive information on the performance of referrals, providing a test of "peer influence," as well as whether referrers work with referrals, thus providing a test of the importance of referrals for team production. In the second

[30] This literature is summarized in Topa (2019).

RCT, the authors hire for a new job four months after the first RCT, making offers to all referrals and non-referrals from the peer influence RCT.

Starting with the first RCT, Pallais & Sands (2016) find limited evidence for peer influence, but some evidence for team production. In other words, referrals are more productive when they work with their referrers. Their most striking evidence comes from the second RCT. When workers are hired four months after the first experiment, referred workers outperform non-referred workers across a range of metrics. This cannot be due to referrals being better matched for the job as the referrals were not referred for the job in question. Rather, the evidence suggests that referrals are of higher ability overall. Overall, this paper illuminates the mechanisms by which referrals and non-referrals differ. Understanding why referrals and non-referral differ remains an important question, and the answer may differ across settings.

Employee referral programs Most work by economists on referrals focuses on whether or not an employee is referred. However, firms do not directly control whether or not someone is referred. Instead, firms decide about employee referral programs, a management practice where workers are explicitly encouraged to refer contacts for jobs. Despite the voluminous literature on referrals, literature on the impact of how practices interact with referrals is scant despite the fact that 69% of firms on CareerBuilder reported using employee referral programs (Friebel *et al.*, 2023).

Friebel *et al.* (2023) examine the impact of employee referral programs by working with a large Eastern European grocery chain. The firm's 238 stores were randomized into one of five RCT arms: a control group and four treatment groups. In the control arm, workers were allowed to make referrals, but there was no explicit encouragement. In the information-only treatment, workers were encouraged to make referrals through posters and letters, but there was no financial incentive. In the three financial incentive arms, workers received bonuses of varying amounts for making a referral.

The financial incentive treatments increased referrals, and, consistent with the observational literature, referrals had better retention than non-referrals. However, there was also a quality-quantity tradeoff: as the financial incentive increased, the relative difference between referrals and non-referrals decreased. The authors also examine the indirect effect of referral programs. They find that simply introducing a referral program reduced employee turnover by approximately 15%, regardless of the number of referrals generated. The effects on attrition are strongest among workers hired before the RCT and remain substantial even in treatment stores where no referrals were made. Surveys and other evidence suggest that the main mechanism for these effects is that workers value being involved in the hiring process, and perceive the referral program as a signal that the firm trusts its workers or values employees' voice.[31]

Recommendations The procedures by which referrals occur are important and differ across circumstances. Economists and sociologists often focus on "word of mouth" hiring, where an employee communicates with their employer or a social contact about a job opportunity. However, information about candidates is often passed along using recommendation letters. A feature of some recommendation letters is that they are physical and potentially re-usable, and do not depend on people being able to have face-to-face conversations.

Recent work has examined the impact of recommendation letters, especially their impacts on disadvantaged workers who may be less likely to be friends with people in high-wage jobs. Abel *et al.* (2020) examine the value of reference letters for disadvantaged workers in South Africa, a setting with high unemployment. Using an audit study, they first show that reference letters increase callbacks. Second, using an experiment that encouraged referrals, the authors find that providing reference letters increases job-finding outcomes. Heller & Kessler (2021) focus on reference letters in New York City. The sample is from participants in a summer jobs program. The authors use feedback from workers' supervisors to create reference letters for workers, and randomly provide letters to some of the workers. They find that having access to such a letter improves both employment prospects and earnings in future years. Together, Abel *et al.* (2020) and Heller & Kessler (2021) show that encouraging reference letters for disadvantaged workers could significantly improve their labor market outcomes.

Lazear's (1998) point about the option value of hiring risky workers raises questions about why firms need recommendations. There are two likely reasons. First, entry-level or summer jobs are likely to be short-term, and firms must balance whether to pay screening costs to uncover workers' quality when the benefits may only be realized for a short period of time. This point is consistent with Tervio's (2009) model showing that there are inefficiencies in talent discovery when firms do

[31] There is a growing body of research on the importance of employee voice in various organizational settings (Adhvaryu *et al.*, 2021a,b).

not get long-term benefits from investing in uncovering information about workers. Second, it is possible that information from external sources is higher quality or more reliable than what firms would generate through their own procedures, especially if those sources are closer to groups of workers that may be different than an employer's typical hires.

Communication between referrers and referrals Another direction is to better understand the nature of communication between referrer and referrals. Barwick *et al.* (2023) explore referral communication patterns using the universe of cellphone records from a Chinese city. Using worker movements to a preexisting friend's workplace as a proxy for referrals, the authors employ event studies to analyze communication around job changes. They find that communication with referrer friends increases substantially around a job change, whereas communication with non-referrer friends does not.

Reputation systems While recommendation letters and referrals communicate private information, some labor market histories (e.g., feedback scores in online labor markets, certifications, or disciplinary records in many professions) are public and may affect the attractiveness of hiring a worker. Public signals can be a benefit for workers' careers, but may serve as a disincentive for firms. For example, Pallais (2014) examines a model similar to Tervio (2009) and shows that firms are reluctant to hire new workers because they cannot internalize the value of the signal that hiring provides about a worker. She runs an experiment where she endows treated workers in an online market with public feedback scores and shows that data on new workers is under-provided by private market participants relative to the planner's optimum. Stanton & Thomas (2016) show how networks, in the form of agencies, can overcome this friction. Joining an agency provides a public signal of connected workers' reputations, and agency formation occurs through network ties. In exchange for connection, an agency head receives a fraction of affiliated workers' earnings. The authors show that high-quality workers benefit from the signaling value of agencies early in their careers, but that independent workers who establish reputations on their own do eventually catch up. These features follow a similar logic to the information channels that make referrals and recommendation letters valuable. However, the public nature of feedback may introduce certain drawbacks. For example, Filippas *et al.* (2022) show that pressure in relationships has led to reputation inflation online, degrading the information content of public feedback mechanisms. As a result, many platforms have begun to use aggregated private ratings to augment the information content of public feedback.

We believe there is substantially more that can be learned about hiring networks and how firms can leverage information about potential employees. For example, how do firms use social media or other information about workers' outside activities when making hiring decisions? Does having access to network data through LinkedIn substitute for referrals, or does the private information content of a referral trump potentially-inflated public praise? Other topics for further research include the broader effects of these policies. For example, how do employee referral programs affect worker diversity? For what types of jobs are employee referral programs most effective? What types of workers most value being involved in hiring?

4.2 How do technology and other procedures influence hiring?

Aside from referrals, firms use other mechanisms to uncover information about job candidates. To reduce the costs associated with finding this information, firms may leverage technology and develop procedures to improve the selection process.

Technology One important and topical approach is to better use technology. Industrial and organizational psychologists have a rich history of researching the development of selection tests, as well as a long body of work on best practices in hiring. However, work in economics has been more limited.

Autor & Scarborough (2008) study hiring outcomes in the context of a retail firm that gradually implemented job testing across its stores. Autor & Scarborough (2008) show that testing increases employee retention, which is an important outcome in their setting. However, testing does not seem to cause racially disparate outcomes in terms of the test disfavoring Black or Hispanic candidates. In addition, the retention benefits of testing are broad across racial groups.

Since Autor & Scarborough (2008), there has been rising interest in applying machine learning and other algorithms to the hiring process. This is true not only in the academic literature (Hoffman *et al.*, 2018; Cowgill, 2019; Li *et al.*, 2021) but also in the context of business providers, where companies increasingly seek to employ machine learning in hiring domains.

More recently, Hoffman *et al.* (2018) consider the impact of hiring algorithms in 15 firms employing low-skill workers, where job testing is gradually introduced. Consistent with Autor & Scarborough (2008), the authors document that introducing job tests improves worker retention. However, their main question is different: how much discretion should managers have to overrule the algorithms? Hoffman *et al.* (2018) show substantial variation in how often HR managers make hiring decisions in line with the test; some overrule test recommendations frequently, whereas others do so much less. Overall, Hoffman *et al.* (2018) find that managers who more frequently overrule the test tend to get worse hires. The results suggest that HR managers overrule job tests not only because they have private information, but also because of bias.

Could managers be more likely to override test results when they share a common trait or background with a candidate? In a rich ethnography covering the on-campus hiring procedure at an elite US university, Rivera (2015) provides evidence that factors like shared leisure interests appear to play an important role in hiring. It is an open question whether using such factors improves the precision of other parts of an evaluation or contributes to bias, favoritism, or unhelpful homophily. However, to the extent that these are the biases that drive the poor performance of managers who override recommendations in Hoffman *et al.* (2018), an important open question is how firms should provide incentives to hiring managers to limit the extent of their own biases.

A more recent analysis of hiring algorithms is provided by Li *et al.* (2021), who focus on the impact of algorithms on diversity. They point out that algorithms have two main uses. The first is exploration, which can be thought of as selecting from less well-represented groups to learn about quality. The second is exploitation, which can be thought of as selecting from groups with proven track records. Li *et al.* (2021) argue that modern hiring algorithms focus on exploitation instead of exploration. They estimate that by orienting hiring algorithms to focus more on exploration, firms can achieve both more diversity and better hiring rate performance.

Hiring procedures The field of industrial/organizational psychology is deeply concerned with optimal procedures for worker selection. For example, there is a rich literature on the optimal way to conduct interviews, often arguing for the importance of structured interviews instead of open-ended interviews. Such topics can help firms improve hiring. They can also point to broader issues in decision-making, such as the tendency to focus on extraneous or productivity-irrelevant characteristics. So far, there has been little research by economists aiming to understand the value of these different hiring procedures.

An important issue in hiring is to what extent decisions should be made by local business unit managers instead of HR. Wu & Liu (2022) analyze this issue using an RCT with a large Chinese retail firm. Half of stores are randomized to have local business unit managers make decisions instead of central HR. Wu & Liu (2022) find that the treatment increases productivity, both by getting better new hires and through positive spillovers of these new hires onto other workers.

Turning from authority to bias, a key mechanism for bias is the subjectivity inherent in some aspects of the hiring process. For example, in a study of students graduating from elite Indian universities, Shukla (2024) finds that low-caste students are more systematically disadvantaged by personal interviews than by more structured stages of the hiring process. Similarly, Mocanu (2023) studies the impact of a public sector hiring reform in Brazil that increased the use of impartial hiring practices, focusing on its gender implications. She finds that the reform substantially increased the number of female hires by increasing women's application rates, evaluation scores, and probability of being hired conditional on applying.

Studying the same public sector hiring reform, Mocanu & Patacchini (2024) investigate how removing discretion in the hiring process affects the utilization of connections. They find that both the prevalence and usefulness of connections decline following the reform, leading to improvements in certain public sector performance measures. This finding contrasts with research in the private sector, where hiring through connections generally improves outcomes.Mocanu & Patacchini (2024) also show that this leads to improvements in some public sector performance measures. One possibility is that the value of using connections in hiring may vary between public and private sector positions. Another possibility is that referrals are less valuable in workplaces using promotion tournaments as workers face incentives not to refer people better than them.

A further issue in hiring processes concerns the nature of decision-makers. Economists have so far studied the demographic characteristics of decision-makers. For example, Benson *et al.* (2024) use data from a large retail firm to study how managers' identities affect discrimination in hiring. They find that retail managers are more likely to hire workers of the same race. However, they also show that productivity may be higher when workers and managers share the same race, making it unclear whether this hiring pattern reflects bias or whether managers simply have better information about racially similar candidates. Shifting from race to gender, Bagues & Esteve-Volart (2010) study hiring for the Corps of the Spanish Judiciary, leveraging the random assignment of candidates to committees with varying numbers of male and female evaluators. Interestingly, they find that candidates are less likely to be hired when they face more evaluators of the same gender. Given the topic's importance and policy relevance, further research is needed to better understand how the demographic characteristics of hiring decision-makers influence hiring outcomes.

Understanding the role of demographic characteristics of decision-makers is an important and policy-relevant research question, and we believe that further research is needed to address it. However, there are also a vast array of other research questions to be answered. For example, some hiring decisions are made by individuals and some by groups—how does the form of decision-maker affect hiring? Should hiring decisions be made one-at-a-time, or is it better to consider all candidates at once and choose among them? Using data on hiring in a large consulting company, Radbruch & Schiprowski (2024) show that people are evaluated more negatively when they are quasi-randomly assigned to be in an interview sequence with higher-quality candidates, especially when the immediately preceding candidate is stronger, suggesting that contrast effects are important in hiring. Another unanswered question is, what information should be made available to decision-makers during the hiring process? For example, should recruiters have access to a worker's past salary history (Agan *et al.*, 2021b; Cowgill *et al.*, 2024)?

A final issue on hiring procedures is the role of firm costs in considering applications.[32] In a simple view of hiring, receiving more applications is beneficial to firms because it provides them with a larger pool of candidates to select from. However, reviewing applications is costly. Using a demand model that accounts for which applicants a hiring manager considers, Stanton & Thomas (2024) estimate that firms using an online labor platform incur a cost of approximately $1.21 to evaluate each additional applicant. Similar search costs may deter hiring. Consistent with this view, Algan *et al.* (2023) consider an RCT in France where the government helped firms screen candidates, including screening out candidates who were not well qualified. Algan *et al.* (2023) find that the intervention substantially increased the number of vacancies firms post and the number of candidates hired These results point to the potential value of improving worker sorting ex-ante as a way to limit firms' wasted effort screening workers who are likely to be a poor fit. In fact, Horton

[32] Our discussion of hiring procedures is far from exhaustive, and there are many other questions and issues. For example, how does emphasizing the pro-sociality of positions affect worker sorting (Ashraf *et al.*, 2020)?

et al. (2021) show that including a simple message in a job post about the desired vertical quality of workers, it helps to segment markets and improve the efficiency of matching. It is an open question how other mechanisms, like instituting costly ordeals during the application process, may drive benefits to firms through better self-selection into applying.

4.3 How do hiring practices affect disadvantaged workers?

Economists are increasingly interested in understanding demand for hiring various types of disadvantaged workers, as well as the returns to doing so. Policymakers are interested in helping disabled workers and workers with a criminal record succeed in the labor market, but in order to craft effective policy, it is important to understand what firms think about such workers, as well as the returns firms receive from hiring them. In addressing these questions, personnel economists may find it natural to partner with coauthors outside of personnel economics, such as non-personnel labor economists, economists studying crime, and criminologists.

Lazear (1998) argues that firms should prefer to hire risky workers (i.e., workers with greater variance in productivity) compared to safe workers, as there is option value to learning about workers, assuming that low-productivity workers can be let go. While Lazear's argument is simple and intuitive, workers with criminal backgrounds and disabilities struggle in many labor markets.

Workers with a criminal background In a classic sociology paper using an in-person audit study where matched workers applied for entry-level jobs, Pager (2003) documents that US workers with a criminal record are substantially less likely to receive a call back. Improving the labor market outcomes of individuals with criminal records is a key policy issue in many countries, but it is particularly pressing in the US, where a sizable share of the workforce has a criminal record. The inability to find work has direct consequences for these individuals and their families, but it also generates broader externalities, as recidivism rates rise during periods of non-employment. A large literature in criminology and labor economics examines policies such as education and training programs aimed at building human capital for individuals with criminal records, yet barriers to employment persist, leaving many of these workers struggling to secure jobs.

A popular policy initiative aimed at addressing this issue is "ban-the-box." Named after the checkbox on job applications requiring applicants to disclose a criminal record, ban-the-box policies restrict firms from inquiring about criminal history in the hiring process. Some versions prohibit employers from asking about records entirely, while others delay the question until later stages to prevent automatic disqualification. Unfortunately, the impact of these policies appears less than promising. Rose (2021) finds that a Seattle law preventing firms from examining criminal records until after an initial screening had little effect on job-finding. Moreover, using different approaches, Agan & Starr (2018) and Doleac & Hansen (2020) both show that ban-the-box laws have an unintended consequence of increasing statistical discrimination against Black men.

Beyond human capital augmentation and ban-the-box, an alternative approach—and a natural one for personnel economists—is to consider the role of firms. What role can policies targeting firms play in increasing demand for workers with a criminal record (Hunt *et al.*, 2018)? How do firms think about the productivity of workers with a criminal record? Recent studies by Cullen *et al.* (2023) and Bushway *et al.* (2023) shed light on these questions.

Cullen *et al.* (2023) collaborate with a large on-demand labor platform to understand firm demand for workers with a criminal background, as well as how various firm-focused policies may shape that demand. Prior to the experiment, workers with a recent criminal background were automatically excluded from being considered by firms on the platform. In the experiment, firms made hypothetical hiring decisions that affected whether individuals with a record would be able to accept jobs from these firms in the future. Bushway *et al.* (2023) conduct survey experiments with HR managers. The experiments include conjoint analysis, where managers make hypothetical hiring decisions between two workers. One characteristic varied is whether a worker has a criminal record.

Despite using different methods and samples, several findings emerge. First, providing relatively modest levels of insurance substantially increases interest in hiring workers with a criminal record. Second, wage subsidies also increase demand for workers with a record—however, in Cullen *et al.* (2023), wage subsidies appear considerably less cost effective than several non-wage policies. Beyond policy implications, the papers broadly support employers having concerns both about everyday productivity and left-tail risk (e.g., the probability of a bad event) from workers with a criminal record, but policies may go a long way in helping overcome these concerns. The concerns about performance occur even though there is evidence that workers with a criminal background are less likely to quit their jobs than workers without a criminal background (Lundquist *et al.*, 2018; Minor *et al.*, 2018).

Beyond risks to firm performance and firm reputation, the left-tail risk from hiring workers with a criminal record can also expose firms to legal liability related to the tort of negligent hiring. Pyle (2023) studies the adoption across US states of reforms to negligent hiring laws. Firms can be held liable for large judgments under negligent hiring laws when an employee with a criminal record commits a criminal act. Pyle (2023) finds that reforms to negligent hiring laws, which decrease firms' liability, increase the hiring of individuals with a criminal background and also decrease crime, reflecting less recidivism from more employment. Thus, negligent hiring laws, which seek to hold firms accountable for employee criminal behavior, can backfire and actually create more criminal behavior. The results on liability affecting hiring (i.e., selection into the firm) echo earlier results that liability costs affect selection out of the firm (Oyer & Schaefer, 2002).

