

NBER WORKING PAPER SERIES

MANAGERS AND THE MANAGEMENT OF ORGANIZATIONS

John Roberts
Kathryn L. Shaw

Working Paper 30730
<http://www.nber.org/papers/w30730>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2022

The authors thank Mitchell Hoffman and Chris Stanton for their insightful comments, Ivy Manna for her dedicated research assistance, and Ken Smeton for his superb assistance in editing. We received no funding for this project. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by John Roberts and Kathryn L. Shaw. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Managers and the Management of Organizations
John Roberts and Kathryn L. Shaw
NBER Working Paper No. 30730
December 2022
JEL No. J3,L2,M5,M50

ABSTRACT

We critically review the emerging literature in Organizational and Personnel Economics concerning the role of managers and management practices. Our focus is on the middle managers who populate the hierarchies between top executives and front-line employees. We are especially concerned with research that brings together theoretical modeling and empirical investigation.

John Roberts
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
jroberts@stanford.edu

Kathryn L. Shaw
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
and NBER
kathryns@stanford.edu

1. INTRODUCTION

Economics has made major contributions to the analysis and practice of management. Striking evidence of this is the large employment of economists in business schools, where they teach and do research not just in economics, but also finance, marketing, accounting, entrepreneurship and other fields. Yet economics has, until recently, paid very little attention to the phenomena of managers and management as research topics.

It is now well understood in economics that firms are substitutes for the market mechanism in allocating resources (Coase 1937). In fact, the vast bulk of economic transactions are carried out within firms, not through markets (McMillan 2003). This means that, just as economists need to understand how markets work, they need to understand the workings of managed organizations. What determines resource allocation by and within firms? How does this come about? What do managers do and what is their effect?

This essay examines managers in firms and the management practices and policies they adopt to affect operations. Our focus is largely on middle managers, who report to more senior managers and may, in turn, oversee other managers. This is not because top executives are uninteresting. Rather, it is because CEOs and other denizens of the C-Suite have already been studied relatively extensively.

Thus, this paper tackles two different, but very interrelated questions: How do middle managers contribute to performance, and how does management contribute? When we use the term “management” we are referring to both particular management practices, like inventory management, supplier relations, or hiring practices, as well as to the overall structure of the firm, such as the shape of the hierarchy and the extent of delegation. While top executives make the overarching decisions directing and affecting the whole organization, the middle managers we study elaborate, complement, adjust, and implement these executive decisions. Thus they affect whether their own units run well. They also may affect the performance of other managers and the need to achieve coordination across the overall business.

Some fundamental empirical studies illuminate the importance of management in determining organizational performance. Ichniowski et al. (1997) was one of the earliest. They demonstrated that differing human resource management practices in steel finishing lines were associated with significant productivity differences across lines with identical technology. Syverson (2011) documented immense Total Factor Productivity differences across firms within US industries, suggesting that managers and management practices do matter. Perhaps most strikingly, Bloom et al. (2013) demonstrated causal effects of management practices on performance via a randomized controlled trial involving Indian cotton weaving firms. Moreover, innovations in management practices have been shown to have macroeconomic effects (Alexopoulos & Tombe 2012).

Our overall goal is to use our professional lens as organizational economists and personnel economists to assess how far we have come and where we might go next.

In the first parts of this essay, we will be especially concerned with papers where theory motivates empirical analysis. In the second half of the paper, we address work that has less theoretical underpinning, but instead tackles management topics with very careful methodology. We introduce the empirical methods these papers use, and then provide evidence from a series of papers. We conclude with sections on where research might go, and a lens on what we have learned.

2. MANAGERS

In this section, we explore what middle managers do, what makes them more or less effective, and their impact on performance.

Until recently, the firm was modelled as a production or cost function. This essay, and all of organizational economics, rejects that view as too simplistic for many purposes. Still, it is useful to ask how managers or management practices fit into a production function framework.

2.1 The Technology of Managers and Management

For simplicity, assume a Cobb-Douglas production function. Taking logs, write out the production function for each establishment as:

$$(1) \log Y = \log A + \alpha \log K + (1 - \alpha) \log L$$

where Y is output for establishment, K and L are measured capital and labor inputs, and A is a Total Factor Productivity term.

How does management enter in this model? In the empirical work below, many papers will imply that management is an intangible that elevates an establishment's output through A :

$$(2) \log A = \beta + \gamma M + P\Gamma + MP\Theta + MX\Lambda + XE + Z\Pi$$

where γM is the performance impact of managers, P is a vector of management practices, the MP is the interaction between the manager and management practices, the X are variables that may interact with P , such as the general state of knowledge or technology shocks, and the Z are workforce or workplace characteristics. Allowing for differences among establishments and firms and over time is handled in the obvious fashion below that estimates these equations. .

The key point is that we introduce M and P . Many of the papers show that the way a firm is run, through its managers and management practices, is crucial to its performance. Note that P is intended to be a vector and thus represent a set of management practices that could be complements, as was done in Ichniowski et al. (1997), Bloom & Van Reenen (2007), Bloom et al. (2013) and many others.

2.2 Value of Managers

Before addressing why managers add value, an obvious question is, how much value do they add? Let's start with the simple estimation of Equations 1 and 2.

In Lazear et al. (2015), good bosses raise the productivity of their team members, implying managers contribute significantly to worker output as in Equation 2. The value of managers is estimated through "boss" (or manager) fixed effects in the worker's productivity regression of Equations 1 and 2. Boss effects can be estimated because workers change bosses frequently. This firm employs technology-based workers doing computer tasks, so the authors have individual measures of productivity. The value of bosses is high: replacing a boss who was in the lower 10th percentile of boss quality with one at the 90th percentile is equivalent to increasing his team's total output by about 11%. Good

workers also match with good bosses, though the effect on productivity is small. Plant visits confirmed that good bosses are good coaches, trainers, and leaders.

Turning to a comprehensive paper on managers' productivity, Metcalfe et al. (2022) follow the managers of thousands of retail stores owned by two large companies. Using managers' movement across stores, they find that a large amount of the productivity differences across retail stores is explained by these manager effects. They also show the mechanisms by which top managers raise store productivity: they increase the labor efficiency and energy efficiency of their stores, as well as other factors.

Other papers demonstrate managerial value in similar ways. Fenizia (2022) shows that in the Italian public sector better managers raised labor productivity significantly, largely by inducing retirements of less productive workers. Englmaier et al. (2021) showed substantial productivity effects from team leaders in a field experiment.

For most mid-level managers, their positive effect as good managers is measured as a reduction in turnover for those who report to them. Hoffman & Tadelis (2021) show this to be true for the managers of skilled engineers, as does Friebel et al. (2017) for the managers of small bakeries.

The value of managers certainly varies across firms or divisions within a firm. In the large bank studied in Bartel et al. (2022), divisions differ as to the need for close managerial oversight. The Risk Management group has close oversight, with 3.5 people reporting to a manager, and the Wealth Management division has little oversight, with eight people reporting to a manager. The difference is that Wealth Management has performance pay, which acts as a substitute for managerial oversight.

2.3 Managerial Attention and Managerial Tasks

What do middle managers do to contribute value to a firm? At the highest level of abstraction, they have to motivate and coordinate their people and they have to pursue operational objectives. But what does that mean in practice?

To become more concrete, recognize that most managers have prescribed jobs. These define managers' powers and responsibilities and correspondingly determine the tasks that they must execute. New technology permits discovery of what managers in particular circumstances are actually expected to do. Bartel et al. (2022) is an excellent example. Using machine learning (ML) to analyze job postings, they show that the Personal Banking division of the bank they study wants managers to be "coaches" and the Risk Management division wants them to have "operational" skills. Though across the bank, with 80,000 people, the tasks of a manager are described by the phrases "lead team, oversight, foster, accountable, leadership model" in the ML analysis of job postings.

An alternative approach to learning how managers add value would be to examine how they spend their time. This data might come from their daily calendars, as has been done increasingly for CEOs (Bandiera, Prat, Hansen, Sadun (2020). Unfortunately, this sort of information does not appear to have been collected for lower-level managers. There would seem to be value in assembling such data. Nevertheless, there is still a rich literature on how managers allocate their limited attention and the resultant effects.

Economic intuition would suggest that devoting more managerial time and attention to a task would

increase performance on the task. So, for example, paying more attention to monitoring workers ought to reduce shirking, and allocating more attention to training ought to increase skills and productivity. A key paper tests this idea.

Friebel et al. (2021) employed a RCT experiment in a large retail chain in Germany to study the effect on performance of shifting managerial attention.. Managers of treatment stores, which were local bakeries, were asked to “do what they could to reduce attrition.” In response, they allocated marginally more of their time to communication with employees and less to trying to increase sales and control costs. The result was a 25% reduction in attrition.

The recent theoretical literature on attention builds much more complicated models and draws out more subtle insights. Two of the earliest theory papers were Halac & Prat (2016) and Dessein et al. (2016). Each offered a rather complicated model that led to subtle and nuanced insights. Unfortunately, neither paper’s predictions have yet been subject to empirical examination.

