NBER WORKING PAPER SERIES

PEOPLE, PRACTICES, AND PRODUCTIVITY: A REVIEW OF NEW ADVANCES IN PERSONNEL ECONOMICS

Mitchell Hoffman
Christopher T. Stanton

Working Paper 32849
http://www.nber.org/papers/w32849

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2024

Preliminary chapter in preparation 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. Comments are very welcome. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

People, Practices, and Productivity: A Review of New Advances in Personnel Economics
Mitchell Hoffman and Christopher T. Stanton
NBER Working Paper No. 32849
August 2024
JEL No. J01,M5

ABSTRACT

This chapter surveys recent advances in personnel economics. We begin by presenting evidence showing substantial and persistent productivity variation among workers in the same roles. We discuss new research on incentives and compensation; hiring practices; the influence of managers and peers; and time use, technology, and training. We emphasize two main themes. First, we seek to illustrate the interplay between these topics and productivity differences between people and work units. Second, we argue that personnel economics has benefited from exploration, which we think of as the willingness to use new data and methods to shed light on existing questions and to raise new ones. As many personnel studies use data from individual firms, we discuss external validity and provide concrete guidance on how to improve discussions of the generalizability of findings from specific contexts.

Mitchell Hoffman
Department of Economics
University of California, Santa Barbara
3016 North Hall
Santa Barbara, CA 93106
and NBER
hoffman2@ucsb.edu

Christopher T. Stanton
210 Rock Center
Harvard University
Harvard Business School
Boston, MA 02163
and NBER
christopher.t.stanton@gmail.com

1 Introduction

Personnel economics is the study of labor economics issues inside firms and organizations, applying labor economics 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. While non-personnel labor economists often focus on issues affecting workers, personnel economists frequently focus instead on returns to firms and the efficacy of organizational practices.

Lazear's research agenda grew out of the Chicago tradition pioneered by Becker (1993) and others in 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." Since the last handbook chapter by Oyer & Schaefer (2011), much has been learned in personnel economics on a wide range of topics. In the spirit of Lazear (2000b), we believe that 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.[1]

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 much greater statistical precision, to exploit exogenous shocks, and to measure variables that were previously unobserved.[2] Together, these features allow researchers to better address old questions and to address new ones.

[1]The engagement of personnel economics with different sources of data, methods, and questions is not new. Personnel economists pioneered the process of "insider econometrics" (Shaw, 2009; Ichniowski & Shaw, 2012) where data are collected in close collaboration with companies. Lazear's research agenda touched on many new areas (Lazear, 1999, 2006, 2012).

[2]For a discussion of different types of data that are being used insider companies for people analytics, including the uses of digital trace data, see Polzer (2022). The study of collaboration patterns represents one example of how new data sources allow new insight. For example, Bernstein & Turban (2018) use data from sensors to study face-to-face interactions and the impact on collaboration after 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.

For example, big data may help researchers address key questions about heterogeneity, such as what types of people respond most to different types of incentives or which workers value different amenities. Sensor technologies and data in digital systems may help to map collaboration patterns or estimate how information or different ways of working diffuse through an organization.

On methods, economists increasingly collect their own data using randomized controlled trials (RCTs) (Duflo, 2020). In personnel economics, this often takes the form of individual firms collaborating with researchers on RCTs, though there are also RCTs across firms, such as those done with workforce platforms. While certain labor topics are often hard to address with RCTs (e.g., the impact of trade liberalization or right-to-work laws), because personnel economics is concerned with outcomes of HR policies, it is often natural to consider single-firm RCTs. Separate from RCTs, economists increasingly collect their own data using surveys, sometimes measuring or shifting beliefs among workers or managers. Finally, 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.[3]

On questions, Lazear argues that personnel economics is primarily positive instead of 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, increasingly research—especially using RCTs—examines 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 gives access to new settings to test new hypotheses. New questions also emerge naturally from new sources of data, e.g., the use of surveys can be useful for understanding why employees take certain decisions or what they believe.[4]

Beyond exploration, a theme throughout the chapter relates to how new data and methods have allowed economists to understand the sources and implications of worker heterogeneity, especially heterogeneity in productivity or output. Studies of frontline workers, managers, and teams all show

[3]Machine learning is also affecting the behavior of HR practitioners, thereby providing opportunities for research.

[4]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.

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.

Our chapter proceeds as follows. Section 2 walks readers through stylized facts showing tremendous differences in productivity or output exist across workers who do the same job. This productivity dispersion appears to be ubiquitous across different types of jobs – ranging from entry level to highly-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 (and managers) for closing gaps or magnifying them. While this theme helps to organize the chapter, the focus on productivity *dispersion* is not the only compelling aspect of personnel economics. Consider the thought experiment where there are no differences in workers' types or backgrounds. In this case, there are still very interesting questions about the effects of contracts that change workers' effort, affecting firm performance. Questions getting at workers' and managers' beliefs, the processes firms use, how studies themselves should be conducted, among others, are all interesting questions in their own right.

Section 3 discusses new work on incentives and compensation. This has historically been the "bread and butter" of personnel economics, and important work continues to be done. Understanding how different incentive schemes influence workers' sorting between firms or jobs continues to be an important area of focus. Other work has examined how incentives change workers' effort while holding fixed the set of workers in a firm. 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, like teaching, medicine, and law. Emerging evidence also suggests that workers' responsiveness to the type of pay varies based on whether they are relatively more or less productive. In particular, recent research suggests that highly productive workers are relatively more responsive to performance pay, while less productive workers are more responsive to the level of pay. 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 the tradeoffs that subjective evaluations can introduce when objective measures provide an incomplete characterization of performance-relevant activities. We apply our discussions of incentives to two recent streams of literature in labor economics writ-large on the

role of demographics and competition in the labor market. An especially important focus area is monopsonistic competition, where firms set wages strategically based on market conditions. While we discuss how market structure may influence incentive provision, we suggest that firms' incentive policies in monopsonistic competition likely depend on effort responses to wage levels (through an efficiency wage channel) and the extent of worker heterogeneity that may drive firms to use output-linked pay as an attraction and retention device for targeting more productive types of workers.

Section 4 discusses research on hiring, an area which has seen substantial growth since the last handbook chapter. We discuss networks and hiring, especially 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 shows that different hiring methods such as employee referrals and job tests often provide relevant information about the quality of candidates. Different hiring methods often matter substantially for worker quality and firm outcomes, but 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 they play and what they do have proven to be quite influential. Many studies now indicate that individual managers strongly influence their subordinates' productivity and that social skills and people skills appear to matter for managerial productivity. 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. Two mechanisms behind peer spillovers are peer pressure and knowledge spillovers. We discuss evidence on the magnitude of each in turn. Frictions appear to prevent peer spillovers in some settings. Some forces, 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 those doing autonomous tasks. While the literature on peer effects tends to isolate spillovers in settings where workers produce individually,

this is just one aspect of what happens in more general team production, raising important questions about the management of teams and how team membership affects 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, in total, have received slightly 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. Some workers appear more productive than others because of how they spend their time. Can firms learn these best practices and disseminate them, possibly using technology? A core issue is how new technologies will interact with productive heterogeneity and the boundaries of the firm. Will AI lift lower tail or performance, as has been shown in early studies, or will it eventually augment incredibly high performers? Early evidence suggests that many generative AI tools appear to differentially lift the productivity of lower-performers. An open question is how tools to wrangle and leverage digital trace data, combined with gen AI, will change the role of managers. Will managers become more central to coaching workers or less (if AI tools can do that coaching)? Technology may also change the boundaries of firms, by enabling access to workers in different labor markets or in remote settings. We discuss several papers about how firms use online labor markets to find workers and manage them, emphasizing that market design choices by platform operators often have profound effects on matching. Finally, the literature on training has only had a limited role for firms – much of it is abstract, yet most HR departments include a learning and development function and firms are estimated to have spent $100 billion in 2022 on training-related expenses. We mention several recent studies that touch on different aspects of firm training, including how to measure returns to firm training expenses and different processes that can be used to implement training programs.

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 a treatment and general equilibrium effects. Overall, we suggest that personnel economics can help to understand aggregate labor market trends, like polarization between superstar and other firms, by better understanding what leading firms and organizations do in practice.

Where can personnel researchers continue to advance? In our view, several important common themes arise throughout the chapter and call for more work:

- Analyses that help understand *heterogeneity*, i.e., why do workers perform differently on the job and what types of workers and firms are most likely to benefit from a practice? For example, what types of workers and firms respond positively to performance pay, and what types do not?
- Analyses that help understand *worker and firm beliefs* about the returns to different practices. For example, what are beliefs about the value of good managers, and about who good managers are?
- 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 the field of personnel is vast, and many studies touch on topics related to firms, we unfortunately have to focus. Our choice of coverage has largely been driven by what is novel in personnel combined with the existence of other chapters in this handbook that overlap with some personnel themes. The biggest omission is related to work at the nexus between personnel economics, discrimination, and diversity, as these topics receive coverage in other handbook chapters. We also provide relatively limited background information on several topics that are the focus of other chapters.

Our chapter is primarily aimed at PhD students and researchers interested in questions through the lens of the role that firm policies and practices have for worker behavior and outcomes. After all, how individuals behave in long-run, high-stakes, repeated interactions is informative about one of the most important relationships in workers' lives.

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 there is substantial, persistent productivity variation across 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 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 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, and how, 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. Productivity differences across workers are found to be comparable in magnitude to those across firms or plants. Productivity variation also appears to dwarf pay differences for workers at the same level, at least in the contexts studied.[5]

Table 1 displays estimates of productivity dispersion at the individual level. Estimates are sorted by the approximate skill level required by the job, in ascending order. The "Setting" column describes the context and productivity measure for each paper. As papers report different statistics to capture productivity dispersion, the third column reports the "Dispersion Statistic" used in the study. The final column reports the dispersion estimate. An asterisk in the final column means that the estimates use a shrinkage or regularization technique to isolate persistent, person-level differences that take into account noise in estimation.

[5]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.

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

Moving beyond relatively routine or more manual occupations, Lazear *et al.* (2015) benchmark worker and manager fixed effects in a technology-based services firm with relatively little incentive pay. They use a manager rotation design and find that the standard deviation of worker fixed effects is approximately 13% of the grand mean of output-per-hour. Assuming normality, this implies a roughly 20% difference in output between the 75th and 25th percentile frontline worker. The standard deviation of managers' fixed effects on output, when scaled over an entire team, are found to be 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 on top of 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% 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 altering hiring or training practices be done at low enough costs such that it would be a net benefit to address productivity dispersion?

On hiring practices, which we discuss in Section 4, evidence varies about whether firms can identify relative performance differences across workers. Sandvik *et al.* (2021a) show that interviewers' pre-hire assessment scores correlate with workers' on-the-job sales. 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 appear to hire workers who are ex-ante likely to be poorly matched to the job. Why? Three possibilities are: a) the mapping between productivity and assessments is not well

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%*

known to recruiters, b) there are agency conflicts or bias issues such that recruiters want to fill seats even with sub-par candidates, or c) firms face constraints from external labor pools such that accepting a lower performer may be better than raising wages to improve the recruiting pool or letting a slot go unfilled.

Although prior work has shown the ability to identify relatively more productive workers in some settings, in other contexts it appears challenging to identify high performers. For example, Staiger & Rockoff (2010) report that a standard deviation change in New York teachers' persistent value added scores are equivalent to 0.15 and 0.12 of a standard deviation gain in student test scores in math and reading, respectively.[6] Based on the link between student test scores and future earnings, differences in teacher output are worth hundreds of thousands of dollars to society. But predicting teacher effectiveness in the hiring process has proven to be enormously difficult. The best predictor seems to be prior teaching experience, but this does not help filter new entrants into the profession.

Moving to higher earning occupations, Chan & Chen (2022) show that the total cost of emergency department visits across Veterans Affairs (VA) hospital providers appears to vary substantially more than differences in provider salaries. Because VA hospitals are captive health organizations,

[6]Value added scores are the part of student test score changes that can be attributable to an individual teacher.

cost differences matter for the efficiency of service delivery. Chan & Chen separately benchmark differences in costs based on quasi-random patient assignment. For both medical doctors (MDs) and Nurse Practitioners (NPs), the difference in spending between the 75th and 25th percentile of the respective distributions totals about $650,000 per year. Additionally, the average MD (who has more intensive training but also has a higher salary) is more cost effective than the average NP; at the mean, the additional pay for doctors is justified through lower costs to the captive 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 amounts of training. Even more stark, the variation within each professional group is larger than the average difference between MDs and NPs. Finally, while emergency department practitioners deal with a variety of cases, there is substantial variation even while holding fixed commonly undertaken tasks. For example, Chan *et al.* (2022) show that the diagnostic accuracy of chest X-ray reads varies substantially between the 90th and 10th percentile of the distribution; they attribute this variation to skill differences and conclude that improving skills is a more effective policy than others that impose treatment or process guidelines.

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 setup is used as motivation, and we will comment through the text when it needs to be expanded, such as when output can only be linked to a group, rather than an individual.