Workers with a disability Individuals with disabilities also face significant challenges in the labor market; they are less likely to be employed, and tend to earn lower wages than other workers. Unlike workers with a criminal background, workers with disabilities may raise fewer left-tail risk concerns for firms, but productivity concerns—whether actual or perceived—may still be a barrier to hiring.

A major effort to improve the lives of disabled Americans was the passage of the Americans with Disabilities Act (ADA), which codified civil rights protections for individuals with disabilities, including the right to reasonable workplace accommodations. However, Acemoglu & Angrist (2001) find that the ADA actually decreased the employment and hiring of workers with disabilities, potentially reflecting employers' cost and litigation concerns. These results are broadly consistent with Pyle (2023), suggesting that increased employer liability tends to reduce hiring. Subsequent work similarly finds that anti-discrimination legislation fails to increase employment for workers with disabilities (Derbyshire *et al.*, 2024).

A policy alternative aimed at increasing employment opportunities for workers with disabilities is labor market quotas, which specify that workers with disabilities comprise a certain share of a firm's workforce, with the quota size varying across countries (Derbyshire *et al.*, 2024). Lalive *et al.* (2013) study the impact of a quota in Austria that applied to firms with 25 or more employees. Comparing affected to unaffected firms, they find that quotas significantly increased employment for disabled workers. More recently, Szerman (2022) studies quotas for disabled workers in Brazil, finding that quotas increase the hiring of workers with disabilities, though much of the increased hiring occurs for lower-paying, less-skilled jobs. Despite this, her estimates imply that the quotas generate aggregate welfare gains.

We believe that personnel economists interested in hiring can probe further into the hiring of workers with disabilities, an area that remains largely a black box. For example, what do firms believe about workers with disabilities? How do employers form and act on beliefs regarding different types of disabilities?

4.4 How do workers decide on firms and what do workers know about firms?

The discussion so far has focused on how firms select workers, but it is equally important to investigate how workers select and value firms. What role do workers play in the matching process? A large literature in labor and macroeconomics analyzes job search, for example, by estimating structural models of the search process, or by analyzing the impact of unemployment benefits. Other work evaluates how workers tradeoff different amenities when choosing between jobs (Mas & Pallais, 2017). But understanding what workers know about the firms they are considering may be important given recent research suggesting that workers have highly imperfect information about the labor market (Jäger *et al.*, 2022).

HR professionals and practitioners appear deeply attuned to the reputation of their firms. Firms are ranked in terms of best places to work, and many firms are concerned about how they are rated by employees on review websites like GlassDoor in the US or Kununu in Germany. To understand the impact of employer reputation on worker flows, Benson *et al.* (2020) create firms on Amazon Mechanical Turk—a website used for simple, short tasks, (e.g., identifying images for medical research)— and randomly endow the firms with different reputations. The authors find that reputation significantly affects job applications, with employers with good reputations receiving twice as much applications as those with bad reputations. Using an audit study, they also show that low-reputation firms are less likely to pay workers on time.

In more recent work, Bryan *et al.* (2023) analyze what workers know about a firm's business model and science quality, rather than just its overall reputation, and analyze how this knowledge affects job applications. The authors specifically focus on science-based startups, which may have fewer quality signals than large established firms (e.g., there may be no Glassdoor page to consult).

Partnering in an RCT with a leading science-based entrepreneurship program, Bryan *et al.* (2023) show that expert ratings of firms' science and business quality substantially affect which firms workers apply to, and that these effects are driven at least in part by workers' beliefs about firm outcomes. Using incentivized belief questions, they also show that workers exhibit substantial overconfidence about firms' right-tail events, highlighting a broader problem of informational deficits among firms.

The challenges that employees face in obtaining information about firms are explored in Sockin & Sojourner (2023), who analyze job ratings on Glassdoor. They find that negative reviews—the most valuable information for workers—are often under-supplied, reflecting concerns about retaliation.

Gee (2019) studies a field experiment on LinkedIn where some workers were able to view information about the number of people who had applied to a job. She finds that workers are more likely to complete a job application when they have access to the number of current applicants, and that the effects are stronger for women than men. Rather than analyzing the impact of additional information, Sockin *et al.* (2024) examine the factors that shape the information available to job seekers. They. They focus on non-disclosure agreements, showing that NDAs limit workers from willingness to share information about unlawful workplace conduct, thereby restricting the information that prospective employees have about the firm.

Flory *et al.* (2015) study how workplace competitiveness influences job entry decisions, focusing particularly on gender differences. The authors advertise for an entry-level position across several US cities, varying different features of the job ad. They find that women are significantly less likely to apply for positions where pay depends on relative performance. Using similar methods, Hedblom *et al.* (2019) and Leibbrandt & List (2018) examine the effects of including statements on social responsibility and equal opportunity in job ads. Hedblom *et al.* (2019) find significant positive benefits on applications received when firms advertise corporate social responsibility, while Leibbrandt & List (2018) find that equal opportunity statements have limited impacts on applicant diversity. Fuchs *et al.* (2024) study how emphasizing flexibility or career advancement affects the quality and quantity of applicants in a European technology firm. They find that highlighting job flexibility increases job applications from both men and women, while emphasizing career advancement only increases applications from men.

4.5 Open questions about hiring

Despite the surge in research on hiring by personnel economists in the last 10 years, important questions remain unanswered. We have outlined many of these questions throughout this section, but here are a few more.

While research has examined the impact of algorithms and AI on hiring outcomes, more descriptive work is needed to understand how firms actually implement these technologies in their hiring processes. How do workers and managers interact with algorithms and AI, and what factors affect the nature of their interaction? Anecdotal evidence suggests that many workers perceive algorithmic rating systems as potentially unfair. One question is whether this perception is driven by the way these algorithms are presented to workers.

We also believe that more research is needed on fairness and employer reputation in hiring. How do workers perceive the fairness of hiring processes, and does this perception matter for firms? Using an RCT, Bapna *et al.* (2021) show that the way firms communicate to candidates that they've been rejected affects whether those candidate apply in the future. How do other aspects of firm reputation—such as workplace culture, employee treatment, or transparency in hiring—affect worker flows and hiring outcomes?

Finally, we would like to see more research on how the returns to hiring practices vary by worker or position type. For example, are job tests more useful for high-skill workers compared to lower-skill workers? How does the return to employee referral programs vary based on the type of worker and the incentives those workers face? Economists are using new diagnostic tools to predict wages for different types of workers (Deming, 2021; Caplin *et al.*, 2023; Weidmann, 2024).

5 Managers, Peers, and Teams

At the time of the last handbook chapter, there was a large body of work on peer effects, including studies examining the impact of peers in the workplace. Since then, important new research has emerged, with a notable expansion in studies focusing on managers. In this section, we first examine the role of managers, and then discuss workplace peer effects.

5.1 Managers

Interest in managers stems from the recognition that management practices play a crucial role in explaining large productivity differences across firms (Syverson, 2011). While research on management practices has expanded, less attention had been given to the role of individual managers as of Oyer and Schaefer's (2011) handbook chapter, particularly below the CEO level (Benson & Shaw, 2024).

Since then, several key findings have emerged about the role of lower-level managers. First, the evidence indicates that lower-level managers significantly impact both their subordinates' productivity and the overall performance of their teams. Second, there is emerging interest in which traits of managers matter, but there is less consensus around this question as the set of potential traits and trait combinations is enormous. Third, we are still learning how and why managers matter, with research pointing in several different directions. It is natural that how and why managers matter could differ by setting, and further research can help to establish systematic patterns and how these patterns might vary in different contexts.

5.1.1 Do managers matter?

While economists and management scholars have long studied the importance of individual managers, earlier research primarily focused on CEOs and other top managers. In a seminal paper, Bertrand & Schoar (2003) utilize CEO movements between companies to estimate regressions of firm outcomes on CEO fixed effects, firm fixed effects, and controls. This approach allows them to assess the impact of individual CEOs on firm performance. Other designs use natural experiments to estimate the impact of individual managers. For example, researchers have examined CEO deaths and hospitalizations (Bennedsen *et al.*, 2020) and the forced removal of Jewish CEOs during the rise of the Nazis (Huber *et al.*, 2021) to estimate the effects of managerial leadership on firm performance.

Below the level of top managers, Lazear *et al.* (2015) examine the impact of frontline supervisors on service sector workers. Using data from a firm where employees provide customer support for technology products, they apply a methodology similar to Abowd *et al.* (1999), estimating productivity as a function of worker fixed effects, supervisor fixed effects, and control variables.[33] They show that differences between higher- and lower-performing frontline supervisors explain a substantial share of within-firm productivity dispersion. Replacing a 10th-percentile manager with a 90th-percentile manager has a greater impact on productivity than adding an average-productivity worker to a typical 9-person team.

Since Lazear *et al.* (2015), the importance of lower-level managers has been documented in a range of settings, including the public sector (Fenizia, 2022; Otero & Munoz, 2022), sales (Benson *et al.*, 2019), finance (Frederiksen *et al.*, 2020), retail (Metcalfe *et al.*, 2023), and manufacturing (Adhvaryu *et al.*, 2022, 2020). These studies have also significantly expand the empirical methodology for estimating manager effects through the use of event studies.

Another approach to estimating manager effects is through laboratory experiments, where workers and managers are randomly assigned to teams. In Weidmann *et al.* (2024), for example, a good manager is defined as someone who increases the productivity of the workers on their team. Their findings reveal substantial variation in managerial effectiveness, but it is not correlated with most observable characteristics. Interestingly, however, those who express a desire to be managers tend to be less effective, suggesting a negative correlation between managerial ambition and actual managerial ability.

It is also useful to understand the circumstances in which managers have little impact. Turning back to CEOs instead of lower-level managers, Janke *et al.* (2019) find that hospital CEO fixed effects are generally negligible in the context of the UK National Health Service. In certain public sector organizations, it may be hard for individual people to influence business practices and policies. Even in private organizations, it is possible that managers may have less of an impact in some settings. For example, in settings with high piece rates and relatively homogeneous work, managerial influence may be less variable, as strong incentives might substitute for leadership in driving worker output.[34] To the extent that frontline managers assist workers or solve complex problems (e.g., Garicano, 2000), their per-worker impact may be lower for more skilled employees who require less guidance. However, in such cases, managerial span (team size) is likely to increase, keeping the

[33]It has now become standard in the literature to use shrinkage procedures to account for sampling variation in estimated fixed effects. See the section in this handbook on empirical bayes methods for additional details.

[34]Note, however, that this conjecture is at odds with evidence in Benson *et al.* (2019) showing that managers matter substantially in sales firms with high performance pay.

manager's overall impact constant. Further research is needed to better understand the settings where managers have greater or lesser influence.

5.1.2 Which manager traits matter?

Given the evidence on the importance of lower-level managers, it is natural to ask which traits of managers matter the most for augmenting their subordinates' performance. While it might be difficult to observe a manager's relative productivity effect during a job interview or promotion discussion, it may be possible to observe and select for particular managerial traits.

A natural characteristic to consider is a manager's skill as an employee. Many firms promote heavily from within and they are often able to observe worker productivity when deciding who to promote to manager. One might assume that top-performing workers naturally become top-performing managers—anecdotally, star performers in some fields transition into successful leadership roles (e.g., Zinedine Zidane moving from star soccer player to coach). However, there is little empirical evidence supporting this intuition.

Instead, Benson *et al.* (2019) analyze the importance of managers' prior performance as employees in a large number of sales firms. They use data from a salesforce management software provider that tracks both individual sales productivity and managerial assignments. They find that more productive sales workers are much more likely to be promoted to manager roles than less productive ones. However, once promoted to manager, there is no positive correlation between a worker's prior performance and their effectiveness as a manager. They interpret these findings through the lens of a tournament model, where firms use the prospect of promotion to motivate salespeople to work hard, despite the fact that the skills required for sales success do not necessarily translate into effective sales management.

Hoffman & Tadelis (2021) analyze the importance of people management skills, defined as managers' social skills with respect to their workers. They exploit the fact that many firms conduct upward feedback surveys, where workers answer survey questions about their managers and their behavior toward workers. Using data from a large high-tech firm, they measure people management skills using responses on upward feedback surveys. They find that people management skills significantly reduce employee attrition, particularly attrition the firm seeks to prevent. However, these skills have little impact on most non-attrition outcomes for employees. For managers, people management skills are generally rewarded, and are associated with higher subjective performance scores, a higher chance of promotion, and larger salary increases. These results are broadly consistent with the importance of social skills in the labor market (Deming, 2017), and support findings that leadership skills are valuable to workers (Kuhn & Weinberger, 2005).

A related approach to identifying which manager traits matter is to measure managerial characteristics through surveys. Adhvaryu *et al.* (2023a) apply this approach in a large garment manufacturer in India, where frontline supervisors oversee production lines. Their main finding is that observable cognitive and non-cognitive skills do not reliably predict managerial effectiveness. However, traits related to managerial attention and controlare strong predictors of managerial skill.

While the papers discussed so far in this section tend to concern managers' interpersonal skills, other manager skills may be important as well. For example, Caplin *et al.* (2023) administer tests and show there is substantial heterogeneity in subjects' allocative skills (a measure of one's ability to assign workers based on comparative advantage to tasks). Workers with better allocative skills likely make better managers to the extent that a managers' role is realizing comparative advantage. Consistent with this view, allocative skills are a good predictor of earnings. Managers' background education and training also appears to be important for their success. Giorcelli (2023) studies business school training for middle managers using the Engineering, Science, and Management War Training during World War II. Entrance into the program was based on an exam with a strictly-enforced cutoff rule. A regression discontinuity design shows that receiving business school training had large positive impacts on managers' careers, making middle managers more likely to get promoted and to engage in entrepreneurship.

There is also interest in interactions between characteristics of workers and managers. Characteristics like race and gender have been extensively studied in other manager-like relationships, such as teacher-student and doctor-patient interactions. More recently, research has begun to examine their role in the manager-employee relationship as well. For example, Cullen & Perez-Truglia (2023a) and Fortin *et al.* (2022) examine the importance of gender concordance between managers and workers. Both papers find strong evidence that gender concordance is important, and that female employees fare better under female supervisors in terms of wages and speed of promotion.

Similarly, Delfino & Espinosa (2024) examine the impact of value dissonance between workers and managers. They analyze survey data from a large multinational bank, where employees were asked about the values they would want to nurture or promote in children. Their findings show that workers perform better when managed by a supervisor whose values align more closely with their own.

5.1.3 How do managers matter?

While there is strong evidence that lower-level managers have a significant impact, how they matter remains less clear. Lazear *et al.* (2015) suggest that managers influence workers primarily through motivation, as most manager effects disappear within six months of a worker receiving a new manager. However, because the decline in manager effects occurs gradually rather than immediately, it is possible that managers provide learning or guidance on factors with a high transitory component, requiring continuous reinforcement to sustain their impact. Using an RCT, Friebel *et al.* (2022) show that managers can matter by focusing on certain workplace issues, such as reducing employee turnover.

In the context of the Italian government, Fenizia (2022) shows that managers operate more on the selection margin than the motivation margin. In her setting, good managers encourage low performers to leave, thereby increasing overall performance. Comparing these findings, managers seem to play a different roles depending on the flexibility of the work environment. In settings with fewer rigidities, they primarily motivate and possibly teach. However, when barriers to firing are high, their ability to make jobs unpleasant for underperformers may become a necessary tool for selecting productive workers.

In addition to motivation and selection, there are many other ways that managers can matter, and there is emerging evidence for these other channels. Using data from a large European multinational, Minni (2023) shows that managers matter via their role in better allocating workers to jobs. Specifically, some managers are much more likely than others to reallocate workers within the firm using horizontal and vertical transfers.

Another way that managers matter is via their role in subjective performance reviews. Frederiksen *et al.* (2020) study manager effects in the context of a large Scandinavian firm in financial services. They estimate Abowd *et al.* (1999) fixed effects, but using subjective performance reviews as the left-hand side variable. They show that managers vary widely in their tendency to award higher subjective performance reviews, and that workers benefit career-wise from having a more lenient rater.

A further way that managers may matter is through what they know. Starting at least with Hayek (1945), there is a long tradition in economics considering the possibility that information in firms is dispersed and not fully accessible to everyone at the same time. Models that consider delegation to managers and their authority take this disparate information seriously (Prendergast, 2002; Dessein, 2002). For example, Friebel *et al.* (2024) show experimentally that managers are likely to have private information about local conditions. They analyze an RCT conducted in a large German bakery chain aimed at reducing workplace control. The firm had been using a large number of checklists in their retail stores, some which were quite detailed, time-consuming, and distasteful to workers (e.g., checklists related to whether the displayed doughnuts are at the correct angle and whether workers use certain phrases when interacting with customers). The RCT consisted of randomly removing two checklists that employees perceived as low-value. Prior to random assignment, for each store the regional manager made a prediction about whether the treatment would be effective. Friebel *et al.* (2024) find that the beneficial effects of checklist removal are highly concentrated in the stores where regional managers predict that the treatment will work, suggesting that these managers have substantial private information about treatment effect heterogeneity.

Models of hierarchies offer a different rationale for what managers do. In these models, managers are more knowledgeable than their subordinates, enabling them to solve exceptional problems, while lower-level workers focus on more routine tasks. The seminal models by Garicano (2000) and Garicano & Rossi-Hansberg (2006) imply that efficient organizations balance communication and helping costs with the costs of knowledge acquisition. Extensions by Caliendo & Rossi-Hansberg (2012) and Caliendo *et al.* (2020) detail how the number of managerial layers, and hence the degree of specialized knowledge held by each worker in a firm, varies with firm scale. Recent tests examining variation in communication costs between layers of an organization provide support for the core mechanism. For example, Gumpert *et al.* (2022) show that the introduction of high-speed rail that lowers communication costs in multi-establishment firms changes their organizational structure in a way consistent with the hierarchies model.