In Dessein & Santos (2021), managers face multiple tasks to which they can devote attention, while having more or less skill at performing each task. Attention and skill both increase the payoff from a task. The authors show the extent of uncertainty in task payoffs determines whether the two are substitutes for one another or complements. Low payoff uncertainty favors attending more to tasks where the manager is relatively unskilled, trying to overcome ignorance. High payoff uncertainty favors “playing to strength,” focusing on what the tasks that manager already knows how to do well. Thus, in turbulent environments managers display what the authors call “style;” paying particular attention to a limited set of tasks, such as marketing versus cost-reduction.

This model presents some hypotheses that can be examined empirically. One paper in particular begins this effort. Lo et al. (2022) adopted a model much like that in Dessein & Santos (2021), where both management attention and expertise are valuable in addressing tasks. They argue that in situations where attention is scarce, perhaps because other tasks place rigid demands on managers’ time, expertise and attention were complements: A manager focuses attention on tasks where she is relatively more expert, which again they call managing with “style.” When attention is abundant, however, attention and expertise become substitutes: The manager shifts attention toward tasks that involve improving where she knows less. Studying managers from a Japanese department store chain and from a Chilean beverage company, the authors collected data on measures of “time stress” among managers, which they interpreted as limiting available attention, and on the actual allocation of the available attention. They found that attention capacity had the predicted effect: when available attention was scarce, managers emphasized what they already did well, and when attention was plentiful, managers sought to offset their weaknesses.

Dessein et al. (2021) explored the relationship between the volatility of a firm’s local environment and its organizational structure. While it might seem that more uncertainty would favor delegating decisions to front-line managers, who are presumably better informed about local conditions, the authors offer a model where this is true if and only if the need for coordination across local units is not too high. If coordination is very important, the center’s insight into different units becomes more valuable. Using micro-level data on the same Japanese department store chain as in Lo et al. (2022), the authors test this hypothesis by following department managers, with confirming results.

The value of managerial attention in permitting reallocation of workers across production lines to improve efficiency is identified via a natural experiment in an Indian garment firm (Adhvaryu et al.

2022). The indoor air quality in the firm's plants depended on outside air quality. The effects of spikes in outside pollution were worse in some areas of a plant than in others. Line managers who paid more attention to their people and the productivity of their lines were more likely to shift especially pollution-sensitive workers to less exposed lines. Doing so reduced individual and line losses of productivity during bad bouts of pollution. The effects were substantial: a manager who devoted one standard deviation more attention than average would largely eliminate the negative productivity effects of increased pollution, saving the firm about US \$43,000.

Overall, the topic of "managerial attention" points to the simple fact that managers ask themselves regularly how best to spend their time to improve the performance of their units or their workers. These theoretical papers, and the few testing the models, have highlighted a series of important topics on what managers should and do consider.

2.4 Styles of Managing

Important elements of the literature on how managers spend their time connects to a notion of "managing with style." The idea is that there are manager fixed effects: some individual managers have sizable, persistent effects on firm or worker performance. These effects are typically associated with differing patterns of choices and, sometimes, can be linked to personal characteristics of the manager. This is the subject of a large literature on CEOs.

This CEO literature begins with the celebrated paper by Bertrand & Schoar (2003). They tracked top managers across firms over time and found strong manager fixed effects explaining firm performance, differences that could be explained by "style" differences across managers. Their "styles" are distinctive patterns of decision-making, such as some managers choosing to take very aggressive approaches to investments or acquisitions and some not.

The concept of "style" is acknowledged as describing relevant differences across CEOs, but it is likely to pertain to mid-level managers as well. For mid-level managers, Dessein & Santos (2021) and Lo et al. (2022) papers, discussed above, endogenized a version of style in a model where individual managers have greater or lesser expertise at specific tasks. In these papers, as discussed, an individual "manages with style" by choosing to devote more attention to tasks at which he is already particularly skilled, where style is the set of tasks that are accentuated. The authors interpret this behavior as generating manager fixed effects, or why managers differ.

2.5 Leadership

Another aspect of managerial behavior that has recently found increasing attention in economics is the phenomenon of leadership. Leadership has, of course, been a major focus in management research, but there it has largely been studied using non-economic methods.

A useful survey of economic research on leadership is Hermalin (2013). While exclusively theoretical, this work did offer an attractive resolution of the thorny issue of what constitutes leadership: Leaders have followers, which means they get them to do things without the formal authority to compel the desired behavior or the ability to design effective contracts to induce it. Leaders may be in positions of authority, but they can get their followers to go "above and beyond." They have traits or behaviors that are inspirational to their followers.

To our knowledge, the first economics paper to model leadership and test predictions is Lazear (2012). He defined leaders as wanting to attract followers and being successful at doing so. What then are the leaders' skills that seem to attract followers? Lazear showed that, in a very large sample of Stanford MBA alumni, corporate leaders tended to have especially broad academic and professional experience: They were "Jacks of all trades."

Carter et al. (2019) sought to study the effect of differences in officers' leadership quality on their subordinate officers' decisions to leave the US Army. They identified strong leaders by their receiving early promotion, because the Army constantly evaluates its officers and gives early promotion to the best leaders. When young subordinate officers were assigned to strong leaders, they were almost 3% less likely to leave when their service commitments were over.

Englmaier et al. (2021) used a field experiment involving escape rooms to assess the value of leadership. Teams were formed and asked to solve a complex problem within one hour. Treatment groups were asked to select a leader before starting, while control groups were not. Groups with leaders completed their task more quickly, with the percent of tasks actually completed on time at 63%, compared to 44% for those without leaders, with no drop in the originality of the solutions.

Rotemberg and Saloner (2000) and Van den Steen (2005) investigated the role of leadership in situations where the leader and prospective followers do not necessarily agree on the best course of action. The basic insight is that the leader can shape the followers' action choices by her own actions, thus providing leadership. Further, it may be optimal for the firm to employ managers who hold particularly strong views about the best choice of action, because this can be especially motivating for employees. Van den Steen (2010) extends this logic to offer a rationale for the existence of firms with their control over assets and low-powered explicit incentives as a mechanism to generate interpersonal authority over employees.

2.6 Characteristics and Skills of Successful Managers.

Very little is known about what makes a manager successful, because data linking individual characteristics of a manager to his or her managerial performance is scarce. However, a few papers do shed light on the issue.

Managerial training can make firms more productive. Bianchi and Giorelli (2022) examined the Training Within Industry (TWI) program, in which the US government provided training to managers and supervisors in firms involved in war production between 1940 and 1945. Due to government resource constraints, only half of 11,575 firms that applied for the program were randomly selected for training, and not all these received all three training modules. After ten years, firms that received training had 16% higher sales and 27% higher Total Factor Productivity than applicants that did not. Moreover, there were complementarities across pairs of apparently distinct types of training.

Grönqvist and Lindqvist (2016) showed that having held a leadership role in the Swedish military had a strong positive impact on subsequently attaining a management position in civilian life. The subjects held ranks as sergeants and second lieutenants, which gave them command over a group of soldiers. The effect of having received the leadership training and experience in the military was to increase the probability of becoming a manager by age 40 from 6.7% in the general population to 12.7%, an increase of 75%.

Huber et al. (2021) analyzed a natural experiment involving the Nazis' driving out Jewish senior managers from large German corporations in the 1930s and replacing them with "Aryans." Firms led by these replacements had persistently lower performance. The authors concluded: "The findings imply that individual managers drive firm performance," in part because the Jewish managers were better connected to other firms and better educated.

Are better managers generally better educated? Are they risk-takers? Queiro (2015) examined the effect of managerial education on firm growth in Portugal. He found that firms created by college-educated managers grew to be much larger, by a factor of 20 times, than those headed by managers with only primary education. However, other studies, mostly on entrepreneurs, do not find such large education effects. Regarding risk taking, using a large sample of German firms, Caliendo et al. (2022) found that risk-averse managers favored providing training to subordinates who were less likely to leave, not those whose productivity might have been most affected by training. Thus, managers may have been making decisions that were not consistent with firm profitability.

Hoffman & Tadelis (2021) examined the effects of a managers' interpersonal skills with subordinates, which they called "people management skills." In a large high-tech firm, they showed that people management skills, as measured by employees' views of their managers, had a strong negative relation to employee turnover.

Adhvaryu et al. (2022) asked which managerial skills, traits, and practices matter most for productivity and whether they were priced into managerial pay. Using productivity data from a large Indian garment firm and survey data on its line managers, they found that several dimensions of managerial quality, like attention, autonomy, and control, were important for learning-by-doing as well as for establishment productivity. Whether these are, in fact, characteristics of the managers, their job assignments, or the choices of management practices seems unclear.

These various studies are finding that certain characteristics are associated with better managers in the firms studied. These papers do not tell us if any traits are universally better, but they do tell us some traits are much more valuable within each firm. This leads to the topic of assortative matching: the idea that some managers could really improve the performance of the particular firm in which they reside, but not improve it in others. This is an important point, but beyond the scope of this paper.