Workers' effort choices given employment depend on the firm's compensation policy function,

$w(y)$, and their individual cost of effort function, $c_i(e)$.[7] Compensation policy may include fixed wages, individual performance pay, group incentives, or tournaments, meaning that the argument in $w(.)$ is general. While effort choices depend on $w(.)$, so too do employment choices.

A huge theoretical literature in contract theory has focused on a maximization problem where firms set $w(.)$ to maximize some overall production function $f(.)$ less wages,

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

Most of the action has focused on how changes in $w(.)$ change e (i.e. incentive compatibility) and potentially change the set of workers present at the firm (i.e. individual rationality). Until recently, there has been much less focus on how hiring practices or the selection and training of managers and workers influence h and T. Hiring and selection of workers and managers is particularly important for how the individual rationality constraint in traditional contracting problems operates in practice. Practices that alter the IR constraint and workers' selection are likely enormously important, as market-level variation across workers is almost surely larger than the within-firm variance that 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 important, practical problem for firms. Although our discussion begins with recent work on incentives, we explore other channels through which firms make important decisions about managing people.

3 Incentives, Compensation, and Labor Markets

Personnel economics research has demonstrated that firms' choices of incentive and compensation policy can affect output. The underpinning for this work often comes from agency theory. One implication of the basic agency model is a risk/incentive tradeoff. Input-based contracts (i.e., a standard fixed wage) may provide only weak motivation for workers to exert effort, while output-

[7] Typically worker heterogeneity enters either through skill heterogeneity or cost of effort heterogeneity, but there may be settings where it is possible to distinguish between these two. For example, a highly skilled plumber may be able to do any job after decades of experience but scaling up effort may prove more difficult compared to a newcomer with less accumulated wear-and-tear.

based contracts (i.e., pay-for-performance) cause workers to bear risk from exogenous variation beyond their control (Lazear, 2018). Due to measurement problems from un-modeled production considerations (like agents' local information advantages, e.g. Prendergast (2002)), Oyer & Schaefer (2011) argue that personnel economists should move beyond studies of risk/incentive tradeoffs and optimal incentive contracts. Since Oyer & Schaefer (2011), a rich literature has provided evidence on the different margins by which incentives may affect worker behavior. This literature has been more empirical and less focused on characterizing optimal contracts. It has examined many "non-standard" responses, like behavioral models, and incentive provision in "non-standard" settings such as compensation for innovators or groups of workers, subjective evaluations, and more. Given its vastness, our discussion will require us to omit some core contributions. In particular, because of excellent recent surveys on executive compensation by Frydman & Jenter (2010) and Edmans *et al.* (2017), we have chosen to focus on incentives for lower-level employees.

Conceptually, the form of incentives can affect productivity through three channels. First, incentives can affect effort. In the baseline model, performance pay (piece rates, tournaments, or subjective bonuses) raises effort relative to a salary or hourly pay since the agent's payoff is linked to output. Second, incentives can affect the human capital or ability of workers at a firm through sorting (e.g. under performance pay, less able workers leave and more able workers 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.[8]

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, there was substantial research that workers in routine jobs (e.g., fruitpickers, tree planters) respond strongly to performance pay. Research over the last 15 years shows that workers in non-routine jobs also respond strongly to performance pay, especially high-performers.

[8]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), but we would welcome additional research on how high-powered incentives influence workers' human capital acquisition.

There are many settings where firms do not write output-based contracts. In the second part of the section, we discuss work on the level of pay as an incentive, subjective performance evaluation and career incentives, and how firms set standards or requirements for workers to supply effort.

Third, standard models tend to take as given outside options for workers, but the literature on monopsony (reviewed elsewhere in this handbook) suggests that an important future area of research is to understand how incentive provision varies with outside labor market conditions in practice. External factors, like aging societies, will also have implications for whether firms will substitute between career incentives and short-term contracts.

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, in which a switch from hourly wages to a piece rate per successfully installed windshield increased productivity by 44 percent, in large part due to sorting (Lazear, 2000a).

We spend some time on recent evidence from other contexts in the spirit triangulating across different workers and settings. For example, theory recognizes that workers who do a variety of different tasks will have multi-task incentives, meaning that incentives for doctors or lawyers or teachers may differ from those whose job focuses on a single dimension of output. Second, workers in different jobs may have varied cognitive skills or attention to financial details: artists may not pay attention to incentives or some workers may not understand obfuscated links between outputs and pay.

To illustrate what has evolved in the literature, we begin with results on performance pay for teachers. Although teachers often work in the public sector, there are several good reasons that studies of their incentive responses should interest personnel economists. First, teacher output measures — value added scores — can be standardized across settings, like schools or districts, enabling economists to see not just who responds to incentive changes in one establishment or school, but to track who moves to different schools and how they respond after moving. Second, teachers'

output measure, based on gains in student test scores, is one that economists understand (even if there are debates about whether value-added measures are the right performance indicator).[9] Third, the bureaucracy of most school systems means that complementary inputs are unlikely to vary much between settings in the same local area or labor market,[10] but like many personnel studies, there are questions about whether findings from areas with certain cultural institutions or behavioral norms translate to other contexts. Finally, and perhaps most important, teaching is a non-routine task that draws in workers who likely have intrinsic or pro-social motivations, two factors that have been examined by studies questioning the validity of results on performance pay in simpler, routine settings.

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

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

The 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 are non-wage amenities for each job.

Differences in output between the fixed and variable pay jobs can be expressed with three terms: sorting on ability, sorting on the individual's own increase in effort under performance pay, and the average effort increase under performance pay. They estimate these effects using a creative 2-part experiment in a chain of private schools in Pakistan. Prior to the experiment, the standard

[9] For readers interested in best practices to deal with gaming (and a discussion of contracts, like pay for percentile, that are often used in practice), see Neal's (2011) excellent handbook chapter on teacher incentives.

[10] The main difference across settings that one might worry about is student or parent quality. While value added measures are based on achievement gains rather than levels, there may be sorting based on match effects where some teachers may have a comparative advantage in targeting certain parts of the achievement distribution.

deviation in teachers' value added, which they use as a measure of 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.

The first phase of the experimental design involved eliciting individual teacher preferences for what fraction of their annual pay raises they would prefer to be either performance-linked or instead based on a flat increase. To make the elicitation incentive compatible, there was a roughly 1-in-3 chance that their contract choice would be implemented. 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 that chose a larger fixed raise. This represents about 1/3 of a standard deviation in baseline productivity.

The second phase of randomization was done across schools. Some schools were randomized to implement the individual teachers' choices. Other schools were randomized into either a fixed raise that increased teachers' base salary by 5% or into a performance-based raise contract. In the performance contract, raises were determined using a pay-for-percentile scheme within each school.[11]

This second phase has two attractive features for inference. First, there is mobility across schools in this network, allowing the authors to estimate how teachers sort to schools with different contracts. After implementing the different contracts in the second phase, teachers who moved from flat pay to performance pay schools had value added scores that were 0.064 standard deviations above the mean. Not everyone moves. In total, re-sorting of teachers across schools was responsible for a 0.022 standard deviation difference in baseline value-added between performance and flat pay schools. These effects happen over a one year horizon and are averaged over movers and stayers, suggesting that the long-run effects are likely bigger.[12]

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

[12]Some other studies of market-level introduction of incentive pay also document positive sorting. Biasi (2021) studies a Wisconsin policy change that enabled districts to implement flexible pay, breaking from lockstep, seniority-based compensation models. After the reform, sorting of high quality teachers to districts with flexible pay increased, raising student achievement. Here the sorting channel appears to explain about two-thirds of the relative increase in achievement between districts that implement flexible pay. Leaver *et al.* (2021) use a similar design to study pay for performance effects for teachers in Rwanda. The first stage of their experiment involved randomization of advertised contracts as pay for performance or fixed wage at the district level. 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 appear larger than sorting on ability; teachers who do sort into the performance pay contract appear to have lower intrinsic motivation, but they increase their on-the-job effort. Contrasting this design to others, it is possible that the lack of contract variation within a district labor market reduces sorting effects because of search, moving, or

Second, the design allows Brown and Andrabi to compare performance changes for those that get their preferred contract versus those that do not. Value added increased by 0.09 standard deviations for teachers who wanted and were assigned the performance raise contract. In contrast, performance pay only increased value added by 0.01 standard deviations for teachers that preferred the fixed raise contract. They interpret the difference as arising from effort response heterogeneity.[13]

Given concerns about multitask incentives and crowdout 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 it likely would have improved aggregate outcomes. Overall, the paper's creative design shows both heterogeneity in preferences and responses to performance-pay contracts. We are aware of few other papers that combine pre-treatment measures of preferences and actual sorting behavior to uncover heterogeneous effects of incentives. Methodologically, the clever experimental design may offer guidance to other researchers about how to advance the frontier of estimating incentive responses while allowing for different workers to respond on different margins. Understanding these margins is likely useful for understanding why different firms might vary their incentive offers to appeal to distinct groups of workers.

Other work on incentives for non-routine jobs shows responsiveness as well. Clemens & Gottlieb (2014) use a 1997 change in the way Medicare (a nearly universal health program for US residents over 65) reimbursed doctors. Prior to 1997, reimbursements rates for a given procedure had a multiplier that varied across 210 payment regions. In 1997, Medicare cut the number of payment regions to 89, meaning that in some areas prices for a procedure rose, while in others prices fell. Although this is not exactly a principal-agent analogy, about 60 percent of doctors were self-employed at the time of the change, while approximately 85 percent of those who worked in group practices

commute costs for workers to take jobs in other areas.

[13]Other work on performance pay for teachers in developing countries show substantial performance gains, but the evidence on pay-for-performance in advanced economies is much less clear. For example, Muralidharan & Sundararaman (2011) find that introducing a bonus of 3% of salary tied to student achievement lifts performance in rural Indian schools by 0.27 and 0.17 standard deviations in math and language tests. Gains appear to come from increased teacher effort and are larger when incentives are based on individual teachers rather than provided at the school level. On the other hand, Fryer *et al.* (2022) argue that most teacher incentive pilots in the US fail to show the same level of achievement gains. Several possibilities exist for these discrepancies, including implementation questions, designs that shut down sorting, and environments that start with higher levels of monitoring and norms for basic good behaviors, like attendance.

had compensation that was tied to patient care revenue. Using this variation in prices, Clemens and Gottlieb find a long-run elasticity of care supplied with respect to prices of 1.5. Procedures increase most in elective categories and in areas where physicians can invest in in-office technologies, like MRI machines, where a relatively intense practice style can be justified. In this case, physicians make fixed cost investments to lower their marginal costs, giving rise to higher margins on care. Impacts on health outcomes after the price changes appear limited.

As Bartel *et al.* (2017) note in a study that reduced the power of incentives for lawyers at a leading international firm, the problem with objective incentives is often that they work too well, crowding out effort on tasks like mentoring or building culture that the firm may want to encourage. 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 reimbursement rates (Clemens & Gottlieb, 2014), high-skilled workers in non-routine tasks respond to incentive pay.

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.

Other tests to reduce medical expenditures through incentives often show evidence of gaming. Alexander makes the point that these tests are often done where health providers can sort patients to the scheme or out, meaning they do not approximate how incentives would look at scale when all facilities and procedures would be subject to the scheme. Alexander (2020) 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 patients—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.

Effects of performance pay on the intensive margin Much of the work on incentive pay has looked at the introduction of incentive pay on the extensive margin of moving from fixed to variable pay. However, the magnitude of incentive pay also likely matters. Sandvik *et al.* (2021b) analyze incentive pay changes on the intensive margin in a quasi-experiment involving a reduction in commissions at a sales call center. One of six divisions reduced commissions. In the one that did, overall pay was predicted to drop by about 7% on average. After implementing the change, turnover increased significantly among the most productive workers in the treated division compared to controls. Effects on turnover were relatively muted for less productive workers. The change in commissions did not impact sales levels or reported effort. The authors speculate that income effects and price effects move in opposite directions, cancelling out any effort adjustment. However, the attrition of the most productive workers meant that the revenue lost from these employees' exits (compared to to replacements) was greater than the wage bill savings for the firm.

In a consulting firm featuring workers with higher levels of formal education and who do less structured work, Krueger & Friebel (2022) study the equalization of pay across divisions. Prior to the change, some divisions had higher base pay and lower bonuses, while others had more bonus pay relative to base pay. Divisions with higher bonus pay saw their bonuses cut, while base pay increased. In contrast to the pure reduction in commissions in Sandvik *et al.* (2021b), effort fell substantially for the most productive workers when variable pay was cut. These productive workers were also most likely to leave the firm. Less productive workers and new entrants to the firm had limited responses. Two possibilities drive the differences in results compared to Sandvik *et al.* (2021b): i) sales workers in a call center have less discretion than those in consulting and ii) the firm in Krueger & Friebel (2022) raised base pay to compensate for the loss of bonuses, potentially muting income effects.[14]

A common theme across studies in this section is that high-powered incentives can help with

[14]Coviello *et al.* (2022a) study an incentive change in a retail call center where commission thresholds increased, requiring more effort to maintain income. In their setting, workers can inappropriately match customers to some items, resulting in returns that hurt 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.

the attraction and retention of workers, especially more productive ones who appear to be most sensitive to the existence and level of performance pay. Beyond the studies of teachers, however, a limitation with these results is that workers' outside options are often unobserved. There is likely an opportunity for future researchers to link together data on worker flows between firms as a function of how firms alter incentives. We discuss this more in Section 3.3 on external labor market conditions.