A final way that managers may matter is via their social interactions with subordinates. For example, Macdonald *et al.* (2023) study the phenomenon of managers dating their employees in

Finland. They show that worker wages increase after a worker starts cohabiting with their manager, but that worker wages decrease and turnover increases following a breakup or when the manager leaves the firm. This finding suggests that managers are biased toward workers they are dating, and that worker careers are potentially shaped by their social interactions with managers.

These studies illustrate a range of possible explanations for how managers matter. These explanations likely differ across settings. For example, in some settings, managers may matter most for their ability to push employees to "get stuff done." In others, managers may matter because of what they do behind the scenes (e.g., allocating resources) or because of what they know. How managers matter may also differ depending on the nature of their role and how high they are in the firm's hierarchy. It would be useful for future work to separate between different explanations, as well as to systematize how managers matter differently in different environments.

5.1.4 Additional questions about managers

The prior three questions—do managers matter, which traits matter, and how do they matter—are tightly logically connected, but there are many other questions related to managers that are interesting and important.

Do managers engage in talent hoarding? Managers play a key role in deciding who gets promoted. Indeed, helping identify talent is widely regarded as a core responsibility for managers at multiple levels of the hierarchy. However, if promoting someone means no longer being able to manage them, managers are faced with a dilemma: Is it worth promoting a highly capable employee if doing so means losing them from your team?

Talent hoarding is explored theoretically by Friebel & Raith (2022), who identify the problem and derive an optimal contract to use within the firm that properly encourages mobility. Empirically, talent hoarding is studied by Haegele (2022) who, using data from a large European multinational, shows that employees are substantially more likely to apply for promotions shortly around the time a manager departs. She argues that this and other evidence indicates that many managers suppress the mobility of employees within the firm.

What do people know or believe about managerial quality? Labor economists are increasingly interested in beliefs. For example, Jäger *et al.* (2022) and Bryan *et al.* (2023) examine worker beliefs, focusing on wages and firm quality, respectively, while Cullen *et al.* (2024) study firm beliefs about how pay affects job filling. Given the strong results that managers matter and vary widely in productivity, it would be interesting for future research to further explore beliefs about managerial effectiveness. How much do firms, workers, and managers believe that managers matter and, perhaps more interestingly, can they identify who the top managers are?

What can firms do to improve the quality of their managers? There is a longstanding debate about whether leaders are born or made, and it is natural to extend this discussion to managers. Is managerial quality fixed or malleable? If malleable, what can firms do to improve the quality of their managers? What can managers do themselves? In particular, we would like to see more work on the role of training and prior experience in shaping managers' styles and productivity. While evidence exists for CEOs, there is little systematic research on middle and lower-level managers. Given that top executives represent a highly selected group, it would be valuable to examine whether similar patterns hold—or diverge—for a broader, less selected sample of lower-level managers.

We would also like to see more work on the portability of managerial success across different firms. For example, Groysberg *et al.* (2008) show that star security analysts' performance is not fully portable across firms when they move, highlighting the importance of matching talent with context. Is the same true for managers? Does managerial success in one firm indicate a high likelihood of continued success in a different firm?

How are managers matched with workers or production units? One might assume that firms would seek to match their best managers with the strongest organizational units or workers. However, some research suggests that this may not always be the case in practice. In fact, several papers find evidence supportive of *negative* assortative matching, where better managers are matched with worse offices (Fenizia, 2022) or worse workers (Adhvaryu *et al.*, 2020; Metcalfe *et al.*, 2023). Given the importance of sorting and the pervasive notion of complements in production, we believe that it would be useful to better understand the pervasiveness and underlying causes of negative assortative matching, and under what conditions we would expect negative assortative matching to be optimal.

5.2 Peer Effects and Team Effectiveness

There has long been recognition that peers can influence work output, productivity, learning, and overall job satisfaction. Both lab and field studies have documented peer effects of similar size. Herbst & Mas (2015) conduct a meta-analysis of estimates of peer effects at work. Across 34 studies, they find that a one-unit increase in a co-worker's productivity leads to an average productivity increase of 0.12 for the focal worker. These large effect sizes may be surprising given studies that find limited evidence of peer effects in wage data (at least for Germany) (Cornelissen *et al.*, 2017). However, not all estimates reviewed by Herbst & Mas are positive, and the standard deviation in effect size estimates across studies is 0.16, suggesting substantial variation in different contexts.

In this section, we discuss evidence on two main channels by which peers may influence coworkers' productivity in organizations: knowledge spillovers and peer pressure.[35] What sets these peer effect studies apart from more general research on team production is that many focus on spillovers among workers performing autonomous tasks, rather than collaborative team output. Team production, which we discuss in Section 5.3, allows for more general production functions, where workers may perform different tasks and have the ability to trade or reallocate tasks among themselves. In this way, our discussion of team-based output and organization accounts for workers' multi-dimensional skills and different backgrounds and perspectives. As part of that discussion, we highlight a tradeoff between worker skill/background heterogeneity in teams and communication costs.

5.2.1 Knowledge spillovers

Beginning with Alfred Marshall, the literature on agglomeration has argued that proximity enables learning from others, increasing productivity (Glaeser, 1999). Empirical evidence supports Marshall's knowledge spillover rationale for the spatial clustering of economic activity. A worker who moves to a denser city can expect to have persistent wage growth (Glaeser & Maré, 2001), while an inventor located in a technology cluster that is 1% larger will produce about 0.07% more patents per year (Moretti, 2021).

However, compared to this broader literature on productivity, peer spillovers in wages within firms appear smaller, especially when considering estimates from knowledge-intensive occupations

[35]For an excellent overview of how social interactions with colleagues influence productivity, see Ashraf & Bandiera (2018).

(Waldinger, 2012; Cornelissen *et al.*, 2017). One possibility is that some degree of commonality in work tasks helps to facilitate knowledge transmission, making peer effects stronger in routine jobs where workers' tasks are more similar. Alternatively, peer effects in knowledge-intensive occupations may be offset by competitive dynamics that limit knowledge-sharing incentives.

Recent work has begun to unpack how exposure to higher quality peers in the workplace influences learning and knowledge exchange within the firm. Just as the research in urban economics has examined productivity and opportunity dispersion across localities, we believe this is an exciting area for exploring how personnel practices interact with intra-firm productivity dispersion.

Using data from the US Patent and Trademark Office (USPTO), Frakes & Wasserman (2021) examine how exposure to peers influences a focal patent examiner's patent granting activity. In their baseline specification, a one standard deviation increase in the grant rate of an examiner's peer group is associated with a 0.15 standard deviation increase in the examiner's own grant rate within the first two years of tenure. Several tests suggest that these effects stem from examiners learning to process applications by observing their peers. In line with this interpretation, the magnitude of spillovers is largest for impressionable, early-career examiners who are still learning on the job. While higher grant rates for patent examiners do not necessarily imply higher productivity, these estimates illustrate how spillovers occur. Relatedly, higher rates of remote work in the peer group reduce the magnitude of the spillover effect. The move to remote work also dampens the likelihood that a group of peer examiners cites common prior art during the evaluation process, suggesting that remote work reduces common knowledge that is shared when workers are co-located.

To disentangle how changes in exposure to different workers influence output, Sandvik *et al.* (2020) conducted a field experiment where sales workers were randomly paired together under different treatment conditions. In one condition, worker pairs met together and shared advice for handling different sales situations. In another, incentives were introduced that tied bonuses to pairs' joint sales growth. A final active treatment combined the meetings and incentives treatments. Sandvik *et al.* find that both the meetings and incentives treatments increased sales during the 4-week experiment. In the meetings treatment, below-median productivity agents who were paired with above-median productivity partners closed 82% of the initial productivity gap between the partners. The sales increase was approximately the same in the meetings treatment and the treatment that combined meetings and pair incentives, indicating that the firm did not have to directly compensate higher-performers for this spillover. Only treatments including meetings led to persistent sales gains after the experiment ended. Output gains in the incentive treatment reverted back to the pre-experimental level after joint incentives were removed. As part of the design, coworkers recorded what happened during their meetings on worksheets, allowing the authors to conclude that knowledge exchange was the likely channel for the observed gains. These results indicate that active interventions can alter intra-firm productivity distributions, but the need for the firm to intervene raises questions about why advice-seeking or knowledge flows did not occur without the firm's proactive involvement.

In a follow-up experiment that labeled the meetings treatment as a mentoring program for new hires, Sandvik *et al.* (Forthcoming) test whether a voluntary version or a mandatory version is more effective. They find that the lowest-productivity new hires are the least likely to participate in mentorship when the program is voluntary, yet these workers have the highest returns from participation. These results suggest a possible explanation for large productivity dispersion within some firms: lower-performing workers may be reluctant to seek help or to take advantage of learning or training opportunities.

5.2.2 What limits knowledge spillovers?

We survey evidence on three possible mechanisms that limit knowledge spillovers in organizations: advice-seeking stigma, limited observability and search frictions, and incentives to limit information flows.

Advice-seeking stigma Two recent papers support the notion that workers' inhibitions in asking for help restrict knowledge flows at work. Chandrasekhar *et al.* (2018) provide several mechanisms for why some individuals may fail to seek advice. They consider a model where individuals with uncertain ability can learn how to do a task from an advisor. If the demand for advice is negatively correlated with ability, then seeking advice may signal low ability. Even when ability is known, feelings of shame, rather than a signaling motive, may hinder advice seeking. The authors test these mechanisms in an experiment involving 1,200 villagers in rural India. The subjects had the opportunity to win a prize by guessing which of two boxes contained it. Participants received an initial endowment of clues that increased their chances of guessing the correct box. In the experiment, seekers—the main decision-makers—were paired with advisors, and had three days to obtain additional clues from advisors, beyond their initial endowment of clues. Seekers and advisors were paired based on previously-observed social ties.

The experiment had two main treatment arms. The first treatment, {*Random*, *Skill*}, varied whether the availability of information, measured by the number of clues provided, was positively correlated with the seeker's cognitive ability, proxied by their score on a Raven's Progressive Matrices test. Seekers in the *Skill* condition knew that the number of clues they were endowed with was positively related to their ability score, but did not know their actual score. The second treatment, {*Private*, *Revealed*}, varied whether the seeker's ability was disclosed to the advisor from the start. The idea was to see if participants in the *Skill*, *Private* condition would refrain from asking questions to avoid signaling low ability. The authors find a 55% decline in advice-seeking when needing to ask for information is negatively correlated with cognitive ability. Comparisons across treatment arms suggest that the signaling channel dominates the shame channel in most contexts, meaning participants were more concerned about revealing low ability to others than experiencing personal embarrassment from asking for help.

Relatedly, Heursen *et al.* (2023) study whether professionals strategically forego advice-seeking to appear more competent. The authors conducted an artefactual field experiment where 2,500 white-collar professionals were paid based on the number of multiple-choice test questions they answered correctly. Participants each answered 10 questions in isolation, providing a baseline measure of their ability. Afterward, they had the option to pay a small fee for computerized advice, which allowed them to revisit their answers after eliminating three-fourths of the incorrect options on the test. The experiment manipulated whether participants were aware that their decision to seek advice would be visible to a manager. Managers were tasked with estimating the worker's competence from a short biographical profile which, in some treatment arms, included whether or not the worker sought advice. About 70% of workers' final maximum total compensation was based on their manager's estimate of their competence.

Advice-seeking decreased by 16% when participants knew their choice was visible to a manager. Incentive-compatible elicitations suggest there is significant heterogeneity in the perceived reputation cost of seeking advice; while some participants expected substantial costs, others expected benefits. In reality, managers' competence estimates were not meaningfully impacted by whether or not the participant sought advice. These results suggest that misperceptions or biased beliefs about how others will judge information-seeking may hinder learning at work. An exciting direction for future research is to dive into mechanisms that might influence workers' beliefs about the costs and benefits of engaging with peers or managers.

Limited observability and distance Emanuel *et al.* (2023) study how the introduction of remote work and the disruption of co-location impact collaboration and feedback patterns among software engineers at a Fortune 500 company. The authors exploit variation in teams' pre-pandemic co-location to examine whether Covid-induced remote work had different effects on workers' communication patterns depending on their initial reliance on in-person versus electronic communication. In particular, some teams had initially been split across different offices located several blocks apart, while others sat together in the same building. When engineers worked on-premises, those who were co-located with their team received 22% more feedback on their code than engineers on teams split between buildings. Remote work eliminated the feedback difference for previously co-located teams, highlighting the importance of proximity for information flows. After the shift, senior engineers became more productive, but younger workers' career advancement suffered as learning opportunities dried up.

Tying these results to workers' beliefs, Roghanizad & Bohns (2022) find in two experiments that people underestimate the effectiveness of asking for help face-to-face compared with asking via zoom, phone, or email. Their results show that in-person requests are significantly more likely to be fulfilled, yet participants fail to anticipate the magnitude of this difference. While remote work increases physical distance between colleagues, long-distance collaboration had already been expanding with the internet. The growing literature on remote work (beyond the scope of this chapter) offers valuable insights into how workers interact in virtual settings. For further discussion, see the overviews by Barrero *et al.* (2023) and Barrero *et al.* (2021).

Adverse effects of incentives Other work suggests that competition for promotions can disrupt the link between knowledge spillovers and peer effects in wages. Follow-up work by Cornelissen *et al.* (2023) finds that a one standard deviation increase in the quality of trained peers leads to a 3.4% wage increase for untrained workers 5 years later. However, an increase in the quality of untrained peers—who are more likely to be competitors—has a negative effect on untrained focal workers. These findings suggest that when promotions are determined by rank-order tournaments, workers may be less likely to advance in the presence of more skilled peers. It is less clear whether knowledge spillovers are nonetheless occurring but are not reflected in wages due to competition, or whether competition actively suppresses knowledge transmission through avenues like sabotage or reduced cooperation (Lazear, 1989).

5.2.3 Coordination

Turning to coordination in a setting requiring rapid and complex decision-making, Battiston *et al.* (2021) provide evidence linking face-to-face communication among co-workers to increased productivity. Using a natural experiment through the Greater Manchester Police Department's Operational Communications Branch, the authors examine how co-location affects workers' efficiency and response times in handling emergency calls. Each call is handled by two different staff members, a "call handler" who speaks with the caller and a "radio operator" who dispatches officers in the field. The staff are distributed across four different rooms within the facility, meaning that some calls are handled by workers who can see one another, while others are managed by workers without a direct line of site. When the two workers are co-located in the same room, response times to incidents fall by 2%. However, this benefit comes with a trade-off: handlers took 2.5% longer to be ready for the next call when co-located with operators, as they spent more time communicating about the incident.

The study also finds that social connections between workers, such as similar age and gender, amplify the benefits of co-location. Workers who had more in common or had worked together previously were more likely to engage in helpful communication, further reducing response times. These findings suggest that workspace design should account for the importance of direct communication, especially in tasks that require rapid and complex decision-making.

5.2.4 Peer pressure

A classic dilemma in workplaces with peer interactions is the free-rider problem: when agents' inputs are imperfectly observed, they may have incentives to shirk (Holmstrom, 1982). Such environments are common in practice. Early research on peer spillovers focused on how peer pressure mitigates free-riding (Kandel & Lazear, 1992; Mas & Moretti, 2009). For example, in a supermarket checker setting, where one slow worker forces others to compensate, Mas & Moretti (2009) find that introducing highly productive workers into a shift generates positive productivity spillovers, driven by peer pressure.

Similarly, exploiting shift-to-shift variation in physical proximity created by an open office plan, Battiston *et al.* (2023) find that public sector police call center handlers take 7% more calls when the two desks adjacent to them are occupied. The authors attribute this effect to peer pressure, as callers are not assigned to a single handler—if one handler works slowly, others must compensate. In this setting, managers also monitor handlers' work in real time. The study finds that peer pressure effects are stronger when the manager is seated further away, suggesting that peer pressure and direct managerial monitoring act as substitutes. When managerial oversight is weaker, peer-driven accountability increases.

5.2.5 Extending research designs and open questions about peers

Many of the papers we have discussed leverage pre-existing differences to study the heterogeneous effects of changes in work practices or conditions. For example, Emanuel *et al.* (2023) exploit variation in team co-location to study how proximity influences collaboration following a single, exogenous shift to remote work due to Covid. Similarly, Espinosa & Stanton (2022) estimate spillovers from a randomized training investment by a Colombian government agency, using email data to measure manager and coworker exposure to workers who eventually receive training. Other studies, like De Grip & Sauermann (2012), examine how the share of treated coworkers affects untreated workers' productivity by varying the timing of treatment. Random saturation designs, in which the proportion of treated individuals varies, provide another approach for detecting spillover effects (Baird *et al.*, 2018). While spillovers are often considered a nuisance in experimental designs due to Stable Unit Treatment Value (SUTVA) assumptions, for personnel economists, they offer valuable insights into production functions and knowledge and communication flows within firms.

Comparing the effects of large-scale organizational changes with smaller, incremental adjustments and naturally occurring variations is likely a promising area for future research. Because natural experiments often rely on small interventions, they may fail to capture the individual re-optimization that would occur in response to larger-scale changes. For example, in a major intervention aimed at engineering peer effects in educational settings, students at the Air Force Academy adjusted their own peer network endogenously after the administration aggressively manipulated group composition, beyond the natural variation observed in the data (Carrell *et al.*, 2013). This suggests that when changes are large enough, individuals may respond in ways that smaller-scale interventions fail to capture. More broadly, we think that future research should explore when and how program or practice choices generate spillovers. For instance, Adhvaryu *et al.* (2023b) examine soft skills training for Indian garment workers and find that returns are highest when trained workers engage in team production or collaborative tasks. Additionally, they show that soft skills training substitutes for managerial attention, suggesting that such interventions may create opportunities for larger-scale organizational changes.