3. WHY ARE THERE MIDDLE MANAGERS? HIERARCHY AND DELEGATION

Almost all firms possess some hierarchy and delegation, in which there are layers of authority relationships and each layer takes permitted decisions and actions and reports to the next higher layer. Key questions raised are in this section are: What controls the extent of delegation, and what controls the amount of horizontal communication?

3.1 Theoretical Models of Hierarchy and Delegation

Several theory papers examine the basis for delegation and the emergence of hierarchy as a solution to whatever problem that the firm is assumed to face.

Holmstrom (1984) was first to examine how decision authority should be allocated, in this case of a

superior and a subordinate with better information but diverging preferences. He found that the solution was to allow the agent to pick among a specified subset of possible actions, with the remainder residing with the superior. This is reminiscent of the standard policy of management by exception.

Garicano (2000) demonstrated that a hierarchy could be an optimal way to arrange learning to solve emergent problems the firm might face. At each level of the hierarchy, the members invest in learning to solve a range of problems, with the problems arrayed by their frequency of arising. Front-line employees, who receive the problems as they arise, invest in learning to solve the most common ones. Each successively higher level invests in learning to solve a range of less common problems. If a front-line worker cannot solve a problem, it is referred up the hierarchy, and continues up until it reaches a level where the ability to solve it resides.

Athey & Roberts (2001) developed a rich model of endogenous hierarchy. The driving force in the model is a multitasking problem: The two agents each must exert effort on their normal duties and offer up investment projects. There is confounding in the available measures of the agents' decisions and actions. The solution under some parameter values is that one agent is given authority over the other's investment choices, with information being transferred at a cost. In addition, it may also be optimal to involve a third agent, who only makes investment decisions. These measures permit giving strong incentives for effort. This paper illustrates the complexity of the organizational design problem and has led to several recent theoretical contributions that study the allocation of decision rights and their interaction with other design choices, particularly information transfer.

Formally delegating a decision does not fully determine how the decision actually is made and by whom. Hierarchic superiors in a firm always have the right to take back or over-rule decisions that had nominally been delegated. While relational contracts (Baker *et al.*, 2002) might limit this, it is important to understand more fully how the delegation decision is made effective. Aghion and Tirole (1997) present a rich model of how delegation is made credible, involving many aspects of organizational design, including the superior's deliberately choosing not to become informed. These studies offer abundant opportunities for empirical research.¹ A major obstacle, as with much research on organizational economics, is the difficulty of obtaining data and then estimating key testable hypothesis. Nevertheless, a few empirical papers address the phenomena of hierarchy and delegation, described next.

3.2. Empirical Models of Hierarchy and Delegation

In empirical studies on hierarchy, most common are those that look at how exogenous factors shape the hierarchy.

For example, Dessein W, et al. (2019) argued that the volatility in the firm's local environment affects its organizational structure. Increased uncertainty might seem to favor delegating decisions to front-line managers, who are presumably better informed about local conditions. This was true in the model

¹ Excellent surveys of the more recent theoretical literature on hierarchy and delegation are Bolton & Dewatripont (2013), Gibbons et al. (2013), Garicano & Van Zandt (2013) and Mookherjee (2013). Among more recent contributions. Rantakari (2011) offered a particularly rich model of the delegation choice interacting with a variety of environmental and internal concerns. Alonso et al. (2015) predicted that increased competition will increase the need for coordination across units and thus lead to more centralization of decisions.

when the need for coordination across local units was not too high. However, if the need for coordination was particularly high, increased uncertainty led to decisions being centralized. Using micro-level data on a large Japanese retailer, the authors tested this hypothesis, with confirming results.

Another source of outside pressure that changes the structure of the firm is competition. Rajan & Wulf (2006) studied a set of large firms from 1986 to 1998. They found a strong shift in organizational structure, producing a larger span of control at the top (see also Smeets et al. 2019). The number of direct reports to the CEO increased on average from 4.4 to 8.2, and the number of layers of management between the CEO and the lowest level managers fell 25%. These changes largely reflect more business managers joined the C-suite. They presented evidence that increasing international competition was a driving force behind these developments.

Note the phenomenon of the flattening firm, removing layers of hierarchy, does not correspond simply to centralization or decentralization. Some decisions formerly taken by managers whose levels disappeared do get pushed down to lower levels, but some rise to be treated higher up. This is exemplified by BP's reorganization of its exploration and production business in 1995 (Roberts 2004). Regional managers were eliminated, and front line managers had increased scope for decision-making, but now reported to a single global executive.

In another example, demand shocks led firms to reorganize and introduce an additional managerial layer. In a rich employer-employee matched data for Portugal, the result was that quantity-based productivity increased by about 6%, while revenue-based productivity dropped by around 3% (Caliendo et al. 2020).

A second set of empirical studies are done in developing countries, where is it especially likely that new methods of management can improve the performance of firms or address corruption problems and where RCTs are less costly to perform.

For example, Bandiera et al. (2021) examined the effects of reallocating final decision authority over purchases from managers down to purchasing officers in the government of Punjab, Pakistan, and of sometimes adding performance pay for the officers. The issue was that prices paid for purchases were perceived to be too high. This was possibly due to corruption by the managers, some of whom would systematically hold up approvals of purchases, a practice known to be associated with corruption. In the RCT, there were three treatments: one group of officers were delegated greater authority over purchasing decisions without manager interference, one group got performance pay, and the third group got both authority and pay. The control group got none of these. However, across all managers, good and bad, the average effect of performance pay alone was negligible, and adding it to delegation reduced the effect on prices: The two measures were substitutes.

A pathbreaking empirical paper on delegation and decentralization is Columbo & Delmastro (2004). They studied a sample of Italian manufacturing firms and an unusually rich set of possible determinants of the choice of whether to allocate decision power to plant managers, including such factors as the horizontal scope of the firm and complexity of the organization. They argued that these were linked to the extent and nature of local information, an issue not studied in the theoretical models. Their results are largely consistent with the predictions they set forth, but their work points to the potential value of richer, deeper modelling.

4. ORGANIZATIONS AS SYSTEMS: STRATEGIC COMPLEMENTARITY OF

MANAGEMENT PRACTICES

A major theme in organizational economics is that economic organizations are systems. They are not collections of randomly chosen people and procedures. Rather, in successful organizations, there is a logic as to how the pieces fit together and how they fit with the overall objectives of the organization and the environment in which it operates. This must be recognized in studying them.

Perhaps the first organizational system studied as such by economists was the Toyota Production System. MacDuffie (1995) explicitly addressed the Toyota System as a system involving human resource, operations, and supplier relations policies and practices that he argued empirically were associated. The most powerful logic as to why particular features might fit together or detract from one another comes from Milgrom & Roberts (1990): some features are *complements* that re-enforce one another and should be adopted together; others are *substitutes* and should not be observed together. In a payoff function, choices are complements if increasing one of them increases the gain from increasing the other; substitutes work the other way. How this logic specifically applies to Toyota System and manufacturing more generally is discussed in Milgrom & Roberts (1995).

The pioneering empirical study of Ichniowski et al. (1997) built upon these formal theoretical models by showing the importance of system effects in steel finishing lines. They identified four clusters of human resource management practices: lines with most innovative system of practices, like careful hiring, performance pay, and teamwork, were the most productive. The model was identified as a natural experiment, because over time manufacturing was switching from old traditional management practices, in which the insights of workers were ignored, and moving towards management methods that had aspects of the Toyota Production System.

A recent example of testing for complementarities in an RCT is Mbiti et al. (2019). In a large-scale experiment across 350 schools in Tanzania, they find strong evidence of complementarities between unconditional grants to schools and teacher incentives based on student performance.

There can also be substitution of management practices within a firm, across divisions or over time. In Bartel et al. (2022), a very large bank chose to use performance pay in one division and close managerial monitoring in another. The two practices do not appear together. As just discussed, Bandiera et al. (2021) showed in an RCT that delegating decision-making authority to lower levels and providing incentive pay could be substitutes in particular circumstances.

Thus, if systems are in place in which management practices are either complements or substitutes, papers that study one practice in isolation are flawed. Increasingly, economic development research does field experiments with three treatment arms – two management policies and their interaction. In reality, if two practices are complements, they ought never be found in isolation unless the firm is introducing changes slowly or is unaware of the importance of the complementary practices.

This logic argues for the desirability of measuring the effects of adjusting managerial choices in concert. Studies are increasingly doing so, and we highlight some that do. To omit complementary practices creates an omitted variable bias. Thus, it is crucial to acknowledge or incorporate the fact that practices are chosen in particular contexts, in concert with other practices. And over time or across divisions, practices could be substitutes (Bartel et al. 2022).

In Section 5.3, we urge scholars to attempt to allow for interaction among practices and for testing for

complementarity when collecting data or in designing experiments.

5. DO MANAGEMENT PRACTICES MATTER? WHY DO THEY MATTER?

The goal of this section, and the next, is to address the question, do management practices matter? Why or how much? For example, would a move towards performance pay increase the productivity of workers (Lazear, 1999)? Why would it? What other practices should be adopted or avoided when adopting performance pay?