3.1.2 Behavioral models and non-monetary motivation

A behavioral literature has tackled questions related to incentives using a wide variety of models outside of the agency-theoretic paradigm. Some results suggest that factors like inequality aversion, social preferences, or stress may reduce the efficacy of high-powered incentives. Other factors, like loss aversion or limited cognition about complexity, may be manipulated by firms to magnify incentive power. The propensity to exhibit behavioral biases is also likely heterogeneous across people, suggesting the possibility that, in the long run, firms will offer different incentive menus that allow workers to sort on behavioral factors, not just abilities and costs of effort. Calibrating the relative importance of behavioral factors for how workers sort into and out of different incentive schemes is a potential area of future research. To understand how these studies might be conducted, we briefly walk through a baseline model and some tests from DellaVigna & Pope (2018) that consider different 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 how one could estimate a particular subject's behavioral biases. We then discuss results of behavioral incentive concepts that have been tested 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 is 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 capture different behavioral responses.

Here we describe some of the extended models:

- Crowd Out: Crowd out is the concept that small monetary incentives get in the way of intrinsic motivation, captured by the first order condition for effort $e^* = c'^{-1}(s + \Delta s_{CO} + p)$. The parameter Δs_{CO} captures any effort change when going from no piece rate to a very low one with limited direct incentive effects.

- Altruism and Warm Glow: 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: 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: 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: 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 their real effort experiment lend support for both classical incentive effects (piece rate pay raises output) and some behavioral factors. Strikingly, even small piece rates appear to generate substantial effort responses, which are larger than the average effects of gift exchange, meaningful work, and social comparisons. The authors find no evidence of crowd out with small incentives, but it is possible that the task they utilized (repeatedly pressing keys) and the setting bias the results against finding crowd out.[15] Finally, the loss-framed treatment is slightly more effective than the gain-framed one, but the main effect of both treatments relative to controls appears to be introduction of a threshold that shifts effort so that workers qualify to earn (or keep) the bonus.

The result on the importance of setting a target suggests that how firms set goals is likely to be an important incentive device which has received very little attention. An exception is Kuhn & Yu (2024), who discuss the voluminous literature in psychology on goal setting. But isolating the importance of goals is challenging because reaching a goal often implies an increased monetary reward. To separate the effect of rewards from goal achievement, Kuhn & Yu (2024) study small retail teams where the pay schedule becomes kinked (but does not have a discrete jump) upon reaching a goal. The kink means that the marginal return to effort just after achieving a goal is higher than the marginal return just before reaching the goal. If the goal is psychologically important, one would expect to observe bunching near the achievement level, whereas if the incentive effect dominates, workers should increase effort beyond the kink because of the higher returns. Kuhn & Yu (2024) find significant bunching, suggesting that teams get intrinsic rewards (or some symbolism, praise, or extra positive attention) from achieving goals.

To what extent can firms leverage other behavioral factors in practice? 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.[16]

Several prominent studies test 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

[15]Other work has suggested that incentives may adversely impact an agent's perception of the task, eroding intrinsic motivation in the long-run (Benabou & Tirole, 2003; Gneezy *et al.*, 2011).

[16]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.

to gain-framed incentives or no incentives. In their study, the loss-framed incentive means that bonuses are pre-paid for achievement gains and then clawed back if teachers fail to improve test scores. Brownback & Sadoff (2020) study loss-framed incentives for community college instructors and find 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) implement a field experiment in which bonuses in a car dealership network are loss-framed. In contrast to their expectations, the treatment effect was a 5 percent reduction in sales. The authors attribute this to gaming behaviors, in which workers neglected uncompensated tasks to attend to the task with loss-framed incentives. In the spirit of connecting to the prior literature on sorting, we would be very interested to see how the introduction of behavioral incentive frames influences sorting into and out of firms.

Other framing strategies may target crowd out, meaning, or intrinsic motivation. Implicit in this work is the notion that some workers will be intrinsically motivated or will find greater meaning in work than others. Machiavelli makes a similar point when writing about mercenaries, 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." Papers in this area often assess whether awards and recognition, communicating task significance, and paying are substitutes or complementary practices. For example, Ashraf *et al.* (2014) show that a "star" 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,[17] but they also show that financial incentives actually boost the performance of the sellers who are likely to have the highest intrinsic motivation. On the other hand, 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 may 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

[17]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).

groups may be less inclined to have intrinsic motivation at baseline.

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 subjects' own expected performance on a multiplication test and then offer 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. Hoffman & Burks (2020) study work overconfidence in the context of long-haul truckers. They show that new workers systematically overpredict their productivity on average, even when given strong incentives for predictions, and they learn much slower than predicted by Bayes' rule. Overprediction is especially consequential since 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. In a field experiment of managers working at a food and beverage store chain that has tournament-like pay based on store performance, Huffman *et al.* (2022) find that managers have overly-positive memories of past performance and make overconfident predictions about future performance. Overconfidence and lack of updating likely means that managers will continue to supply effort in tournaments even when it may be optimal to give up.[18]

More complicated incentives may also lead to varied reactions across different workers. For example, Abeler *et al.* (2023) find that differences in cognitive ability affect responses to incentive schemes. A field experiment on warehouse workers shows that the introduction of a dynamic incentive scheme with ratchet effects caused workers to increase effort, even though, by design, it was optimal to decrease effort. However, complex schemes may lack salience, and prior work has shown that reminders about a complex scheme can get workers to increase effort (Englmaier *et al.*, 2017). This work suggests that complexity is an important design consideration for workplace incentives that has not received enough attention in field studies.

[18]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).

3.1.3 Incentives for teams and innovators

Team incentives: The prevalence of team activities in the workplace has increased, with estimates that workers' time spent collaborating has increased by 50 percent over the past 2 decades (Cross *et al.*, 2016). Over the period 1980 to 2012, jobs requiring high levels of social skill grew by nearly 12 percentage points as a share of the U.S. labor force, suggesting that the importance of team-based work is increasing (Deming, 2017). Evidence from scientific papers and patents shows that teams increasingly dominate knowledge production (Wuchty *et al.*, 2007). As teamwork is becoming more important, personnel economists may need to ask whether individuals or teams are the right unit of analysis for future studies. Forming effective teams—and making inference on effective teams—is a complex problem due to the potential number of combinations that exist across coworkers. While in section 5.2 we focus on aspects of multi-agent production and team effectiveness, such as peer learning, mutual monitoring, and coordination, here we focus on team-based incentives.

Workers in teams face an inherent tradeoff between spending time on individual work and spending time helping others (Lazear, 2018). A potential solution 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 using individual incentives because of free-riding (Holmstrom, 1982). In linear production functions, individual benefits from effort provision are proportional to the inverse of team size, while each individual bears their own costs of supplying additional effort.

Empirically, however, team incentives have proven to be effective in various field conditions. A famous example comes from Knez & Simester (2001), finding that the introduction of a firm-wide bonus to all 35,000 hourly employees of an airline raised employee performance. The organization of employees into autonomous work groups enabled the firm to induce mutual monitoring among employees in smaller groups, suggesting the importance of the interaction between social and team incentives.

Bandiera *et al.* (2013) make this link clear in a study of team-based rank incentives and tournament pay in a soft fruit producer. The authors motivate a field experiment with a model that nests 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

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

The weight $\pi_{ik} \in [0,1]$ which worker i places on the earnings of worker k in her team effectively boosts the incentive power p as each worker internalizes the effect her effort has on the other team member. 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.* (2013) organized workers into teams of five; they were free to change teams once a week, enabling the authors to study how 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 were displayed, meaning that the absolute productivity level of each team was posted 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, but the introduction of tournament incentives increased productivity by 24 percent. Both treatments led to significant changes in team composition relative to the control condition, primarily arising from sorting on ability. In the rank condition, ability rather than friendship sorting weakened social incentives for average teams, reducing productivity. In the control group, these friendship ties allowed even average groups to internalize free-riding spillovers, but these ties broke apart when rank information was displayed. In contrast, the tournament regime was sufficiently powerful to elicit greater effort, enough to offset any increase in free riding within teams.

Friebel *et al.* (2017) explore the efficacy of team-based bonuses when the firm determines team composition. In a retail bakery chain, shops employed an average team of seven employees, who

carried out a variety of interconnected tasks such as 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 effect of the bonus on total effort decreases with the share of non-incentivized team members, suggesting that monetary incentives explain the effect, not peer pressure.

What is surprising about this literature is how such small incentives matter in practice. We would like to see more empirical work on the effect of team incentives on productivity, especially in service industries and knowledge work, which constitute an increasing share of economic activity, but for which measurement problems make inference challenging.

Incentives for innovation Motivating innovation or encouraging creativity tends to involve writing long-term contracts that eschew short-term performance measures. Emerging evidence supports the view that innovators should be given long leashes to experiment and fail early on because innovation requires exploration of new approaches that have 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 exhibits substantial tolerance for early failure and rewards for long-term success. In a laboratory experiment, Ederer & Manso (2012) find that subjects are more likely to innovate under such a failure-tolerant incentive scheme compared to fixed-wage and standard pay-for-performance contracts. In the experiment, subjects make decisions about the location of a lemonade stand, selection of products, and price for each of 20 periods. They observe profits and customer feedback after each period. In addition, subjects 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. Subjects threatened with termination exhibit decreased exploration activities, but a severance payment treatment help to alleviate these negative effects.

These contract features appear to matter 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, rewards long-term success, and gives its appointees greater freedom to experiment. This is some of the best evidence on how incentives influence innovation in a high-stakes regime. Understanding how funding agencies set incentives for science and innovation, either explicit or implicit, is likely to have profound consequences.

The time it takes to innovate in failure-prone settings can also change the sorting into innovative activities. For example, Venture Capital investments often leave founders holding risky equity and earning low salaries at the beginning of a startup's lifecycle, while hitting innovation or product milestones leads to competitive salaries and rewards. Ewens *et al.* (2023) use the fact that the introduction of cloud computing sped the time to resolve uncertainty in some industries but not others. They show that increasing the speed with which entrepreneurs realize whether an innovation will succeed increases the likelihood that highly-compensated, highly-skilled non-entrepreneurs will select into innovative activities.

Some intermediate reflections The literature on performance pay and output-linked incentives is vast, possibly reflecting somewhat diffuse priors about the applicability of different levers managers may pull in different settings. We refer readers to recent work by Georgiadis & Powell (2022) who re-frame the problem by asking what experiment a manager can run to improve upon an existing contract. Their results and setup will likely be useful for researchers who want to design experiments to understand incentive responses. Their paper provides a (relatively technical) characterization of what economic primitives can be recovered with a/b tests of incentive contracts.

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, many firms that could write output-based contracts do not, suggesting the level of pay also provides incentives. Important work has been done 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 device and comment on constraints that firms may face when setting pay levels, including those stemming from perceptions of inequity across workers. We then discuss subjective performance evaluations and incentive-bias tradeoffs. Career incentives raise questions of

how firms learn about workers and make promotion decisions.

3.2.1 Pay levels

In the prior subsection, the evidence points to highly productive workers being most sensitive to performance pay. How does this compare to the level of base pay? Lazear *et al.* (2016) analyze this question in the context of fixed hourly pay and find that the least productive workers are likely most sensitive to the level of fixed pay. That is, raising pay levels likely improves output and retention of workers who are marginal to the firm. They reach this conclusion by comparing the output of workers doing the same job across states with different unemployment rates during the great recession. A higher unemployment rate reduces workers' alternate opportunities, making the current job more attractive. They find that lower productivity workers increase their effort most when the unemployment rate increases. As a result, average output per worker rises, with the greatest increase at the bottom tail of the productivity distribution.

Coviello *et al.* (2022b) examine the effect of the minimum wage on worker effort across the distribution of worker productivity. In an extension to efficiency wage models, their theory allows for the firm to offer both a fixed wage component and performance-linked pay. 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 most, dominating the reduced power of pay-for-performance incentives. However, the efficiency wage channel only appears to matter in cases where there is a high supervisor-to-worker ratio, consistent with the ability to monitor. Although workers become more productive and retention improves, the increase in costs outweighs the improved productivity for their 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 likely outside option. It is also possible that improving job quality can reduce stress, which has been shown to increase productivity in some settings. For example, Kaur *et al.* (2021) find a significant effect of workers' financial concerns on their productivity. The authors stagger when expected wage payments are

distributed among a sample of 408 male low-income Indian piece rate manufacturers, 86% of whom report experiencing at least some financial worry. This design eases financial strain while holding constant both wealth and the incentive to work. Paying workers earlier allows them to pay off debts and purchase essentials, increasing output by 7.1% (0.12 SDs) and reducing the likelihood of attention-related mistakes by 0.08 SDs. These effects are concentrated among more financially-constrained workers, whose productivity increases by 0.22 SDs, and whose attention-related mistakes decrease by 0.23 SDs.[19]

Given the heterogeneous response of different employees to the level of pay, firms may find it profitable to attempt to target those workers who are most responsive to different types of incentives. However, firms may be constrained when adjusting pay 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 significant effects of peer wage changes on separations, largely driven by increased turnover when a coworker becomes relatively more highly paid. Similarly, Breza *et al.* (2018) utilize data from a month-long field experiment of manufacturing workers. They show that wage inequality reduces output by 0.24 standard deviations and attendance by 12% when productivity differences are difficult to observe between coworkers. However, in settings where individual output was easily observable, pay inequality had no discernible effect on these metrics, suggesting that within-firm inequality is more acceptable to workers when they can also observe performance differences.