5.3 Team Production and Effective Teams

The literature on teams is relatively sparse compared to the research on peers. One possible reason is statistical power: while studies of peer effects often focus on the influence of other workers for a focal worker doing an individual task, output measures for teams are often only observed for a group, limiting access to large samples in all but the largest organizations. Issues with endogenous sorting into teams can lead to selection on unobservables, potentially confounding inference in certain contexts. This challenge is particularly relevant when team formation differs across studies, such as when teams are assigned in some cases and voluntarily formed in others. Still, understanding teams and team production is of huge practical importance, and we hope to see more work on this topic in the future.

5.3.1 Rationales for team production

Hamilton *et al.* (2003) provide one of the first empirical economics papers on the introduction of team production. They find that introducing team production in a garment factory increased productivity by approximately 18%. Contrary to concerns that less productive workers would join teams to free-ride, the results show that more productive workers preferred team settings. In particular, teams appeared to "expand production possibilities by utilizing collaborative skills ... [that] ... differ from and are not necessarily perfectly correlated with the more technical ability associated with individual piece rate production." Follow-up work by Hamilton *et al.* (2012) lays out several ways that teams can alter production possibilities:

- More productive workers can teach others how to do tasks more efficiently.

- Teams can assign workers to tasks based on comparative advantage.
- Teams may be more effective at discovering new methods or engaging in process innovation by "putting together the teammates' idiosyncratic information."

Lazear & Shaw (2007) elucidate how specialization can justify team production, using an example where workers have different endowments of skills. When workers each have absolute advantages in a different skill, team production where workers ask one another questions is likely close to optimal. By contrast, when one worker has an absolute advantage in all skills, a hierarchy where that worker becomes a problem solver for all is likely advantageous. In this framing, the rationale for teaming is based on worker heterogeneity across multi-dimensional skills.

In other contexts, the objective of teaming may instead be choosing an idea, innovating on a problem, or aggregating information. In this vein, work on contests and crowdsourcing shows that drawing ideas from multiple individuals can lead to extreme-value outcomes, with exceptional solutions often coming from those with diverse or non-traditional backgrounds in a subject (Boudreau *et al.*, 2011; Boudreau & Lakhani, 2013). Diverse perspectives may also tilt the direction of innovation or improve understanding of user behavior (Koning *et al.*, 2020).

5.3.2 Costs of team production

A central theoretical issue in team design is how diverse a team should be (Prat, 2002). Lyons (2017) finds that while diverse teams can offer production benefits, these gains do not necessarily outweigh the communication costs associated with bringing together workers from dissimilar backgrounds. In her experiment, Lyons randomly assigned contractors from an online labor market to complete a task that required JavaScript and PHP programming. Workers were assigned to groups of two, and worked either as a team or individually. Some teams consisted of two workers from the same country, while others paired workers from different countries. Teams of workers from the same country outperformed individuals by about 30% on output metrics, yet teams with workers from different countries under-performed individuals by about 34%. Lyons finds evidence that the results are driven by communication challenges, rather than same-nationality preferences.

Hjort (2014) studies teams and diversity in the context of flower production in Kenya. At the time of the study, Kenyan society was primarily divided along two main ethnic lines, with multiple tribes grouped into each. Flower production involved one upstream worker providing flowers to two downstream workers. Hjort (2014) finds that multi-ethnicity teams performed worse than more homogeneous teams, and that this gap widened following a disputed election that triggered ethnic violence. However, introducing team-based incentives reduced the gap between heterogeneous and homogeneous teams.

5.3.3 What makes teams effective

Although studying teams is difficult, organizational behavior scholars have started to open up key questions using data from real firms. Famously, Google's Project Aristotle tried to identify the key determinants of effective teams using observational data. Their findings highlight psychological safety as the primary factor driving team effectiveness within the company (Duhigg, 2016). In contrast, they found little evidence that differences in team composition significantly impact performance.

Other studies yield more evidence of the tradeoff between the benefits of diverse skill backgrounds and the communication challenges that arise from them. Di Fang & Iglesias (2023) study how the overlap of expertise for teams of doctors in Brazil influences patient outcomes. They find that having physicians with more similar specialties on a team results in lower mortality, shorter hospital stays, and fewer other expenses. The effects are larger in more complex cases and when teams have less accumulated experience working together, suggesting that communication frictions reduce some of the benefits of having diverse skill sets. We would be very interested in further studies examining workers' and managers' beliefs about what makes an effective team, paired with studies of workers' preferences over different team arrangements.

6 Time Use, Technology at Work, and Training

This section considers three topics in personnel — time use, technology, and training — that are closely connected to agency theory and incentives, but also represent important and separate areas of inquiry. Our goal here is not to present a comprehensive review of these topics, but instead to sketch some interesting areas where personnel economics has made progress and to suggest where additional work seems promising.

6.1 Time use

Time is arguably the scarcest commodity in the workplace. Ever since the pioneering work of Becker (1965), economists have been deeply interested in how people spend their time. However, time use has not been a central part of personnel economics, in large part because it has been very difficult to measure and track. Recent advances have made it possible to study, in much greater granularity, how people use their time and accomplish tasks at work.

One standard approach is to use surveys, such as the American Time Use Survey. As summarized by Hamermesh (2019), this and other surveys have illuminated long-term trends in how people allocate their time between work and non-work activities. Research by Bandiera *et al.* (2020) represents a significant advancement in measuring time use at work. Their focus is on how CEOs spend their time, a topic traditionally challenging to study through surveys. Instead of using a survey, Bandiera *et al.* (2020) enlist CEOs' personal assistants to track how CEOs spend their time over a week. Based on these accounts, researchers can observe how often CEOs meet with different parties in the organization, how often they take breaks, and at what parts of the day they are working. The authors then use a machine learning model to classify CEOs into two distinct types: leaders and managers. Leaders focus on higher-level, multi-function meetings, whereas managers focus on individual functions. The authors find that leader-type CEOs outperform manager-types, though the performance gap takes several years to emerge. The data also reveal that professional CEOs put in more work hours than family CEOs (Bandiera *et al.*, 2018).

Another issue related to time use at work is how workers manage competing tasks. For example, many people face multiple competing deadlines at work and, given limited time, have to decide on the order for completing tasks. Coviello *et al.* (2014) present a theory of "task juggling," and Coviello *et al.* (2015) test the theory using data on Italian judges. They find that greater task juggling—measured by dealing with multiple cases simultaneously— results in longer average case completion times, indicating a tradeoff between multitasking and efficiency.

Hoffman & Lyons (2020) study the relation between pay levels and time use for US state legislators. The pay for US state legislators varies substantially across states, even after accounting for the wide variation in the nature of legislators' responsibilities. They find that legislators with higher salaries spend less time on legislative activity, on average. Instead, they allocate more time to fundraising, which may offer greater returns for reelection prospects than legislative work.

We believe that the future is ripe for exploration on workplace time use by personnel economists. Substantial amounts of time use data are being tracked electronically, such as when workers swipe in and out of the building (Cullen & Perez-Truglia, 2023a), or the meetings that they take (Impink *et al.*, 2024; Dillon *et al.*, 2025). Understanding and measuring how employees use time also can help illuminate priorities and preferences, whether set by the organization or chosen by the employee.

6.2 Technology at work

Labor economists have long been interested in the impact of technology in the labor market. A separate handbook chapter covers many facets of this topic. However, we believe that certain aspects are especially well-suited for personnel economists, namely: 1) how workers use technology on the job and how new technologies affect productivity levels and distributions, and 2) how technological advancements affect firm boundaries, influencing labor relationships, like with the rise of platforms like Uber?

Technology adoption and productivity A classic example on the use of technology and productivity is the study by Bartel *et al.* (2007) on valve manufacturing.[36] Bartel *et al.* (2007) focus on the introduction of computer numerically controlled machines in valve manufacturing, showing that the technology causes firms to change their business strategy, increase skill demands for workers, and increase efficiency at different stages of production.

Artificial intelligence (AI) has the potential to alter productivity dramatically. Early studies have shown that using generative AI increases average output while decreasing production heterogeneity across workers. In an RCT, Noy & Zhang (2023) randomly provided generative AI to some workers performing a writing task, and found that the access to AI decreased task completion time by 40%, increased output quality by 18%, and reduced the variance of output between workers. Brynjolfsson *et al.* (2023) study the impact of generative AI in the context of a firm operating a large number of contact centers. They gradually introduced AI technology that generated text replies to customers. They found an increase in productivity concentrated among less productive and less experienced contact center agents. Workers in the bottom 20% of the pre-experiment productivity distribution

[36] Another early example is the work by Autor *et al.* (2002) on the usage of computers at work.

had a 35% increase in output, while those in the top 20% had no change in output. Both customer satisfaction and employee retention increased after technology adoption. These results suggest that generative AI will narrow productivity dispersion between people, but the results may be task or context dependent. For innovative tasks or entrepreneurship, the impact of AI appears to exacerbate production differences. For example, Otis *et al.* (2024) analyze the impact of AI no entrepreneurs in developing countries and find that the experimental introduction of AI helps the most productive entrepreneurs.

Understanding how new AI technologies alter the level and distribution of productivity will likely be important given our motivating evidence on production heterogeneity. This is especially true in light of rapid diffusion, although adoption appears to vary across demographic groups of workers. For example, Humlum & Vestergaard (2024) study the adoption of ChatGPT among workers in Denmark. They find quick and substantial adoption of the technology, but that men are much more likely to adopt than women.

Other work suggests that the adoption of AI will require complementary management inputs. In an experiment studying how the early wave of AI altered time use at work, Dillon *et al.* (2025) show that workers treated with Microsoft's Copilot tools spent less time on email but did not alter collaboration patterns or shift work processes that would require coordination. Dell'Acqua *et al.* (2023) study the effects of generative AI in an artefactual field experiment with workers at the Boston Consulting Group. They categorize tasks into two types: those where generative AI is "inside the frontier," performing well, and those where it falls "outside the frontier," producing unacceptable solutions. Productivity increased for consultants, and especially lower-performers, who used AI for tasks inside the frontier. In contrast, productivity fell for tasks outside the frontier (e.g., combining quantitative and qualitative analysis). The authors suggest that consultants struggled to discern when generative AI was useful and inadvertently applied it to unsuitable tasks, effectively "falling asleep at the wheel." These studies suggest that process reorganization is likely needed to unlock AI's benefits beyond the individual level, while managerial guidance to deploy AI tools for appropriate tasks may be required.

Generative AI may also have important effects on the behavior of jobseekers, substantially reducing the amount of time it takes to prepare cover letters for job applications. This could reduce a jobseeker's application cost, paralleling earlier technologies like the internet that were believed to potentially reduce application costs (Kuhn & Skuterud, 2004). On the firm side, Wiles & Horton (2024) analyze the impact of ChatGPT on firms' job ads. They conduct an RCT where some firms were randomly offered AI-written first drafts of their job ads. They find that treated firms are more likely to post jobs, but they are no more likely to hire.

Firm organization, boundaries, and digitization of work Blinder *et al.* (2009) used data from O*Net to classify jobs into four categories of offshorability based on their task content. They estimated that between 26 and 29 percent of the jobs that existed in the US economy in 2004 could be offshored, with the declining information and communication technologies technology (ICT) costs increasing the attractiveness of offshoring service sector jobs.[37] Estimating the extent of offshoring is challenging, but the forces that could drive it also impact the organization of work. Two relevant questions for personnel economists are the extent to which work outside of traditional employment arrangements (often through digital platforms) has altered the labor market and how platform work differs from that done through traditional arrangements.

Reliably estimating the extent of platform work has proven challenging because the transactions often occur through private intermediaries. Collins *et al.* (2019) overcome some of these challenges by using tax data for the United States. Between 2000 and 2016, self-employment, as measured by 1099 tax filings, increased by 1.9%. At least half this growth can be attributed to the rise of online platforms.[38] Cross-border transactions, where workers are located outside of the United States, are even more difficult to observe. The Oxford Internet Institute estimates that 160 million global workers use online labor markets, suggesting that the supply of those interested in these new working models may be vast (Kässi *et al.*, 2021). Agrawal *et al.* (2015) and Horton *et al.* (2017) show that the majority of transactions in online labor platforms involve clients in high-income countries while workers tend to supply labor from developing countries.

A concern with online labor markets is that the ability to hire from anywhere could put downward pressure on wages in high-income countries. Yet accumulating evidence suggests that the payoff to

[37]Other work, like Autor *et al.* (2013) has looked at the offshoring of manufacturing jobs which comes from liberalization of trade restrictions. While falling costs of ICT likely complement using distant manufacturing facilities, the offshoring of services jobs is arguably distinct because the output from these jobs is not subject to formal trade barriers.

[38]This estimate is likely understated. Small transactions under 1099 reporting thresholds are not captured in this measure, and using self-employment as the reference frame misses contracting arrangements where workers receive a W-2 from a third-party intermediary, as is common for many large enterprises. See, for example, the discussion of Google's use of contractors who are paid W-2s in (Kerr & Kreitzberg, 2019).

workers from online work is skewed in large part due to information frictions, and these frictions prevent downward pressure on wages by making platform workers imperfect substitutes for local workers. Pallais (2014) and Stanton & Thomas (2016) show that there are substantial information frictions related to workers' reputations. Specifically, the studies show that providing workers with public, positive feedback alters their career and earnings trajectories by allowing them to break into the market. Workers who lack these online reputations struggle to secure work because they are riskier hires and firms cannot capture upside from uncovering quality talent. In contrast, firms are likely better able to utilize information in resumes about local hires in high-income countries (Agrawal *et al.*, 2016). In practice, this means that firms using online labor markets need to invest in screening job candidates and parsing their work histories and credentials. Stanton & Thomas (2024) suggest that this investment limits the extent of clients' search, preserving surplus for workers who are hired despite the huge supply of workers online.

The functioning of online labor markets is affected by market design, as features like worker recommendations and algorithmic matching lower firms' transaction costs in finding talent (Horton, 2017; Barach *et al.*, 2020). Similarly, how tasks are specified may reduce contracting frictions. For example, when examining outsouring of domestic tasks on the platform TaskRabbit, Cullen & Farronato (2021) show that the degree of task standardization influences matching efficiency. In rideshare markets, where the platform mediates matching, Hall *et al.* (2023) shows that it is difficult to change equilibrium earnings through market design interventions because of supply responses when tasks are undifferentiated.

6.3 Training

There is a large literature on public sector investments in training and active labor market programs, but there has been less focus on training in firms.[39] The limited work on firms' investment and participation in training is unfortunate because, at the labor market level, there is evidence that firms' participation in training improves program effectiveness (Dustmann & Schoenberg, 2012). We argue that more work on firm training is likely to be timely and topical, as firms' training expenditures are substantial[40] and growing due to several background trends, like digitization,

[39]For a discussion firm training and measurement issues, see an excellent recent chapter by Black *et al.* (2023).

[40]For example, Bartel (1995) reports that firms averaged spending about \$385 per worker on training in 1989 (roughly \$1,000 in 2024 dollars). Sandvik *et al.* (Forthcoming) report estimates that US firm training expenditures

electrification of vehicles, and public investment in chip manufacturing that require new skills and workforce capabilities. In fact, evidence from Switzerland suggests that changes in apprentices' training curricula can facilitate changes in firms' technology adoption (Schultheiss & Backes-Gellner, 2024), suggesting training may be a pre-requisite for many firms' investments or practice changes.

Training and contracts. A firm's ability to maintain long-term relationships with workers is crucial for investing in training. Garicano & Rayo (2017) present a model in which firms cannot use formal contracts to prevent workers from leaving, making it difficult to capture the returns from training. As a result, they show that knowledge is not provided at once, and the time it takes workers to learn is inefficiently long. The reason is that firms slow knowledge transmission to prevent workers from leaving, giving them partial skills with the promise of future knowledge in exchange for intermediate work. This inefficiency arises because workers are credit constrained and have no assets to pledge to pay for training. As a result, in this and other models, credit market imperfections provide a rationale for firms to provide training, but this introduces distortions because of the inability to bind workers to the firm.

An active area of empirical research considers how the use of contracts and imperfections in the external labor market affect the investment in and returns to training (Acemoglu & Pischke, 1998, 1999b). For example, firms often use training repayment agreements, also known as training contracts. These contracts impose penalties for workers who leave the firm within a set period of time after training, and have been used across occupations, including police officers, pilots, truck drivers, and teachers (Kraus, 1993, 2008). Hoffman & Burks (2017) analyze the impact of training contracts in the context of a large US trucking firm. Exploiting the staggered introduction of a 12-month training contract and the later staggered replacement by an 18-month contract, they show that training contracts substantially affect worker turnover and firm returns from training.

Another contractual mechanism affecting training provision is non-compete agreements. Starr (2019) studies how the enforceability of non-compete agreements and consideration laws impact training.[41] He uses data on training from the Survey of Income and Program Participation (SIPP) for 1996, 2001, 2006, and 2008. 21% percent of respondents in his sample report participating in

topped $100 billion in 2022. Konings & Vanormelingen (2015) report estimates for a 1997 to 2006 Belgian panel, where about 17% of workers in large firms (averaging over 50 employees or meeting revenue or asset thresholds) receive formal training annually at a per-capita cost of 1,500 euros.

[41] Consideration laws require that workers receive something in return for signing a non-compete agreement, such as training, for the agreement to be enforceable.

training at some point in the prior year, about 90% of which was firm-sponsored.[42] Starr finds significant effects of non-compete agreements on training. In occupations where non-competes are widely used, 27% of workers received firm-sponsored training in the past year, compared to only 14% in occupations with low non-compete usage. He then estimates how training provision varies as a function of state-level non-compete enforcement, interacted with the use of non-competes at the occupation level. The results suggest that non-compete enforceability increases training provision by about 14% relative to the mean.[43]

Which firms train? While the number of workers receiving firm-sponsored training appears large, which firms provide training is less clear across different contexts. In the United States, little data exist at the firm level on training provision. In Belgium, larger firms are more likely than others to provide training (Konings & Vanormelingen, 2015), consistent with the existence of setup or fixed costs that can be spread over multiple workers. Settings with active labor market programs, like apprenticeships, also provide data on which firms engage in training. In Germany and Switzerland, large firms appear more likely to engage in apprenticeship training and, in the Swiss case, they are more likely to report positive net benefits from doing so (Muehlemann *et al.*, 2007). In very small firms, the indirect costs of training, like pulling an owner away from production or operations to conduct training, appear to stifle training provision (Alfonsi *et al.*, 2020). Additionally, Caicedo *et al.* (2022) show that Colombian firms in high-skill sectors are less likely to take on apprentices, potentially due to higher training costs or lower perceived benefits.