The above sections follow the pattern of introducing theoretical models and then the empirical papers that arise from them. Sections 5 and 6 differ in approach. Rather than using theory to motivate empirical work, we first present the empirical methodologies, in Section 5, that enables researchers to best test the impact of management practices on worker performance. Then, in Section 6, we discuss a selection of key research papers that follow these methodologies in examining management practices.

5.1 How do we know that management practices matter? Key identification strategies.

We may think of recent empirical work as having three alternative methodologies – papers utilizing “field experiments,” “natural experiments,” or “observational data.” The top half of Table 1 defines these methodologies, showing the nature of the data and how the researcher obtained it. The bottom half of Table 1 provides examples of research papers that will be discussed in Section 6, categorized by the methodology of the research paper.

INSERT TABLE 1 HERE

Nearly all of the papers in Table 1 use data from experiments with one or two firms. Exceptions are when data comes from the internet (Bianchi & Giorcelli 2022) or from a survey done by the researchers (Bloom et al. 2014), or from a field experiment involving many firms (Bloom et al. 2013) or a field experiment with many people using LinkedIn data (Gee 2019). Increasingly, researchers are using employer-employee matched data, often collected by governments, to study natural experiments. Using this data requires matching it with information on management practices (as in Bender et al. 2018 and Cornwell et al. 2021).

Experiments are obviously attractive sources of data, especially when they can be supplemented with knowledge gained from spending time in the subject firm. Yet not all management practices can be subject to experiments; firms will not radically change their entire operating practices to please researchers. The phrase “observational data” has been coined to denote the many papers using firms’ existing data and then modeling how they operate. This tradition goes back to Lazear (1992) and Baker et al. (1994), producing the first powerful papers of the job structure of large firms.

Observational data can also contain some quasi-experiments. For example, Lazear et al. (2015) identified the value of a boss by obtaining boss fixed effects on worker productivity using workers mobility across bosses.

5.2 Why do preferred management practices change over time? Why do firms adopt particular practices?

Changing practices allow identification of the effect of management practice on performance. The question posed in this section is, if a management method is valuable, why was the firm or its competitors not already using it? Table 2 presents three alternative answers.

INSERT TABLE 2 HERE

In Section 2.1, we argued that management and management practices should be in the production function. Therefore, the left column of Table 2 presents innovations in management practices over time, making these a management technology shock. The Toyota production process's spreading internationally, introducing many "modern" human resource practices, is a clear example that was documented in the steel lines of Ichniowski & Shaw, 1995, 1999. Similar innovations also arose outside manufacturing (Lemieux et al. 2009).²

We focus on human resource (HR) practices in much of what follows. This is driven both by our interests and by the weight of research. An important lacuna is that we do not, and cannot, address practices that are not documented, but are instead subject to "relational contracts," unwritten understandings among members of the firm --- management and employees --- about how things will be done. See Gibbons & Henderson (2013) for a rich discussion of relational contracting.

Recent examples of HR innovations include the increasing use of referrals (in which existing workers refer a friend to their employer (Burks et al. 2015), of the "engagement survey," in which workers appraise the value of their bosses (Hoffman & Tadelis 2021), and online job matching of workers to jobs (Stanton & Thomas 2016).

The second type of innovation is digital shocks, in column 2, where "digital" refers to a large range of innovations, such as computerized methods of production or communication. When new technology enters a firm, optimal management practices also change. For example, in Lazear (2000), the firm adopted performance pay, but the reason for doing so was that it had begun keeping computerized records of each employee's performance. Gosnell et al. (2020) do a careful analysis of airline pilot productivity, a study enabled by new developments in aircraft data processing that captures precise flight-level measures of pilot fuel use productivity. See also Bartel et al. (2007) and Bresnahan et al. (2002), which show that digital innovations were combined with HR changes. Today's digital changes, of AI, large storage, and cloud computing, are discussed in Sections 6 and 9.

Firms may change their production methods so that they adopt new management technologies at the same time as digital technologies: often they are complements, as, for example, computerization required modern HR practices as workflows changed. This occurred in Lazear (2000), Bartel et al. (2007), Bresnahan et al. (2003), and Bloom et al. (2014b).

² Bartel et al. (2007) note that the share of large firms with at least 20% of workers covered by some kind of performance pay, even small bonuses, rose from 38% to 67% over 1987 to 1996. Lemieux et al. (2009) show dramatic increases in performance pay from the late 1970s to the early 1990s.

The third reason for adopting new management practices, in column 3, is that firms become informed about already available management practices. Either they did not know of the practices, as appears common in developing countries (Hyland et al. 2019), or managers did know about the practices but (often wrongly) believed they were not right for their firms (Bloom et al. 2013).

A final reason for adoption, not displayed in Table 2, is that the strategy of the firm changes. Few existing economics research papers have information on changing strategies that result in changes in management practices. In reality, however, adjustments in strategy requiring changes in management should be common.

5.3 Challenges in estimation methods, and the evolution of “insider econometrics”

Empirical methods have evolved over the last 30 years, and will continue to evolve: at one time, researchers did not ask what firms did, and the use of today’s experimental methods is very recent. Challenges and opportunities will be part of the future.

In Section 4, the point was made that individual HR practices are often complements to each other. A challenge is that, taking the theory seriously, researchers should not be able to test for complementarity, or for a positive interaction between practices, because a firm should never have one practice and not the other if they are complements. Using observational data, tests for complementarity arise only when a firm makes a “mistake” and does not include all the best management practices, or is transitioning and does not have all the system elements in place.

Complementarity can be incorporated in field experiments, though probably not in natural experiments, which reflect past decisions. Field experiments can introduce multiple treatment arms, including treatments that interact management practices. In this case, field experiments are very much like medical experiments: the researcher may know from focus groups that the two treatments should be adopted together, but permits the relative failure of some subjects by giving them only one treatment element.

The methods used in all this work can be considered as “insider econometrics.” Ichniowski and Shaw created the term to describe their research method of going inside firms to gain understanding from managers and employees that would inform and enrich their statistical work (Shaw 2009; Ichniowski & Shaw 2013).

Today’s papers using “insider” methods are improving over those of the past. Originally, field experiments were very simple, as in Esther Duflo’s early development treatments. Papers today with a field experiment as their cornerstone now often feature multiple treatments and interactions, surveys of employees before and after the treatment, quotations from the company, and a careful discussion of the context. The field study begins by talking to managers, to get their views, but then the paper tests managers’ implicit hypotheses. All of this makes for high quality research, in which the reader is best able to understand not only the treatment effects, but also why and how they occurred. And, honestly, a research paper containing detailed inside information makes the paper more memorable to readers.

6. MANAGEMENT PRACTICES: WHAT ARE SOME KEY RESULTS, AND WHY ARE THEY IMPORTANT?

Research papers about management practices are producing illuminating results using the empirical frameworks described above. In this section, we consider the three topics --- performance pay, hiring, and teamwork --- but do not have the space to describe the methodology used by each paper. Many other topics have excellent research papers. But in contrast to these topics, some management practices, like inventory management, supplier relations, or operational performance management, do not have the same wealth of empirical research.

6.1 Performance Pay

The topic of performance pay is certainly both of great importance and long studied. Because there are good reviews of this literature, even though not very recent, we keep this section short. See Gibbons & Roberts (2013), Prendergast (2011), Lazear & Shaw (2007), Syverson (2011).

6.1.1 Piece rate pay. Lazear (2000) models the shift from hourly to piece-rate pay at the Safelite. While there was a direct performance effect as workers work harder, the key insight was that the introduction of performance pay caused workers to self-select, so that those who were good at the job, and thus would now earn the most, were attracted to the firm. This “selection effect” has been displayed in many papers since. See for example, Sandvik et al. (2021).

Ongoing work on performance pay focuses on details of how it is best implemented in specific contexts. Note that past researchers have studied manufacturing, fruit picking, or sales workers, where worker performance data is easily available (Ku 2019, Larkin 2014, Sandvik et al 2021). Experiments are now finding data covering a broader range of jobs.

Another approach to increased productivity is to simply monitor the employees and give them feedback (Gosnell et al. 2020). While this is common in call centers, an excellent experiment with airline pilots shows the extent to which it can increase performance. Using very detailed data on how airline pilots perform – such as meeting fuel load and taxi targets – performance in meeting these targets, and thus in increasing the productivity of the pilots, rose very markedly when pilots knew they were being monitored.

6.1.2 Tournaments. A second type of performance pay is tournaments. The idea is that employees perform better in hopes of getting promoted. The promotion mechanism is like a tournament, as in golf or other sports, because there is only one job at the top and limited places below that (Lazear & Rosen 1981). Employees work hard to earn the chance to be promoted to better paying jobs and to keep being considered for further promotion – so the tournament structure of pay and promotions is effort inducing, just as piece rate pay is.

The voluminous empirical tests of tournaments began immediately after Lazear and Rosen (1981), so the basic theoretical model may not need to be tested again. In one compelling lab experiment, the tournament elicited considerably more effort from the players than a piece rate pay system, suggesting research on tournaments is warranted. However, changes may be occurring: as tenure is falling at firms, do promotion tournaments remain powerful?