3.2.2 Subjective evaluation, career concerns, and employer learning

In many settings, an individual worker's performance is challenging to quantify, necessitating the use of other types of incentives. This is especially true in jobs that involve complex, non-routine tasks and require a high degree of creativity, interpersonal skills, and/or judgment such as management, research and development, and professional services.[20] 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 discuss recent evidence on alternatives to objective

[19] Other work on the timing of pay, like Parsons & Van Wesep (2013), analyzes workers' self-control problems and suggests that firms can save on wage costs by timing pay to match workers' consumption.

[20] 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.

performance measures. We review some recent work on tradeoffs and suggest pathways for future work, especially as text-as-data techniques become more 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.[21] Two opposing forces—the introduction of bias and the ability to provide performance-based incentives—compete to determine the net effects of subjective evaluation. The literature has identified two sources of bias. One is leniency, where an agent may inflate ratings, like a teacher with grade inflation. In these cases, relative performance mechanisms like a forced curve can recover accurate rankings (Frankel, 2014),[22] but comparisons between raters or across different divisions may be challenging. A different form of bias reflects differential treatment between workers such that some types are less likely to receive high ratings given the same performance measure. Many firms do not use relative mechanisms to address leniency, suggesting many value the ability to make absolute comparisons across people or teams.

Frederiksen *et al.* (2020) benchmark the degree of leniency bias in subjective performance evaluations across supervisors. An empirical difficulty is that subjective evaluations are often used because of the absence of individual performance measures, which makes separating supervisor performance effects from leniency bias challenging. To overcome this problem, they use data from a large Scandinavian services firm that has 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's subjective rating. Mapping ratings to earnings, moving from a supervisor at the 10th percentile to the 90th percentile of the leniency distribution would increase the present discounted value of lifetime earnings at the firm by 6 to 12%. Differences across supervisors may reflect styles or other factors, but such a large magnitude of bias suggests that firms likely find it important to get at least *some* performance-relevant metrics about workers, even if they are imperfect. We suspect an interesting area of future work will be to understand how text-processing tools and the use of digital trace data will be used by firms to

[21] Relative to using objective measures, however, subjective performance evaluation leads to more compressed pay Macleod (2003).

[22] One could think of a tournament as consistent with this mechanism, in which only some workers are rewarded. There are other issues with relative performance evaluation, including sabotage, that may prevent these mechanisms from being used in some some settings.

understand worker activities and correct leniency bias.

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 percent. 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.

Benson *et al.* (2023) use data from a large retail chain to illustrate tradeoffs with the second type of bias, differential treatment. They investigate the impact of subjective assessments of employee short-term performance and long-term potential on gender gaps in promotion and pay.[23] Despite receiving higher performance evaluations on current job tasks, women were 13% less likely to be promoted than men. Women receive lower assessments of future potential conditional on their current performance, but these low potential ratings do not accurately predict women's future performance upon promotion. To show differential treatment, they establish that the marginal promoted woman outperforms the marginal promoted man. At their study firm, the job hierarchy explains 70% of the gender wage gap. Biased promotion rates for women account for a substantial portion of the overall gender wage gap at the firm. Biased potential ratings raise questions about how firms deploy ex-post incentives (either 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 used in many settings, but one of the original motivations was to describe contests for promotions as a form of career incentive. Waldman (2013) discusses the evidence on promotion tournaments and emphasizes the view that promotions have a market-based role in leveling up wages. Ke *et al.* (2018) discuss constraints that firms have for using promotion-based incentives, like the availability of slots, and show there is a tradeoff between using promotions to motivate and firm-level efficiency. Other work has tested behavioral responses to variation in tournament composition, like

[23]Although our focus is on how subjective evaluation interacts with bias inside firms, the general issue of credit attribution has been studied in other contexts (Sarsons, 2017; Sarsons *et al.*, 2021).

the presence of strong competitors in the field or in future rounds (Brown, 2011; Brown & Minor, 2014), which suggests that the degree of heterogeneity in a workplace may impact the efficacy of tournament incentives. 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).

The response to tournament or promotion incentives also likely depends on what workers know about future rewards. Cullen & Perez-Truglia (2022) show that workers have limited information about their likely earnings upon promotion, as evidenced by surveys of workers at an Asian bank. Many workers under-predict the earnings of their managers. When these workers learn that their managers earn more than they thought, they work longer days and generate more revenue. These effects appear consistent with career concern incentives becoming more salient or pronounced. In a related paper, Cullen & Perez-Truglia (2023b) explore why salary information does not diffuse, 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 that may be available.

On the firm side, the question is how firms learn about workers' types in the first place (Kahn, 2013), which is an input into decision rules about subjective evaluation and promotion. 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 employer learning models. Heterogeneous growth means that promotion rules, and how to use promotion incentives, can become extremely complicated.

Even when objective measures of performance are available in some jobs, like in sales, firms appear to put excessive weight on these measures to determine promotions where different skills may be required. Benson *et al.* (2019) show that across many sales firms, the best lower-level

salespeople tend to get promoted based on their prior performance, but that prior performance has little predictive value for managerial success. It is not clear whether firms believe that the incentive value of promotion is so great that they can tolerate the misallocation or whether they are making mistakes. We know of few other studies on how firms learn and use information about worker heterogeneity for the purposes of incentive provision and career planning, but we feel additional work in this area would be very useful.

Both managers and personnel economists will likely begin using text-processing algorithms and digital trace data to understand what individual workers are doing on the job and who they collaborate with. 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. How these tools will interact with subjective ratings, the rationales for giving those ratings, and whether ratings map to different activities or actions workers conduct will likely be an exciting area for the future. However, the applicability of these new measures, and the conditions under which they should be used, likely vary across settings.

3.3 External markets and firms' incentive responses

Sorting and equilibrium incentive contracts have generally been examined with single firm studies, taking external labor market opportunities as given. The exceptions are studies, like Biasi (2021), that look at changes in labor market institutions that introduce incentive pay (in this case for teachers). Given an enormous literature that has emerged in other areas of labor economics around the interactions of incentives and external labor market conditions, we believe that personnel economists' expertise in modeling the effects of incentive contracts can add nuance to our understanding of monopsony and labor market power, while accounting for some of these questions may add realism to personnel economists' contracting models. In this section, we review some recent evidence about how incentives interact with external labor market competition. Beyond incentive contracts, accounting for labor market features has implications for an exciting new wave of personnel research on how firms invest in non-wage aspects of employer relationships, like training or other amenities, that may be hard for workers to finance on an individual basis.

Perfect competition for workers implies that the labor supply curve facing the firm is infinitely elastic. But 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) surveys two classes of frictions that give rise to firms' labor market power: search costs and job differentiation. Job search and temporary unemployment are costly to workers; in the theory, firms anticipate these frictions and mark down wages. Job differentiation also gives firms some power over workers, as the interaction of firm-specific attributes, like location, and workers' preferences differentiate employers and create local market power.

Accounting for imperfectly competitive labor markets is likely to add richness and predictive power to studies in personnel economics. We see three areas where additional exploration will likely prove fruitful. First, studies of labor market power often use models with fixed wages and limited incentive effects. How does building incentive effects 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 only depends 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. Emanuel and Harrington then allow worker productivity (either through effort or sorting) to depend on wages, capturing this relationship using the function $e(w)$. When there are incentive effects from higher wages,i.e. 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 take the theory to data by studying the imposition of a voluntary minimum wage. They show that, based on the elasticity estimates, the optimal wage is very close to having no markdown. In other models, efficiency wage incentive effects come from labor market thinness, such that workers' effort incentives increase as it becomes more difficult to find employment at a comparable wage. 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 that limit the labor supply elasticity. If firms anticipate efficiency wage effects, 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 the 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 behind wage setting, which we return to in Section 3.4.[24]

Second, personnel economists may wish to explore how incentive provision differs when firms face different labor market structures. Many compensation contracting models implicitly have some notion of firms' market power built in when writing down a worker's participation constraint, but few models or empirical tests explicitly consider how labor market structure and competition influence the form of incentive contracts observed in equilibrium.[25] An exception is Bénabou & Tirole (2016), who have a model predicting that performance pay will be more prevalent in competitive labor markets. In their model, there are two worker types, indexed by high and low productivity, who exert effort across two dimensions, an observed/contractible dimension and an unobserved/uncompensated dimension. Firms compete for workers by offering two contracts designed to attract high and low type workers, respectively. Each contract consists of a fixed wage component and a variable pay component. The core prediction for personnel economists is that labor market competition lifts the rate of performance pay in both contract types across the ability distribution because competing employers seek to attract high-productivity workers by offering a high bonus rate. By contrast, a monopsonistic firm does not need to compete for high-productivity workers and can instead offer

[24]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).

[25]A literature exists on the effect of competition in the product market on the form of firm pay (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 performance pay has the strongest effect on sales in a bakery chain under moderate competition but then becomes detrimental on the margin under intense competition.

a lower fixed wage to extract rents from low types. 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.

Finally, the setup in Bénabou & Tirole (2016) with different worker types raises the question of how performance heterogeneity within a job affects firms' markdowns and market power. These issues are not new, as Kuhn (2004) argues 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 the way that these differences in productivity affect the way that workers compete in the market, will likely be crucial to building more realistic models of monopsony and for inferring market power.

Friedrich (2016) makes some progress on related questions by analyzing how information frictions influence individual firms' use of internal labor markets and external hiring to fill managerial positions. In his model, each firm considers its position in the productivity distribution across competitors, which in turn determines its available talent pool. Asymmetric employer learning about internal candidates' ability and their accumulation of firm-specific human capital over time makes external hiring of managers a risky strategy. To mitigate these concerns, firms can invest in hiring young talent and building an internal pipeline of promotions into management. Using evidence from Denmark, this strategy of hiring young workers and using internal promotions is most common among highly productive firms that compete for elite graduates; other firms are more likely to use market-based hiring for managers. We would applaud further work in this area, especially as it relates to heterogeneity on the firm and worker side and what that means for market versus internal talent practices.

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 How do firms set compensation and goals/standards in practice?

We would be very interested in papers that conduct surveys of what firms do to set compensation or to set incentives in practice. In particular, there has been a rich literature in corporate finance on how CFOs behave, use models, and make decisions (Graham & Harvey, 2001). A recent approach using large-scale surveys in Germany by Caldwell *et al.* (2024) recovers firms' wage bargaining strategies, but to our knowledge, very few surveys or clinical cases uncover managers' beliefs about incentive effects or how pay policies affect productivity. We think a central unanswered question is about what managers believe about incentive pay and whether those beliefs map to practice.

Indirect evidence, however, suggests that human resource managers use heuristics like salary benchmarks to set pay. Cullen *et al.* (2022) exploit variation in firms' access to salary benchmarking information, finding large and significant effects of salary compression by skill level, 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. Similarly, Hazell *et al.* (2022) find that 40 to 50 percent of firms set wages nationally, so that wages are the same for jobs in the same firm across different locations.

Finally, 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.

3.5 Core takeaways

- Performance-linked incentives change effort in non-routine jobs (e.g. for teachers, doctors, and lawyers), suggesting that prior work on the effects of performance pay in relatively routine jobs is likely to generalize to other types of tasks and jobs.
- High-performing workers appear most sensitive to performance pay on the attraction/retention and effort margins.
- Responses to group-based incentive pay can often be positive, despite predictions that free-riding should dampen the effectiveness of group incentives. Responsiveness to group incentives may depend the composition of teams or coworkers and their social connections.
- The level of pay provides incentives to the extent that the firm can monitor and set goals.[26] Lower-performing workers appear most sensitive to variation in effective pay levels.
- The link between firm-level practices and external labor market conditions is an exciting area for future research.

3.6 Open questions about incentives

While much has been learned since the past handbook chapter, there is still fruitful work to be done in this area. Understanding the extent to which firms can personalize incentives— through gamification or individual targets—and whether they should do so is a question that has received some attention. Our view of the literature also suggests that how firms set goals for workers is just starting to be appreciated as an important part of incentive provision. Understanding how firms set objectives and communicate goals will likely be an area for future research that will leverage new text-as-data techniques and digital records. Future researchers may also link this data to how relationships, and relational incentives, work in practice (see Gibbons (1998) for background).