Differences in firm responses may reflect varying abilities to capture value from training. In practice, however, we wonder if the intellectual history behind Becker's general skills training agenda has led firms to under-invest in providing worker skills. In many cases, experiments where firms provide skills training point to positive payoffs (see e.g., Adhvaryu *et al.* (2023b) on soft skills training for managers in India), raising questions of how firms make decisions about what training investments they can support. Do firms under-provide skills training, and, if so, what do managers believe about value creation and value capture from training provision?

Little research has covered how firms decide on training investments, but several papers suggest

[42] Although the SIPP is a panel survey, training questions were asked only once. Respondents were first asked: "During the past year, has [the respondent] received any kind of training intended to improve skill in one's current or most recent job?" Respondents who answered yes were then queried about what type of training they received and who paid for the training.

[43] There is also a rich literature on how non-compete agreements affect other aspects of worker and firm behavior.

individual worker-level measures of training returns may be understated. Training is closely tied to human capital and knowledge accumulation in firms and knowledge spillovers between people may raise firms' returns from skill investment. De Grip & Sauermann (2012) estimate these spillovers for workers at the same level, and (Espinosa & Stanton, 2022) show spillovers exist across layers of a hierarchy where training increases the productivity of managers. More work is needed to understand how firms forecast investment returns from training and how managers account for spillovers in practice.

Other topics related to training. Within firms, emerging research suggests that lower-level managers may hinder or accelerate the effectiveness of training programs. They may hinder training efficacy through either talent hoarding or enhance it through encouragement (Diaz *et al.*, 2024). Other potential barriers to training effectiveness come from the worker side, as emerging evidence from both active labor market programs and training within firms suggest the greatest beneficiaries of training investments may be the least likely workers to participate (Sandvik *et al.*, Forthcoming; Delfino *et al.*, 2024).

A promising area for further research is the extent to which training enhances workers' social and managerial skills. Bianchi & Giorcelli (2022) examine the impact of industry training programs during World War II, when the U.S. government provided various types of managerial training to support the war effort. The training covered job instructions, job relations, and job methods. The authors find substantial returns to managerial training, as well as complementarities between some training topics.

7 Three Big Challenges and Opportunities in Personnel: External Validity, Scaling, and General Equilibrium

7.1 External Validity

Given that personnel economists often focus on individual firms and industries, external validity is often a central question. How do we know if the results of one study are specific to a particular trucking, fruitpicking, or courier firm, or whether they apply to truckers, fruitpickers, and couriers overall? And how would we know if the results would apply to workers more generally?

These are important and difficult questions. One response to this criticism is to acknowledge that such concerns are not unique to personnel economics, and are common in other fields of economics, such as industrial organization and marketing. By accumulating evidence from different jobs or industries, one can examine whether results are specific to particular industries, or whether they are robust across different contexts. Several results in personnel economics have been shown to be present across different contexts, such as the finding that referred workers are more likely to be hired and less likely to quit than non-referred workers.

Another response is to acknowledge that some estimates are specific to the firm or context studied, but the particular focus is warranted. For example, Nicholas (2023) studies the emergence of professional management at General Electric, a special case that warrants broad interest because GE's prominence and example for other firms. With the growing evidence on firm effects on wages and productivity (Abowd *et al.*, 1999; Card *et al.*, 2013; Song *et al.*, 2019; Autor *et al.*, 2020), understanding what good firms do and how practices affect individual workers is likely to be of independent interest even if these practices do not generalize to less productive firms.

That said, building on discussions of external validity in economics (List, 2020), we do believe that there are steps that personnel economists can take to clarify and potentially reduce concerns about external validity.

1. **Choose workers & firm(s) that are well-suited for your question.** It is critical that personnel economists do not study a firm simply because it is convenient. The firms and workers studied need to be well-suited for the research question, and this should be explained to the reader. One simple test of relevance is whether the research question is of first-order importance for those particular workers and firms. For example, long-distance truckers seem well-suited for studying the impact of monitoring technologies, as monitoring is used for many truckers and other blue-collar workers, but would be less well-suited for studying the impact of patent bonuses.

2. **What type of firm or type of work is it?** To address external validity, personnel economists can provide meaningful details about the firm or type of work they are studying. While firms often wish to remain anonymous as a condition of collaboration on research, researchers can still provide important context about the firm. Beyond basic metrics like the

number of employees or the level of revenue, additional details can be useful. Does the firm follow a low-cost strategy, or is it focused on product differentiation or quality? Is the firm a typical performer in its industry, or is it an outlier? Based on employee reviews from websites like GlassDoor (for the US) or Kununu (Germany), does the firm have a typical employer reputation, or is it exceptional in some way? This information can provide useful context for thinking about treatment effects and other results.

Further, we encourage researchers to provide summary statistics comparing workers in their sample to other populations of workers. Do the firm's employees resemble others in the industry demographically? In addition, the effectiveness of workplace interventions often depend on complementary factors. For example, employee referral programs may be more effective when workers value having strong teammates, such as in environments where workers are paid group bonuses, but may be less effective in highly competitive settings. Researchers should provide details about other practices in place that may interact with treatments.

Other experiments focus on specific tasks or jobs, raising questions about whether results generalize to similar work settings. Surveys may help on this front. Englmaier *et al.* (2024) study team performance in an escape room game, using surveys to classify the task as a non-routine analytical task. They then survey one group of experts about the effects of group incentives in non-routine analytical tasks and another group about group incentives in escape rooms. By comparing these responses, they assess whether experts view the experimental task as representative of non-routine analytical tasks more broadly.

3. **Show results separately for different types of workers.** While personnel economists often have data from one firm, they may have data on different types of workers within that firm. For example, a tech firm may have engineers and business workers. A grocery store may have store workers, truckers, and back office workers. By presenting results separately for different types of workers, one can examine how results vary across different types of occupations, tasks, or background characteristics of workers. Such patterns can inform how we might expect results to vary across other firms or other types of work in general.

4. **Explain why the firm is working with you on research.** The results of a personnel economics study can make a firm look good or bad. As such, it is natural to question

why a firm is choosing to be a subject for research. Explicitly answering this question can provide important context on how it was possible for the experiment or study to take place.[44] Furthermore, discussing why the firm did the study can be useful for addressing the issue of multiple hypothesis testing. Multiple hypothesis testing is traditionally addressed using pre-registration and pre-analysis plans—and we support the use of both—but we think it can be further bolstered by explaining what the firm hoped to learn. Finally, explaining why the firm works with researchers can help elucidate other events or shocks the firm is experiencing that could bear on the research results.

5. **Examine results under different conditions for the same population.** Beyond the particular firm, it is important to consider whether effects are tied to a particular country or geography. In addition, studies conducted during unique periods, such as during the COVID pandemic, raise questions about how results would generalize under different circumstances. When studies are large enough, researchers can analyze results by geography or business cycle. For example, Friebel *et al.* (2023) examine how the effects of employee referral programs vary with local labor market conditions. Since personnel economics questions may be culturally specific – such as how treatment effects depend on local attitudes toward managerial authority – researchers studying multinational firms can compare results across countries. Researchers are sometimes able to examine the impact of treatments or policies in a context where the firm itself is undergoing a separate change, allowing one to examine whether the impact of a treatment is complementary or contingent with regard to another organizational change (Blader *et al.*, 2020).

6. **What happened after the research?** As personnel economists collaborate frequently with firms, firms are often very interested in the results of the study, whether observational or experimental. How did the firm react to the results? Did it find the results surprising or expected? Did the firm change anything as as result of the study? In the same way that researchers increasingly ask other economists to predict experimental results to gauge how estimates shift beliefs (DellaVigna & Pope, 2018), we believe it can be useful to report similarly

[44]For example, in studying the impact of injecting charter school practices in public schools, Fryer (2014) explains that several public schools were underachieving. The state actually demanded that the lowest performing two schools be treated, while randomization was performed on the next highest 18 ranked schools (above the two lowest scoring ones).

belief changes for the firm being studied.

7. **Return to theory.** Theoretical work in the economics of organizations predicts how different types of firms will adopt different practices, and how practices are complementary to one another. By using economic theory, researchers can make predictions on how empirical results may vary based on the type of worker or firm studied, or based on what practices the firm is using. That is, while personnel economists often use theory to make core empirical predictions, theory can also be used to inform external validity.

8. **Explain why the firm didn't do it already.** Economists have long puzzled over why firms fail to adopt seemingly beneficial practices. If a treatment proves successful, researchers can help contextualize the findings by explaining why the firm hadn't implemented it earlier. Was it due to lack of awareness? If so, what does that imply about overall manager quality? Did managers expect that the treatment wouldn't work, or that it wouldn't be cost effect? If so, where do those beliefs come from? We highlight here that many businesses run "experiments" that look very different than the RCTs conducted in medicine or social science (Banerjee *et al.*, 2020). Instead, the small tests run by managers are likely to be optimized for fast feedback and rapid testing. The downside is that this approach may generate false negatives that limit beneficial practice adoption. Researchers should report on any prior "experiments" that inform leaders' beliefs about treatment returns.

The items listed here are not intended to constitute a rigid checklist that researchers need to answer in full. Instead, we think these are useful points for researchers to consider reporting in order to better situate papers and address external validity concerns.

7.2 Scaling a Treatment

In RCTs, treatments are often studied while holding fixed other practices. What does this mean for inference? Consider the following example: Firm A exclusively pays salaries and Firm B mostly pays through piece rates or sales commissions. To compensate for the lack of performance pay, Firm A sets a minimum standard for workers to retain employment. For simplicity, consider what happens when each firm introduces a pilot treatment to a small subset of workers that includes resources, like training or a new technology, designed to make workers' jobs easier. At Firm B, since pay is performance-linked, we would expect output to increase among treated workers. Standard intuition would suggest that the output increase at Firm A would be much weaker, as the firm likely did not raise the minimum standard required of workers as part of treatment. If output is the experiment's endpoint, a test run at these two firms would likely reach different conclusions about the effects of training or technology, with a null effect being much more likely at Firm A. However, this null effect at Firm A is likely to be a false negative. If Firm A were applying the treatment beyond the experiment, they would likely increase their minimum performance standard, inducing an effort response from workers. A simple RCT done at a small scale or as a pilot would miss this effort response. However, such an RCT might detect improved worker retention, suggesting that the firm could feasibly raise standards if the program were adopted more broadly.

We go through this example to make two points. First, experiments done in the absence of complementary practice changes – in this case, where the firm raises standards after investment – may affect inference. Researchers need to think carefully about what else a firm might put in place in the event an experimental treatment is successful. This requires reverse engineering and a detailed understanding of the setting where a test is run. If done well, a design that is sufficiently powered could in principle test individual and complementary practice changes jointly, providing compelling evidence on mechanisms that drive results. The key is to understand and forecast, ex-ante, what the firm would want to do in the event they observe a successful test of the main dimension of interest.

Second, even in the absence of a design that tests complementary treatments, researchers should consider multiple end points and what they mean economically. In our example, increased retention likely means the firm has more scope to make other changes (like increasing the effort requirement or slowing wage growth). Examining a variety of endpoints that map to different behavioral responses is crucial for holistic inferences about programs.

7.3 General Equilibrium Effects

Personnel economists often consider the impact of a single firm adopting a policy, e.g., performance pay. However, such studies, often conducted with one firm, may not reveal what would happen if many or all firms adopt such a policy. Likewise, the impact of a single firm adopting a technology like artificial intelligence may be different from the impact of all firms adopting a technology. This contrasts with many studies in labor economics that examine the impact of a policy (e.g., changes in firing costs) that affects many or all firms at once. How should personnel economists address this issue?

The key is to clearly distinguish between: (1) whether a policy or practice affects one firm or many, and (2) how prevalent the policy is in the broader population. For example, questions about (i) the impact of performance pay for an individual firm, and (ii) the impact when all firms adopt performance pay are both important but distinct. Researchers should specify which question they are addressing, and keep in mind the appeal to different audiences. Business practitioners are likely most interested in question (i), which is also highly relevant to economic theory. Policymakers, on the other hand, may be more concerned with question (ii).

8 Conclusion

Personnel economics is an exciting and growing subfield of labor economics focused on worker-related issues inside of firms. In this chapter, we have sketched out advancements in the field and outlined where we expect future research can continue to advance. Since the last handbook chapter, personnel economics studies have documented robust and persistent productivity differences across people. Researchers have broken new ground on how workers respond to and sort based on incentives, identifying treatment effect heterogeneity that often varies with workers' productivity. There has also been an explosion of work on new topics, from hiring to peer and manager effects. Better data, better methods, and new questions have enabled this growth, pointing to a bright future for personnel economics research. At the same time, we would continue to urge researchers to dig deeper to connect results with practice. We would welcome studies on managers' and executives' beliefs about the efficacy of various practices and the factors that slow adoption in some firms.

Recent research in labor economics, organizational economics, and industrial organization emphasizes the large and growing importance of firm-level differences in pay and productivity. One possible driver of this variation is that certain firms have adopted bundles of different HR practices that enhance performance. Personnel economists can play a key role in understanding how these practices contribute to firm-level performance variation. Pushing further, personnel economists may reflect on whether the returns to practices used by leading firms would be similar if adopted by laggards.

We conclude by highlighting several limitations of our chapter. As mentioned in the introduction, this chapter includes very little coverage of personnel topics connected to diversity. This is an extremely active research area, but we have mostly vied away from it due to its coverage in other handbook chapters. We have also devoted little coverage to research on amenities and monopsony even though those are also active areas in personnel economics. Furthermore, this chapter has focused primarily on empirical work, though there is also substantial theoretical work conducted on personnel topics in organizations.

References

Abel, Martin, Burger, Rulof, & Piraino, Patrizio. 2020. The Value of Reference Letters: Experimental Evidence from South Africa. *American Economic Journal: Applied Economics*, **12**(3), 40–71.

Abeler, Johannes, Huffman, David, & Raymond, Collin. 2023. Incentive Complexity, Bounded Rationality and Effort Provision. *working paper*.

Abowd, John M., Kramarz, Francis, & Margolis, David N. 1999. High Wage Workers and High Wage Firms. *Econometrica*, **67**(2), pp. 251–333.

Acemoglu, Daron, & Angrist, Joshua D. 2001. Consequences of Employment Protection? The Case of the Americans with Disabilities Act. *Journal of Political Economy*, **109**(5), 915–957.

Acemoglu, Daron, & Pischke, Jorn-Steffen. 1998. Why Do Firms Train? Theory And Evidence. *Quarterly Journal of Economics*, **113**(1), 78–118.

Acemoglu, Daron, & Pischke, Jorn-Steffen. 1999b. The Structure of Wages and Investment in General Training. *Journal of Political Economy*, **107**(3), 539–572.

Adhvaryu, Achyuta, Bassi, Vittorio, Nyshadham, Anant, & Tamayo, Jorge A. 2020. *No line left behind: Assortative matching inside the firm*. Tech. rept. National Bureau of Economic Research.

Adhvaryu, Achyuta, Molina, Teresa, & Nyshadham, Anant. 2021a. *Expectations, Wage Hikes, and Worker Voice*. UMichigan working paper.

Adhvaryu, Achyuta, Gade, Smit, Molina, Teresa, & Nyshadham, Anant. 2021b. Sotto Voce: The Impacts of Technology to Enhance Worker Voice.

Adhvaryu, Achyuta, Kala, Namrata, & Nyshadham, Anant. 2022. Management and shocks to worker productivity. *Journal of Political Economy*, **130**(1), 1–47.

Adhvaryu, Achyuta, Nyshadham, Anant, & Tamayo, Jorge. 2023a. Managerial Quality and Productivity Dynamics. *Review of Economic Studies*, **Forthcoming**.

Adhvaryu, Achyuta, Kala, Namrata, & Nyshadham, Anant. 2023b. Returns to on-the-job soft skills training. *Journal of Political Economy*, **131**(8), 2165–2208.

Agan, Amanda, & Starr, Sonja. 2018. Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *Quarterly Journal of Economics*, **133**(1), 191–235.

Agan, Amanda, Freedman, Matthew, & Owens, Emily. 2021a. Is your lawyer a lemon? Incentives and selection in the public provision of criminal defense. *Review of Economics and Statistics*, **103**(2), 294–309.

Agan, Amanda Y, Cowgill, Bo, & Gee, Laura K. 2021b. *Salary history and employer demand: Evidence from a two-sided audit*. Tech. rept. National Bureau of Economic Research.

Agrawal, Ajay, Horton, John, Lacetera, Nicola, & Lyons, Elizabeth. 2015. Digitization and the contract labor market. *Economic analysis of the digital economy*, **219**, 89–110.

Agrawal, Ajay, Lacetera, Nicola, & Lyons, Elizabeth. 2016. Does standardized information in online markets disproportionately benefit job applicants from less developed countries? *Journal of international Economics*, **103**, 1–12.

Alexander, Diane. 2020. How do doctors respond to incentives? unintended consequences of paying doctors to reduce costs. *Journal of Political Economy*, **128**(11), 4046–4096.

Alfonsi, Livia, Bandiera, Oriana, Bassi, Vittorio, Burgess, Robin, Rasul, Imran, Sulaiman, Munshi, & Vitali, Anna. 2020. Tackling youth unemployment: Evidence from a labor market experiment in Uganda. *Econometrica*, **88**(6), 2369–2414.

Algan, Yann, Crépon, Bruno, & Glover, Dylan. 2023. Are Active Labor Market Policies Directed at Firms Effective? Evidence from a Randomized Evaluation with Local Employment Agencies. *J-PAL working paper*.

Ashraf, Nava, & Bandiera, Oriana. 2018. Social Incentives in Organizations. *Annual Review of Economics*, **10**, 439–463.