6.2 Hiring

It has long been said that hiring is the most important management practice that firms face. Yet,

economists have written relatively little about it. They have rarely had the data needed because firms themselves do not keep information on those whom they considered and rejected. In addition, once a person is hired, information like educational achievement is typically not included in personnel records. All this prevents using observational data such as described in Column 3 of Table 2. In response, some research has focused on experiments, while other research uses firms' data on hiring practices, such as the use of referrals, for which they do keep data.

In any case, the firm does not know the counterfactual, how their rejects would have performed if they had been hired.

Note that promotions and hiring are very similar processes: they represent experience good search. That is, when hiring or promoting, the applicants' value to the firm can only be determined by placing the worker in the new position (DeVaro & Morita 2013). Both processes face asymmetric information. In addition, a matching problem is at work: it is not simply that the firm does not know the person's skills, but it does not know the quality of the job match (DeVaro & Waldman, 2012). There is uncertainty as to whether the person's current skills and character will make him successful in the next position (Benson et al. 2019), and whether his success was due to luck that will be repeated in the next job (Lazear 2004). Meanwhile, in contrast, promotions have the additional role of acting as an incentive (Milgrom & Roberts 1992, p. 364).

6.2.1 Screening job applicants – traditional methods and AI. Several studies examine standard hiring methods. Hoffman et al. (2018) follow the introduction of applicant testing in 15 firms and find that the hiring managers who ignored applicant's test results made worse hiring decisions than those who followed testing recommendations. In the former case, managers thought they had superior "insight," but for the low-skilled jobs being offered, testing was actually better. In contrast, in Wu & Liu (2022) local hiring by managers was better than using recommendations from HR headquarters, because local managers had better private information. Another option is to use outside recruiters, which firms may even prefer for contract work, like programming projects, because recruiters build their own knowledge of top performers who post their availability on online matching markets (Stanton and Thomas 2016). In other conditions, recruiters may produce poor recommendations for the firm, because recruiters themselves have conflicting career concerns (Cowgill & Perkowski 2021).

For large firms that receive too many job applications, AI or another form of data analytics are increasingly used to screen the resumes. Using algorithms to sort and recommend job applicants to employers substantially increases employer's hiring rates from this smaller job pool. This is shown in an experiment done on an online labor market (Horton 2017). Glaeser et al. (2021) ran a field experiment showing that AI algorithms can be better than human judgment because the algorithms use hard data, and the managers think their intuition is better than data.

The technology of AI is improving daily, so hiring using AI models is also improving. Older AI models, described as "supervised learning," produced recommended job candidates based on the traits of the people who had been hiring over past years. This is because the AI model was trained to make recommendations based on historical data, producing results that have "algorithmic bias.." As firms look to hire more diverse workers, innovative AI models that are based on "unsupervised learning" are exploring a wider range of options, thus producing better job candidates for the future of the firm (Li et al. 2020).

Regarding hiring, a key research topic today is algorithmic fairness when using AI. In addition to the

study just mentioned, in a study of software engineers in a large firm, AI algorithms can remove the biases humans introduced into the historical hiring data if the data was noisy, or not highly accurate (Cowgill, 2020). Cowgill & Tucker (2019) discuss management practices for reducing AI algorithmic bias.

6.2.2 Referrals. Firms increasingly use referrals, where a current employee recommends someone to be hired. Using data from nine large diverse firms, Burks et al. (2015) found that referred employees were likely to stay longer and be more productive compared to new hires who had not been referred. There was an association with higher profitability too. Friebel et al. (2022) ran a field experiment in a grocery chain and also found that a referral program reduced attrition, even among referrers and other workers in treated stores, who valued being involved in hiring. The field experiment of Pallais & Sands (2016) showed that referrals contain information about valuable unobserved characteristics of the job candidates. While referrals increase the quality of employer-employee matches, they have raised questions about the effects on diversity.

6.2.3 Using salary histories. Many firms now ask applicants to report the wages they earned on their previous jobs, so that firms know what pay to offer and can undertake pay negotiations with more information. Between 2016 and 2019, 21 states and many localities banned the practice, believing it could perpetuate gender or race inequities, but many applicants still volunteer their past salary data (Agan et al. 2022, Barach & Horton, 2021). In a field experiment, researchers tested the effects of voluntary salary disclosure on job offers. They hired real recruiters to review 2000 fake job applications, in an audit study experiment. They found that employers who viewed the voluntary salary history of a candidate thought the salary history was valuable, and candidates who did not voluntarily reveal it got lower wage offers. However, many candidates did not want to reveal their histories: in a survey, 47% preferred not to volunteer their salary, but many did so, especially women, because they thought not revealing would be a negative signal (Cowgill et al. 2021).

Another legal change surrounding pay and hiring has occurred since 2004, when state-level “pay transparency” laws began to make it legal for employees to disclose their pay to their co-workers (Cullen & Pakzad-Hurson 2021). Where these laws were passed, wages fall about 2%, with hiring reduced among workers with high outside pay options.

6.3 Teamwork

Teamwork is extremely common in firms; and there are really two types of teams, problem-solving teams and production teams.

Four questions are addressed here: why form teams; is there evidence of successful problem solving; is team-based incentive pay valuable; and is there an alternative to team incentives? The over-riding question is, how does a firm obtain the greatest value from teams?

6.3.1 Problem-solving teams. In many circumstances, firms form temporary teams to solve a particular business problem. A firm may want to introduce a new product and forms a team from marketing, design, and production to plan it. Alternatively, in manufacturing a plant manager may address a new production problem by forming problem-solving teams of people from different positions on the line.

Data on white collar teams is rare but emerging. Shangguan et al. (2022) follow a white collar team

responsible for undertaking new construction jobs: team members do the drafting, planning, and running of very large commercial jobs. This paper builds upon the point that members of the team need complementary skills, but at the same time, need low costs of communication (Lazear 1999). These conditions are important as firms are introducing more diversity among team members.

Of course, not all establishments need problem-solving teams. They are best adopted when production is complex (Boning et al. 2007); otherwise engineers or managers may solve the problems best. And problem solving teams can also be modelled using social network analysis, to see who talks to who on a daily basis, and thus whether the teamwork is among equals or whether the team is really run by the manager at the top. Social network graphs make clear, for example, that problem-solving on steel lines works best when production workers talk to each other (Gant et al. 2004).

6.3.2 Production teams. The production team, in which all members are working on one overall job, such as sewing clothes or assembling cars, is another form of team. In this case, there is diversity across people. Importantly, some workers are more skilled than others. Ideally they become teachers, and team productivity exceeds individual productivity. Hamilton et al. (2003; 2012) have a natural experiment in which workers sewing clothes were switched from individual piece-rate pay to team pay, and plant productivity rose substantially. The teams had heterogeneous worker abilities, and the star workers volunteered to teach others to improve, so group productivity rose. In another example, nurses on a hospital ward were all doing the same job of taking care of patients, but they coordinated care as a group, so if a nurse left, group productivity fell (Bartel et al. 2014). In auto assembly, when a plant switched from workers' doing only one task to rotating across five or six related tasks a day, productivity rose as workers learned from each other and perhaps enjoyed less tedium. (Levitt et al. 2013).

Across problem-solving teams or production teams, when workers share knowledge, performance increases. Sandvik et al. (2020) show that a treatment in which co-workers are encouraged to share information in meetings increases individual sales productivity by over 15%.

6.3.3 Performance pay for teams Researchers have not often tested for productivity-enhancing effects of team-based incentive pay, which is actually more common than individual incentive pay. Friebel et al. (2017) conducted a field experiment in which they introduced group bonuses in 193 small bakery shops. The teams responded to small incentives and the bakeries became more profitable. Fredericksen et al. (2022) also found sizable productivity gains when team-based pay was introduced into heavy construction equipment plants. In the field experiment of Burgess et al. (2010), incentives for teams raised individual performance, as managers selected the most efficient workers. In their field experiment introducing team incentives, Bandiera et al (2013) found workers wanted to choose their best co-workers, rather than forming teams of relatives. As in Lazear (2000), in these cases when incentive pay is added, selecting the right workers raised group performance.

For teams, there should be a classic free-rider problem, even with team-based pay: Some workers are slackers while others do the work. One way of reducing free-riding is to use peer pressure (Kandel & Lazear 1992). Being watched by your peers, or watching them, raises worker performance, but in lab settings or grocery store or white collar, the effects are economically small (Mas & Moretti 2009, Falk & Ichino 2006, Cornelissen et al. 2017). Peer pressure effects were large when less-skilled workers used social pressure (Bandiera et al. 2013; Cornelissen et al. 2017).

7. WHAT IS TO COME? STUDYING ARTIFICIAL INTELLIGENCE AND WORKING FROM HOME

If nothing were changing in the world, so that firms' strategies and technologies were stagnant, then there would be much less need for the study of managers and management. But technologies are changing, and management processes are responding. We cover two elements here.

