NBER WORKING PAPER SERIES

INSIDER ECONOMETRICS: EMPIRICAL STUDIES OF HOW MANAGEMENT MATTERS

Casey Ichniowski
Kathryn L. Shaw

Working Paper 15618
http://www.nber.org/papers/w15618

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2009

Insider Econometrics: Empirical Studies of How Management Matters
Casey Ichniowski and Kathryn L. Shaw
NBER Working Paper No. 15618
December 2009
JEL No. D2,J01,L2

ABSTRACT

This paper describes an approach for conducting empirical research into three interrelated questions that are fundamental to the field of organizational economics:

1.Why do firms in the same industry adopt different management practices?

2.Does the adoption of a new management practice raise productivity?

3.If so, why does the new management practice raise productivity?

This research approach, which we term insider econometrics, addresses these questions by combining insights from industry insiders with rigorous econometric tests about the adoption and productivity effects of new management practices using rich industry-specific data. Understanding the selectivity in the adoption and coverage of different management practices within a single industry is central to this empirical research methodology. The paper considers a number of studies to illustrate persuasive features of insider econometric research and summarizes a number of themes emerging from this line of research.

Casey Ichniowski
Graduate School of Business
3022 Broadway Street, 713 Uris Hall
Columbia University
New York, NY 10027
and NBER
bei1@columbia.edu

Kathryn L. Shaw
GSB, Littlefield 339
Stanford University
Stanford, CA 94305-5015
and NBER
kathryns@gsb.stanford.edu

I. Introduction

Survey data and casual observation suggest that there are striking differences in management practices among firms and establishments that operate within the same industry. Human resource management practices such as types of compensation, teamwork, job design, or training activities vary across firms. Decisions about the scope of the firm, such as extent of outsourcing or vertical integration, and about adopting new technologies are among the many management policies that vary across firms.[1] Industries are not populated by homogeneous firms having a single set of optimal management practices. Why do some firms within the same industry look so different from one another? Are some management practices more productive than others? Do differences in management practices help explain large differences in productivity that exist among firms and establishments within the same industry?[2] Can one set of management practices be the right choice for some firms, while a different set of practices is appropriate for other firms that compete with them?

Researchers are now going inside firms and industries to answer these questions. With firms increasing their use of new software to track productivity of their employees and operations, managers in these firms are looking for new ways of using these data to elevate the performance of their firms. As economists obtain access to these data, they can identify more detailed determinants of productivity and test richer theories of the firm than were previously possible. In this chapter, we describe an empirical research strategy for investigating these questions about the effects of management practices on productivity and the determinants of the choice of management practices. We refer to this research strategy as *insider econometrics*. The term *insider* refers to the use of rich micro-level data on workers or work groups inside firms that share a common production function. The term *insider* also refers to the use of insights from insiders – from managers or employees – that inform almost every facet of the research. The term *econometrics* refers to the use of rigorous statistical tests of the effects of management practices on

---

[1] See Osterman (1994, 2000), Lawler, Mohrman, and Ledford (1995), and Lawler, Medford, and Benson (2001) for surveys on HRM practices. See Bresnahan, Brynjolfsson and Hitt (2002) for data on computer technologies and other management practices.

[2] Haltwanger (2008) documents very large differences in productivity across establishments within narrowly defined industry categories.

productivity, or tests of why some firms adopt one set of management practices rather than an alternative set. In sum, insider econometrics research combines insiders' insights with econometrics techniques applied to the right data to reveal when and why management practices matter.

II. The Distinctive Characteristics of Insider Econometrics Research

Five characteristics are common in insider econometric studies. We first list these features and then describe them in more detail. Insider econometric studies:

1. Estimate a productivity regression in which productivity is a function of some management practice.
2. Identify why management practices raise productivity and where the practice has larger and smaller effects on productivity
3. Model the adoption of the management practices.
4. Analyze micro-level data on production units, such as individual employees, teams of employees, or larger work groups such as establishments, that share a common production process
5. Use field research and interviews from industry insiders to formulate testable hypotheses and interpret results.

The first three features identify the objectives of insider research – what questions do insider econometric studies try to answer? The last two features are two defining methodological features of insider econometric research – what are the main methods used to answer these questions?

The first question insider econometric research tries to answer is *does a new management practice raise productivity*. Thus, the first step is the estimation of the treatment effect in a productivity regression, where the treatment is a new management practice and the treatment effect is the impact of that practice on productivity. All insider econometric research draws upon the literature of treatment effects. Treatment effect literature in economics is often concerned with estimating the effects of changes in public policies, such as the effect of an increase in unemployment compensation on labor supply.[3] In insider econometric studies, the policy change is the adoption of a new management practice. However, applying treatment effect methodologies to behavior of firms and their workers offers new opportunities and new challenges to researchers.

[3] For a discussion of the evolution of treatment effect research in economics, see Levitt and List (2008)

Throughout this chapter, we refer to the estimation of "productivity" regressions, but this term is really a short hand for the estimation of performance regressions. The performance outcome is more accurately defined as any variable that the firm monitors and the worker controls that affects the firm's profits. Therefore, performance is rarely output per hour, but might instead be product quality, or production line downtime, number of customers processed per hour, or worker absenteeism. The performance variable is also rarely profits, because profits are measured at the level of the firm and management practices are typically not the same across all workers in a firm. Production workers are not covered by the same practices as managers, and employees in one site may be covered by different practices than those at another site.

The second question addressed by insider econometric research is a natural extension of the first. *Why does the new management practice raise productivity?* To build richer theories of the firm, researchers must find out why a management practice is or isn't effective. One way insider studies address this question is by identifying reasons why the treatment effect of a management practice varies across workers, work groups, or establishments in the same industry. Productivity may increase considerably after a new practice is adopted among some workers or work groups in the study while for others the effect is negligible. This variation can help identify the mechanisms and behaviors that explain the productivity increases. Insider studies allow workers and work groups to respond differently to a management practice and therefore estimate the production function with heterogeneity in the management treatment effect.

The third question that insider econometric research addresses is *why is the new management practice adopted?* In insider studies that examine data from multiple firms, the same management practices are not adopted by all firms for all work groups. One reason for differences in the adoption of some management practice across workers or establishments in an industry is that the productivity impacts vary. Insider econometric studies also try to identify additional costs and benefits of adoption (beyond the expected productivity gains) that explain differences in adoption. By modelling adoption accurately, the researcher can identify any selection bias in the estimation of the treatment effect in the productivity regressions. Even if the researcher does not have the data to estimate the full adoption equation, reports on managers' views of the reasons for adoption can be very helpful in interpreting results from the productivity equations [4] These questions about the effects of management practices on productivity and the reasons for adopting different management practices are the focus of insider econometric research. The last two features in the list are distinctive methodological features of insider econometric research.

The fourth element of insider econometric research is analysis of micro-level data from one narrowly defined production process. The observations that form the data set should be the natural "production units" for the single production process being studied. These production units can be individual workers if employees work alone, small groups of employees as in problem-solving or project-development teams, or relatively large groups of employees such as retail stores or production lines in manufacturing plants. This feature offers a number of obvious advantages in addressing the three questions listed above. The focus on one specific production process helps isolate the productivity effects of new management practices, reduces concerns about omitted variable bias in the productivity regression, and allows the researcher to build and estimate an accurate production function. For example, by modelling one specific production process, the researcher can choose a logical dependent variable for that process – sales volume for salesmen, calls for call center workers, flight delays for airline carriers, or downtime for a production line – and also identify determinants for that specific measure of performance.

The last feature in the list is perhaps the defining feature of insider econometric research – consultation with industry insiders who have direct experience in the production process. These insider insights help identify: practices and policies that are relevant to managers in the industry, the most appropriate measures of productivity, other determinants of productivity that might be correlated with the management treatment, reasons why a management practice was adopted for one work group or plant but not another, how employees responded to the new management policy, and so on. Insiders therefore are instrumental in identifying meaningful hypotheses about the effects of practices on productivity outcomes and on worker behaviors, and about the determinants of adoption of the practices. They also can help identify what factors need to be included

[4] However, many insider studies reviewed below do try to add up the broader economic and welfare effects of a given management practice beyond its effects on productivity outcomes, such as the effects on overall firm profitability or the prospects for future growth in adoption of the management policies.

in an accurate model of productivity or adoption and therefore what kinds of data are needed to estimate those models.

Insider studies usually cover four of the five steps listed above. Specifically, insider data sets often do not permit the researcher to estimate both the productivity regression and the adoption regression. As the examples below show, researchers will often estimate the production function, but carefully describe the likely conditions for adoption based on interview evidence. Other examples below show that researchers estimate the adoption regression, and use interviews to describe the production function underlying the adoption decisions. Overall, all insider studies aim to marshal micro-level data and insights from industry insiders to study: the effects of management practices on productivity; changes in worker behavior when new practices are adopted; the costs of the practices; and ultimately, the reasons for the adoption of the practices.

III. Illustrating the Analytical Challenges of Insider Econometric Research

Researchers face several analytical challenges, and many unique opportunities, in undertaking insider studies. Firms are making optimizing decisions when they choose management practices, and workers are making optimizing decisions when they respond to management practices. Therefore, our models of the effects of these practices on productivity face all the traditional problems of using non-experimental data. There is likely selection bias and endogeneity in the choices of workers and managers. There is omitted variable bias in the production function. We can't know the unobserved counterfactuals about what would have happened if non-adopting firms adopted some new management practice or if adopting firms had not adopted. However, many things make insider studies quite different from other treatment effect studies. The quality of the data, the information from industry insiders who adopt the practice or work under it, and the types of comparisons the researcher wants to make with the data, all differ from more traditional treatment effect research.