Ashraf, Nava, Bandiera, Oriana, & Jack, B Kelsey. 2014. No margin, no mission? A field experiment on incentives for public service delivery. *Journal of public economics*, **120**, 1–17.

Ashraf, Nava, Bandiera, Oriana, Davenport, Edward, & Lee, Scott S. 2020. Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. *American Economic Review*, **110**(5), 1355–94.

Autor, David, Dorn, David, Katz, Lawrence F, Patterson, Christina, & Van Reenen, John. 2020. The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, **135**(2), 645–709.

Autor, David H., & Scarborough, David. 2008. Does Job Testing Harm Minority Workers? Evidence from Retail Establishments. *Quarterly Journal of Economics*, **123**(1), 219–277.

Autor, David H, Levy, Frank, & Murnane, Richard J. 2002. Upstairs, downstairs: Computers and skills on two floors of a large bank. *ILR Review*, **55**(3), 432–447.

Autor, David H, Dorn, David, & Hanson, Gordon H. 2013. The China syndrome: Local labor market effects of import competition in the United States. *American economic review*, **103**(6), 2121–2168.

Azoulay, Pierre, Zivin, Joshua S. Graff, & Manso, Gustavo. 2011. Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics*, **42**(3), 527â554.

Bagues, Manuel F., & Esteve-Volart, Berta. 2010. Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment. *Review of Economic Studies*, **77**(4), 1301–1328.

Baird, Sarah, Bohren, J Aislinn, McIntosh, Craig, & Özler, Berk. 2018. Optimal design of experiments in the presence of interference. *Review of Economics and Statistics*, **100**(5), 844–860.

Bandiera, Oriana, Barankay, Iwan, & Rasul, Imran. 2013. Team Incentives: Evidence from a Firm Level Experiment. *Journal of the European Economic Association*, **11**(5), 1079â1114.

Bandiera, Oriana, Lemos, Renata, Prat, Andrea, & Sadun, Raffaella. 2018. Managing the family firm: Evidence from CEOs at work. *Review of Financial Studies*, **31**(5), 1605–1653.

Bandiera, Oriana, Prat, Andrea, Hansen, Stephen, & Sadun, Raffaella. 2020. CEO Behavior and Firm Performance. *Journal of Political Economy*, **128**(4), 1325–1369.

Banerjee, Abhijit V, Chassang, Sylvain, Montero, Sergio, & Snowberg, Erik. 2020. A theory of experimenters: Robustness, randomization, and balance. *American Economic Review*, **110**(4), 1206–1230.

Bapna, Sofia, Benson, Alan, & Funk, Russell. 2021. Rejection Communication and Women's Job-search Persistence. *Available at SSRN 3953695*.

Barach, Moshe A, Golden, Joseph M, & Horton, John J. 2020. Steering in online markets: the role of platform incentives and credibility. *Management Science*, **66**(9), 4047–4070.

Barrero, Jose Maria, Bloom, Nicholas, & Davis, Steven J. 2021. Why Working from Home Will Stick. Apr. DOI: 10.3386/w28731.

Barrero, Jose Maria, Bloom, Nicholas, & Davis, Steven J. 2023. The Evolution of Work from Home. *Journal of Economic Perspectives*, **37**(4), 23â50.

Bartel, Ann, Ichniowski, Casey, & Shaw, Kathryn. 2007. How Does Information Technology Affect Productivity? Plant-level Comparisons of Product Innovation, Process Improvement, and Worker Skills. *Quarterly Journal of Economics*, **122**(4), 1721–1758.

Bartel, Ann P. 1995. Training, Wage Growth, and Job Performance: Evidence from a Company Database. *Journal of Labor Economics*, **13**(3), 401–25.

Bartel, Ann P., Cardiff-Hicks, Brianna, & Shaw, Kathryn. 2017. Incentives for Lawyers: Moving Away from "Eat What You Kill". *ILR Review*, **70**(2), 336–358.

Barwick, Panle Jia, Liu, Yanyan, Patacchini, Eleonora, & Wu, Qi. 2023. Information, Mobile Communication, and Referral Effects. *American Economic Review*, **113**(5), 1170–1207.

Battiston, Diego, Blanes i Vidal, Jordi, & Kirchmaier, Tom. 2021. Face-to-face Communication in Organizations. *Review of Economic Studies*, **88**(2), 574–609.

Battiston, Diego Ezequiel, Vidal, Blanes I, Kirchmaier, Tom, Szemeredi, Katalin, *et al.* 2023. Peer pressure and manager pressure in organisations.

Bayer, Patrick, Ross, Stephen L., & Topa, Giorgio. 2008. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *JPE*, **116**(6), 1150–1196.

Becker, Gary S. 1965. A Theory of the Allocation of Time. *The economic journal*, **75**(299), 493–517.

Becker, Gary S. 1993. Nobel Lecture: The Economic Way of Looking at Behavior. *Journal of Political Economy*, **101**(3), 385–409.

Benabou, Roland, & Tirole, Jean. 2003. Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, **70**(3), 489â520.

Bénabou, Roland, & Tirole, Jean. 2016. Bonus Culture: Competitive Pay, Screening, and Multitasking. *Journal of Political Economy*, **124**(2), 305–370. Publisher: The University of Chicago Press.

Bennedsen, Morten, Pérez-González, Francisco, & Wolfenzon, Daniel. 2020. Do CEOs Matter? Evidence from Hospitalization Events. *Journal of Finance*, **75**(4), 1877–1911.

Benson, Alan, & Shaw, Kathryn. 2024. *Thinking About What Managers Do.* Working Paper, Stanford University.

Benson, Alan, Li, Danielle, & Shue, Kelly. 2019. Promotions and the Peter Principle. *Quarterly Journal of Economics*, **134**(4), 2085–2134.

Benson, Alan, Sojourner, Aaron, & Umyarov, Akhmed. 2020. Can Reputation Discipline the Gig Economy? Experimental Evidence from an Online Labor Market. *Management Science*, **66**(5), 1802–1825.

Benson, Alan, Board, Simon, & Meyer-ter-Vehn, Moritz. 2024. Discrimination in Hiring: Evidence from Retail Sales. *Review of Economic Studies*, **Forthcoming**.

Benson, Alan M., Li, Danielle, & Shue, Kelly. 2023. Potential and the Gender Promotion Gap. *Academy of Management Proceedings*, **2023**(1), 19580.

Bernstein, Ethan S, & Turban, Stephen. 2018. The impact of the âopenâworkspace on human collaboration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **373**(1753), 20170239.

Bernstein, Shai, Colonnelli, Emanuele, Hoffman, Mitchell, & Iverson, Benjamin. 2023 (February). *Life After Death: A Field Experiment with Small Businesses on Information Frictions, Stigma, and Bankruptcy.* Working Paper 30933. National Bureau of Economic Research.

Bertrand, Marianne, & Schoar, Antoinette. 2003. Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics*, **118**(4), 1169–1208.

Bessone, Pedro, Rao, Gautam, Schilbach, Frank, Schofield, Heather, & Toma, Mattie. 2021. The Economic Consequences of Increasing Sleep Among the Urban Poor. *Quarterly Journal of Economics*, **136**(3), 1887–1941.

Bianchi, Nicola, & Giorcelli, Michela. 2022. The Dynamics and Spillovers of Management Interventions: Evidence from the Training within Industry Program. *Journal of Political Economy*, **130**(6), 1630–1675.

Bianchi, Nicola, & Paradisi, Matteo. 2022. Countries for old men: An analysis of the age wage gap. *Available at SSRN 3880501.*

Bianchi, Nicola, Bovini, Giulia, Li, Jin, Paradisi, Matteo, & Powell, Michael. 2023. Career Spillovers in Internal Labour Markets. *The Review of Economic Studies*, **90**(4), 1800–1831.

Biasi, Barbara. 2021. The labor market for teachers under different pay schemes. *American Economic Journal: Economic Policy*, **13**(3), 63–102.

Black, Dan A, Skipper, Lars, & Smith, Jeffrey A. 2023. Firm training. *Handbook of the Economics of Education*, 287.

Blader, Steven, Gartenberg, Claudine, & Prat, Andrea. 2020. The Contingent Effect of Management Practices. *Review of Economic Studies*, **87**(2), 721–749.

Blanes i Vidal, Jordi, & Nossol, Mareike. 2011. Tournaments without prizes: Evidence from personnel records. *Management science*, **57**(10), 1721–1736.

Blinder, Alan S, *et al.* 2009. How many US jobs might be offshorable? *World Economics*, **10**(2), 41.

Bloom, Nicholas, & Van Reenen, John. 2007. Measuring and Explaining Management Practices Across Firms and Countries. *Quarterly Journal of Economics*, **122**(4), 1351–1408.

Bloom, Nicholas, Eifert, Benn, Mahajan, Aprajit, McKenzie, David, & Roberts, John. 2013. Does Management Matter? Evidence from India. *Quarterly Journal of Economics*, **128**(1), 1–51.

Bloom, Nicholas, Lemos, Renata, Sadun, Raffaella, Scur, Daniela, & Van Reenen, John. 2014. The New Empirical Economics of Management. *JEEA*, **12**, 835–876.

Bloom, Nicholas, Sadun, Raffaella, & Reenen, John Van. 2016. *Management as a Technology?* Working Paper 22327. National Bureau of Economic Research.

Bloom, Nicholas, Brynjolfsson, Erik, Foster, Lucia, Jarmin, Ron, Patnaik, Megha, Saporta-Eksten, Itay, & Van Reenen, John. 2019. What Drives Differences in Management Practices? *American Economic Review*, **109**(5), 1648–83.

Boudreau, Kevin J, & Lakhani, Karim R. 2013. Using the crowd as an innovation partner. *Harvard business review*, **91**(4), 60–9.

Boudreau, Kevin J, Lacetera, Nicola, & Lakhani, Karim R. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management science*, **57**(5), 843–863.

Breza, Emily, Kaur, Supreet, & Shamdasani, Yogita. 2018. The Morale Effects of Pay Inequality. *Quarterly Journal of Economics*, **133**(2), 611â663.

Brown, Christina, & Andrabi, Tahir. 2023. Inducing Positive Sorting through Performance Pay: Experimental Evidence from Pakistani Schools. *Working Paper*.

Brown, Jennifer. 2011. Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy*, **119**(5), 982–1013.

Brown, Jennifer, & Minor, Dylan B. 2014. Selecting the Best? Spillover and Shadows in Elimination Tournaments. *Management Science*, **60**(12), 3087–3102.

Brown, Meta, Setren, Elizabeth, & Topa, Giorgio. 2016. Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System. *Journal of Labor Econonomics*, **34**, 161–209.

Brownback, Andy, & Sadoff, Sally. 2020. Improving college instruction through incentives. *Journal of Political Economy*, **128**(8), 2925–2972.

Bryan, Kevin A, Hoffman, Mitchell, & Sariri, Amir. 2023. *Information Frictions and Employee Sorting Between Startups*. Tech. rept. National Bureau of Economic Research.

Brynjolfsson, Erik, Li, Danielle, & Raymond, Lindsey R. 2023. *Generative AI at work*. Tech. rept. National Bureau of Economic Research.

Burks, Stephen V., Cowgill, Bo, Hoffman, Mitchell, & Housman, Michael. 2015. The Value of Hiring through Employee Referrals. *Quarterly Journal of Economics*, **130**(2), 805–839.

Bushway, Shawn, Pickett, Justin T, & Social, RAND. 2023. *Direct Incentives May Increase Employment of People With Criminal Records*. RAND.

Caicedo, Santiago, Espinosa, Miguel, & Seibold, Arthur. 2022. Unwilling to train?âFirm responses to the Colombian apprenticeship regulation. *Econometrica*, **90**(2), 507–550.

Caldwell, Sydnee, Haegele, Ingrid, & Heining, Jörg. 2024. Bargaining in the Labor Market.

Caliendo, Lorenzo, & Rossi-Hansberg, Esteban. 2012. The impact of trade on organization and productivity. *The quarterly journal of economics*, **127**(3), 1393–1467.

Caliendo, Lorenzo, Mion, Giordano, Opromolla, Luca David, & Rossi-Hansberg, Esteban. 2020. Productivity and organization in Portuguese firms. *Journal of Political Economy*, **128**(11), 4211–4257.

Caplin, Andrew, Deming, David J, Leth-Petersen, Søren, & Weidmann, Ben. 2023. *Allocative Skill*. Tech. rept. National Bureau of Economic Research.

Card, David, Heining, Joerg, & Kline, Patrick. 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics*, **128**(3), 967–1015.

Carrell, Scott E, Sacerdote, Bruce I, & West, James E. 2013. From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica*, **81**(3), 855–882.

Castilla, Emilio J. 2005. Social Networks and Employee Performance in a Call Center. *American Journal of Sociology*, **110**(5), 1243–1283.

Castro-Pires, Henrique, & Georgiadis, George. 2025. Incentives and Selection.

Chan, David C, & Chen, Yiqun. 2022. *The productivity of professions: evidence from the emergency department*. Tech. rept. National bureau of economic research.

Chan, David C, Gentzkow, Matthew, & Yu, Chuan. 2022. Selection with variation in diagnostic skill: Evidence from radiologists. *The Quarterly Journal of Economics*, **137**(2), 729–783.

Chandrasekhar, Arun G, Golub, Benjamin, & Yang, He. 2018. *Signaling, shame, and silence in social learning*. Tech. rept. National Bureau of Economic Research.

Chapman, Jonathan, Snowberg, Erik, Wang, Stephanie W, & Camerer, Colin. 2022. *Looming large or seeming small? Attitudes towards losses in a representative sample*. Tech. rept. National Bureau of Economic Research.

Checchi, Daniele, De Fraja, Gianni, & Verzillo, Stefano. 2021. Incentives and careers in academia: theory and empirical analysis. *Review of Economics and Statistics*, **103**(4), 786–802.

Clemens, Jeffrey, & Gottlieb, Joshua D. 2014. Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, **104**(4), 1320–1349.

Collins, Brett, Garin, Andrew, Jackson, Emilie, Koustas, Dmitri, & Payne, Mark. 2019. Is gig work replacing traditional employment? Evidence from two decades of tax returns. *Unpublished paper, IRS SOI Joint Statistical Research Program*.

Cornelissen, Thomas, Dustmann, Christian, & Schönberg, Uta. 2017. Peer effects in the workplace. *American Economic Review*, **107**(2), 425–456.

Cornelissen, Thomas, Dustmann, Christian, & Schönberg, Uta. 2023. *Knowledge Spillovers, Competition, and Individual Careers*. mimeo.

Coviello, Decio, Ichino, Andrea, & Persico, Nicola. 2014. Time Allocation and Task Juggling. *American Economic Review*, **104**(2), 609–23.

Coviello, Decio, Ichino, Andrea, & Persico, Nicola. 2015. The Inefficiency of Worker Time Use. *Journal of the European Economic Association*, **13**(5), 906–947.

Coviello, Decio, Deserranno, Erika, & Persico, Nicola. 2022a. Counterproductive worker behavior after a pay cut. *Journal of the European Economic Association*, **20**(1), 222–263.

Coviello, Decio, Deserranno, Erika, & Persico, Nicola. 2022b. Minimum Wage and Individual Worker Productivity: Evidence from a Large US Retailer. *Journal of Political Economy*, **130**(9), 2315–2360.

Cowgill, Bo. 2019. Bias and productivity in humans and machines. *Columbia Business School Research Paper Forthcoming*.

Cowgill, Bo, Agan, Amanda, & Gee, Laura Katherine. 2024. The gender disclosure gap: Salary history bans unravel when men volunteer their income. *Organization Science*.

Cross, Rob, Rebele, Reb, & Grant, Adam. 2016. Collaborative Overload. *Harvard Business Review*, Jan.

Cuñat, Vicente, & Guadalupe, Maria. 2005. How Does Product Market Competition Shape Incentive Contracts? *Journal of the European Economic Association*, **3**(5), 1058–1082.

Cullen, Zoë, & Farronato, Chiara. 2021. Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms. *Management Science*, **67**(7), 3985–4003.

Cullen, Zoë, & Perez-Truglia, Ricardo. 2022. How Much Does Your Boss Make? The Effects of

Salary Comparisons. *Journal of Political Economy*, **130**(3), 766–822.

Cullen, Zoe, & Perez-Truglia, Ricardo. 2023a. The old boys' club: Schmoozing and the gender gap. *American Economic Review*, **113**(7), 1703–1740.

Cullen, Zoë, & Perez-Truglia, Ricardo. 2023b. The salary taboo privacy norms and the diffusion of information. *Journal of Public Economics*, **222**, 104890.

Cullen, Zoe, Dobbie, Will, & Hoffman, Mitchell. 2023. Increasing the Demand for Workers with a Criminal Record. *Quarterly Journal of Economics*, **138**(1), 103–150.

Cullen, Zoe, Hoffman, Mitchell, & Koenig, Felix. 2024. Wage-surging: Labor Shortages and Firm Search. *Working paper*.

Cullen, Zoe B., Li, Shengwu, & Perez-Truglia, Ricardo. 2022. What's My Employee Worth? The Effects of Salary Benchmarking. Oct.

Dahlstrand, Amanda. 2025. Defying Distance The Provision of Services in the Digital Age.

De Bondt, Werner, & Thaler, Richard. 1995. Financial Decision-making in Markets and Firms: a Behavioral Perspective. *Pages 385–410 of:* Jarrow, R. A., Maksimovic, V., & Ziemba, W. T. (eds), *Handbooks in Operations Research and Management Science*. North Holland.

De Grip, Andries, & Sauermann, Jan. 2012. The effects of training on own and co-worker productivity: Evidence from a field experiment. *The Economic Journal*, **122**(560), 376–399.

De Ree, Joppe, Muralidharan, Karthik, Pradhan, Menno, & Rogers, Halsey. 2018. Double for nothing? Experimental evidence on an unconditional teacher salary increase in Indonesia. *The Quarterly Journal of Economics*, **133**(2), 993–1039.

Dean, Joshua. 2022. *Noise, Cognitive Function, and Worker Productivity*. Mimeo, University of Chicago.