A contentious current issue, referenced earlier in this essay, is the practice of working from home. The Bloom et al. (2015) field experiment with call-center operators showed that productivity, output and job satisfaction rose significantly for those who chose to work from home. However, there are real questions about whether this generalizes to other types of employees in other industries and to what extent it will survive Covid. Barrero et al. (2022) addresses many of these questions and their implications for workers and management .

Turning to AI, it is revolutionizing the way that we work, the processes and the products of firms. It is considered by many to be the next General Purpose Technology, so just as the steam engine, electricity and computerization changed all "everything," dramatic changes will follow from AI. There are a growing number of case studies on the impact of AI on the changing roles within the firm for workers and managers (Jedras & Shaw, 2021, Golden, Jedras, & Shaw, 2022). Researchers are just beginning to do find natural experiments, in which we can model the quality-enhancing or productivity enhancing effects of AI. A new research paper demonstrates how AI is used to target wildfire risks to homes, enabling firms and homeowners to mitigate the risk by cutting foliage (Jedras & Shaw, 2022a, 2022b). While previous papers on the impact of AI have focused on the substitution of AI for people, new work is emphasizing the ways in which AI is also "augmenting" work. This is particularly true for managers.

Thus, an important for researchers in this field to understand the ubiquity of AI and to be aware of the many ways it is already affecting both firms and consumers. As describing in the Hiring subsection, researchers have begun to study AI impacts in HR functions, where AI was introduced some time ago. It is now being used in many other aspects of management. Economists who gain the cooperation of firms using AI will have rich research opportunities.

8. THE LENS FOR INTERPRETING EMPIRICAL PAPERS: STRATEGY AND TREATMENT-OF-THE-TREATED

For much of the methods of empirical work with individual firms, a key question is whether the results "generalize." If the paper is not making broadly applicable points, then why publish it?

Experiments, whether conducted by the researcher or natural, are usually done with one firm. Therefore, the reader must be convinced that what works in that firm probably also matters in similar firms. Experiments done with multiple firms are partial exceptions. Examples are the studies of Indian textile manufacturing and American steel production. Still, it would be good if the reader could believe that results from one industry were likely to generalize to other industries.

At the same time, empirical results probably should not be generalizable to all firms. It is very unlikely that there is any management practice that is uniquely best in all circumstances.

At a minimum, firms are in different industries, with distinct technologies, suppliers, customers and competitors. They must create individual strategies to give sustainable competitive advantage in their particular context, and management practices must match the strategy of the firm. For example, a strategy like that of Southwest Airlines', where the need to turn around flights quickly led to an emphasis on cabin-crew teamwork, led to contrasting practices to those in the airlines that focused on exquisite service. Differences in regulation provide another important source of heterogeneity.

Because a management practice should only be adopted when it is aligned with the firm's strategy, the econometric exercise of estimating the practice's impact on productivity is one in which the empirical model is identifying the "treatment of the treated." Firms should be expected only to permit experiments that might have a positive impact. That researchers are not observing fully randomized choices and corresponding effects needs to be recognized in their work, and especially in interpreting results.

9. MANAGERS AND MANAGEMENT PRACTICES: IS THERE A LINK?

This paper has two themes: The effectiveness of middle managers and of management practices. Is there a link between the two?

The basic fact is that senior executives decide on company-wide practices.³ Yet, while middle managers' scope for affecting practice selection is limited, it is not always negligible. For example, the plant managers in the Indian textile mills studied in Bloom et al. (2013) had considerable discretion in adopting many of the operational practices advocated by the consultants. This conclusion was reinforced by new plant managers' dropping practices that had been well established (Bloom et al. 2020). Unfortunately, it was not possible to measure manager quality in these studies. Nor was it possible to identify the effects of individual practices. Thus, looking for linkage was not possible. However, it may be possible to gather such data in other studies, and it would be worthwhile to do so.

In some cases, middle managers have little impact choosing management practices, but the management practices and manager quality themselves both directly affect worker productivity. For example, in Ichniowski et al. (1997), great area line managers have a sizable impact on the productivity of their mills, as shown in the significance of manager fixed effects and management quality variables. However, innovative HR practices also had a very strong effect in elevating mill productivity. Good managers and good practices did not interact, reaching the conclusion that managers and management practices have large independent effects on productivity.

Two other papers get a different result. In these cases, managerial quality does interact with management practice scores, where these scores include people management and operational practices in the World Management Survey (WMS). Bender et al. (2018) and Cornwell et al. (2021)⁴ found that firms with higher WMS scores had higher quality managers, so that accounting for managerial quality substantially reduced the measured impact of practices on productivity.

³ A recent contribution in this line by Baltrunaite et al. (2021) found a significant linkage between top management quality and the adoption of management practices in Italian firms. Better managers and better practices both improved performance, and the two together had an effect greater than the sum of the two individually: They were complements.

⁴ The latter paper replicated the former using data from Brazil rather than Germany and employing a better measure of which employees were managers.

Reconciling these different results, the line managers in steel mills were lower level managers: the plant manager would choose the HR practices, not the line managers. In contrast, in the smaller manufacturing firms on Bender et al. (2018) and Cornwell et al. (2021), their managers were running the entire plant.

10. CONCLUSION AND AGENDA

The goal of this paper has been to explore rigorously and critically the economics literature on managers and management practices. Much of this literature is recent, but it is really first-rate and we find it inspiring.

We have suggested that there are largely three possible empirical frameworks for this sort of work: Data can come from field experiments, natural experiments, and firms' records. We have identified outstanding work using each of these approaches. However, at several points, we have advised caution. Existing research papers rarely point out that a particular management practice is only of value to a firm if the practice is aligned with the strategy of the firm, and thus the effects of management practices on productivity are "treatment-of-the-treated" effects and should be thought of that way. A management practice is also of more value if it is complementary to other practices, and this fact is not recognized in most field experiments estimating practice effects.

What is to come? New management practices are introduced for two reasons. One is that there was a "technology shock" such as computerization or AI, or simply better management ideas. A second is that firms are not well run, and thus in field experiments, researchers can introduce practices that are new to the firm. Both these offer rich opportunities for modeling and estimation.

Meanwhile, there are also fundamental issues surrounding managers, such as the origins of their success or failure and the determinants of the choice of what matters to delegate, that call for theoretical analyses and for empirical investigation of the new models' predictions. Managers make firms work, and understanding them and what they do is really important.

Thus, we see our subject of managers and management as an exciting, worthwhile area for future research.

References

- Adhvaryu A, Kala N, Nyshadham A. 2022. Management and Shocks to Worker Productivity. *J. Political Econ.* 130(1):1-47
- Agan A, Cowgill B, Gee L. 2022. *Salary History and Employer Demand: Evidence from a Two-Sided Audit*. Working paper, Columbia University
- Aghion P, Tirole J. 1977. Formal and Real Authority in Organizations. *J. Political Economy.* 105(1):1-29
- Alexopoulos M, Tombe T. 2012. Management Matters. *J. Monetary Econ.* 59(3):269-85
- Alonso R, Dessein W, Matouschek N. 2015. Organizing to Adapt and Compete. *Am. Econ. J. Microecon.* 7:158-87
- Athey S, Roberts J. 2001. Organizational Design: Decision Rights and Incentive Contracts. *Am. Econ. Rev.* 91(2):200-05
- Baker G, Gibbons R, Murphy K. 2002. Relational Contracts and the Theory of the Firm. *Quarterly Journal of Economics.* 117(1):39-84
- Baker G, Gibbs M, Holmstrom, B. 1994. The Internal Economics of the Firm: Evidence from Personnel Data. *The Quarterly Journal of Economics.* 109(4):881-919
- Bandiera O, Barankay I, Rasul I. 2013. Team Incentives: Evidence From a Firm Level Experiment. *J. Euro. Econ. Assoc.* 11(5):1079-114
- Bandiera O, Prat A, Hansen S, Sadun R. 2020. CEO Behavior and Firm Performance. *J.P.E.* 28(4):1325-1369
- Bandiera O, Barankay I, Rasul I. 2013. Team Incentives: Evidence From a Firm Level Experiment. *J. Euro. Econ. Assoc.* 11(5):1079-114
- Barach M, Horton J. 2021. How Do Employers Use Compensation History? Evidence From a Field Experiment. *J. Labor Econ.* 39(1)
- Barrero J, Bloom N, Davis S. 2022. *Why Working from Home Will Stick*, NBER Work. Pap. 28731
- Bartel A, Beaulieu N, Phibbs C, Stone P. 2014. Human Capital and Productivity in a Team Environment: Evidence from the Healthcare Sector. *Am. Econ. J: App. Econ.* 6(2):231-59
- Bartel, A, Ichniowski C, Shaw K. 2007. How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills. *Q. J. Econ.* 122(4):1721-58
- Bartel A, Jedras J, Pastorino E, Shaw K. 2022. *The Value of Managers*. Work. Pap.
- Bender S, Bloom N, Card D, Van Reenen J, Wolter S. 2018. Managing practices, workforce selection, and productivity. *J. Labor Econ.* 36(S1):S371-S409
- Bertrand M, Schoar A. 2003. Managing with Style: The Effect of Managers on Firm Policies. *Q. J. Econ.* 118(4):1169-1208
- Bianchi N, Giorcelli M. 2022. The Dynamics and Spillovers of management Interventions: Evidence from the Training within Industry Program. *J. Political Econ.* 130(6):1630-75
- Bloom N, Van Reenen J. 2007. Measuring and Explaining Management Practices Across Firms and Countries. *Q. J. Econ.* 122(4):1351-1408
- Bloom N, Eifert B, Mahajan A, McKenzie D, Roberts J. 2013. Does Management Matter? Evidence from India. *Q. J. Econ.* 128(1):1-51
- Bloom N, Mahajan A, McKenzie D, Roberts J. 2020. Do Management Interventions Last? Evidence from India. *Q. J. Econ. American Economic Journal: Applied Economics.* 12(2):198-219
- Bloom N, Garicano L, Sadun R, Van Reenen J. 2014. The Distinct Effects of Information Technology and Communication Technology on Firm Organization. *Mgt. Sci.* 60(12):2859-85