To illustrate the challenges that insider econometric research confronts, the production functions and patterns of adoption are displayed graphically in Figure 1 for a hypothetical insider study. Figure 1 shows two age-productivity profiles. For convenience, we refer to these productivity profiles as the profiles for two groups of establishments, such as plants in a manufacturing industry or stores in a retail industry.

However, insider data sets focus on single production function, and thus the data is comprised of observations on workers, work teams, or parts of an establishment.

FIGURE 1 HERE

These two hypothetical productivity profiles in Figure 1 reflect several assumptions that are made or tested in insider econometric studies. First, despite being in the same industry and even having the same basic production process, differences still exist between the two sets of plants; the Type 1 group has a higher level of productivity than the Type 2 group. Second, both productivity profiles are upward sloping, reflecting the effects of factors like learning-by-doing and experience. In this illustration, we let the growth rate in productivity be the same for the two types of establishments. Third, at t*, management in the high productivity Type 1 establishments adopts some new practice, for example a new work practice or technological enhancement in their operations. Management in the low productivity Type 2 establishments does not adopt. The new management practice causes a one-time increase in productivity by an amount equal to ΔP^T. Finally, productivity continues to grow in both types of establishments after t* due to continued effects of experience and learning.

Figure 1 also displays assumptions about two unobserved counterfactuals. First, we assume that productivity among Type 2 non-adopting establishments had they adopted (ΔP^N) would have also been positive but smaller than what was observed for the Type 1 adopters. Second, the post-t* profile for the adopters would have continued along the same pre-t* trajectory but without the one time increase.

The hypothetical example in Figure 1 highlights the basic empirical challenges. Consider two types of data that could be used to estimate the "treatment effect" of the new management practice on productivity. First, consider the case where the researcher obtains true experimental data – data from a sample of organizations with random adoption of the new practice at time t*. Figure 1 does not depict this case but rather the case of non-random adoption: organizations that would experience bigger productivity gains are the adopters in the figure.

If adoption of the management practice were randomly assigned across both Type 1 and Type 2 establishments as part of an experiment, the estimated gain in productivity from the new practice will be an average of ΔP^T and ΔP^N since the group of adopters in an experimental design would include both types of establishments. As with any experiment, this estimated gain is an unbiased estimate of the average effect of the practice on productivity across all plants in this industry. However, in the real world, this estimate of the change in productivity due to the management practice will never be observed within any single establishment: a firm will never randomly adopt a new management practice. From the perspective of a manager in any of the organizations in the data set, the experimental estimate of the productivity gain due to the adoption of the practice will not answer his exact question: what happens to productivity in *my* business if I adopt this new management practice?

Managers, as well as the economists who study their firms, want to understand three fundamental issues: (1) the economic reasons why competing organizations would and would not adopt some productivity-altering practice; (2) what the differential impact of the practice on productivity would be between organizations that do and do not adopt the practice; and (3) why the practice was effective in raising productivity among the adopters. Without additional attention to the factors that differentiate the Type 1 from the Type 2 establishment, the experimentalist who calculates the average effect of the management treatment across the full sample does not address this set of questions. An experiment with random assignment of the "management treatment" to firms eliminates selection bias, and tells the manager what he and his competitors could achieve on average if all adopted. However, random assignment of the treatment does not offer all the methodological advantages one desires in insider studies.[5] The manager wants to know what the expected productivity gain is for his operations and why it will happen.

To address these questions, the researcher must rely on non-experimental data generated by the actual selection process governing the adoption of the management

[5] Management does not adopt practices randomly so insider studies typically confront cases like the example shown in Figure 1 with selective adoption. However, the case of random adoption is relevant for some insider applications, such as those studying personnel records for a single firm that adopts a new work practice for all employees since universally applied practices are not applied selectively (for that firm's population). Also, in a recent example, Bloom and VanReneen (2007) have obtained management permission to institute new HR practices in a sample of firms. By eliminating the question of selective adoption of management practices, such a study can focus on the question of whether the various management practices raise productivity for an average business in the sample. Doyle, Ewer and Wagner (2009) examine the performance of different kinds of physicians through random assignment of patients.

practices. Figure 1 illustrates one possible kind of selection bias: Type 1 establishments have higher levels of productivity and higher expected gains from the new management practice than Type 2 establishments. The manager might not be concerned about this selection bias. If the manager knows, for example, that his establishment belongs to the group of establishments that matches the group that adopts, then with certain assumptions (e.g., that the pre-t* productivity trend for Type 1 would continue in the post-t* period) he could predict his expected gain in productivity with confidence by examining the subsample of data for just the adopters.

While a manager with inside information about how his own operations differ from others in his industry might know if his is a Type 1 or Type 2 establishment, the practical difficulty that the outside researcher faces is that all establishments that compete within an industry using a common production technology can look very similar. An outside researcher is not likely to know the reasons why some establishments adopt new practices and some don't. Without the insights of industry insiders, he might not be able to identify what characteristics differentiate Type 1 and Type 2 establishments,[6] which in turn limits his ability to model the selectivity in adoption accurately.

What do we learn from Figure 1? It provides a clear picture of what insider studies want to model – adoption of management practices and their varying effects on performance outcomes. Figure 1 also highlights several assumptions we must make and the econometric issues we face in conducting an insider study. In the next section, we review six insider studies that confront these issues. As these examples will illustrate, the two distinctive methodological features of insider studies – analysis of rich micro-level data about one specific production process and access to insiders who are knowledgeable about their management practices and operations – are instrumental in constructing convincing empirical tests of the effects of management practices on productivity.

[6] While we describe the Figure 1 as example involving data on establishments or other multi-employee production units, the structure of Figure 1 is also useful when the data describe production by individual workers. Many of our examples are insider econometric studies with worker-level data sets, and we will apply the structure of Figure 1 to interpret these worker-level studies as well.

IV. Examples to Illustrate the Methods of Insider Econometric Research

In this section, we review six examples of insider econometric research to illustrate the common features of this research and to give the new researcher a true feel for both the methods and results from this body of work. After reading these examples, it is much easier for the researcher to understand the econometric methods which we describe next in section V and the key decisions about more specific elements of the research design of an insider study (beyond the use of micro-level data on a single production process) which we describe in section VII.

The review of each example follows a pattern: we describe the intra-industry sample and the production process; the nature of the management practices; the basic empirical findings about adoption of the practice and its effects on productivity; and the results on why productivity does or doesn't rise due to management practices. Variants of the Figure 1 profiles are used to summarize each of these studies. We provide a broader review of insider studies at the end of this section, where we categorize a large number of insider studies according to the type of management practices analyzed.

A. Two Single-Firm Worker-Level Studies on the Productivity Effects of Incentive Pay

The first two studies both use worker-level observations from a single firm. In both, the new practice adopted by management is an incentive pay practice that covers all employees in the firm. Despite these similarities, both confront their own unique methodological issues and each shows the importance of different insider insights.

Incentive Pay at Safelite Glass. Lazear (2000) studies the productivity of workers within one firm before and after the introduction of a new incentive pay plan. Lazear's study models automobile windshield installation for the Safelite Company. In this company, each employee drives a truck to the homes of people who need a new car windshield. The production function is worker-specific and measurable: employees work alone installing about two or three windshields a day. The production unit is the worker. The data are monthly productivity data for some 3000 workers for nineteen months. During this period, the firm shifts from hourly pay to incentive piece rate pay, where pay is a function of the number of windshields the worker installs that day. After the move to piece rate pay, productivity per worker rises by 44%. Workers are highly sensitive to monetary incentives.

Why is this interesting? The reader is not interested in windshield installation. The management practice – piece rate pay – is interesting, because, even though few firms use piece rate pay, many use some form of individual performance pay. Therefore, the evidence that monetary incentives raise workers' productivity is in itself valuable, though perhaps not surprising.[7] What makes the result valuable is that the study identifies *why* productivity goes up. First, as expected, some of the overall productivity increase is due to the same worker increasing his or her effort and output. Second, about half of the increase in the mean productivity of all workers comes from the self-sorting by workers: those workers who have low productivity levels leave the firm (or are not hired by the firm) after incentive pay is adopted. The more interesting result of the paper is that the firm's choice of its incentive pay practice induces sorting by workers, and the sorting itself has a big productivity impact.

Figure 2a illustrates this. The productivity response to the incentive pay is *heterogeneous* across workers. Analogous to Figure 1, the Safelite workers are classified as types – here three types. The pay plan is changed at time t*. The response to the treatment is heterogeneous across the three types. Type 1 workers are employed before and after the new pay plan and these workers raise their productivity (by 22%). The Type 2 workers quit in response to the new incentive pay because their expected pay gain is less than the disutility of their extra effort to be more productive. The Type 3 worker is hired after t*. He has higher productivity and earns more under incentives than those who left would have earned had they stayed. Together, productivity rises by an additional 22 percent more from the effect of losing Type 2 and hiring of Type 3.[8]

FIGURE 2a HERE

----

[7] The adoption of the incentive plan of course cannot be considered exogenous. Safelite adopts the practice because it expects productivity to increase. In section VI below, we consider the question of why "optimal management practices" can change over time.

[8] The figure is not drawn to scale to capture this size of gain. Shaw and Lazear (2007) displays this figure using the actual Safelite data and offers more detail on differences between employees who leave and join this firm after introduction of the new pay plan.