Delfino, Alexia, & Espinosa, Miguel. 2024. *Value Dissonance at Work*. Mimeo, Bocconi.

Delfino, Alexia, Garnero, Andrea, Inferrera, Sergio, Leonardi, Marco, & Sadun, Raffaella. 2024. *Unwilling to Reskill? Evidence from a Survey Experiment with Jobseekers*.

Dell'Acqua, Fabrizio, McFowland, Edward, Mollick, Ethan R, Lifshitz-Assaf, Hila, Kellogg, Katherine, Rajendran, Saran, Krayer, Lisa, Candelon, François, & Lakhani, Karim R. 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*.

DellaVigna, Stefano, & Pope, Devin. 2018. Predicting experimental results: who knows what? *Journal of Political Economy*, **126**(6), 2410–2456.

Deming, David. 2017. The Growing Importance of Social Skills. *Quarterly Journal of Economics*, **132**(4), 1593–1640.

Deming, David J. 2021 (April). *The Growing Importance of Decision-Making on the Job*. Working Paper 28733. National Bureau of Economic Research.

Derbyshire, Daniel W., Jeanes, Emma, Khedmati Morasae, Esmaeil, Reh, Susan, & Rogers, Morwenna. 2024. Employer-focused interventions targeting disability employment: A systematic review. *Social Science Medicine*, **347**, 116742.

Deserranno, Erika. 2019. Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda. *American Economic Journal: Applied Economics*, **11**(1), 277–317.

Deserranno, Erika, & León-Ciliotta, Gianmarco. 2025. Promotions and Productivity: The Role of Meritocracy and Pay Progression in the Public Sector. *AER: Insights*.

Deserranno, Erika, Caria, Stefano, Leon-Ciliotta, Gianmarco, & Kastrau, Philipp. 2023. *The Allocation of Incentives in Multi-Layered Organizations: Evidence from a Community Health Program in Sierra Leone*. Tech. rept. CEPR Discussion Papers.

Dessein, Wouter. 2002. Authority and communication in organizations. *Review of Economic Studies*, **69**(4), 811–838.

Di Fang, Danyelle Branco Bladimir Carrillo, & Iglesias, Wilman. 2023. Expertise Overlap and Team Productivity: Evidence from the Hospital Industry.

Diaz, Brayan, Neyra-Nazarrett, Andrea, Ramirez, Julian, Sadun, Raffaella, & Tamayo, Jorge. 2024. Training within Firms.

Dillon, Eleanor Wiske, Jaffe, Sonia, Immorlica, Nicole, & Stanton, Christopher T. 2025. Shifting Work Patterns with Generative AI. *arXiv preprint arXiv:2504.11436*.

Doleac, Jennifer L, & Hansen, Benjamin. 2020. The Unintended Consequences of "Ban the Box": Statistical Discrimination and Employment Outcomes when Criminal Histories are Hidden. *Journal of Labor Economics*, **38**(2), 321–374.

Dube, Arindrajit, Giuliano, Laura, & Leonard, Jonathan. 2019. Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior. *American Economic Review*, **109**(2), 620â663.

Duflo, Esther. 2020. Field Experiments and the Practice of Policy. *American Economic Review*, **110**(7), 1952–1973.

Duhigg, Charles. 2016. What Google learned from its quest to build the perfect team. *The New York Times Magazine*, **26**(2016), 2016.

Dustmann, Christian, & Schoenberg, Uta. 2012. What Makes Firm-Based Vocational Training Schemes Successful? The Role of Commitment. *American Economic Journal: Applied Economics*, **4**(2), 36–61.

Ederer, Florian, & Manso, Gustavo. 2012. Is Pay-for-Performance Detrimental to Innovation? Oct.

Edmans, Alex, Gabaix, Xavier, & Jenter, Dirk. 2017. Executive compensation: A survey of theory and evidence. *The handbook of the economics of corporate governance*, **1**, 383–539.

Emanuel, Natalia, & Harrington, Emma. 2020. *The Payoffs of Higher Pay*. Tech. rept. Working Paper.

Emanuel, Natalia, Harrington, Emma, & Pallais, Amanda. 2023. *The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?* Harvard WP.

Englmaier, Florian, Roider, Andreas, & Sunde, Uwe. 2017. The role of communication of performance schemes: Evidence from a field experiment. *Management Science*, **63**(12), 4061–4080.

Englmaier, Florian, Grimm, Stefan, Grothe, Dominik, Schindler, David, & Schudy, Simeon. 2024. The Effect of Incentives in Nonroutine Analytical Team Tasks. *Journal of Political Economy*, **132**(8), Forthcoming.

Espinosa, Miguel, & Stanton, Christopher T. 2022. *Training, communications patterns, and spillovers inside organizations*. Tech. rept. National Bureau of Economic Research.

Ewens, Michael, Nanda, Ramana, & Stanton, Christopher T. 2023. Founder-CEO compensation and selection into venture capital-backed entrepreneurship. *Journal of Finance*, **Forthcoming**.

Fenizia, Alessandra. 2022. Managers and Productivity in the Public Sector. *Econometrica*, **90**(3), 1063–1084.

Fernandez, Roberto M., & Weinberg, Nancy. 1997. Sifting and Sorting: Personal Contacts and Hiring in a Retail Bank. *American Sociological Review*, **62**(6), pp. 883–902.

Fernandez, Roberto M., Castilla, Emilio J., & Moore, Paul. 2000. Social Capital at Work: Networks and Employment at a Phone Center. *American J. Sociology*, **105**, 1288–1356.

Filippas, Apostolos, Horton, John J, & Golden, Joseph M. 2022. Reputation Inflation. *Marketing Science*, **41**(4), 733–745.

Flory, Jeffrey A., Leibbrandt, Andreas, & List, John A. 2015. Do Competitive Workplaces Deter Female Workers? A Large-scale Natural Field Experiment on Job Entry Decisions. *Review of Economic Studies*, **82**(1), 122–155.

Fortin, Nicole M, Markevych, Mila, & Rehavi, Marit. 2022. *Closing the gender pay gap in the us federal service: the role of new managers*. Tech. rept. Working Paper.

Frakes, Michael D, & Wasserman, Melissa F. 2021. Knowledge spillovers, peer effects, and telecommuting: Evidence from the US Patent Office. *Journal of Public Economics*, **198**, 104425.

Frankel, Alexander. 2014. Aligned delegation. *American Economic Review*, **104**(1), 66–83.

Frederiksen, Anders, Kahn, Lisa B., & Lange, Fabian. 2020. Supervisors and Performance Management Systems. *Journal of Political Economy*, **128**(6), 2123–2187.

Frederiksen, Anders, Hansen, Daniel Baltzer Schjødt, & Manchester, Colleen Flaherty. 2024. Group-Based Incentives and Individual Performance: A Study of the Effort Response. *ILR Review*, **77**(2), 273–293.

Friebel, Guido, & Raith, Michael. 2022. Talent management: The role of bosses. *Available at SSRN 4162631*.

Friebel, Guido, Heinz, Matthias, Krueger, Miriam, & Zubanov, Nikolay. 2017. Team Incentives and Performance: Evidence from a Retail Chain. *American Economic Review*, **107**(8), 2168–2203.

Friebel, Guido, Heinz, Matthias, & Zubanov, Nikolay. 2022. Middle Managers, Personnel Turnover, and Performance: A Long-Term Field Experiment in a Retail Chain. *Management Science*, **68**(1), 211–229.

Friebel, Guido, Heinz, Matthias, Hoffman, Mitchell, & Zubanov, Nick. 2023. What Do Employee Referral Programs Do? Measuring the Direct and Overall Effects of a Management Practice. *Journal of Political Economy*, **131**(3), 633–686.

Friebel, Guido, Heinz, Matthias, Hoffman, Mitchell, Kretschmer, Tobias, & Zubanov, Nick. 2024. Is This Really Kneaded? Identifying and Eliminating Potentially Harmful Forms of Workplace Control. *ECONtribute Discussion Paper*.

Friedrich, Benjamin. 2016. Internal labor markets and the competition for managerial talent. *Unpublished manuscript, Department of Economics, Yale University*.

Frydman, Carola, & Jenter, Dirk. 2010. CEO compensation. *Annu. Rev. Financ. Econ.*, **2**(1), 75–102.

Fryer, Roland. 2014. INJECTING CHARTER SCHOOL BEST PRACTICES INTO TRADITIONAL PUBLIC SCHOOLS: EVIDENCE FROM FIELD EXPERIMENTS. *The Quarterly Journal of Economics*, **1355**, 1407.

Fryer, Jr., Roland G., Levitt, Steven D., List, John, & Sadoff, Sally. 2022. Enhancing the Efficacy of Teacher Incentives through Framing: A Field Experiment. *American Economic Journal: Economic Policy*, **14**(4), 269–99.

Fuchs, Larissa, Heinz, Matthias, Pinger, Pia, & Thon, Max. 2024. *How to Attract Talents? Field-Experimental Evidence on Emphasizing Flexibility and Career Opportunities in Job Advertisements*. ECONtribute Discussion Papers Series 332. University of Bonn and University of Cologne, Germany.

Gallus, Jana. 2017. Fostering Public Good Contributions with Symbolic Awards: A large-scale Natural Field Experiment at Wikipedia. *Management Science*, **63**(12), 3999–4015.

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

Garicano, Luis, & Rayo, Luis. 2017. Relational knowledge transfers. *American Economic Review*, **107**(9), 2695–2730.

Garicano, Luis, & Rossi-Hansberg, Esteban. 2006. Organization and inequality in a knowledge economy. *The Quarterly journal of economics*, **121**(4), 1383–1435.

Gee, Laura K. 2019. The More You Know: Information Effects on Job Application Rates in a Large Field Experiment. *Management Science*, **65**(5), 2077–2094.

Georgiadis, George, & Powell, Michael. 2022. A/b contracts. *American Economic Review*, **112**(1), 267–303.

Gibbons, Robert. 1998. Incentives in organizations. *Journal of economic perspectives*, **12**(4), 115–132.

Gibbons, Robert, & Henderson, Rebecca. 2012. *What do Managers Do?: Exploring Persistent Performance Differences among Seemingly Similar Enterprises*. Mimeo, Harvard Business School.

Gibbons, Robert, & Waldman, Michael. 1999. A Theory of Wage and Promotion Dynamics Inside Firms. *Quarterly Journal of Economics*, **114**(4), 1321–1358.

Giorcelli, Michela. 2023. The Effects of Business School Education on Manager Career Outcomes. *Available at SSRN 4468594*.

Glaeser, Edward L. 1999. Learning in cities. *Journal of urban Economics*, **46**(2), 254–277.

Glaeser, Edward L, & Maré, David C. 2001. Cities and skills. *Journal of labor economics*, **19**(2), 316–342.

Gneezy, Uri, Meier, Stephan, & Rey-Biel, Pedro. 2011. When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, **25**(4), 191–209.

Graham, John R, & Harvey, Campbell R. 2001. The theory and practice of corporate finance: Evidence from the field. *Journal of financial economics*, **60**(2-3), 187–243.

Granovetter, Mark. 1973. Strength of Weak Ties. *The American Journal of Sociology*, **78**(6), 1360–1380.

Groysberg, Boris. 2010. *Chasing Stars: The Myth of Talent and the Portability of Performance*. Princeton University Press.

Groysberg, Boris, Lee, Linda-Eling, & Nanda, Ashish. 2008. Can they take it with them? The portability of star knowledge workers' performance. *Management Science*, **54**(7), 1213–1230.

Gumpert, Anna, Steimer, Henrike, & Antoni, Manfred. 2022. Firm organization with multiple establishments. *The Quarterly Journal of Economics*, **137**(2), 1091–1138.

Haegele, Ingrid. 2022. Talent Hoarding in Organizations. *arXiv preprint arXiv:2206.15098*.

Hall, Jonathan V, Horton, John J, & Knoepfle, Daniel T. 2023. *Ride-sharing markets re-equilibrate*. Tech. rept. National Bureau of Economic Research.

Hamermesh, Daniel S. 2019. *Spending time: The most valuable resource*. Oxford University Press.

Hamilton, Barton H, Nickerson, Jack A, & Owan, Hideo. 2003. Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, **111**(3), 465–497.

Hamilton, Barton H, Nickerson, Jack A, & Owan, Hideo. 2012. Diversity and productivity in production teams. *Pages 99–138 of: Advances in the Economic Analysis of participatory and Labor-managed Firms*. Emerald Group Publishing Limited.

Hart, Oliver, & Zingales, Luigi. 2017. Companies Should Maximize Shareholder Welfare Not Market Value. *Journal of Law, Finance, and Accounting*, **2**(2), 247–275.

Hayek, Friedrich August. 1945. The Use of Knowledge in Society. *American Economic Review*, **35**(4), 519–530.

Hazell, Jonathon, Patterson, Christina, Sarsons, Heather, & Taska, Bledi. 2022. National Wage Setting. Nov. DOI: 10.3386/w30623.

Hedblom, Daniel, Hickman, Brent R, & List, John A. 2019. *Toward an understanding of corporate social responsibility: Theory and field experimental evidence*. Tech. rept. National Bureau of Economic Research.

Heller, Sara B, & Kessler, Judd B. 2021. *Information Frictions and Skill Signaling in the Youth Labor Market*. Tech. rept. National Bureau of Economic Research.

Hensvik, Lena, & Skans, Oskar Nordström. 2016. Social networks, employee selection, and labor market outcomes. *Journal of Labor Economics*, **34**(4), 825–867.

Herbst, Daniel, & Mas, Alexandre. 2015. Peer effects on worker output in the laboratory generalize to the field. *Science*, **350**(6260), 545–549.

Heursen, Lea, Friess, Svenja, & Chugunova, Marina. 2023. Reputational Concerns and Advice-Seeking at Work.

Hjort, Jonas. 2014. Ethnic Divisions and Production in Firms. *Quarterly Journal of Economics*, **129**(4), 1899–1946.

Hoffman, Mitchell. 2017. The Value of Hiring through Employee Referrals in Developed Countries. *IZA World of Labor*, **369**, 1–8.

Hoffman, Mitchell, & Burks, Stephen V. 2017. *Training contracts, employee turnover, and the returns from firm-sponsored general training*. Tech. rept. No. w23247, National Bureau of Economic Research.

Hoffman, Mitchell, & Burks, Stephen V. 2020. Worker Overconfidence: Field Evidence and Implications for Employee Turnover and Firm Profits. *Quantitative Economics*, **11**(1), 315–348.

Hoffman, Mitchell, & Lyons, Elizabeth. 2020. A Time to Make Laws and a Time to Fundraise? On the Relation Between Salaries and Time Use for State Politicians. *Canadian Journal of Economics*, **53**(3), 1318–1358.

Hoffman, Mitchell, & Tadelis, Steven. 2021. People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis. *Journal of Political Economy*, **129**(1), 243–285.

Hoffman, Mitchell, Kahn, Lisa B., & Li, Danielle. 2018. Discretion in Hiring. *Quarterly Journal of Economics*, **133**(2), 765–800.

Holmstrom, Bengt. 1982. Moral Hazard in Teams. *The Bell Journal of Economics*, 324–340.

Holmström, Bengt. 1999. Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies*, **66**(1), 169–182.

Holmstrom, Bengt, & Milgrom, Paul. 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization*, **7**(0), 24–52.

Horton, John. 2017. The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment. *Journal of Labor Economics*, **35**(2), 345–385.

Horton, John, Kerr, William R, & Stanton, Christopher. 2017. *Digital labor markets and global talent flows*. Tech. rept. National Bureau of Economic Research Cambridge, MA.

Horton, John J, Johari, Ramesh, & Kircher, Philipp. 2021. *Cheap Talk Messages for Market Design: Theory and Evidence from a Labor Market with Directed Search*. Tech. rept. National Bureau of Economic Research.

Huber, Kilian, Lindenthal, Volker, & Waldinger, Fabian. 2021. Discrimination, managers, and firm performance: Evidence from "aryanizations"' in nazi germany. *Journal of Political Economy*, **129**(9), 2455–2503.

Huffman, David, Raymond, Collin, & Shvets, Julia. 2022. Persistent Overconfidence and Biased Memory: Evidence from Managers. *American Economic Review*, **112**(10), 3141–75.

Humlum, Anders, & Vestergaard, Emilie. 2024. *The Adoption of ChatGPT*. Mimeo, University of Chicago.

Hunt, Priscillia, Smart, Rosanna, Jonsson, Lisa, & Tsang, Flavia. 2018. *Breaking Down Barriers Experiments into Policies That Might Incentivize Employers to Hire Ex-Offenders*. RAND Report.

Ichniowski, Casey, & Shaw, Kathryn. 2012. Insider Econometrics: A Roadmap for Estimating Empirical Models of Organizational Design and Performance. *R. Gibbons & J. Roberts, Handbook of Organizational Econ.*

Impink, Stephen Michael, Prat, Andrea, & Sadun, Raffaella. 2024. Communication within firms: evidence from CEO turnovers. *Management Science*.

Jäger, Simon, Roth, Christopher, Roussille, Nina, & Schoefer, Benjamin. 2022. *Worker Beliefs about Outside Options*. NBER Working Paper 29623.

Janke, Katharina, Propper, Carol, & Sadun, Raffaella. 2019. *The Impact of CEOs in the Public Sector: Evidence from the English NHS*. Tech. rept. National Bureau of Economic Research.

Johnson, Erin M, & Rehavi, M Marit. 2016. Physicians treating physicians: Information and incentives in childbirth. *American Economic Journal: Economic Policy*, **8**(1), 115–141.

KAHN, LISA. 2013. Asymmetric Information between Employers. *American Economic Journal: Applied Economics*, **5**(4), 165–205.

KAHN, LISA B, & LANGE, FABIAN. 2014. Employer learning, productivity, and the earnings distribution: Evidence from performance measures. *The Review of Economic Studies*, **81**(4), 1575–1613.

KANDEL, EUGENE, & LAZEAR, EDWARD P. 1992. Peer Pressure and Partnerships. *Journal of Political Economy*, **100**(4), 801–817.