- Bloom N, Lemos R, Sadun R, Scur D, Van Reenen J. 2014. JEEA-FBBVA Lecture 2013: The New Empirical Economics of Management. *J. Euro. Econ. Assoc.* 12(4):835-76
- Bloom N, Liang J, Roberts J, Ying A. 2015. Does Working from Home Work? Evidence from a Chinese Experiment. *Q. J. Econ.* 130(1):165-218
- Bolton P, Dewatripont M. 2013. Authority in Organizations in Gibbons R, Roberts J, eds. *Handbook Org. Econ.* Princeton: Princeton University Press. 442-72
- Boning B, Ichniowski C, Shaw K. 2007. Opportunity counts: Teams and the effectiveness of production incentives. *J. Labor Econ.* 25(4):613-50
- Bresnahan T, Brynjolfsson E, Hitt L. 2002. Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Q. J. Econ.* 117(1):339-76
- Brynjolfsson E, McElheran K. 2016. *Data in Action: Data-Driven Decision Making in U.S. Manufacturing.* Rotman Work. Pap. 2722502
- Burgess S, Propper C, Ratto M, von Hinke S, Scholder K, Tominey E. 2010. Smarter task assignment or greater effort: The impact of incentives on team performance. *Econ. J.* 120:968-89
- Burks S, Cowgill B, Hoffman M, Housman M. 2015. The Value of Hiring through Employee Referrals. *Q. J. Econ.* 130(2):805-39
- Caliendo M, Opromolla L, Rossi-Hansberg E. 2020. Productivity and Organization in Portuguese Firms. *J. Poli. Econ.* 128(11):4211-57
- Caliendo M, Cobb-Clark D, Pfeifer H, Uhlendorff A, Wehner C. 2022. *Managers' Risk Preferences and Firm Training Investments.* IZA Disc. Pap. 15043
- Carter S, Dudley W, Lyle D, Smith J. 2019. Who's The Boss? The Effects of Strong Leadership on Employee Turnover. *J. Econ. Behavior and Org.* 159:323-43
- Coase R. 1937. The Nature of the Firm. *Economica.* 4(16):386-405
- Columbo M, Delmastro M. 2004. Delegation of Authority in Business Organizations: An Empirical Test. *J. Industrial Econ.* 52(1):53-80
- Cornelissen T, Dustmann C, Schönberg U. 2017. Peer Effects in the Workplace. *Am. Econ. Rev.* 107(2):425-56
- Cornwell C, Schmutte I, Scur D. 2021. Building a Productive Workforce: The Role of Structured Management Practices. *Mgmt. Sci.* 67(12):7291-50
- Cowgill B, Tucker C. 2019. Economics, Fairness and Algorithmic Bias. In preparation for *J. Econ. Perspectives*
- Cowgill B. 2020. *Bias and Productivity in Humans and Algorithms: Theory and Evidence from Resume Screening.* Columbia University
- Cowgill B, Agan A, Gee L. 2021. *Salary History Bans and Unravelling: Psychological Costs and Coarse Beliefs.* NBER Work. Pap.
- Cowgill B, Perkowski P. 2022. *Delegation in Hiring: Evidence from a Two-Sided Audit.* Work. Pap.
- Cullen Z, Pakzad-Hurson B. 2021. *Equilibrium Effects of Pay Transparency.* NBER Work. Pap. 28903
- Dessein W, Santos T, Galeotti A. 2016. Rational Inattention and Organizational Focus. *Am. Econ. Rev.* 106:1522-36
- Dessein W, Santos T. 2021. Managerial Style and Attention. *Am. Econ. J. Microecon.* 13(3):372-403
- Dessein W, Lo D, Minami C. Forthcoming. *Coordination and Organization Design: Theory and Micro-Evidence.* Am. Econ. J. Microecon
- Gibbons R, Waldman M. 1999. A Theory of Wage and Promotion Dynamics Inside Firms. *Q. J. Econ.* 114(4):1321-58
- DeVaro J, Morita H. 2013. Internal Promotion and External Recruitment: A Theoretical

- and Empirical Analysis. *J. Labor Econ.* 31(2):227-69
- Dustmann C, Lindner A, Schönberg U, Umkehrer M, Vom Berge P. 2022. Reallocation Effects of the Minimum Wage. *Q. J. Econ.* 137(1):267-328
- Englmaier F, Griss S, Grothe D, Schindler D, Schudy S. 2021. *The Value of Leadership: Evidence from a Large-Scale Field Experiment*. CESifo Work. Pap. 9273
- Falk A, Ichino A. 2006. Clean Evidence on Peer Effects. *J. Labor Econ.* 24(1):39-57
- Fenzia A. 2022. Managers and Productivity in the Public Sector. *Econometrica.* 90(3):1063-84
- Frederiksen A, Hansen D, Manchester, C. 2022. *Does Group-Based Incentive Pay Lead To Higher Productivity? Evidence from a Complex and Interdependent Industrial Production Process*. IZA Disc. Pap. 14986
- Friebel G, Heinz M, Krueger M, Zubanov N. 2017. Team incentives and performance: Evidence from a retail chain. *Am. Econ. Rev.* 107(8):2168-203
- Friebel G, Heinz M, Hoffman M, Zubanov N. 2022. What Do Employee Referral Programs Do? Measuring the Direct and Overall Effects of a Management Practice. *J. Poli. Econ.*
- Friebel G, Heinz M, Zubanov N. 2021. Middle Managers, Personnel Turnover, and Performance: A Long-Term Field Experiment in a Retail Chain. *Mgt. Sci.* 68(1):211-229
- Gant J, Ichniowski C, Shaw K. 2004. Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations. *J. Econ. Mgt. Strat.* 11(2):289-328
- Garciano L. 2000. Hierarchies and the Organization of Knowledge in Production. *J. Poli. Econ.* 108(5):874-904
- Garicano L, Van Zandt T. 2013. Hierarchies and the Division of Labor in Gibbons R, Roberts J, eds. *Handbook Org. Econ.* Princeton: Princeton University Press. 604-51
- Gee L. 2019. The More You Know: Information Effects on Job Application Rates in a Large Field Experiment. *Mgt. Sci.* 65(5):1949-2443
- Gibbons R, Waldman M. 1999. A Theory of Wage and Promotion Dynamics Inside Firms. *Q. J. Econ.* 114(4):1321-58
- Gibbons R, Henderson R. 2013. What Do Managers Do? in Gibbons R, Roberts J, eds. *Handbook Org. Econ.* Princeton: Princeton University Press:680-731
- Gibbons R, Matouschek N, Roberts J. 2013. Decisions in Organizations in Gibbons R, Roberts J, eds. *Handbook Org. Econ.* Princeton: Princeton University Press:373-431
- Gibbons R, Roberts J. 2013. Economic Theories of Incentives in Organizations in Gibbons R, Roberts J, ed. *Handbook Org. Econ.* Princeton: Princeton University Press:56-99
- Glaeser E, Hillis A, Kim H, Kominers S, Luca M. 2021. *Decision Authority and the Returns to Algorithms*. Harvard Work. Pap.
- Golden J, Jedras J, Shaw K. 2022. CAPE Analytics. GSB Stanford E-791
- Gosnell G, List J, Metcalfe R. 2020. The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains. *J. Political Econ.* 128(4): 1195-1233
- Grönqvist E, Lindqvist E. 2016. The Making of a Manager: Evidence from Military Officer Training. *J. Labor Econ.* 34(4):869-98
- Halac M, Prat A. 2016. Managerial Attention and Worker Performance. *Am Econ. Rev.* 106(10):3104-32
- Hamilton B, Nickerson J, Owan H. 2003. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *J. Political Econ.* 111(3):456-97
- Hamilton, B, Nickerson J, Owan H. 2012. *Diversity and productivity in production teams*. Advances in the Econ. Analysis of Participatory and Labor-Managed Firms. Alex Bryson, editor
- Hermalin B. 2013. Leadership and Corporate Culture in Gibbons R, Roberts J eds. *Handbook Org. Econ.* Princeton. Princeton University Press. 432-78