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

Finally, there is an under-appreciated aspect of incentives that has received relatively little attention to date: do higher powered incentives increase the speed of workers' human capital investments on the job? Do some types of incentives, potentially like relative performance or tournament incentives, slow learning-by-doing due to sabotage? A well-known literature has looked at the returns to general experience and firm-specific tenure, but relatively little work has considered how different incentive schemes influence the rate of learning. Finally, 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.

4 Hiring

Hiring is the process by which workers and firms match with one another. As argued in the previous handbook chapter by Oyer & Schaefer (2011), hiring had been substantially understudied by personnel economists, especially relative to the work on incentives. Since the last handbook, there has been substantial research by personnel economists on hiring. In the spirit of exploration, we believe that there are several angles by which it is useful to better understanding 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 viewed as understanding the determinants of h_{ij}, as well as the consequences of various firm policies used to try to increase h_{ij}.

4.1 Networks and information

Workers may find jobs by reading a newspaper or searching the internet (Kuhn & Mansour, 2014; Kuhn & Shen, 2013), yet a large share of jobs are found informally via referrals. In the pioneering

work of Granovetter (1973), who studies job-finding in Massachusetts, over half of jobs are found via referrals.

A large literature in sociology and economics documents the prevalence of referrals in the job-finding process. Historically, much of this literature focuses on hiring from the perspective of employees. For example, Bayer *et al.* (2008) document that individuals are more likely to find jobs with other people from their census block compared to people from nearby census blocks, interpreting this as evidence of word-of-mouth hiring. This literature is summarized in Topa (2019).

A central question is whether firms benefit from the use of informal networks in hiring. Growing evidence examines referrals from the perspective of firms. 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). Two consistent findings across a number of 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).

It is difficult to answer by referred and non-referred workers' experiences differ. These patterns often appear in observational studies, but it is possible they arise through endogenous hiring, i.e., where firms set different hiring standards for non-referrals or tolerate lower match quality for non-referrals because of unobserved availability of referrals. As a result, observational studies may fail to provide guidance to firms on what to do with 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?

Progress on this issue is made by Pallais & Sands (2016), who run their own firm and hire oDesk contractors. The methodological contribution is to break through the endogeneity of the hiring process by hiring all referred and non-referred applicants. There are two RCTs in Pallais & Sands (2016). In the first RCT, the authors 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 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, i.e., referrals are more productive when they work with their referrers. Their most striking evidence comes in the second RCT. When workers are hired four months after the first RCT, referrals outperform non-referrals 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 a first-order topic, 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) work with a large Eastern European grocery chain to examine the impact of employee referral programs. The firm's 238 stores are randomized into one of five RCT arms, one control and four treatment. 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 using posters and letters, but there was no financial incentive. In the three financial incentive arms, workers received bonuses of different levels in order to make a referral.

The financial incentive treatments increased referrals, and referrals had better retention relative to non-referrals, consistent with the observational literature on referrals. However, there was also a quality-quantity tradeoff: as the financial incentive increased, the relative difference between referrals and non-referrals decreased. In addition, the paper examines the indirect effect of referral programs. Friebel *et al.* (2023) find that simply having an employee referral program led to a drop of employee turnover of roughly 15%. These effects are indirect as they do not primarily occur from more referrals. Instead, effects on attrition occur strongly among workers hired before the RCT and are sizable in treatment stores where no referrals are made. Surveys and other evidence suggest that the main mechanism for these effects is that workers value being involved in the hiring process.[27]

[27]There is growing work on the importance of employee voice in different organizational settings (Adhvaryu *et al.*,

An employee referral program can be a credible signal that a firm trusts its workers or values their voice, as a firm may not wish to involve its workers in hiring if it does not trust their judgment.

Recommendations. The procedures by which referrals occur are important and differ across circumstances. Economists and sociologists often focus on "word of mouth" hiring, where one person says something to their employer or a social contact. However, information about candidates is often passed along using recommendation letters. An interesting 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 examines the impact of recommendation letters, especially in the context of 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. The paper shows first using an audit study that reference letters increase callbacks. Second, using an experiment that encouraged referrals, Abel *et al.* (2020) 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. Having access to such a letter improves employment 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 outcomes.

Returning to Lazear's (1998) point about hiring risky workers, why do 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 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.

2021a,b).

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 using the universe of cellphone records from a Chinese city. Proxying referrals by whether a worker moves to a preexisting friend's workplace, Barwick *et al.* (2023) use event studies to examine the nature of communication around job changes. They find that communication to referrer friends increases substantially around a job change, whereas communication to non-referrer friends does not.

Reputation systems. While recommendation letters and referrals communicate private information, some labor market histories, like feedback scores, are public and may affect the attractiveness of hiring a worker. For example, Pallais (2014) examines a model similar to Tervio (2009) and shows that firms are reluctant to hire new workers because they are unable to internalize the value of information they provide 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 groups of workers called agencies, can overcome this friction. Workers connected to other experienced workers can join an agency that provides a public signal of all connected workers' reputations; in exchange, an agency head will get a fraction of affiliated workers' earnings. They show that high-quality workers benefit from this information revelation early in their careers but independent workers who do manage to break into the market and establish their own reputations eventually catch up. These features appear to have a similar logic to the information channels underlying the benefits of referrals and recommendation letters, but the public nature of feedback may ultimately have some 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, platforms have moved to use aggregated private ratings (that are closer in space to private recommendation letters than a feedback score linked to an individual) to augment the information content of public feedback.

We believe there is substantially more that can be learned about networks in hiring and how firms can tap into the information that workers have about potential employees. For example, how do firms use social media information 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 public praise that might be cheap talk? Other questions likely concern the broader effects of these policies. For example, how do employee referral programs affect worker diversity? For what type of jobs are employee referral programs most effective? What types of workers value being involved in hiring?

4.2 How do technology and other procedures influence hiring?

Hiring is often viewed as a problem of information. Firms have imperfect information about the traits of job candidates. What can be done to improve this 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 different 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 the context of 15 firms employing low-skill workers where job testing is gradually introduced. Consistent with Autor & Scarborough (2008), Hoffman *et al.* (2018) document first that introducing testing leads to an improvement in worker retention. However, their main question is different: how much discretion should managers have to overrule the algorithms? Hoffman *et al.* (2018) show that HR managers vary substantially in how often they obey the test. Some overrule test recommendations frequently, whereas others do so much less. Hoffman *et al.* (2018) find that managers who overrule more 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 it be that tests are sometimes overruled when a manager and candidate share something in common? 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 especially on the impact of algorithms on diversity. Li *et al.* (2021) 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 of conducting 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 tendencies to focus on extraneous or productivity-irrelevant characteristics, or to the presence of local information within the firm. So far, economists have performed relatively little work in understanding the value of adopting 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, Mocanu (2023) studies the impact of a reform for public sector hiring in Brazil that increased the use of impartial hiring practices, focusing on its gender implications. Mocanu (2023) finds that the reform substantially increases female hires by increasing women's application rates, evaluation scores, and probability of being hired conditional on applying. Shukla (2024) studies hiring for students graduating from Elite Indian universities. He finds that low-caste students are systematically disadvantaged by personal interviews, whereas they are much less disadvantaged by earlier parts of the hiring process.

Mocanu & Patacchini (2024) use the same reform as Mocanu (2023) to study how removing discretion affects the use of connections in public sector hiring. They document that the use and value of connections in hiring decreases following the reform. 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. Benson *et al.* (2024) use data from a large retail firm to study discrimination in hiring. They show that retail managers are more likely to hire workers of the same race, but also that there is some evidence that productivity is higher when workers and managers are of the same race. There may be bias in managers' tendency to hire workers who are racially similar to themselves, but it may also reflect managers having better information about such candidates. Turning from race to gender, Bagues & Esteve-Volart (2010) study hiring for the Corps of the Spanish Judiciary, exploiting that candidates are randomly assigned to face committees with different numbers of male and female evaluators. Interestingly, they find that candidates are less likely to get hired when facing more evaluators of the same gender.

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.[28] In a simple view of hiring, more applications is good because it gives firms more candidates to consider. However, reviewing applications is costly. Stanton & Thomas (2024) estimate the cost of considering an additional applicant for firms using an online labor platform. Using a demand model that accounts for which applicants a hiring manager considers, they estimate that it costs about $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, which would limit firms' wasted effort screening workers who are likely to be a poor fit. In fact, Horton *et al.* (2021) show that when firms add a simple message in a job post about the vertical quality of workers desired, it helps to segment markets and improve the efficiency of matching. It is an open question how other mechanisms, like instituting costly ordeals in the application process, may drive benefits to firms because of 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

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

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 study 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. This is an important policy issue in many countries, but especially in the US, where a sizable shore of the workforce has a criminal record. It not only affects the lives of people with a criminal record and their families, but also has broader externalities, as people without jobs are more likely to return to crime. A large literature in criminology and labor economics studies various policies to help build the human capital of individuals with criminal records, such as education and training programs, but workers with records continue to struggle.

A popular policy pursued in many areas is "ban-the-box." Referring to the box on job applications that ask workers about criminal histories, ban-the-box policies restrict what firms can ask workers regarding their criminal records, whether outright or limiting it until later in the application process (e.g., so that workers with a criminal record are not automatically screened out, and can instead explain their record during an in-person interview). Unfortunately, evidence indicates less than promising success for ban-the-box laws. Rose (2021) shows that a prominent Seattle law banning firms from examining criminal records until after an initial screening had little effect on job-finding. 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. Another group facing substantial challenges in the labor market is workers with disabilities. Workers with disabilities are less likely to be employed, and earn less than other workers. Relative to workers with a criminal background, workers with a disability may present less concern to firms in regard to left-tail risk, but concerns about productivity or perceived productivity may also be present.

A major step to try to improve the lives of disabled Americans is the Americans with Disabilities Act (ADA), which codified civil rights based on disability, including the right to receive reasonable accommodations at work. Acemoglu & Angrist (2001) find that the ADA decreased the employment and hiring of workers with disabilities, potentially reflecting cost and litigation concerns. These results are broadly consistent with Pyle (2023), in that more liability for firms leads to less hiring. Subsequent work also finds that anti-discrimination legislation fails to increase employment for workers with disabilities (Derbyshire *et al.*, 2024).

An alternative policy from civil rights protections is labor market quotas, which specify that workers with disabilities comprise a certain share of employees, with the share ranging across countries (Derbyshire *et al.*, 2024). Lalive *et al.* (2013) study the impact of a quota in Austria that applied to firms employing 25 or more employees. Comparing affected firms with 25+ employees to those unaffected, Lalive *et al.* (2013) find that quotas had a positive and significant effect on disabled worker employment. 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 black box regarding the hiring of workers with disabilities. For example, what do firms believe about workers with disabilities? How do firms behave or form beliefs with respect to different disabilities?

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

The discussion so far has focused on firm selection of workers, but what do workers value about particular 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 showing 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. Benson *et al.* (2020) find that reputation significantly affects job applications, with good reputation employers attracting twice as much labor as bad reputation employers. They also show using an audit study 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 firm business model and science quality (as opposed to a broad reputation) and how this affects job applications. Bryan *et al.* (2023) focus on 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 effects are driven at least in part by workers' beliefs about firm outcomes. Using incentivized belief questions, Bryan *et al.* (2023) also show that workers exhibit substantial overconfidence about firms' right-tail events, showing 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. Sockin & Sojourner (2023) provide evidence that the information most valuable to workers, negative reviews, is often under-supplied, reflecting concerns about retaliation.

Gee (2019) studies a field experiment on LinkedIn where some workers can view information about the number of people who have applied to jobs. Gee (2019) finds that showing the number of current applicants increases the chance that a worker completes a job application, and that effects are stronger for women than men. Rather than analyzing the impact of additional information, Sockin *et al.* (2024) study what affects the information that jobseekers have access to. They focus

on non-disclosure agreements, showing that non-disclosure limit workers' willingness to information about unlawful workplace conduct.

Flory *et al.* (2015) study how the competitiveness of workplaces affect job entry decisions, focusing on gender. The authors advertise for an entry level position across several US cities and vary various features of the job ad. When it is advertised that pay for a position will depend on relative performance, women are significantly less likely to apply. Turning from competitiveness of a firm to its commitment to corporate social responsibility and equal opportunity statements, Hedblom *et al.* (2019) and Leibbrandt & List (2018) of having job ads emphasize these features. Hedblom *et al.* (2019) find significant positive benefits of firms advertising corporate social responsibility, while Leibbrandt & List (2018) find that equal opportunity statements provide limited benefits in terms of applicant diversity.

4.5 Open questions about hiring

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

Research has focused on the impact of algorithms and AI on hiring outcomes, and these are key issues, though we also think much more needs to be done descriptively about how firms use these technologies to hire. How do workers and managers interact with algorithms and AI, and what factors affect the nature of their interaction? Anecdotally, many workers seem to regard the process of being rated by algorithms as potentially unfair. Is this driven by the way that algorithms are presented to workers?

We also believe that more research is needed in general about fairness and employer reputation in hiring. How do workers regard the fairness of hiring process? Does this matter for firms? Using an RCT, Bapna *et al.* (2021) show that the way firm communicate to candidates that they've been rejected affects whether those candidate apply in the future. How do other aspects of firm reputations affect worker flows in hiring?