Obviously, not all firms would get this exact productivity gain. This is just one firm that chose the practice probably expecting it to be effective.[9] While one cannot expect the estimated productivity gains of this incentive pay plan to apply to other firms and production processes, the design of this study has several advantages that make the estimates especially convincing for this setting by avoiding typical econometric problems associated with estimating treatment effects. Because the production function is very simple with workers working alone, the unobserved counterfactual productivity paths are very easy to model (based on the observed productivity paths, as shown in Figure 2a). There is little likelihood of omitted variable bias, and thus typical selection bias and endogeneity bias are not relevant to the study. In a study of the impact of management practices on performance using a more heterogeneous sample of production units, we would worry about the possibility of unmeasured factors causing a selection bias with some firms choosing the practice while others do not.

At the same time, in this study, there is still a different source of selection "bias" in the estimated treatment effects – workers decide whether to work at the firm or not and that optimizing decision by workers affects the size of the productivity gain from incentive pay. But the study identifies and measures this self selection, leading to the key theoretical point of the paper. An overall change in productivity can be decomposed into two components that are analogous to the intrinsic and extrinsic margins of each employee's labor supply decision. Some workers respond to the new incentive pay at the intrinsic margin by increasing the amount of labor (in this case effort rather than hours) they supply. Other workers respond at the extrinsic margin by leaving their jobs with different workers taking their place. The broader point for economic theory applies beyond the scope of the single firm in this study. When the firm selects its HR practices, it must consider not only their effects as a motivation device but also their effects as a signaling device to sort specific kinds of workers into and out of the firm.

[9] Note that the reason that the firm chose to adopt incentive pay at this time is that the firm had introduced new information technology software that kept computerized records of each employee's productivity. Thus, Lazear is estimating the joint impact of a change in the HR practice given that the firm made a change in its IT infrastructure. Other firms that introduce piece rate pay may have lower performance gains, and not find the use of piece rate pay optimal (see Freeman and Kleiner, 2005, for evidence from shoe manufacturing).

Incentive Pay Among Fruit Pickers. Bandiera, Barankay, and Rasul (2005) analyze productivity of employees who pick fruit. The workforce in this study is hired only during the summer harvest season. The firm in this study undertakes a series of experiments that change the payment schemes for these employees.[10] In one experiment, the firm introduces new piece rate pay, as in the Lazear paper above. However, in Lazear, the move was from hourly pay to piece rate pay; here the move is from a performance pay plan based on workers' relative performance to the piece rate pay. In the relative output pay plan, the firm fixes the average pay for the field *ex ante*, but pays each worker based on how much he or she does relative to other workers in his or her group (but keeping the mean pay fixed). In the second half of the picking season, the farm switches to simple piece rates in which pay per unit of output is fixed *ex ante* and does not vary as a function of how co-workers perform. The panel data in this study cover 142 workers for 108 days spanning the periods in which different compensation plans are in place. After the switch to simple piece rates, the average productivity of workers rises by 58%. Nearly all workers increase their effort, and the variance of output also rises markedly with a jump in the number of high performers.

FIGURE 2b HERE

Figure 2b illustrates the unique results of this study. The productivity response to the incentive pay is *heterogeneous* across workers. Analogous to Figure 1, there are Type 1 and Type 2 workers with a new incentive policy adopted at t*. Unlike Figure 1, both types of workers are covered by the new incentive pay. Type 1 workers exhibit a bigger productivity jump after t* than do Type 2 workers. The study identifies who the Type 1 workers are – employees who had been working with their friends under the initial relative pay plan. Prior to the piece rate pay plan, these workers withheld effort because they knew that if they worked hard, they would lower their friends' pay. The Type 2 workers have lower productivity gains because they were not working with their friends and thus had been more productive than the Type 1 workers, when they were working under the initial relative pay plan. The researchers reach this conclusion about

[10] For papers on their other experiments, on managerial pay or on teamwork, see Bandiera, Barankay, and Rasul (2007, 2009).

why productivity increases because they gather additional data on friendships among workers.

What makes this study interesting? The readers of this paper are not intrigued by either the occupation of fruit picking or even this specific case of piece rate pay. Unlike the setting for Lazear (2000) where workers came and went, all workers in this study stay for the short fruit picking season covered in the data. Thus, we are looking only at workers who are employed before and after the change in pay policy (in the Figure 2a illustration of the Safelite study, these would all be Type 1 workers). Here however, we are examining the heterogeneity in productivity responses among these workers. The more fundamental economic insight in this study of fruit pickers is that workers with certain characteristics internalize an externality – the externality that their effort harms others – during the relative payment scheme. Thus social relationships determine productivity. More broadly, this example identifies the importance of peer effects and social networks in the determination of worker productivity. Other insider style studies make related points about peer and network effects including: Ichino and Maggi (2000) for banking; Mas and Moretti (2009) for grocery stores; Giorgi, Pellizzari and Woolston (2009) for classrooms in education; and Bartel, Phibbs, Beaulieu, and Stone (2009) for nursing; and Ingram and Roberts (2000) for hotels.

In this study, there is no selectivity within the firm in the adoption of the practice since all employees are covered by the change in pay policy at t*.[11] As in Lazear (2000) the treatment effects estimated are effects conditional on being employed in this firm. But the design of this study is illuminating because it solves other econometric problems associated with estimating treatment effects. In the econometrics treatment effects literature, it is well acknowledged that treatment effects should vary across people. In managing a firm, managers know that not all workers respond equally to a treatment. Often, a study would conclude that the variance in the treatment effects is from some unobserved abilities of the workers, but here the authors go inside the company and

[11] Even though the management policy change covers all employees in the data set, there are still important unobserved counterfactuals; but in this case, these are outside the scope of the study's data. The workers' productivity profiles in Figure 2b come from only a subset of farms. As in the case of Figure 1, there are other farms, not in this particular study's data set, that did not introduce incentive pay, and these farms are analogous to the "untreated" establishments in the bottom half of Figure 1. The fruit-pickers study shows us what would happen to the farming fields that adopt; we don't know anything about the non-adopters that are not covered in this study's data.

gather the data on friendships to identify the source of variance displayed across Type 1 and Type 2 workers. While the data pertain to the limited setting of one industry and firm, the logic of the economic model applies more broadly. In settings where workers might withhold their effort because that effort can harm their co-workers, managers need to consider these peer effects in the design of compensation plans.

B. Two Cross-Company Within-Industry Studies of the Productivity Effects of Innovative HR Practices

In the next two examples, we contrast two studies that both examine the productivity effects of new HR practices inside the steel industry. In both, the production process is a production line involving multiple employees. Despite these similarities, these studies analyze different segments of the steel industry and reach different conclusions about why these HRM treatments are adopted or not.

Human Resource Systems in Integrated Steel Mills. In a third example of insider econometrics research, two companion studies by Ichniowski, Shaw, and Prennushi (1997) and Gant, Ichniowski, and Shaw (2002) analyze productivity of finishing lines within integrated steel mills. The management innovation is the adoption of entire sets of new human resource management practices in integrated steel mills, where the new HR practices cover recruiting and selection policies, compensation, teamwork, communication, and employee training. The authors model production of one specific type of production line that finishes steel. These steel lines take a coil of very thin steel sheet, unwind it, and then chemically treat and coat the steel for use in products like auto bodies. In the 1997 study, the authors collect panel data on productivity outcomes for the finishing lines. The monthly productivity data for the rolling mill work groups within 36 steel mills cover several years for each line. The researchers identify four “systems” of HR practices, ranging from the most innovative HR system to the most traditional system with no innovative HR practices.

The authors show that the mills with complete systems of highly innovative HR practices have the highest productivity levels. Yet, some mills adopt highly innovative HR systems and some don’t. What accounts for the difference in adoption among a set of finishing lines that are competitors within a very narrowly defined industry and that are making very homogeneous products using the same machinery? Site visits to the finishing lines did not reveal any differences in product attributes, technologies, or demographic characteristics of the workers. Site visits did reveal that, even though older and newer vintage lines adopted the innovative practices, the only kind of older lines that adopted the innovative practices were old "reconstituted" lines – lines that had been temporarily shut down and restarted by new owners. Based on this subtle distinction, the authors hypothesize that the difference in the adoption of innovative HRM systems was transition costs. Older continuously operating lines had higher transition costs; newer lines and those older mills that were restarted after a purchase by new owners did not have transition costs associated with changing their HRM systems (Ichniowski and Shaw, 1997). The authors argue that the productivity of older continuously operating lines would also increase were they to adopt innovative HRM systems, but the high transition costs that only these lines would experience produce lower profits from innovative HR practices compared to the newer or reconstituted mills.

What makes this insider study interesting? Once again, readers of the paper do not have to be interested in steel production. The authors show that HRM practices matter – that innovations in HR practices produce tangible performance gains. However, even more interesting is that the authors show *how* HR practices matter and *why* productivity goes up. Productivity increases only when a firm adopts a set of complementary HR practices that together reinforce productive worker behaviors. Insider insights about why HRM practices are complementary lead the researchers to measure the "management treatment" as a set of practices rather than a single HRM practice.