KARUNA, CHRISTO. 2007. Industry product market competition and managerial incentives. *Journal of Accounting and Economics*, **43**(2), 275â297.

KÄSSI, OTTO, LEHDONVIRTA, VILI, & STEPHANY, FABIAN. 2021. How many online workers are there in the world? A data-driven assessment. *A Data-Driven Assessment (March 23, 2021)*.

KAUR, SUPREET, MULLAINATHAN, SENDHIL, OH, SUANNA, & SCHILBACH, FRANK. 2021. *Do financial concerns make workers less productive?* Tech. rept. National Bureau of Economic Research.

KE, RONGZHU, LI, JIN, & POWELL, MICHAEL. 2018. Managing careers in organizations. *Journal of Labor Economics*, **36**(1), 197–252.

KERR, WILLIAM, & KREITZBERG, CARL. 2019. Google: To TVC or Not to TVC? *Harvard Business School, September*.

KHASHABI, POOYAN, HEINZ, MATTHIAS, ZUBANOV, NICK, KRETSCHMER, TOBIAS, & FRIEBEL, GUIDO. 2021. Market Competition and the Effectiveness of Performance Pay. *Organization Science*, **32**(2), 334–351.

KNEZ, MARC, & SIMESTER, DUNCAN. 2001. Firm-wide Incentives and Mutual Monitoring at Continental Airlines. *Journal of Labor Economics*, **19**(4), 743–772.

KONING, REMBRAND, SAMILA, SAMPSA, & FERGUSON, JOHN-PAUL. 2020. Inventor gender and the direction of invention. *Pages 250–254 of: AEA Papers and Proceedings*, vol. 110. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

KONINGS, JOZEF, & VANORMELINGEN, STIJN. 2015. The impact of training on productivity and wages: firm-level evidence. *Review of Economics and Statistics*, **97**(2), 485–497.

KRAUS, ANTHONY W. 1993. Repayment Agreements for Employee Training Costs. *Labor Law Journal*, **44**(1), 49–55.

KRAUS, ANTHONY W. 2008. Employee Agreements for Repayment of Training Costs: The Emerging Case Law. *Labor Law Journal*, 213–226.

KRUEGER, MIRIAM, & FRIEBEL, GUIDO. 2022. A pay change and its long-term consequences. *Journal of Labor Economics*, **40**(3), 543–572.

KU, HYEJIN. 2022. Does Minimum Wage Increase Labor Productivity? Evidence from Piece Rate Workers. *Journal of Labor Economics*, **40**(2), 325–359.

KUHN, PETER. 2004. Is monopsony the right way to model labor markets? A review of Alan Manning's monopsony in motion. *International Journal of the Economics of Business*, **11**(3), 369–378.

KUHN, PETER, & MANSOUR, HANI. 2014. Is internet job search still ineffective? *The Economic Journal*, **124**(581), 1213–1233.

KUHN, PETER, & SHEN, KAILING. 2013. Gender discrimination in job ads: Evidence from china. *The Quarterly Journal of Economics*, **128**(1), 287–336.

KUHN, PETER, & SKUTERUD, MIKAL. 2004. Internet job search and unemployment durations. *American Economic Review*, **94**(1), 218–232.

KUHN, PETER, & WEINBERGER, CATHERINE. 2005. Leadership skills and wages. *Journal of Labor Economics*, **23**(3), 395–436.

KUHN, PETER J, & YU, LIZI. 2024. *Kinks as goals: Accelerating commissions and the performance of sales teams*. Tech. rept. National Bureau of Economic Research.

LALIVE, RAFAEL, WUELLRICH, JEAN-PHILIPPE, & ZWEIMÜLLER, JOSEF. 2013. Do Financial Incentives

Affect Firms' Demand for Disabled Workers? *Journal of the European Economic Association*, **11**(1), 25–58.

Landers, Renee, Rebitzer, James B., & Taylor, Lowell J. 1996. Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms. *AER*, 329–348.

Larkin, Ian, & Leider, Stephen. 2012. Incentive Schemes, Sorting and Behavorial Biases of Employees: Experimental Evidence. *American Economic Journal: Microeconomics*, **4**(2), 184–214.

Lazear, Edward. 2000a. Performance Pay and Productivity. *American Economic Review*, **90**(5), 1346–1361.

Lazear, Edward P. 1989. Pay equality and industrial politics. *Journal of political economy*, **97**(3), 561–580.

Lazear, Edward P. 1998. Hiring Risky Workers. *Pages 143–158 of: Internal Labour Earkets, Incentives and Employment*. Springer.

Lazear, Edward P. 1999. Culture and language. *Journal of political Economy*, **107**(S6), S95–S126.

Lazear, Edward P. 2000b. Economic Imperialism. *Quarterly Journal of Economics*, **115**(1), 99–146.

Lazear, Edward P. 2006. Speeding, Terrorism, and Teaching to the Test. *Quarterly Journal of Economics*, **121**(3), 1029–1061.

Lazear, Edward P. 2012. Leadership: A Personnel Economics Approach. *Labour Economics*, **19**(1), 92–101.

Lazear, Edward P. 2018. Compensation and Incentives in the Workplace. *Journal of Economic Perspectives*, **32**(3), 195â214.

Lazear, Edward P, & Oyer, Paul. 2012. *The Handbook of Organizational Economics, chapter Chapter 12: Personnel Economics*.

Lazear, Edward P, & Shaw, Kathryn L. 2007. Personnel economics: The economist's view of human resources. *Journal of economic perspectives*, **21**(4), 91–114.

Lazear, Edward P., Shaw, Kathryn, & Stanton, Christopher. 2015. The Value of Bosses. *Journal of Labor Economics*, **33**(4), 823–861.

Lazear, Edward P, Shaw, Kathryn L, & Stanton, Christopher. 2016. Making do with less: Working harder during recessions. *Journal of Labor Economics*, **34**(S1), S333–S360.

Leaver, Clare, Ozier, Owen, Serneels, Pieter, & Zeitlin, Andrew. 2021. Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from Rwandan primary schools. *American economic review*, **111**(7), 2213–2246.

Leibbrandt, Andreas, & List, John A. 2018. *Do Equal Employment Opportunity Statements Backfire? Evidence from a Natural Field Experiment on Job-Entry Decisions*. Tech. rept. National Bureau of Economic Research.

Lerner, Josh, Manley, Henry, Stein, Carolyn, & Williams, Heidi. 2024. *The wandering scholars: Understanding the heterogeneity of university commercialization*. Tech. rept. National Bureau of Economic Research.

Levin, Sharon G, & Stephan, Paula E. 1991. Research productivity over the life cycle: Evidence for academic scientists. *The American economic review*, 114–132.

Li, Danielle. 2017. Expertise versus Bias in Evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics*, **9**(2), 60–92.

Li, Danielle, Raymond, Lindsey R, & Bergman, Peter. 2021 (August). *Hiring as Exploration*. Working Paper 27736. National Bureau of Economic Research.

List, John A. 2020. *Non Est Disputandum de Generalizability? A Glimpse into the External Validity Trial*. Tech. rept. National Bureau of Economic Research.

Lundquist, Jennifer Hickes, Pager, Devah, & Strader, Eiko. 2018. Does a Criminal Past Predict Worker Performance? Evidence From One of America's Largest Employers. *Social Forces*, **96**(3), 1039–1068.

Lyons, Elizabeth. 2017. Team production in international labor markets: Experimental evidence from the field. *American Economic Journal: Applied Economics*, **9**(3), 70–104.

Macdonald, David, Montonen, Jerry, & Nix, Emily. 2023. DATING AND BREAKING UP WITH THE BOSS: BENEFITS, COSTS, AND SPILLOVERS.

Macleod, W. Bentley. 2003. Optimal Contracting with Subjective Evaluation. *American Economic Review*, **93**(1), 216â240.

Manning, Alan. 2021. Monopsony in Labor Markets: A Review. *ILR Review*, **74**(1), 3–26.

Manso, Gustavo. 2011. Motivating Innovation.

Mas, Alexandre, & Moretti, Enrico. 2009. Peers at Work. *AER*, **99**(1), 112–45.

Mas, Alexandre, & Pallais, Amanda. 2017. Valuing alternative work arrangements. *American Economic Review*, **107**(12), 3722–59.

McAfee, Andrew, & Brynjolfsson, Erik. 2012. Big Data: The Management Revolution. *Harvard Business Review*.

Metcalfe, Robert D, Sollaci, Alexandre B, & Syverson, Chad. 2023. *Managers and Productivity in Retail*. Tech. rept. National Bureau of Economic Research.

Minni, Virgnia. 2023. Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs. *Working paper, Chicago Booth*.

Minor, Dylan, Persico, Nicola, & Weiss, Deborah M. 2018. Criminal Background and Job Performance. *IZA Journal of Labor Policy*, **7**, 1–49.

Mocanu, Tatiana. 2023. *Designing Gender Equity: Evidence from Hiring Practices and Committees*. UCL working paper.

Mocanu, Tatiana, & Patacchini, Eleonora. 2024. *Personal Connections and Hiring Decisions in the Public Sector*. Mimeo, Columbia.

Molitor, David. 2018. The evolution of physician practice styles: evidence from cardiologist migration. *American Economic Journal: Economic Policy*, **10**(1), 326–356.

Moretti, Enrico. 2021. The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, **111**(10), 3328–3375.

Muehlemann, Samuel, Schweri, Juerg, Winkelmann, Rainer, & Wolter, Stefan C. 2007. An empirical analysis of the decision to train apprentices. *Labour*, **21**(3), 419–441.

Muralidharan, Karthik, & Sundararaman, Venkatesh. 2011. Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, **119**(1), 39–77.

Neal, Derek. 2011. The design of performance pay in education. *Pages 495–550 of: Handbook of the Economics of Education*, vol. 4. Elsevier.

Nicholas, Tom. 2023. Human Capital and the Managerial Revolution in the United States: Evidence from General Electric. *Review of Economics and Statistics*, 1–47.

Niederle, Muriel, & Vesterlund, Lise. 2007. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, **122**(3), 1067–1101.

Noy, Shakked, & Zhang, Whitney. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, **381**(6654), 187–192.

Otero, Cristobal, & Munoz, Pablo. 2022. *Managers and Public Hospital Performance*. Tech. rept. WP, UC Berkeley.

Otis, Nicholas, Clarke, Rowan, Delecourt, Solene, Holtz, David, & Koning, Rembrand. 2024. The uneven impact of generative AI on entrepreneurial performance.

Oyer, Paul, & Schaefer, Scott. 2002. Litigation Costs and Returns to Experience. *American Economic Review*, **92**(3), 683–705.

Oyer, Paul, & Schaefer, Scott. 2011. Personnel Economics: Hiring and Incentives. *Handbook of Labor Economics*.

Pager, Devah. 2003. The Mark of a Criminal Record. *American Journal of Sociology*, **108**(5), 937–975.

PALLAIS, AMANDA. 2014. Inefficient Hiring in Entry-level Labor Markets. *American Economic Review*, **104**(11), 3565–99.

PALLAIS, AMANDA, & SANDS, EMILY GLASSBERG. 2016. Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, **124**(6), 1793–1828.

PARSONS, CHRISTOPHER A, & VAN WESEP, EDWARD D. 2013. The timing of pay. *Journal of Financial Economics*, **109**(2), 373–397.

PIERCE, LAMAR, REES-JONES, ALEX, & BLANK, CHARLOTTE. 2020. The Negative Consequences of Loss-Framed Performance Incentives.

POLZER, JEFFREY T. 2022. The rise of people analytics and the future of organizational research. *Research in Organizational Behavior*, **42**, 100181.

PRAT, ANDREA. 2002. Should a team be homogeneous? *European Economic Review*, **46**(7), 1187–1207.

PRENDERGAST, CANICE. 2002. The tenuous trade-off between risk and incentives. *Journal of political Economy*, **110**(5), 1071–1102.

PYLE, BENJAMIN DAVID. 2023. Negligent Hiring: Recidivism and Employment with a Criminal Record.

RADBRUCH, JONAS, & SCHIPROWSKI, AMELIE. 2024. Interview Sequences and the Formation of Subjective Assessments. *Review of Economic Studies*, **Forthcoming**.

RAITH, MICHAEL. 2003. Competition, Risk, and Managerial Incentives. *American Economic Review*, **93**(4), 1425–1436.

REIFF, JOSEPH S, ZHANG, JUSTIN C, GALLUS, JANA, DAI, HENGCHEN, PEDLEY, NATHANIEL M, VANGALA, SITARAM, LEUCHTER, RICHARD K, GOSHGARIAN, GREGORY, FOX, CRAIG R, HAN, MARIA, *ET AL.* 2022. When peer comparison information harms physician well-being. *Proceedings of the National Academy of Sciences*, **119**(29), e2121730119.

RIVERA, LAUREN A. 2015. *Pedigree: How Elite Students Get Elite Jobs.* Princeton University Press.

ROGHANIZAD, M. MAHDI, & BOHNS, VANESSA K. 2022. Should I Ask Over Zoom, Phone, Email, or In-Person? Communication Channel and Predicted Versus Actual Compliance. *Social Psychological and Personality Science*, **13**(7), 1163â1172.

ROSE, EVAN K. 2021. Does Banning the Box Help Ex-offenders Get Jobs? Evaluating the Effects of a Prominent Example. *Journal of Labor Economics*, **39**(1), 79–113.

SANDVIK, JASON, SAOUMA, RICHARD, SEEGERT, NATHAN, & STANTON, CHRISTOPHER. 2021. Employee responses to compensation changes: Evidence from a sales firm. *Management Science*, **67**(12), 7687–7707.

SANDVIK, JASON, SAOUMA, RICHARD, SEEGERT, NATHAN, & STANTON, CHRISTOPHER T. Forthcoming. Should Human Capital Development Programs Be Voluntary or Mandatory? Evidence from a Field Experiment on Mentorship on Mentorship. *Management Science*.

SANDVIK, JASON J., SAOUMA, RICHARD E., SEEGERT, NATHAN T., & STANTON, CHRISTOPHER T. 2020. Workplace Knowledge Flows. *Quarterly Journal of Economics*, **135**(3), 1635–1680.

SCHULTHEISS, TOBIAS, & BACKES-GELLNER, USCHI. 2024. Does updating education curricula accelerate technology adoption in the workplace? Evidence from dual vocational education and training curricula in Switzerland. *The Journal of Technology Transfer*, **49**(1), 191–235.

SHAW, KATHRYN. 2009. Insider Econometrics: A Roadmap with Stops Along the Way. *Labour Economics*, **16**(6), 607–617.

SHUKLA, SOUMITRA. 2024. Making the Elite: Discrimination at Top Firms. *Working paper, Federal Reserve Board*.

SOCKIN, JASON, & SOJOURNER, AARON. 2023. What's the Inside Scoop? Challenges in the Supply and Demand for Information on Employers. *Journal of Labor Economics*, **41**(4), 1041–1079.

SOCKIN, JASON, SOJOURNER, AARON, & STARR, EVAN. 2024 (August). *Non-Disclosure Agreements and Externalities from Silence.* Tech. rept. 3900285. Social Science Research Network.

SOETEVENT, ADRIAAN, & ROMENSEN, GERT-JAN. 2024. The Short and Long Term Effects of In-Person Performance Feedback: Evidence from a Large Bus Driver Coaching Program.

SONG, JAE, PRICE, DAVID J, GUVENEN, FATIH, BLOOM, NICHOLAS, & VON WACHTER, TILL. 2019. Firming up inequality. *The Quarterly journal of economics*, **134**(1), 1–50.

STAIGER, DOUGLAS O, & ROCKOFF, JONAH E. 2010. Searching for effective teachers with imperfect information. *Journal of Economic perspectives*, **24**(3), 97–118.

STANTON, CHRISTOPHER T, & THOMAS, CATHERINE. 2016. Landing the First Job: The Value of Intermediaries in Online Hiring. *The Review of Economic Studies*, **83**(2), 810–854.

STANTON, CHRISTOPHER T, & THOMAS, CATHERINE. 2024. *Who Benefits from Online Gig Economy Platforms?* Tech. rept. National Bureau of Economic Research.

STARR, EVAN. 2019. Consider this: Training, wages, and the enforceability of covenants not to compete. *ILR Review*, **72**(4), 783–817.

SYVERSON, CHAD. 2004. Product Substitutability and Productivity Dispersion. *Review of Economics and Statistics*, **86**(2), 534–550.

SYVERSON, CHAD. 2011. What Determines Productivity? *Journal of Economic Literature*, **49**(2), 326–65.

SZERMAN, CHRISTIANE. 2022. The Labor Market Effects of Disability Hiring Quotas. *Available at SSRN 4267622*.

TERVIO, MARKO. 2009. Superstars and Mediocrities: Market Failure in the Discovery of Talent. *Review of Economic Studies*, **76**(2), 829–850.

TOPA, GIORGIO. 2019. Social and Spatial Networks in Labour Markets. *Oxford Rev. Econ. Policy*, **35**, 722–745.

WALDINGER, FABIAN. 2012. Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *Review of economic studies*, **79**(2), 838–861.

WALDMAN, MICHAEL. 2013. Classic promotion tournaments versus market-based tournaments. *International Journal of Industrial Organization*, **31**(3), 198–210.

WEIDMANN, BEN. 2024. Economic Decision-Making Skill Predicts Income in Two Countries.

WEIDMANN, BEN, VECCI, JOSEPH, SAID, FARAH, DEMING, DAVID J, & BHALOTRA, SONIA R. 2024. *How Do You Find a Good Manager?* Tech. rept. National Bureau of Economic Research.

WILES, EMMA, & HORTON, JOHN. 2024. *Misdirected Search and Artificial Intellgience*. Working paper, Boston University.

WU, HUGH, & LIU, SHANNON. 2022. *The Trade-offs of Letting Local Managers Make Hiring Decisions*. Working paper, WashU Olin.

WUCHTY, STEFAN, JONES, BENJAMIN F., & UZZI, BRIAN. 2007. The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316**(5827), 1036â1039.