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 or less 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 faced? 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 research on peer effects in the workplace. There has been important research since then. One area of significant research growth has been on managers. We will first discuss managers, followed by peers.

5.1 Managers

Interest in managers grows out of the general awareness that management practices are important for understanding the large productivity differences across firms (Syverson, 2011). And while there is growing consideration of management practices, there had been much less attention paid to the role of individual managers, especially outside the level of the CEO (Benson & Shaw, 2024).

Several key findings have emerged about the role of lower-level managers. First, the evidence indicates that lower-level managers matter substantially for the productivity of their subordinates or those on their team. Second, there is much less conclusive evidence regarding which traits of managers matter. Third, we are still learning how and why managers matter, as research points in different directions. It is natural that how and why managers matter could differ by setting, and further research can help better understand systematic patterns and how these patterns might vary in different contexts.

5.1.1 Do managers matter?

Economists and management scholars have studied the importance of individual managers for some time now, though much of it in the past had focused on the importance of CEOs and other top managers. In a seminal paper, Bertrand & Schoar (2003) use the movement of CEOs between companies to estimate regressions of firm outcomes on CEO fixed effects, firm fixed effects, and controls, thereby allowing one to assess the importance of individual CEOs. Other designs use natural experiments, such as exploiting CEO deaths and hospitalizations (Bennedsen *et al.*, 2020) or the firing of Jewish CEOs due the rise of the Nazis (Huber *et al.*, 2021), to estimate the importance

of individual managers.

Lazear *et al.* (2015) study the impact of frontline supervisors for service sector workers. They use data from a firm where workers complete tasks that involve customer support for technology products. They use an Abowd *et al.* (1999) methodology to analyze productivity as a function of worker fixed effects, supervisor fixed effects, and control variables.[29] Lazear *et al.* (2015) show that differences between better and worse frontline supervisors explains a substantial amount of within-firm productivity dispersion. Replacing a manager at the 10th percentile of manager productivity with one at the 90th percentile has a bigger effect than adding an additional average-productivity worker to a typical 9-person team.

Since Lazear *et al.* (2015), evidence that lower level managers matter substantially has been observed 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 papers also significantly expand the empirical methodology for estimating manager effects through the use of event studies.

Another approach to estimating manager effects is to go to the lab and to randomly assign workers and managers together (Weidmann *et al.*, 2024). In Weidmann *et al.* (2024), a good manager is someone who increases the productivity of the workers on their team. They show that there is wide variation in managerial productivity, but it is not correlated with most observable characteristics. Interestingly, however, it is negatively correlated with whether individuals wish to be managers.

It is also useful to understand when managers do not matter. 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 impact in some settings. For example, settings with high piece rates and relatively homogeneous work may have less variability with respect to managers, as strong incentives may substitute for who is leading a group.[30] To the extent that frontline managers provide help or solve difficult problems (e.g. Garicano (2000)), their per-capita impact on workers

[29]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.

[30]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.

may be lower as workers become more skilled. However, in this case, the managerial span (team size) likely increases, holding fixed managers' total impact. There is likely more work to be done on the settings where managers matter more or less.

5.1.2 Which manager traits matter?

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

A natural characteristic to focus on is a manager's skill as an employee. Many firms promote heavily from within and they are often able to observe the productivity of workers when deciding who to promote to manager. One might imagine that the best workers would naturally make the best managers, e.g., if ability within a firm is roughly unidimensional. Anecdotally, one observes in some industries how some star workers go on to become star managers (e.g., the star soccer-player turned coach Zinedine Zidane), but we have relatively little hard evidence on this.

Benson *et al.* (2019) analyze the importance of managers' skill as employees in a large number of sales firms. They use data from a provider of salesforce management software that allows them to observe sales productivity for workers, as well as who is managing whom. Benson *et al.* (2019) find that more productive sales workers are much more likely to be promoted to manager roles than less productive ones. However, after becoming a manager, there is no positive relation between former skill as a worker and managerial productivity. They interpret these results through the lens of a tournament model where firms motivate salespeople to work hard via the prospect of a promotion even though the skills that make someone a good salesperson are not the same ones that make someone a good sales manager.

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 regarding 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 have strong negative effects on employee attrition, especially attrition that the firm is keen to avoid,

though people management skills do not affect most non-attrition outcomes. People management skills are also rewarded within the firm, being 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 identify which manager traits matter is to measure managers' traits using surveys. Adhvaryu *et al.* (2023a) do this in the context of a large garment manufacturer in India. Garment manufacturing occurs in lines, where a frontline supervisor manages a team of workers. Their main finding is that it is difficult to predict managerial skill from observable cognitive and non-cognitive skills. However, factors related to managerial attention and control do predict managerial skill.

The papers discussed so far tend to be about interpersonal dynamics on teams, but 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 skill (a measure of one's ability to assign workers based on comparative advantage to tasks). Workers with better allocative skill likely make better managers to the extent that a managers' role is realizing comparative advantage. Consistent with this view, allocative skills predict 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 more likely to engage in entrepreneurship.

Separate from overall characteristics of managers, there is also interest in interactions between characteristics of managers and workers. Natural ones to examine are race and gender. These characteristics have been examined extensively in other manager-like relationships, such as the relationship between teachers/professors and students and between doctors and patients, but now are also being studied in the context of managers. 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, with female employees faring better

under female supervisors in terms of speed of promotion and wages.

Delfino & Espinosa (2024) examine value dissonance between workers and managers. They use survey data collected with a large multinational bank where employees are asked about the type of value they would wish to nurture or promote in children. They find that workers perform better when they are managed by a supervisor who has values more congruent with theirs.

5.1.3 How do managers matter?

While there is strong evidence that lower-level managers matter, how they matter is less clear. Lazear *et al.* (2015) present evidence that managers matter by motivating workers, as the majority of manager effects go away within 6 months of the worker receiving a new manager. However, the drop in manager effects is gradual rather than instantaneous, so it is possible that managers teach about factors that have a high transitory component, necessitating constant learning.

In the context of Italian government, Fenizia (2022) shows that managers operate more on the selection margin instead of the motivation margin. In her setting, good managers encourage low performers to leave, thereby increasing overall performance. Comparing these results, it appears that managers need to motivate and possibly teach in settings with limited rigidities, but the ability to make jobs unpleasant for under-performers may be required when there are substantial barriers to firing 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 careerwise from having a more lenient rater.

A further way that managers may matter is in terms of 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). Friebel *et al.* (2024) show experimentally that managers are likely to have information about local conditions. They analyze an RCT conducted in a large German bakery chain aimed at reducing workplace control. The firm was using a large number of checklists in their retail stores, some of them 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 consists of randomly removing two checklists perceived by employees to be of low value. Prior to random assignment, for each store the regional manager make a prediction about whether the treatment will be effective. Friebel *et al.* (2024) find that the beneficial effects of checklist removal are highly concentrated in the stores where regional managers predict 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, manager are more knowledgeable than subordinates, enabling enable them to solve exceptional problems, while lower-level workers focus on more routine tasks. Seminal models by Garicano (2000) and Garicano & Rossi-Hansberg (2006) lay out the logic 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) show how the number of managerial layers, and hence the degree of specialized knowledge each worker in a firm has, varies with firm scale. Recent tests examining variation in communication costs between layers of an organization also 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. 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 suggests managerial bias toward workers that managers are dating, and that worker

careers are potentially be shaped via their social interactions with managers. Using an RCT, Friebel *et al.* (2022) show that managers can matter by encouraging workers to focus on certain workplace issues, such as reducing employee turnover.

Thus, recent work indicates different possibilities regarding how managers matter. Managers may matter in different ways in different settings, e.g., in some settings managers may matter most for pushing people to "get stuff done," managers may matter more behind the scenes (e.g., in allocating resources) or in what managers know. How managers matter may also differ depending on the nature of a manger's role and how high they are in the 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 are tightly logically connected (do managers matter, which traits matter, how do they matter), but we believe there are many interesting and important questions related to managers that are distinct.

Do managers engage in talent hoarding? Managers play a key role in who gets promoted. Indeed, helping identify talent is regarded as a key role for managers at different 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). Using data from a large European multinational, Haegele (2022) shows that employees are substantially more likely to make applications for promotions in the time shortly before a manager departs. She provides this and other evidence indicating that many managers suppress the mobility of employees within the firm to other positions.

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) study worker beliefs about 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 to understand beliefs on this matter. 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 to managers. Is managerial quality fixed or is it malleable? What can firms do to improve the quality of their managers? What can managers do themselves to become better managers? In particular, we would like to see more work on the role of training and prior experience in shaping managers' styles and productivity. This evidence exists for CEOs, but systematic evidence for middle and lower level managers is not present. Given the highly selected sample of top executives, we would be interested in seeing how effects differ (or not) for a 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, suggesting an important role for the matching of talent with context. Is the same true for managers? Does hiring a successful manager from a different firm indicate a high likelihood of success in a new setting?

How are managers matched with workers or production units? One might naturally imagine that firms might wish to match the best managers with the best organizational units or best workers. However, some research suggests that this may not be the case. 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). We believe it is useful to understand how pervasive negative assortative matching for managers may be and what the causes are.

5.2 Peers and Teams

There has long been appreciation that peers can influence work output, productivity, learning, and happiness. Herbst & Mas (2015) conduct a meta-analysis of estimates of peer effects at work. Across 34 studies, they find that the mean productivity increase for a focal worker with respect

to a unit increase in a co-worker's productivity is 0.12. Lab studies and field studies have similar distributions of effect sizes, suggesting that lab evidence on spillovers generalizes to the field. These large effect sizes may be surprising in light of studies showing limited evidence for the importance 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 across studies in effect size estimates is 0.16.

In this section, we discuss evidence on two main channels by which peers may influence coworkers' productivity in organizations: knowledge spillovers and peer pressure.[31]

What distinguishes these studies of peer effects from more general team production per-se is that many of the studies of peer effects isolate spillovers for workers doing autonomous tasks. Team production, which we discuss in Section 5.3, allows for more general production functions where worker may do different tasks, potentially trading them amongst one another. In this way, our discussion of team-based output and organization allows for workers to have multi-dimensional skills or different backgrounds and perspectives. 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) and an inventor located in a technology cluster that is 1 percent larger will patent about 0.07 percent more per year (Moretti, 2021).

Yet relative to this literature, peer spillovers in wages within firms appear smaller in knowledge-intensive occupations compared to relatively routine ones where on-the-job learning is less important (Waldinger, 2012; Cornelissen *et al.*, 2017). It is possible that some degree of commonality in work tasks may help with knowledge transmission, or that other forces offset peer effects in knowledge-intensive occupations, which tend to be more competitive.

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

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

Frakes & Wasserman (2021) examine how peer exposure for patent examiners influences a focal examiner's patent granting activity at the US Patent and Trademark Office (USPTO). 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 at the USPTO. Several tests suggest these effects are driven by how examiners learn to process applications from observing those around them. In particular, the magnitude of spillovers is largest for impressionable, early-career examiners who are still learning on the job. In addition, higher rates of remote work in the peer group reduces 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 was shared among co-located workers. Despite compelling evidence that knowledge spillovers lead to convergence in how peers perform their jobs, higher grant rates for patent examiners do not necessarily imply higher productivity. Instead, these estimates show how spillovers occur.

To disentangle how changes in exposure to different workers influences output, Sandvik *et al.* (2020) conducted a field experiment where sales workers were randomly paired together under different treatment conditions. One condition had worker-pairs meet with one another and share advice about handling different sales situations. A different condition introduced incentives based on pairs' joint sales growth. A final active treatment combined the meetings and incentives treatment. Sandvik *et al.* find that both the meetings and incentives treatments increased sales during the 4-week experiment. However, the meeting treatment led to persistent sales gains that averaged more than 15% long after the experiment ended. Workers' output in the incentive treatment reverted back to the pre-experimental level. As part of the design, coworkers recorded what happened during their meetings on worksheets, allowing the authors to conclude that the likely channel for the gains was knowledge exchange.

In the meetings treatment, below-median productivity agents who were paired with above-median productivity partners closed 82% of an initial productivity gap that averaged 61% between these

higher- and lower-performers. The firm did not have to directly compensate higher-performers for this spillover, as the treatment effect in the combined group that included incentives was no different from the main meetings treatment that had no incentives. 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.* (2021a) 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 from coworkers when the program is voluntary, yet these workers have the highest returns from participation. These results point to a possible reason that productivity dispersion is large in some firms: lower performing workers may be reluctant to seek out help or to take advantage of learning from coworkers.

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 adverse effects of incentives.

Advice-seeking stigma. Two recent experimental papers support the notion that workers' inhibitions in asking for help may contribute to suboptimal 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 about 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. Participants in the experiment were 1,200 villagers in rural India, who had the opportunity to win a prize by guessing which of two boxes contained it. Participants receive an initial endowment of clues that increase 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, the number of clues, 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 score. The second treatment, {*Private*, *Revealed*}, varied whether the seeker's ability was made known 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 the need to ask for information is negatively correlated with cognitive ability. A comparison between treatment arms suggest that the signaling channel dominates the shame channel in most contexts.