Moreover, research inside the finishing lines revealed why these bundles of HRM practices improve productivity. The innovative systems of HR practices lead to very different behaviors of the workers themselves. Gant, Ichniowski, and Shaw (2002) collect and analyze data on the communications patterns of workers within these lines and show that workers interact much more with each other as part of problem-solving activity in the high productivity lines with innovative HR practices. The lines with innovative HR practices have much denser social networks and worker interactions, or higher levels of "connective capital" in the firm (Ichniowski and Shaw, 2009). This difference in worker behavior provides another type of evidence that helps explain why the treatment effect due to innovative HRM systems exists. Firms need to think about the right bundle of HR practices and not just the individual HR practices. Firms also need to provide workers with opportunities to change their behavior and form the networks and teams needed for problem solving when new incentives are adopted. The costs of switching practices and of changing patterns of worker behavior are transition costs that startups do not incur, and thus the startups are more likely to have the innovative HR systems.

Problem Solving Teams in Steel Minimills. Boning, Ichniowski, Shaw (2007) also study productivity in the steel industry, but within production lines in the minimill sector that use electric arc furnaces. Minimill lines produce steel bar products, like re-bar used in construction and highways or large beams used in constructing buildings. Minimills all belong to one narrowly defined industrial classification, and all have the same production process and technologies. The mills melt steel scrap and cast it into the bar products. As in Ichniowski, Prennushi, Shaw (1997), the data here also come from one particular type of rolling line in these mills, and thus pertain to very homogeneous production processes. The management innovation in this study is the adoption of problem-solving teams. Workers use problem solving skills to increase the quality of the bars coming out of the mill by watching the production process as the bars roll through the mills, and correcting problems as they arise. The authors collect panel data: monthly productivity data for the rolling mill work groups within 34 steel mills followed over about five years. In the beginning of the data, about 10 percent of the rolling mills had problem-solving teams; by the end, about half of the mills had teams.

Productivity regressions show that there are significant productivity gains after teams are adopted in minimills. However, some mills adopt teams and some don't. Again, without insider insights, the reasons for the differences in adoption are not obvious. Each minimill has the same production machinery, and the demographic characteristics of workers like education, experience, and occupation are also very comparable. However, once inside the mills, the reason for the difference was apparent. Some lines producing basic commodity bar products, like rebar steel, have fewer problems to solve compared to other lines making more complex products like thin steel wire or steel with intricate shapes. Using measures of product complexity, the authors show that 100% of the lines producing the most complex products adopt teams, and 23%

of the lines producing the least complex products adopt teams. The productivity regression results reinforce this conclusion: the productivity impact of teams is greater in the complex lines than it is among commodity lines (where teams are rarely adopted).

What makes this insider study interesting is that it identifies two broader economic principles. First, the study shows that product strategy decisions dictate the appropriate HR or organizational design decisions. A strategy of making customized complex products requires HR practices that foster more problem solving than does a strategy of making standardized products. The firm's product strategy determines its HR practices. Second, firms do not benefit equally from teamwork. Here, it is complex production lines that gain the most; while lines making simpler commodity products do not benefit from problem solving.[12] Economic theory suggests that teams will not be adopted by all firms. Without the insider insights about the importance of differences in product complexity, one could have erroneously concluded that "the team effect" observed among complex production lines applied to other lines that seemed to be identical, and that teams were equally valuable across all manufacturers.

We contrast the findings of these two establishment-level studies of the effects of similar HRM practices within the same industry in Figures 3a and 3b. Figure 3a summarizes the conclusions reached by Ichniowski, Shaw and Prennushi (1997) for lines in integrated steel mills. The adoption of innovative HRM systems at t* raises productivity in these lines. But unlike the selection bias shown in Figure 1 in Section III, the authors here argue that the productivity gains would be the same if non-adopters instituted the new HRM practices. The authors conclude that non-adoption is explained by differences in the transition costs of adopting the practices rather than by any differences in the expected productivity gains for non-adopters relative to the adopters.[13]Despite studying similar HRM practices in a different part of the same

[12] The conclusion that 'complex' production processes value teamwork is also generalizable. Recent researchers emphasize that firms in developed countries have a comparative advantage at producing sophisticated products, rather than commodities, so organizational innovations like teamwork are more likely to be adopted in the U.S. than in other countries across all firms (Bloom and Van Reenen, 2007; Bloom, Draca, and Van Reenen, 2008). Bartel, Ichniowski, Shaw and Correa (2009) document a similar conclusion in the valve making industries of the U.S. and U.K.

[13] See Ichniowski and Shaw (1995) for evidence on the transition costs of non-adopters. An alternative way to think about the productivity profiles in this study is to return to Figure 1 as a relevant summary, but instead of considering "HRM systems" as the managerial treatment, let a single HRM practice like

industry, the minimill study of Boning, Ichniowski and Shaw (2007) reaches different conclusions that the study of integrated mills by Ichniowski, Shaw and Prennushi (1997). Figure 3b shows the selection bias in the adoption of teams at t*. Complex lines have larger productivity gains from teams than do the commodity lines.

FIGURE 3a HERE

FIGURE 3b HERE

In both steel studies, an understanding of selectivity in adoption of the management treatment is critical. For the finishing lines, Ichniowski, Shaw and Prennushi (1997) conclude that the new HR innovation would raise productivity among non-adopters but differences in costs of adoption are higher in older continuously operating lines. For minimills, Boning, Ichniowski and Shaw (2007) conclude that teams have fundamentally different productivity effects in different kinds of lines since the benefits of problem solving activity differ. The careful insider econometric analyses uncover different economic insights. In one segment of the industry, complementary HR practices raise productivity but the costs of adopting these practices differs for startups and continuously operating lines. In the other segment of the steel industry, teamwork is more valuable for manufacturers of complex products. There is no one single “team effect.” Differences in productivity benefits across the two types of lines explain adoption.

C. Two Within-Industry Studies of the Adoption of Management Practices that Affect Productivity

In the previous examples, the production processes were relatively well understood. The examples of fruit pickers and windshield installers were individual-level production processes, and the manufacturing processes in steel, while highly

---

incentive pay be the treatment. Type 1 lines could be the lines that adopt incentive pay as part of a larger system of practices and thus experience a large increase in productivity like the Type 1 establishments in Figure 1. In contrast, lines that adopt incentive pay without other elements of the innovative HRM system experience a smaller post-t* increase in productivity like the Type 2 establishments in Figure 1. Viewed this way, it is the existence of other HRM System policies that differentiate Type 1 and Type 2 lines.

technical, are relatively well understood by engineers. In some sense, these settings were the laboratories in which to study the effects of incentive pay, problem solving teams, or new HRM systems on productivity outcomes. In other settings, such as the productivity of many research activities, the outcomes from the production process are much less predictable even by the expert insiders. Yet innovations from basic research are fundamental to economic growth. The next two examples of insider econometric research are studies that shed light on determinants of productivity in research activity in the pharmaceutical industry. The focus of these papers examine fundamental issues in industrial organization economics related to the size and boundaries of the firm – managerial decisions about the number of projects to pursue and about contracting out work versus doing the tasks in house.[14]

The Effects of Scale and Scope on Productivity in Pharmaceuticals. Henderson and Cockburn (1996) study productivity of early stage research activity in the pharmaceutical industry. Research personnel in this phase try to find a chemical compound that has a desirable effect within a laboratory process that mimics aspects of a human disease. In this study, productivity is measured by patents and the management treatments are changes in the scale of the firm's research activity (expenditures per project, per firm, etc.) and scope of research (extent of other research projects in similar classes of compounds in the firm, in other firms, etc.). The authors analyze 5,000 yearly observations on research projects. The data come from ten firms with up to 38 projects per year and up to 30 annual observations on any one project. The authors describe the data set assembly as an extensive iterative process with knowledgeable scientists in the firm that results in accurate and detailed measures of research outputs and inputs.

[14] Industrial organization research offers many other examples of empirical studies that follow an insider econometrics approach going deeply inside firms in one industry. In this chapter, we define insider econometrics by its focus on the internal management practices of the firm within the context of a single production process. Some empirical IO research that examine firms in a single industry would not fit this description; for example, studies of a change in firms' pricing policies (Chu, Leslie, and Sorensen, 2009) or in how information is revealed to customers (Bollinger, Leslie, and Sorensen, 2009; Leslie, 2004), or studies using data on products rather than production processes. See other chapters in this volume (Gibbons and Roberts, forthcoming) for research on these topics using micro-level data. Other examples of empirical industrial organizational research would fit this chapter's description of insider econometrics more closely; for example, whether the internal management practices of stores in an industry look different if the store is run by a franchisee versus an owner-operator (Lafontaine and Shaw, 1999; 2005).

The authors find that knowledge spillovers across research projects within the same firm are critical determinants of productivity in any single project. The scope of the firm, as measured by the number of large research projects, improves the chances of success in a research project. More specifically, spillovers across research on similar classes of drugs within a single firm and spillovers between firms working on the same drug class both raise research productivity. The authors also document a positive affect of the scale of the research project, as measured by research expenditures on the given project, on research productivity.[15]

What makes this insider study interesting? The authors document how "firm size" matters for research productivity. The authors aggregate their project-level data back up to the level of the firm, and replicate a common result in prior firm-level pharmaceutical studies – that measures of firm size have little impact on patenting. When they then analyze the project-level data set they constructed after extensive field investigations of research productivity in these firms, they show just how misleading the conclusion based on firm-level analysis is. New insights surface only when you correctly model a specific research production function at the level of individual projects. Furthermore, their insider insights allow them to identify what aspects of "firm size" matter most – not simply the total amount of research expenditures on a project but the amount of knowledge on related projects. Finally, this study puts the micro-level productivity results into the larger public policy discussion for regulation of the industry. Allowing larger firms in the industry may in fact lead to concerns about monopoly power, but these concerns must be balanced with efficiency gains that only come with the increase in size that comes with running more projects.[16]

Vertical Integration Decisions in the Pharmaceutical Industry. There are still many steps left in the production process that generates a new drug after basic research

[15] In this study, the authors do not model why the treatment varies – that is, why the firms have different numbers of projects or different expenditures per project. However, the authors do use detailed insider insights about the causal nature of the link between the scale and scope measures and research productivity. While their panel data estimates document increased patenting productivity when measures of scope change, they acknowledge that the question of what is the optimal research portfolio for any single firm in the data set is beyond what the scope of their analysis.