- Hoffman M, Kahn L, Li D. 2018. Discretion in Hiring. *Q. J. Econ.* 133(2):765-800
- Hoffman M, Tadelis S. 2021. People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis. *J. Poli. Econ.* 129(1):243-85
- Holmstrom B. 1984. On the Theory of Delegation in Boyer M, Kihlstrom R eds. *Bayesian Models in Economic Theory*. North-Holland Publishing Co. Amsterdam
- Horton J. 2017. The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment. *J. Labor Econ.* 35(2)
- Huber K, Lindenthal V, Waldinger F. 2021. Discrimination, Managers, and Firm Performance: Evidence from “Aryanizations” in Nazi Germany. *J. Poli. Econ.* 129(9):2455-2503
- Hyland M, Francis D, Meza J. 2019. *Are Management Practices Failing or Aiding the Private Sector in South America?* World Bank. Policy Research Work. Pap. 8783
- Ichniowski C, Shaw K. 1995. Old Dogs and New Tricks: Determinants of the Adoption of Productivity-Enhancing Work Practices. *Brookings Papers on Economic Activity: Microecon.* 1-65
- Ichniowski C, Shaw K. 1999. The Effects of Human Resource Systems on Productivity: An International Comparison of U.S. and Japanese Plants. *Mgt. Sci.* 45(5):704-22
- Ichniowski C, Shaw K. 2013. Insider Econometrics: Empirical Studies of How Management Matters. *The Handbook of Organizational Economics*, R. Gibbons and J. Roberts., eds: 263-312
- Ichniowski C, Shaw K, Prennushi G. 1997. The Effects of Human Resource Management practices on Productivity: A Study of Steel Finishing Lines. *Am. Econ. Rev.* 87(3):291-313
- Jedras J, Shaw K. 2021. *Focal Systems: The AI Automation of Brick and Mortar Retail*. GSB Stanford E-64
- Jedras J, Shaw K. 2022. *Cape Analytics: Fighting Wildfire Losses with AI*. Work. Pap.
- Kandel E, Lazear E. 1992. Peer Pressure and Partnerships. *J. Poli. Econ.* 100(4):801-17
- Ku H. 2019. The effect of wage subsidies on piece rate workers: Evidence from the Penny Per Pound program in Florida. *J. Dev. Econ.* 139:12-34
- Larkin I. 2014. The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales. *Journal of Labor Economics.* 32(2): 1-29
- Lazear E. 1992. The Job as a Concept in Bruns J, ed, *Performance Measurement, Evaluation, and Incentives*. Harvard Business School Press:183-215
- Lazear E. 1999. Globalisation and the Market for Team-Mates. *Econ. J.* 109(Mar):C15-40
- Lazear E. 2000. Performance Pay and Productivity. *Am. Econ. Rev.* 90(5):1346-61
- Lazear E. 2012. Leadership: A Personnel Economics Approach. *Labour Econ.* 19(1):92-101
- Lazear E, Rosen S. 1981. Rank-Order Tournaments as Optimum Labor Contracts. *J. Poli. Econ.* 89(5):841-64
- Lazear E, Shaw K. 2007. Personnel Economics: The Economist’s View of Human Resources. *J. Econ. Perspectives.* 21(4):91-114
- Lazear E, Shaw K, Stanton C. 2015. The Value of Bosses. *J. Labor Econ.* 33(4):823-61
- Lemieux T, MacLeod W, Parent D. 2009. Performance Pay and Wage Inequality. *Q. J. Econ.* 124(1):1-49
- Levitt S, List J, Syverson C. 2013. Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant. *J. Poli. Econ.* 121(4):643-81
- Li D, Raymond L, Bergman P. 2020. *Hiring as Exploration*. NBER Work. Pap. 27736
- Lo D, Brahm F, Dessein W, Minami C. 2022. Managing with Style? Micro-evidence on the Allocation of managerial Attention. *Mgt. Sci. Online*
- MacDuffie J. 1995. Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry. *Indus. Labor Rel. Rev.* 48(2):197-221
- Mas A, Moretti E. 2009. Peers at Work. *Am. Econ. Rev.* 99(1):112-45

- Mbiti I, Muralidharan K, Romero M, Schipper Y, Manda C, Rajani R. 2019. Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania. *Q. J. Econ.* 134(3):1627-1673
- Metcalf R, A Sollaci, Syverson C. 2022. Managers and Productivity in Retail. Working paper.
- McMillan J. 2003. *Reinventing the Bazaar: A Natural History of Markets*. New York. Norton
- Milgrom P, Roberts J. 1990. The Economics of Modern Manufacturing: Technology, Strategy and Organization. *Am. Econ. Rev.* 80:511-28
- Milgrom P, Roberts J. 1992. *Economics, Organization, and Management*. Prentice-Hall Intl.
- Milgrom P, Roberts J. 1995. Complementarities and Fit: Strategy, Structure and Organizational Change in Manufacturing. *J. Acct. Econ.* 19(2-3):179-208
- Mookerjee D. 2013. Incentives in Hierarchies in Gibbons R, Roberts J, eds. *Handbook Org. Econ.* Princeton: Princeton University Press. 764-98
- Pallais A, Sands E. 2016. Why the Referential Treatment? Evidence from Field Experiments on Referrals. *J. Political Econ.* 124:1793–1828
- Prendergast C. 1999. The Provision of Incentives in Firms. *J. Econ. Lit.* 37(1):7-63
- Queiro F. 2015. *The Effect of Manager Education on Firm Growth*. Harvard Work. Pap.
- Rajan R, Wulf J. 2006. The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies. *Rev. Econ. Stat.* 88(4):759-73
- Rantakari H. 2011. Organizational Design and Environmental Volatility. *J. Law Econ. And Org.* 29(3):569-607
- Roberts J. 2004. *The Modern Firm: Organizational Design for Performance and Growth*. Clarendon Lectures Mgt. Stud. Oxford: Oxford University Press
- Rotemberg J, Saloner G. 2000. Visionaries, Managers, and Strategic Direction. *The RAND J. Econ.* 31 (4): 693-716
- Sandvik J, Saouma R, Seegert N, Stanton C. 2020. Workplace Knowledge Flows. *Q. J. Econ.* 135(3):1635-80
- Sandvik J, Saouma R, Seegert N, Stanton C. 2021. Employee Responses to Compensation Changes: Evidence from a Sales Firm. *Mgt. Sci.* 67(12):7291-950
- Shangguan R, Devaro J, Owan H. 2021. *Enhancing Team Productivity through Shorter Working Hours: Evidence from the Great Recession*. RIETI Work. Pap.
- Shaw K. 2009. Insider econometrics: A roadmap with stops along the way. *Labour Econ.* 16(6):607-17
- Smeets V, Waldman M, Warzynski F. 2019. Performance, Career Dynamics, and Span of Control. *J. Labor Econ.* 37(4):1183-1213
- Stanton C, Thomas C. 2016. Landing the First Job: The Value of Intermediaries in Online Hiring. *Rev. Econ Studies.* 83:810-54
- Syverson C. 2011. What Determines Productivity? *Journal of Economic Literature.* 49(2): 326-65
- Wu H, Liu S. 2022. *The Trade-offs of Letting Local Managers Make Hiring Decisions*. SSRN Work. Pap.
- Van den Steen E. 2010. Interpersonal Authority in a Theory of the Firm. *Am. Econ. Review.* 100 (1):466-490
- Van den Steen E. 2005. Organizational Beliefs and Managerial Vision. *J. Law, Econ. and Organizations.* 21 (1): 256-280

Table 1 Sources of identification of empirical models of management effects

Field experiment	Natural experiment	Observational data
a) Source of data		
Economist designs experiment. Firm cooperates in implementing it.	Economist learns of experiment that occurred in the past within the firm. Firm provides historical data.	There is no experiment. The ongoing operations of the firm are studied. The researcher may uncover data that produces a quasi-experiment (see below).
b) Research examples		
Bloom, et al. (2013) Pallais & Sands (2016) Sandvik, et al. (2020) Cullen & Pakzad-Hurson (2021) Friebel, et al. (forthcoming)	Ichniowski, et al. (1997) Lazear (2000) Hamilton, et al. (2003) Dustmann, et al. (2022) Bianchi & Giorcelli (2022)	Bloom & Van Reenen (2007) Lazear, et al. (2015) Benson, et al. (2019) Hoffman & Tedalis (2021) Bartel, et al. (2022)

Table 2 Model identification: What changes over time?

a) Type of shock		
Management technology shocks	Digital technology shocks	Firm becomes aware of new technologies, or learns that new technologies are right for them.
b) Research examples		
Ichniowski, et al. (1997) Lemieux, et al. (2009) Burks, et al. (2015) Bloom, et al. (2016) Stanton & Thomas (2016) Hoffman & Tedalis (2021)	Bartel, et al. (2007) Lazear (2000) Bresnahan, et al. (2002) Stanton & Thomas (2016)	Ichniowski, et al. (1995) Bloom, et al. (2014b) For World Bank measures: Hyland, et al. (2019)