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 answered multiple-choice test questions correctly for pay. Subjects each answered 10 questions in isolation, providing a baseline measure of their ability. Afterward, they could decide to seek computerized advice for a small fee, which provided the opportunity to revisit their answers after eliminating three-fourths of the incorrect options on the test. The experimental design manipulated whether subjects knew that advice-seeking would be visible to a manager. Managers were tasked with estimating the worker's competence from a short biographical profile and, in some treatment arms, the profile included whether the worker sought advice. About 70% of workers' maximum final total compensation was based on the manager's estimates of their competence.

Advice-seeking decreases by 16 percent when the choice is visible to a manager. Incentive-compatible elicitation suggests there is significant heterogeneity in the expected reputation cost of seeking advice, ranging from some subjects who perceive substantial costs to those who expect benefits. Managers in the experiment did not harbor any economically meaningful reputation cost or benefit associated with seeking advice. These results suggest that misperceptions or biased beliefs about how others will judge seeking information may hinder learning at work. An exciting area for future research is to dive into mechanisms that might alter workers' beliefs about the costs and benefits of interacting with peers.

Limited observability and distance. Emanuel *et al.* (2023) study how the introduction of remote work and breaking co-location affects collaboration and feedback patterns among software

engineers in a Fortune 500 company. The authors exploit variation in the level of co-location on a team prior to the Covid-induced remote work shift at the company, allowing them to trace out how remote work altered patterns among workers who were more like to interact in person and those that had higher levels of electronic interaction pre-pandemic. In particular, some teams were split between different offices that were several blocks apart, whereas other teams sat together in the same building. When engineers worked on-premises, those who were co-located with their team received 22 percent more feedback on their code than engineers on teams split between buildings. Remote work eliminated the feedback difference for previously co-located teams, suggesting an important role for proximity in information flows. After the shift, senior engineers became more productive, but younger workers' career advancement suffered as learning opportunities dried up.

In two experiments, Roghanizad & Bohns (2022) find that people underestimate the relative advantage of asking for help face-to-face compared with asking through zoom, phone, or email. The authors show that help is more likely to materialize when it is requested in person, yet participants do not expect the magnitude of the effect. Distance and proximity have surely increased with remote work, but even before then the internet allowed collaboration across different places. The burgeoning literature on remote work (beyond the scope of this chapter) is an excellent place to look for further experiments or to gain insights into how workers interact with one another — see Barrero *et al.* (2023) and Barrero *et al.* (2021) for overviews.

Adverse effects of incentives. Other work suggests that competition for promotions may break the link between knowledge spillovers and peer effects in wages. Follow-up work by Cornelissen *et al.* (2023) shows that a standard deviation increase in the quality of trained peers leads to a 3.4 percent 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. The interpretation of these results is that workers are less likely to get ahead in the presence of more able peers when promotions are determined by rank order tournaments. It is less clear whether knowledge spillovers are present but not detected in the presence of competition or whether competition shutters knowledge transmission, possibly due to 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, the authors show how co-location of workers affects efficiency and response times in handling emergency calls. The study utilizes detailed data from the Greater Manchester Police's Operational Communications Branch, where workers are tasked with handling emergency calls. Every call is handled by two different staff members, a "call handler" who deals with the caller and a "radio operator" who dispatches officers in the field. As the staff are spread across four different rooms in the facility, sometimes calls are handled by two workers who can see one another, while at other times the workers lack a line of site. When the two workers are co-located in the same room, response times to incidents fall by 2 percent. However, this benefit came with a trade-off: handlers took 2.5 percent longer to be ready for the next call when they were co-located with the operators, as they spent more time communicating about the incident.

The authors demonstrate that face-to-face communication facilitates quicker and more effective information transfer, which directly enhances productivity. 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. The research suggests that the design of workspaces should consider the importance of direct communication, especially in tasks requiring rapid and complex decision-making. While remote work and digital communication tools have become more prevalent, this study highlights the irreplaceable value of face-to-face interactions in certain contexts.

5.2.4 Peer pressure

In settings with peers, a classical dilemma is the free-rider problem: if agents' inputs are imperfectly observed, they may face incentives to shirk (Holmstrom, 1982). Such environments are prevalent empirically. The early literature on peer spillovers focused on the role of peer pressure in mitigating the free rider problem (Kandel & Lazear, 1992; Mas & Moretti, 2009). In a supermarket checker setting—where if one checker is working slowly, then others have to pick up the slack—Mas

& Moretti (2009) find that the introduction of highly productive workers into a shift generates positive productivity spillovers via the peer pressure channel.

Similarly, in the context of a public sector police services call center, Battiston *et al.* (2023) find that handlers take 7 percent more calls when the two desks adjacent to them become occupied. Callers are not committed to a single handler, such that if one handler is working slowly, others will have to pick up the slack. Handlers work in an open-plan office, creating variation in the physical proximity of peers from shift to shift. In addition, the manager observes handlers' work contemporaneously. The authors find that the peer pressure effect is stronger when the manager is sat further away—i.e., when their ability to monitor a given handler is weaker—suggesting that peer pressure and direct monitoring from the manager are substitutes, to some extent.

5.2.5 Extending research designs and open questions about peers

Many of the papers we have discussed successfully exploit pre-existing differences to study the heterogeneous effects of practice or working condition changes. Emanuel *et al.* (2023) leverage pre-existing differences in co-location within teams to examine the effects of proximity on collaboration to study the effects of a single change (remote work due to Covid) on collaboration. Espinosa & Stanton (2022) estimate spillovers from a randomized training investment by a Colombian government agency, using emails to measure coworkers' and higher-level managers' exposure to workers who eventually get trained. Other studies, like De Grip & Sauermann (2012), vary the timing of treatment and examine how the share of treated coworkers influence untreated workers' productivity. Random saturation designs, where the share of individuals treated vary, can be used to detect spillover effects (Baird *et al.*, 2018). Usually spillovers are a nuisance in experimental designs because of Stable Unit Treatment Value assumptions, but for personnel economists, these spillovers can yield powerful insights about production functions and knowledge and communication flows inside of firms.

Comparing large-scale changes to the light-touch variation and natural experiments in the literature is likely an area of fruitful work. Natural experiments typically utilize interventions that may be too small to induce individual re-optimization that would occur under larger-scale organizational changes. In a famous large-scale intervention to try to engineer peer effects in educational production, students at the Air Force Academy changed their sorting behavior when the group composition of student peers was manipulated aggressively by the administration, outside

of the natural variation in the data (Carrell *et al.*, 2013). More generally, we think future work on the conditions under which program or practice choices can generate spillovers should be an area of focus. For example, Adhvaryu *et al.* (2023b) study soft skills training for Indian garment workers. They show that the returns are largest when trained workers engage in team production or collaborative tasks. Moreover, this training substitutes for managerial attention, which may lead to larger-scale opportunities for organizational changes.

5.3 Team Production and Effective Teams

The literature on teams is relatively sparse compared to the focus 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 the ability to get large samples in all but the largest organizations. Issues with endogenous sorting into teams also give rise to selection on unobservables that may confound inference for some questions, especially if teams are assigned or voluntarily formed in different studies. 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) is one of the first empirical economics papers on the introduction of team production. They show that the introduction of team production in a garment factory increased productivity by about 18% and, contrary to the view that less productive workers would select into teams to free-ride, that more productive workers wanted to be on teams. 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 in 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 ways or working 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 makes sense. By contrast, when one worker has an absolute advantage in all skills, a hierarchy where that worker becomes a problem solver for all likely makes more sense. In this framing, worker heterogeneity across multi-dimensional skills provides a rationale for teaming.

In other contexts, choosing an idea, innovating on a problem, or aggregating information may be the objective. In this vein, work on contests and crowdsourcing shows that getting many draws of 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 should a team be (Prat, 2002). Lyons (2017) shows that production benefits from having diverse teams do not necessarily outweigh communication costs when joining workers together from dissimilar backgrounds. 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 that either worked as a team or worked individually. Some teams had workers from the same country, while others had one worker from two 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 by about 34%. Lyons finds evidence that nationally diverse teams had communication challenges; she fails to detect evidence that preferences for same-nationality coworkers drives the findings.

Hjort (2014) study teams and diversity in the context of flower production in Kenya. At the time of the study, there were two main ethnic cleavages in Kenyan society, with several tribes identified in one group and several in another. The flower production consists of one upstream

worker providing flowers to two downstream workers. Hjort (2014) finds that mixed ethnic teams perform worse than more homogeneous teams, defining mixture based on whether workers are from the same group or not. This difference grew larger in the context of a disputed election that gave rise to ethnic violence. However, providing team incentives led the difference between heterogeneous and homogeneous teams to decline.

5.3.3 What makes teams effective

Although studying teams is difficult, organizational behavior scholars have started to open these questions up using data from real firms. Famously, Google's Project Aristotle tried to identify the determinants of effective teams using observational data. They found support for pyschological safety as the main explanation for effective teams inside the company (Duhigg, 2016). Google found little support for the notion that variation in team composition drives effectiveness.

Other studies yield more support for the tradeoff between variation in background skill and communication challenges. Di Fang & Iglesias (2023) study how the overlap of expertise for teams of doctors in Brazil influences patient outcomes. They find that teams of doctors with more similar specialties result in lower mortality, hospital stays, and other expenses. The effects are larger in more complex cases and when teams have less accumulated experience working together, suggesting that communication frictions likely reduce some benefits from having diverse skill sets. We would be very interested in further studies around 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 be comprehensive, but instead to sketch some interesting areas where personnel economics has made progress, and 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 survey and others have illuminated important trends over time in the share of time people spend working versus engaging in non-work activities. Bandiera *et al.* (2020) represents a large advancement in measuring time use at work. Their focus is on how CEOs spend their time, which seems hard to study since CEOs are unlikely to devote a large share of their work day to doing surveys. Instead, Bandiera *et al.* (2020) enlist the personal assistants of CEOs to track how they spend their time over a week in 15-minute intervals. The 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. Using this rich data, they apply a machine learning model and classify CEOs into two types, namely, leaders and managers. Leaders focus on higher-level, multi-function meetings, whereas managers meet more with individual functions. The authors find that leaders achieve better performance than managers, though it takes a few years for differences to occur. The data also reveal that professional CEOs work harder than family CEOs (Bandiera *et al.*, 2018).

A different 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. Coviello *et al.* (2015) find that greater task juggling, as measured by dealing with multiple cases at the same time, leads to slower average time of completing cases.

Hoffman & Lyons (2020) study the relation between the level of pay and time use for US state legislators. The pay for US state legislators varies substantially over states even accounting for the wide variation in the nature of legislators' responsibilities across states. They find that larger

salary is actually associated with less time on legislative activity. Instead, higher salary predicts legislators spending more time on fundraising, which may have greater returns for legislators in terms of getting reelected.

We believe that the future is ripe for exploration on workplace time use by personnel economists. So much more data is 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). 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. There is a separate handbook chapter on this topic. We believe that there are aspects of this topic that are especially well-suited for personnel economists, namely: 1) how do people use technology at work and how do new technologies affect productivity levels and distributions? 2) how do technologies change the boundaries of firms, affecting whether labor relationships may change, like with the advent of platforms like Uber?

Technology adoption and productivity A classic example on the use of technology and productive effects is the study by Bartel *et al.* (2007) on valve manufacturing.[32] 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.

A transformative recent technology is artificial intelligence (AI), with generative AI showing the potential to alter productivity. In particular, several early studies show that average output increases, while production heterogeneity across workers falls. An early RCT by Noy & Zhang (2023) where they randomly provide generative AI to some workers doing a writing task shows that the average time taken decreased by 40% and output quality rose by 18%. Between-worker output variance also declined for the group that received the generative AI treatment. Brynjolfsson *et al.* (2023) study the impact of generative AI in the context of a firm operating a large number of contact centers. Technology is gradually rolled out that provides contact center agents with text of

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

different replies to customers. Productivity rises, with effects concentrated among less-productive and more inexperienced contact center agents. Workers in the bottom 20% of the pre-experiment productivity distribution 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.

Dell'Acqua *et al.* (2023) study the effects of generative AI in an artefactual field experiment with workers at the Boston Consulting Group. They split tasks into two types: those where generative AI is "inside the frontier" and can perform well and those where generative AI is "outside the frontier" of acceptable solutions. Productivity increased for consultants on tasks inside the frontier, especially for those who were lower performers in pre-testing on similar tasks. By contrast, productivity fell for tasks that were presumed outside the frontier at the time, like combining quantitative and qualitative analysis. The authors suggest that the consultants had difficulty parsing where generative AI was useful versus not and "fell asleep at the wheel" while employing the technology on inappropriate tasks.

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, one of the most prominent examples of generative AI, among workers in Denmark. They find quick and substantial adoption of the technology, but that men are much more likely to adopt than women.