[16] Other insider studies listed below provide insights about the determinants of productivity in other knowledge-based or research-intensive industries such as bio-technology, education, health care, law, and finance services and venture capital.

identifies promising new chemical compounds. Several stages of clinical trials take place after a promising drug compound is identified but prior to FDA approval of a new drug. Azoulay (2004) offers insider econometric research about management policy decisions in this phase of drug research.

In this study, the treatment is the drug company's decision whether to outsource clinical trials to contract research organizations (CRO's) or not. The hypothesis is that clinical trials that are very data intensive are likely to be outsourced, while those that are knowledge-intensive are allocated to internal teams. In-house trials will involve more subjective performance evaluations and less sensitive pay for performance, while the trials contracted to CRO's will involve a narrower set of more-easily monitored tasks that can be covered by explicit pay-for-performance contracts. The data are project-level data from 6826 projects spanning all pharmaceutical firms from 1991-1999. Coupling these rich micro-level data with insider insights from extensive field work and interviews, Azoulay estimates models predicting the firm's outsourcing decisions and finds strong support for the main hypothesis. Outsourcing is the better decision when the testing tasks can be reduced to standardized data collection tasks, while testing that involves the generation of new findings and knowledge are more likely kept in-house.

Why is this paper interesting? First, the results of the Azoulay study deepen our understanding of a basic economic question – what determines the boundaries of the firm's activities. Here, rich industry-specific insight identifies specific characteristics that tend to keep work tasks inside or outside the firm's boundaries. Furthermore, this industry-specific example lends strong support to the broader economic theory that multi-task jobs (ones involving both knowledge generation and data generation) where at least one task is not easily monitored are better managed with low-powered incentives and more intensive monitoring. Conversely, when the work involves only data-generation tasks that are easier to monitor, explicit high-powered incentive arrangements (that can be outsourced) are more appropriate (Holmstrom, 1999; Holmstrom and Milgrom, 1994). This vertical integration decision in pharmaceuticals is therefore similar to franchising: the in-house contractors will have more subjective performance evaluations and less sensitive pay for performance, while the outside contractor will have more explicit pay for performance (or high powered incentives) based on a narrow set of easy-to-monitor tasks. Thus, the results from a single-industry study provide results for a theoretical proposition that can be applied more broadly.

In this insider study, data are not available on the success, or productivity, of tests done under alternate treatment regimes. Here, the empirical analysis focuses exclusively on the determinants of the "treatment" – the decision to outsource or not. The conclusion that knowledge-generating tests would be less successful and productive were they to be outsourced to firms using high powered incentive contracts is implicit. While we cannot contrast the patterns of the productivity profiles of these two pharmaceutical industry studies, the two studies taken together make important points about insider research. Even inside the same industry, production processes are highly varied. Henderson and Cockburn (1996, p. 35) show aggregating data on different research projects up to the level of the firm produces misleading results. Insider insights suggested that productivity in this part of the pharmaceutical industry be studied at the level of individual projects so that cross-project productivity could be identified. Azoulay's (2004) study then goes on to reveal important sources of heterogeneity even within the drug testing activities in this industry with different tests leading to different decisions about whether to keep the work in-house or not. Thus, one reason for heterogeneity in practices within a single industry is the fact that firms are managing many different production processes at the same time.

D. Additional Insider Studies

The previous six examples are designed to illustrate some of the differences in research designs that can occur in insider studies because management treatments are or are not applied universally in the sample, observations are for one firm or many firms in an industry, longitudinal data on the choice of practices is or is not limited, and so on. To illustrate a broader range of topics, practices, and data analyzed with this kind of methodology, we present a longer list of studies in Table 1. This table summarizes the management practices studied, the nature of the data set, and a short description of conclusions about the economic mechanisms that explain why productivity changes when the practice is adopted. The six studies reviewed above include several early examples that helped develop the field of insider econometrics. The studies reviewed in Table 1 suggest that the methods of these studies have been followed in studies of many other industries. We will refer to details of the studies of Table 1 in several additional sections below on emerging themes from insider econometric research (Section XIII) and sources of data (Section IX).

TABLE 1 HERE

V. Estimation of Treatment Effects and the Econometrics of Insider Studies

The kinds of productivity profiles shown in Figures 1-3 are typically estimated by analyzing panel data on workers or work teams. In this section, we consider issues that arise in applying econometric methods to obtain accurate estimates of treatment effects. Start with the simplest possible regression for estimating the impact of some management policy on productivity:

(1) $Y_{it} = \alpha D_{it} + \beta x_{it} + \eta_i + \theta_t + \varepsilon_{it}$

where i and t are subscripts for the worker- or work-group-specific observations and for time, Y_{it} is productivity, x_{it} is a set of production function control variables, D_{it} is a dummy variable equal to 1 for the presence of the management practice, η_i is the worker-specific (or work-group) fixed effect, θ_t is the common time period effect, and ε_{it} is the transitory worker-specific effect. How should the insider econometrician elaborate on this simple model to estimate treatment effects of management practices on productivity?

A hallmark of insider econometric research is that it models the heterogeneity in workers' responses to the managerial practice. Even for workers within one occupation, there is heterogeneity in workers' responses to managerial treatments; e.g., not all workers respond the same to incentive pay. This differential response drives the reasons why management practices are or are not effective. Translating this into the productivity regression, rewrite equation (1) to allow the treatment effect to vary across observations:

(2) $Y_{it} = \alpha_i D_{it} + \beta x_{it} + \eta_i + \theta_t + \varepsilon_{it}$

The adoption of the management treatment, D_t, at the time t* is determined by its profitability according to:

(3) $D_{it} = 1$ if $\pi_{it} > 0$ and $t > t^*$

= 0 otherwise

(4) $\pi_{it} = \Gamma Z_{it} + v_{it}$

where the index π_{it} measures the expected profits from adopting the new practice which is a function of observables Z_{it} and an error term.

Equations (3) and (4) imply that at time t*, the treatment, some new management policy, is introduced. The treated group adopts (D=1) when expected profits from adoption exceed zero. The non-treated group (D=0) does not adopt since expected profits after adoption are not sufficiently high. The treated group either has a different underlying production function, or different underlying transition costs, than the non-treated group. Since the adoption of the treatment is not random, selection bias arises.

A. Estimating Treatment Effects

Insider econometricians typically estimate productivity regression (2) – a panel data regression with worker and time fixed effects – but what are the properties of the treatment effect estimated in this regression? To answer this question, rewrite the production function in a more general functional form, in a switching regression framework.[17] The treated group, D=1 has production function (5a), with superscript 1 to indicate D=1 for treatment. The non-treated group has production function (5b), with superscript 0 to indicate D=0 for non-treatment.

(5a) $Y_{it}^{1} = g_{t}^{1}(x_{i}) + u_{it}^{1}$

(5b) $Y_{it}^{0} = g_{t}^{0}(x_{i}) + u_{it}^{0}$

where x_i are basic production function variables that are not affected by the treatment and are uncorrelated with error term u_{it} .

Define the “treatment effect” as the productivity gain due to treatment, defined as $\alpha_{it} \equiv Y_{it}^{1} - Y_{it}^{0}$. Given the productivity functions (5a) and (5b), the treatment effect is defined as.

(6) $\alpha_{it}(x_{i}) \equiv g_{t}^{1}(x_{i}) - g_{t}^{0}(x_{i}) + (u_{it}^{1} - u_{it}^{0}).$

Intuitively, the treatment effect is the shift up in the production function after adoption, relative to before adoption. The estimated treatment effect, given real world data with

[17] This structure for the specification of treatment effects follows Blundell and Dias (2002). For further discussions of treatment effects and their estimation, see Heckman, Lalonde and Smith (1999), Imbens (2004), List, Sadoff, and Wagner (2008) and Meyer (1995) and citations therein.

non-random adoption of the treatment, is biased if the difference in the residual in (6) is not zero.

Now, rewrite the productivity equations to identify types of selection biases in the treatment effect. To form one productivity equation across the treated and non-treated groups, combine (5a) and (5b) by weighting them by the treatment dummy D_{it} to get

$$Y_{it} = D_{it}Y_{it}^{1} + (1 - D_{it})Y_{it}^{0}$$

Then, substitute into this equation, the definition of the treatment effect as the expected value of (6), $E(\alpha_{it}(x_i)) \equiv g_t^1(x_i) - g_t^0(x_i)$, and obtain:

$$(7) \quad Y_{it} = g^0(x_{it}) + \alpha_{it}D_{it} + [u_{it}^0 + D_{it}(u_{it}^1 - u_{it}^0)]$$

Equation (7) identifies the potential selection biases in the estimated treatment effect. The expected value of the treatment is:

$$(8) \quad E(\hat{\alpha}_t) = \alpha_t + E(u_t^1 \mid x_t, D_t = 1) + (E(u_t^0 \mid x_t, D_t = 0) \text{ with } t>t^*$$

And the two expected error terms are non-zero due to selection bias into the treatment—only those firms that gain the most from treatment will adopt.