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 instead on firms' job ads. They conduct an RCT where some firms are randomly offered AI-written first drafts of their job ads. They find that treated firms are more likely to post jobs, but 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. He estimated that between 26 and 29 percent of the jobs that existed in the US economy in 2004 could be offshored, with the falling price of information and communication technologies (ICT) driving

the potential attractiveness of offshoring service sector jobs.[33] Estimating the extent of offshoring is challenging, but the forces that could drive it also impact the organization of work. Two relevant questions 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 in 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%. More than half the growth can be attributed to the rise of online platforms, but the estimate potentially understates platform growth. Small transactions under 1099 reporting thresholds are not captured. Using self-employment as the reference frame also misses contracting arrangements where workers receive a W-2 from a third-part intermediary, as is common for many large enterprises.[34] 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 worry with these markets is that the ability to hire from anywhere could put downward pressure on wages in high-income countries. Yet the accumulating evidence suggests that the payoff to workers from online labor markets is skewed in large part due to information frictions. Pallais (2014); Stanton & Thomas (2016) shows that providing workers with feedback alters their career and earnings trajectories by allowing them to break into the market. These information frictions likely prevent downward pressure on wages, as the bulk of new workers are viewed as imperfect substitutes for those with reputations or for those who come from higher-income countries where credentials are likely easier to parse for clients in similar labor markets (Agrawal *et al.*, 2016). Clients likely need to invest in screening job candidates and parsing their work histories and credentials before hiring,

[33]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.

[34]See, for example, discussion of Google's use of contractor (Kerr & Kreitzberg, 2019).

preventing workers from competing away surplus. Stanton & Thomas (2024) suggest this limits the extent of clients' search, preserving surplus for workers who are hired despite the huge supply of them online. Market design features also appear to matter, such that recommended workers appear to gain credibility (Horton, 2017; Barach *et al.*, 2020), such that improvements in algorithmic matching may lower the transaction costs for firms to use external labor.

An important caveat to these results is how standardized the matching process is. 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 design changes because supply responses can undo the platforms' efforts.

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.[35] 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 active programs improves their 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[36] and growing due to several background trends, like digitization, electrification of vehicles, and public investment in chip manufacturing.

Training and contracts. The ability to sustain a relationship with a worker is a core mechanism that enables firms to invest in training. Garicano & Rayo (2017) present a model of firm training under relational contracts, where there are no contractual instruments to enable firms to capture value from training. 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

[35] For a discussion of the measurement issues and an overview of what is known about firm training, see an excellent recent chapter by Black *et al.* (2023).

[36] For example, Bartel (1995) reports that firms averaged spending about $385 per worker on training in 1989 (roughly $1,000 in 2024 dollars). Other estimates suggest firm training expenditure is substantial. Sandvik *et al.* (2021a) report estimates that US firm training expenditures 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.

intermediate work. The reason for this inefficiency, rather than immediate knowledge transmission, is that 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 provide training, but this rationale introduces distortions because of the inability to bind workers to the firm.

An active area of empirical research considers how contracts and external labor markets affect training returns and investment, as explicit contracts and market imperfections may provide incentives for firms to train (Acemoglu & Pischke, 1998, 1999b). One common contract that firms use is training repayment agreements, also known as training contracts. Such contracts specify penalties for workers if they exit the firm within a certain period of time after training, and have been used for many workers, including policemen, pilots, truckers, and teachers (Kraus, 1993, 2008). Hoffman & Burks (2017) analyze training contracts in the context of a large US trucking firm. Exploiting the staggered introduction of a 12-month training contract within the firm, as well as its staggered replacement by an 18-month training contract, Hoffman & Burks (2017) show that training contracts substantially affect worker turnover and firm returns from training.

Another contractual mechanism that influences training provision is non-compete agreements. Starr (2019) studies how non-compete agreement enforceability and consideration laws affect training.[37] He uses Survey of Income and Program Participation (SIPP) data for 1996, 2001, 2006, and 2008 to measure training. Twenty-one percent of SIPP respondents in his sample report participating in training at some point in the prior year, and about 90% of those who report training indicate it was firm-sponsored.[38] Starr finds that 27% of workers received firm-sponsored training over the last year in occupations that heavily use non-competes, whereas only 14% did in low non-compete occupations. He then estimates how training provision varies as a function of state-level non-compete enforcement interacted with occupation-level use of non-competes. The results suggest that non-compete enforceability increases training provision by about 14% relative to the mean.[39]

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

[37] Consideration laws condition the ability to enforce a non-compete on the worker getting something in return for signing a non-compete agreement, which could include training.

[38] Although the SIPP is a panel, training questions are asked once and are phrased: "During the past year, has [the respondent] received any kind of training intended to improve skill in oneâs current or most recent job?" Respondents who answer yes are then queried about what type of training they received and who paid for the training."

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

data exist at the firm level on training provisions. In Belgium, larger firms are more likely than others to provide training (Konings & Vanormelingen, 2015), consistent with there being some setup 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, 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). Caicedo *et al.* (2022) show that firms in high-skill sectors are less likely to take on apprentices in Colombia, potentially due to higher training costs or lower perceived benefits.

Differences in firm responses may reflect varying ability to capture value from training across settings. 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 contexts, 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.

Little work has covered how firms decide on training investments, but several papers suggest that the measurement of returns at the individual worker level may be understated. This is because training is closely connected to human capital and knowledge accumulation in firms. As such, the possibility of knowledge spillovers from one worker to another may influence the firms' return from training. While De Grip & Sauermann (2012) show these spillovers for workers at the same level, (Espinosa & Stanton, 2022) show spillovers exist across layers of a hierarchy such that training can increase 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 the firm, emerging work suggests that lower-level managers may hinder or accelerate the effectiveness of training programs through either talent hording (negatively) or through encouragement channel (positively) (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

programs may be least likely to participate (Sandvik *et al.*, 2021a; Delfino *et al.*, 2024).

A further topic of interest that we believe warrants further attention is to what extent can firms and governments improve the social skills or managerial skills of workers. Bianchi & Giorcelli (2022) study the impact of the training within industry program. During World War II, the US government provided several types of training to managers at different US firms in order to help the war effort. There were three types of training: job instructions, job relations, and job methods. The authors show that there are substantial returns to managerial training and that there are also complementarities observed between some of the training.

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, a central question is one of external validity. How do we know if the results are specific to a particular trucking, fruitpicking, or courier firm, or 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 that referred workers are more likely to be hired and less likely to quit than non-referred workers.

A different response is to acknowledge that some estimates are local to the firm or context studied, while making the case that sometimes special firms require attention (e.g. see Nicholas's (2023) study of the emergence of professional management at General Electric, which served as a focal example for other firms due to GE's prominence). Given the recent explosion of 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, at least not yet.

That said, we do believe that there are steps that personnel economists can take to clarify and potentially reduce concerns about external validity. We build on discussion in general about external validity (List, 2020).

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 relevance 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 more information about the firm or type of work they are studying. While firms often wish to remain anonymous as a condition of collaboration on research, personnel economists should still endeavor to provide important information about the firm. This is often done in the context of number of employees or overall revenue, but we think a bit more detail can be useful. Is this a low-cost firm or a firm focused on product differentiation or product quality? Is the firm considered a typical performer in its industry, or is it an exceptional performer? 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. In particular, we would encourage researchers to provide summary statistics comparing workers in their sample to other populations of workers. Would we expect, based on demographics, that the workers at the firm would look like others in the industry? In addition, the returns to practices may vary in important ways based on the usage of other practices (i.e. practice complementarity). For example, soliciting employee referrals may work

better when workers value good teammates (such as when workers are paid group bonuses), but may work less well when workers compete with one another.

Other experiments are framed around tasks or jobs, raising questions about whether results generalize to similar work settings. Surveys may help on this front. In studying the performance of teams in an escape room game, Englmaier *et al.* (2024) show that surveys suggest the task is a non-routine analytical task. They then survey one set of experts about the effects of group incentives in non-routine analytical tasks and a different set of experts on group incentives incentives for escape rooms; they are then able to compare whether experts believe the experimental task is similar to the likely effect of non-routine analytical tasks more generally.

3. **Show results separately for different types of workers.** While personnel economists often have data from one firm, one firm does not mean one type of worker. Often, researchers collaborating with a firm will have access to multiple types of workers. For example, a tech firm may have engineers and business workers, or a grocery company 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.** Some results in personnel economics can make a firm look good or bad. Fryer (2011) points out that one may be concerned about partnering with a firm to study racial discrimination at the firm, as that firm may be different in important ways from other firms due to its willingness to be studied. This point can be address by explaining why the firm is working with the research team. Answering the point can also provide important context on how it was possible for the experiment or study to take place.[40] 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 strongly support the use of

[40]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).

both—but we think it can be further bolstered by explaining what the firm hoped to learn. Finally, thinking about why the firm works with researchers can help think about other events or shocks that could be happening for the firm that could bear on the research results.

5. **Examine results under different conditions for the same population.** Even beyond the particular firm studied, one wonders whether effects are specific to a particular country or geography. In addition, when studies are conducted at particular points in time, such as during the COVID pandemic, one wonders how effects would apply more generally. When studies are large enough, it is possible to examine results by geography or business cycle. For example, Friebel *et al.* (2023) examine how the effects of employee referral programs vary by local labor market conditions. As it is natural to wonder whether answers to personnel economics questions may be culturally-specific (e.g., if treatment effects may vary based on local attitudes regarding managerial authority), researchers can also present how results vary across countries, for cases where the researcher is studying a multinational firm. 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 more frequently with firms, firms are often very interested in the results, whether observational or experimental. How did the firm react to the results of the study? Did it find the results surprising or expected? Did the firm change anything as as result of the study? Researchers increasingly ask other economists to predict experimental results to gauge if results shifted priors (DellaVigna & Pope, 2018). We believe it can be useful to report similarly for the firm being studied or for other firms.

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,

it can also be used to examine external validity.

8. **Explain why the firm didn't do it already.** Non-adoption of certain beneficial practices has long been a puzzle for economists. If a treatment appears successful, authors can help readers contextualize the validity question by explaining why a firm didn't institute a practice to begin with. Is it lack of awareness? If so, what does that say about overall manager quality? Is it beliefs that a treatment will not work (relative to costs)? 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 to satisfy an adversarial audience of skeptics (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 to better situated 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 with piece rates or sales commissions. To compensate for 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, performance-linked pay means 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 trial 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, the null effect at Firm A is likely to be a false negative. Firm A could increase their minimum standard upon

adoption of the practice at scale—and likely would—but a simple RCT done at small scale or as a pilot would miss the effort response. It is possible, however, that this RCT would find improved retention, suggesting some scope for the firm to raise standards if the program is adopted at scale.

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, offering a compelling way to provide evidence on mechanisms about what factors 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 think about different 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), so 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 is different from 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. What should personnel economists say or do about this issue?

We believe the most important thing is to be very clear about (1) whether the policy or practice is affecting one firm or many firms and (2) the prevalence of the policy in the population. The two questions, (i) what is the impact of performance pay for an individual firm, and (ii) what happens when all firms adopt performance pay, are both important and distinct. Researchers should seek to

clarify which question they answer. Business practitioners are likely keenly interested in question (i) and question (i) is also often of very strong interest for economic theory. On the other hand, policymakers may be very interested in question (ii).

8 Conclusion

Personnel economics is an exciting and growing subfield of labor economics focused on labor issues inside of firms. In this chapter, we have sketched ways by which advancements have occurred and will continue to occur. Since the last handbook chapter, personnel economics studies have benchmarked 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 for more and less productive workers. 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, leading to a bright future for personnel economics research. At the same time, we would continue to urge researchers to dig deeper to connect these results with practice. We would welcome studies on manager and leader beliefs about the efficacy of different practices or, for laggards in adoption, concerns about implementing practice changes.

Recent research in labor economics, organizational economics, and industrial organization emphasizes the large and growing importance of firm differences in pay and productivity. One explanation for this is that certain firms have adopted bundles of different HR practices. Personnel economists can play a key role in helping understanding to what extent the adoption of various practices helps explain the growing variation in firm effects. 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.

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 boysâ 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.

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.

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.

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.

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. 2011. Racial inequality in the 21st century: The declining significance of discrimination. *Pages 855–971 of: Handbook of Labor Economics*, vol. 4. Elsevier.

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.

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, 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.

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.

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.

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.

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.

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 T. 2021a. *Can Firms Use Self-Selection to Improve the Efficacy of Human Capital Investment? Evidence from a Field Eexperiment*. Tech. rept. National Bureau of Economic Research.

Sandvik, Jason, Saouma, Richard, Seegert, Nathan, & Stanton, Christopher. 2021b. Employee responses to compensation changes: Evidence from a sales firm. *Management Science*, **67**(12), 7687–7707.

Sandvik, Jason J., Saouma, Richard E., Seegert, Nathan T., & Stanton, Christopher T. 2020. Workplace Knowledge Flows. *Quarterly Journal of Economics*, **135**(3), 1635–1680.

Sarsons, Heather. 2017. Recognition for group work: Gender differences in academia. *American Economic Review*, **107**(5), 141–145.

Sarsons, Heather, Gërxhani, Klarita, Reuben, Ernesto, & Schram, Arthur. 2021. Gender differences in recognition for group work. *Journal of Political economy*, **129**(1), 101–147.

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.