Equation (8) and Figure 1 can be combined to easily picture the three types of treatment effects that the insider econometrician might want to estimate with insider data.

1. The *"treatment of the treated effect"* (TTE) is the expected treatment effect among treated observations; or the mean value of the observation-specific treatment effects across all observations conditional on treatment occurring (i.e., D=1) and on the production function control variables x_{it}.

$$(9) \quad \alpha^{TTE} \equiv E(\alpha_{it} \mid x = x_{it}, D_t = 1) \text{ with t>t*s}$$

In the Figure 1, this hypothetical treatment effect is $\alpha^{TTE} = \Delta P^T$.

2. The *"treatment of the non-treated effect"* (NTE) is expected value for those groups who are never treated.

$$(10) \quad \alpha^{NTE} \equiv E(\alpha_{it} \mid x = x_{it}, D_t = 0) \text{ with t>t*}$$

In the Figure 1, this hypothetical treatment effect, is represented by $\alpha^{NTE} = \Delta P^N$. To illustrate potential selection bias, we assumed in Figure 1 that the treated group has bigger expected gains than the non-treated group would have.

3. The *"average treatment effect"* (ATE) is the average gain in productivity if the treatment were applied randomly across all i.

(11) $\alpha^{ATE} \equiv E(\alpha_{it} \mid x = x_{it})$

In Figure 1, it is the weighted average (with weights equal to the percentages of treated and non-treated) of α^{TTE} and α^{NTE}.

Thus, the TTE, NTE, and ATE treatment effects are mean effects among the ever-treated observations, never-treated observations, or randomly treated observations respectively.

B. The Estimation of Treatment Effects with Panel Data in Insider Studies

How do insider studies typically estimate the alternative treatment effects using panel data? First, we describe the estimation methods for alternative treatment effects, and then provide examples from the literature for research using these alternative estimators.

Rewrite the production function (7) as:

(12) $Y_{it} = g^0(x_{it}) + \alpha_{it} D_{it} + \eta_i + \theta_t + \varepsilon_{it}$

where η_i is the worker-specific fixed effect, θ_i is the common time period effect, and ε_{it} is the transitory worker-specific effect. Equation (12) matches the typical panel data regression presented above as (2).

Depending on the exact nature of the panel data set and the specific details of the treatment effect the researcher wants to estimate, three common options for estimating treatment effects in panel data are:

1. First Differences: Here, the researcher estimates the treatment of the treated, α^{TTE}, using data on only the treated group, before and after the treatment, or the top productivity-profiles in Figure 1. Estimate the production function (12) in first differences, by introducing controls for the worker-fixed effects using only the before-after data for the treated group to produce α^{TTE} (for the Type 1 top panel in Figure 1). This is equal to

(13) $\hat{\alpha}^{TTE} = [\overline{Y}^1_{post-t*} - \overline{Y}^1_{pre-t*}]$

which states that the treatment of the treated effect is the difference in the conditional means of the treated group before (the "pre-t*" period), and after (the "post-t*" period), the treatment. Assuming we know the right functional form and specification of $g^1(x_i)$, the unobserved counterfactual for the treated group in the post-t* time period can be estimated. Avoiding the use of the non-treated sample eliminates the need to make assumptions about that subsample – about why it was not treated and about its unobserved counterfactuals. An obvious drawback is that there is less data when lacking the information on the non-treated data, and we assume that time effects θ_t are uncorrelated with treatment D_{it} .

2. Difference-in-Differences Estimator: Researchers estimate the α^{TTE} effect using difference-in-differences when the treatment is random for a subpopulation of workers, plants, or stores, or the assumption that the only difference between the control group and the treated group lies in the fixed effect η_i.[18] Given this assumption, the researcher has data on both the treated group, before and after the treatment, and on the non-treated group in Figure 1. Equation (12) is estimated with both samples, over time, so this is double differencing, or "difference-in-differences" because we difference over time (before-after treatment) and cross-sectionally (for the treated and non-treated groups). This double differencing amends the estimated treatment effect (13) to be:

$$(14) \quad \hat{\alpha}^{ATE}(x) = [\overline{Y}^1_{post-t*} - \overline{Y}^1_{pre-t*}] - [\overline{Y}^0_{post-t*} - \overline{Y}^0_{pre-t*}]$$

which states that the treatment of the treated effect is the difference in the conditional means of the treated group before (the "pre-t*" period), and after (the "post-t*" period), the treatment, relative to that difference for the non-treated group. The estimation method is the same as for first differences – we are differencing out the fixed effect η_i , and then differencing out the effect θ_t, but here the differencing occurs over a different sample, of treated and non-treated workers.

The basic advantage of this double-differencing compared to the first differencing, is that we now have a control group, the non-treated group. Adding the control group adds information in two ways. First, if we can assume that the time shocks θ_t are common between the two samples, then the non-treated control group lets us control for any time effects that altered productivity after t* but are not due to

[18] Examples of this given above are Lazear (2000), Ichniowski, Shaw, and Prennushi (1997), and Griffith and Neely (forthcoming).

the treatment. Second, the control group gives us more information on which to estimate the underlying production function $g_t^0(x_i)$.

3. Difference-in-differences with matching estimator: Here, the analyst again uses a difference-in-differences estimator, and has the same data as in Figure 1 for difference-in-differences, but the matching estimator is estimating the production function non-parametrically. Each treated outcome, $g_t^1(x_i)$, is paired with a non-treated outcome, $g_t^0(x_i)$, and the difference is the treatment effect. Given the extra data needed for non-parametric estimation, few insider econometrics studies use this approach.

4. Instrumental Variables or Semi-parametric Methods of Correcting for Selection Bias. If there are non-random (selection bias) differences between control and treatment – or between the Type 1 and Type 2 panels in Figures 1 – then the researcher must address these differences. If there is selection bias, then the ε_{it} is correlated with D_{it}, and we need to add IV- or matching- estimator techniques that model the selection of the treatment. We do not review these methods here.

Under the assumption of selective adoption of the treatment (as modeled in equation (8) above), these alternative estimators that all include controls for observation-specific fixed effects generate estimates of the treatment of the treated (TTE) effect – the effect of some new management practice on productivity given that firms chose to adopt.

C. Types of Insider Data Sets and Treatment Effects Estimated

The answers to the questions about which treatment effects can be estimated and which estimation methods can be applied depend on the nature of the insider data sets. An important distinction here is the difference between insider that use worker-level data, often personnel records from one firm, versus those that use work-group level data on employee teams, project groups, production lines, stores or establishments, often from several firms. When the data are worker-level observations from one firm, the treatment is randomly assigned to the worker (since all relevant workers in the firm are covered), so the estimator is a "conditional average treatment effect" or conditional ATE. It is an ATE effect because assignment is random. We also refer to it as "conditional" since it depends on being in the firm adopting the practice. Estimates of productivity effects of incentive pay in Bandiera, Barankay, and Rasul (2005) that do not make any distinction in the size of the treatment effect for different types of workers is one such example.

When the data is work-group level from several firms where some groups have the treatment and others don't, the estimator is a treatment of the treated effect. In these cases, we want to know the expected treatment effect for the specific type of firm that adopts the treatment, not for a random firm[19] The steel mill studies of Ichniowski, Shaw and Prennushi (1999) and Boning, Ichniowski, and Shaw (2007) are examples. In these studies where some observations are adopters and others are not, the researcher may add the estimation of the adoption equation (3) as in Boning, Ichniowski and Shaw (2007) to model and explain adoption.

Table 1 gives many more examples, classified according to the management practices being studied. However, the table also describes the type of data, and thus implies the likely type of estimator of the treatment effect. For example, Griffith and Neely (forthcoming) have data on an incentive pay treatment, before and after the treatment and for a treated and control group of retail stores, so they use difference-in-differences with a matching estimator to estimate at TTE effect.

D. **Estimating Heterogeneous Treatment Effects in Insider Studies**

Insider researchers aim to estimate heterogeneity in the treatment effect, as shown in equation (2). In some insider studies, this feature is reflected in the idea that the productivity impact of some management treatment among adopters (the TTE effect) differs from the productivity impact that non-adopters would be expected to experience (the NTE effect). In worker-level studies within a single firm that covers all workers under a new policy, this feature of insider studies is reflected in analyses that identify specific subgroups of workers with relatively large and small productivity effects; for

[19] Some studies analyzing randomly assigned managerial treatments are being conducted in developing countries. Bloom, Eifert, Mahajan, McKenzie and Roberts (2009) are randomly imposing lean manufacturing practices on a random set of small manufacturing firms in India who don't currently have those practices. Thus, they are estimating the average treatment effect of lean manufacturing, but not the reasons why some firms adopt and some don't. Duflo and Rema (2005) introduces incentive pay for teachers randomly across a set of teachers and a control group. Presumably, in studies using random assignment of a management practice across workers or plants, identifying observations that enjoy relatively large and small treatment effects could offer clues about where adoption makes more economic sense. See Levitt and List (2008) for a review of the new field experiment literature.

example, the productivity effects of incentive pay among employees who are (or are not) working with friends in Bandiera, Barankay, and Rasul (2005), or the productivity effects of piece rate pay among employees who stay (or leave) the Safelite firm in Lazear (2000).

In introducing heterogeneity in treatment effects, there are two econometric issues. First, when permitting the treatment effect to vary across subgroups of the population, extrapolation beyond those subgroups is impossible. When there are multiple treatment effects, we cannot extrapolate to say what the treatment would be for other types of worker subgroups that are not considered.[20] The second econometric point is that when researchers permit the treatment effect to vary across workers, they often need to gather more data beyond the production function inputs and outputs. Examples illustrating this point include:

- Bandiera, Barankay, Rasul (2005) gather data on friendships among the work crew members, so that they can estimate the effect of piece rate for employees who do (and do not) have friends in their work crew.
- Mas and Moretti (2009) gather data on the physical location of workers in the grocery store, so they can specify when the checkout clerks are (and are not) in a position to observe the co-worker to test how monitoring by high and low performing clerks affects co-worker productivity.
- Gant, Ichniowski and Shaw (2007) collect detailed data on workers' communications networks in steel production lines to show that production lines with denser social networks are more likely to be the lines with innovative management practices and higher levels of productivity.

In sum, this section highlights how insider econometric studies must be careful not to assume that untreated observations would respond the way that treated observations responded to a practice. TTE effects are not the same as (unobservable) NTE effects, and this difference can be the simplest explanation for lack of adoption among the untreated. Second, the magnitude of the treatment effect can vary systematically across treated observations and this variation can provide important economic insights. The insider researcher might be able to identify certain subgroups of

[20] In econometric terms, we cannot extrapolate beyond the "common support," which is the subspace of worker characteristics that are in the treated and non-treated samples.

adopters with relatively small productivity treatment effects, and these patterns can offer clues that help explain the non-adoption among other observations. Third, the cross-sectional variation in insider panel data sets is still important. In studies where some observations are covered by the treatment and others are not, the data on non-adopters is required to construct difference-in-differences estimators with or without a matching estimator. In worker-level studies about a single firm that introduces a new management policy for all workers at one time, the cross-section variation is again important. Here, the cross-section variation is needed to test for differences in the magnitude of the treatment effects across observations.

VI. Why Do Firms Adopt New Management Practices?

The treatments in insider econometric studies are new management practices. In the examples in section IV, these management decisions include changes in: compensation plans, teamwork and related HRM practices, the scope and size of research projects, and decisions to outsource certain work. An important question to acknowledge is why firms would ever change their management practices? For example, why didn't Safelite adopt incentives for their windshield installers, or minimills with complex products put in their problem solving teams sooner than they did? Shouldn't these firms have always had their optimal practices? To answer these questions, we must know what causes the adoption of a new 'treatment' to occur among real firms.

Several sources of disequilibrium are likely to cause 'shocks' to optimal management practices. First, there can be shocks to the prices of capital inputs in the production process. If management practices are complements with these capital inputs, the management practices too would change. Rapidly declining prices in information technologies offer one obvious recent example of such a shock. Several studies illustrate how, as information technology and processing become cheaper in firms, HRM and other management practices need to change to complement the changing way work gets done.[21] The adoption of new information technology in one industry may also cause changes in management practices in related industries.[22]

Second, there are technology shocks to management knowledge as best practices in management improve over time. It takes time for firms to learn about these changes and before they adopt new best practices. For example, the success of a distinctive set of management practices – like problem solving teams, other innovative HR practices, and lean manufacturing practices – among large Japanese manufacturers was a source of productivity and quality improvements in these firms prior to their broader adoption in the U.S. and elsewhere. These management practices in Japanese firms were then "imported" to other developed countries as the lessons about how and why these practices worked spread throughout manufacturing in developed countries.[23] Developing countries are beginning to introduce these kinds of management practices, though adoption in these countries still lags behind the adoption in developed countries.[24]

Third, there are shocks to firms' internal labor markets or firms may decide to experiment with new practices. For example, stock options may no longer be paying off, or an aging workforce within the firm may need different incentives, such as pensions. The external labor market may have an inflow of workers with different levels of human capital. For example, if immigration offers the firm an increased supply of workers with higher (lower) levels of human capital than the existing workforce, training needs may therefore decline (increase).

Fourth, the firm's product market strategy may change, causing needed adjustments to optimal management practices for the firm. When firms enter an industry with new entrepreneurial ideas about the industry's products or processes, existing firms within their industry may need to change product market strategies or processes. For example, as imports from Chinese firms enter U.S. markets, the U.S. firms may move

[21] For evidence, see Breshnahan, Brynjolffson and Hitt (2001) and Lemieux, MacLeod and Parent (2009). Jorgenson, Ho and Stiroh (2003) and Oliner and Sichel (2000) present macro-level evidence that productivity has risen over time in industries are computer-using industries.

[22] When large retailers adopted information technologies that tracked product sales and sent those sales data directly to suppliers, the suppliers adopted new team-based HR practices that allowed them to restock the retailer with smaller batches on shorter notice. (Dunlop and Weil, 2000)

[23] For more on the notion of "HR technology shocks," see Lazear and Shaw (2007) and Bloom and VanReenen. For more on the transfer of Japanese HR practices to U.S. firms and the effects of these practices on productivity among manufacturers in the two countries, see Ichniowski and Shaw (1999).

[24] See Bloom and VanReenen (2007) and Bloom, Eifert, Mahajan, McKenzie, and Roberts (2009).

from the production of commodity-like products to customized products.[25] If management practices are complementary with the new business strategies, firms would need to adjust their practices. These changes in product market strategy can require changes in management practices, such as an increased use of problem-solving skills for producers of customized products as in the case of minimills.[26] These within-industry shocks arise from new product development or new research by firms over time.

Section VII. A Practical Roadmap for Designing Insider Econometric Studies

The insider econometrician is, in many ways, designing his empirical experiment. Few insider econometricians are conducting an experiment with random assignment of treatments, but the insider econometrician like the experimentalist must craft the research design carefully. He must choose the right firm to study, the questions to ask, hypotheses to test, the type of data and the variables to collect, and then the specific econometric methods to test the hypotheses.

A. Key Research Design Decisions for Insider Econometric Studies

In this section, we provide advice about how to answer some of the more specific methodological questions that naturally arise when designing and executing an insider econometric study. At the outset of this paper, we offered a list of five features that characterize insider econometrics studies. These five features tell us that insider studies use micro-level data on individual workers or groups of workers to study the productivity effects of a new management practice and the reasons for the adoption of the practice. Still, many more decisions must be made about specific features of an insider study. What management practices should I study? What should be the structure of the data set? What can I do to address the concern that the data are specific to only one particular production process? This section on key research design decisions delves more deeply into these decisions about research design to describe the features of studies that will produce the most convincing insider econometric results.

---

[25] Bloom, Draca and Van Reenen (2008) offer empirical evidence that reduced trade barriers increased the adoption of new HR practices.
[26] Bartel, Ichniowski, and Shaw (2007) document the effects of product market strategy on HRM practice adoption in another industry.

Key Design Decision 1: Identify a Treatment and Measure It Accurately

Identifying a management treatment is the first necessary condition for an insider study. The researcher must identify a treatment variable that changes. Most firms that have interesting management practices never change their practices, and thus there is no variation for the treatment to test its effects. The treatment must also be interesting economically because of its potential effects on productivity outcomes, and the effects of the practice on workers must be modeled. The researcher may have a theory in mind and been searching for data to test it, or the researcher may be offered access to interesting data and explores possible hypotheses by talking to insiders. Whether discussions with insiders spark ideas about relevant theoretical models, or a theoretical idea sparks interest in having discussions with insiders, there must be an economic logic as to why the treatment would impact performance outcomes.

How does the researcher get the treatment data? Many data sets have information on the productivity of workers or work groups, but these data sets do not link information on management practices to the productivity data. However, as in most examples of insider studies considered in Table 1, management practices are discrete events that change infrequently within firms. The researcher can therefore interview insiders and build a data set on the time path of management practices within or across firms.

The treatment needs to be specified accurately. First, insider studies have several advantages over traditional surveys, and one is that the researcher can use interviews to measure exactly what the management practices are. In contrast, respondents to a survey question that asks, "do you have teamwork," may interpret the meaning of the question in many different ways. Insiders' insights help measure treatments accurately.

Second, the researcher needs to measure all the dimensions of the management treatment. An important contribution of existing insider studies is that "management treatments" often involve the simultaneous adoption of multiple complementary management practices. Insider econometricians who have been inside the firms often find that the firm is changing not just one practice to improve productivity, but is instead adopting a cluster of complementary practices.[27]

[27] MacDuffie (1995) finds that productivity is higher in auto plants that adopt several new HRM practices together with lean manufacturing methods and just-in-time inventory policies. Ichniowski, Shaw and

Finally, the "treatment" need not always be a change in a formal management policy. Clever researchers are finding "natural experiments" within firms involving other kinds of "treatments." For example, Mas and Moretti (2008) model the importance of peer effects on productivity. They identify the peer effects not from a change in management practices surrounding peer groups, but instead from a change in which employees are assigned to peer groups. That is, because workers are assigned randomly to different peer groups, the authors can test whether peers influence productivity.[28]

Key Design Decision 2: Test a Generalizable Principle by Modeling Economic Behavior

The economist's approach to studying organizations has its own distinctive set of building blocks that form the foundation for empirical studies in personnel or industrial economics (Lazear and Shaw, 2007). Central to these building blocks is the idea that workers, managers and firms are optimizing agents. While insider studies are carefully crafted empirical tests of hypotheses about management practices, the hypotheses emanate from models of a more general economic principle about how these optimizing agents really behave when they are at work inside their firms. The data may come from one firm or from several companies within one industry, but the theoretical model generating the hypotheses is not specific to the single industry context. This means that ultimately the broader goal of the insider study is to test a more general economic principle that will enhance our understanding of the theory of the firm.

As examples, in our review of six insider studies in Section III, we asked the question, "why is this study interesting" for each study. The answer was always that the results test a model of fundamental economic principle about the behavior of workers in firms. Certainly, the principle will not be relevant for all firms in all industries, but in settings where the key features of the more general model are relevant, the principle may

---

Prennushi (1997) document improvements in productivity only when lines adopt entire systems of innovative HRM practices. Bartel, Ichniowski and Shaw (2007) find that new computer-aided production equipment that raises productivity among customized valve manufacturers coincides with the adoption of complementary HRM practices. Azoulay (2004) identifies how the management decision to outsource drug trials always means the firm is putting the trial under a different compensation policy. See Milgrom and Roberts (1995) for discussions of complementarity.

[28] Mas and Moretti (2009) find that any individual checkout clerk in their study of grocery stores is more productive when he works with faster co-workers, and argue that differences in shift composition are not endogenous and that changes in shift composition cause changes in co-worker productivity.

also be at work. That is, even though only one firm or industry is studied, the testable hypothesis and results appear to be applicable to a wider set of workers or firms because the paper models fundamental optimizing behavior.[29]

Put differently, without insider data sets with rich details about the workers and production processes in a specific industry setting, convincing econometric tests of these kinds of models would not be possible. The insider studies reviewed in this chapter document important productivity effects of many factors that are routinely neglected in standard specifications of a firm's production function: friendship among co-workers, co-worker monitoring, complementarities among multiple HRM practices, interactions of HR practices and product attributes, knowledge spillovers across workers and projects, and many more. Insider data sets that describe rich details about the context for a single production process are precisely the kind of data that one needs to construct a convincing test of these novel hypotheses. An insider study cannot on its own identify the boundary conditions that determine exactly where the model does and does not apply. However, the models of optimizing behavior by firms and workers in insider studies is generalizable, and one needs detailed firm- or industry-specific data to conduct convincing econometric tests of whether the predictions of the more general model are supported.[30]

Key Design Decision 3: Balance Homogeneity and Heterogeneity in the Sample

Prior to insider studies, production functions would be estimated with aggregate industry-level data or with firm-level data or establishment-level data in broad cross-industry samples. While these data might be appropriate for estimating the effects of

[29] In the examples we review in Section IV, we find models in which: friendship can affect worker utility (Bandiera, Baranky and Rasul, 2005); HR practices are sorting devices (Lazear, 2000); HR practices have complementary effects on productivity (Ichniowski, Shaw and Prennushi, 1997); the value of work on production activities versus problem-solving activities can vary across plants (Boning, Ichniowski and Shaw, 2007); the ability to measure and monitor work determines decisions about incentives and outsourcing (Azoulay, 2004); and knowledge gained in one area of the firm increases productivity elsewhere (Henderson and Cockburn, 1996).

[30] Consider the paper of Bandiera, Barankay, and Rasul (2005) in which friendships influence a worker's response to incentive pay (Example 2 of Section III). Their conclusion is not simply that friends influence people's behavior, but that when economists write a model of incentive pay, that model should include an externality – that when I maximize my utility in response to incentive pay, my response internalizes the externality of my friends' feelings. Economic theorists have introduced peer pressure and social norms into their theories of the firm (Kandel and Lazear, 1992; and Akerlof and Kranston, 2000; 2005).

factors like capital-labor ratios on productivity, they will be less helpful for identifying the effects of management practices on productivity, for isolating the economic reasons why similar organizations or workers are covered by different practices, or for identifying factors that cause the productivity effects of a management practice to vary across workers. (see the discussion in Syverson, Goolsbee and Levitt, 2009). With more aggregate data, there is not one common production process that generates the data, and therefore the number of variables that might be correlated with the management treatment variable is exceptionally large.[31].[32]

To mitigate the issues of unmeasured heterogeneity, insider studies seek data from production units – whether workers or work groups – that are very homogeneous. The homogeneity of these workers is that they all share one common production function that the research can model empirically -- on-site car windshield installation, fruit picking in an orchard, finishing lines or rolling lines in steelmaking, and research projects or clinical testing processes in pharmaceuticals.

While sample homogeneity is therefore important in insider studies, the insider data set cannot have perfect homogeneity of the production units. The adoption of the treatment must vary to investigate why, how and how much management practices affect productivity. The researcher wants a homogeneous production function, but also subtle differences in the workers or organizations that leads to heterogeneity in the adoption of the treatment or heterogeneity in the response to a common treatment. Overall, the goal of sample design decisions in insider studies is to strike a purposeful balance between

[31] Differences in production equipment, worker and manager quality, occupational mix, and many other factors all could be correlated with differences in management practices and thus confound any attempts to use such industry-level data to generate a convincing estimate of a treatment effect due to any management practices.

[32] The methodological issues in generating a persuasive estimate of the effect of some management practice on productivity of course extend well beyond issues of unmeasured heterogeneity and omitted variables. While measurement error in the dependent variable does not have to introduce bias in the estimated treatment effect in a production function, production functions estimated with more aggregate data typically must rely on value added data for the dependent variable, even though value added contains product mix changes that are not part of 'production' per se, and thus can differ dramatically from the variable the researcher hopes to be measuring. See Syverson, Foster, and Haltiwanger (2008) for a discussion of this point.

homogeneity in the production function and heterogeneity in either the adoption of the treatments or the magnitudes of the treatment effects. [33]

Key Design Decision 4: Collect Information on Why the Treatment Was Adopted

Once the researcher identifies a management innovation that varies across observations, he must understand why the treatment was adopted. Perhaps the most fundamental point about understanding and modeling the adoption process is that, because the management treatments are rarely natural experiments, the researcher needs to collect information about the determinants of the adoption of the management practice being studied. While this point is perhaps obvious, unless it is fully appreciated prior to data collection efforts, the researcher may focus only on data related to the production function and management practices, and miss opportunities to obtain data on the factors that cause adoption of the practices.

The researcher can also explore whether the treatment should be considered exogenous or endogenous with respect to productivity outcomes. Even though most researchers do not have data from natural experiments in which the adoption is randomly and exogenously imposed, there are some cases in which the treatment is exogenous. When data come from a single firm, the management innovation is likely to be exogenous with respect to productivity outcomes when productivity is measured at the worker level. In these cases, the new management treatment is imposed by the firm on all of its workers or establishments. In the Safelite (Lazear, 2000) and fruitpickers studies (Bandiera, Barankay, and Rasul, 2005), the productivity data are on workers, and adoption is clearly exogenous to the worker. This is true for other studies of retail stores in a single firm (Lafontaine and Srinivasan, 2008; Griffith and Neely, forthcoming), or for teams of workers within an apparel firm (Hamilton, Jackson, and Owan, 2003). At the same time, as described in Section IV, there are still potential sources of selectivity in the application of the exogenous management practice: workers can leave or join the firm (and enter or leave the sample) after the firm imposes the practice (as in Lazear, 2000); or

[33] List, Sadoff, and Wagner (2008) make this point about the tradeoff between homogeneity and heterogeneity in designing optimal field experiments.

workers might be the ones who decide to opt in or out of the new practice (as in the study of work teams by Hamilton, Jackson, and Owan, 2003).

The management innovation can also be exogenous with respect to productivity outcomes across plants, but endogenous to the firm. In the integrated steel mills, Ichniowski, Shaw and Prennushi (1997) conclude that expected productivity gains for non-adopters are about the same as those observed for adopters, and adoption differences across lines are not a function of expected productivity differences. Rather, they are determined by differences in the costs to the firm of transitioning to new HR practices.[34]

In other insider studies that use panel data across firms, the analysts conclude that the adoption is endogenous to productivity outcomes, and the authors gather data on the adoption regression. In the study of minimills by Boning, Ichniowski, and Shaw (2007), productivity gains from teamwork are greater when a mill produces complex instead of commodity products. The authors show that adoption is more likely among firms making complex products. In Bartel, Ichniowski, and Shaw (2007), the authors also show that adoption of information technologies in valve industry plants is driven by the complexity of the product mix. As described in Section V, in these cases where adoption is endogenous, the treatment effect estimated in the productivity regression is the "treatment of the treated" – the expected gain conditional on optimal adoption of the treatment. By understanding and documenting the reasons for adoption, the researcher can more clearly identify the nature of the treatment effect he is estimating.

Key Design Decision 5: Collect information on how behavior changes and why the treatment was successful

Even though insider econometrics researchers use many different types of data sets and methodologies, all insider econometric studies document *why* the treatment was effective. In the broader treatment effect literature, this is typically not the case. When researchers study the effects of tax cuts on labor supply, researchers rarely collect data to explain why people increase their labor supply. However, in working with workers or establishments of firms, evidence on why the management practice is or is not effective is

[34] With detailed observation-specific data, it is possible to examine other reasons that could account for differences in the adoption of new HR practices in this study. In the study of steel lines for example, explanations related to differences in managerial quality, or plant age, or prior productivity, are tested and rejected, thus lending greater support to the conclusion about transition costs.

both available and essential. Models of the effects of management practices on productivity routinely model why a practice is effective. Insiders can get the data to test these more detailed mechanisms and provide additional tests of the underlying model.

Consider some examples of the data researchers gather to test why management practices are effective. Bandiera, Baranky and Rasul (2005) gather data on friendships among employees to show that a large gain in productivity after the firm switches from relative pay to piece rate pay is concentrated among employees who worked with friends and who were holding back their effort under the old pay plan. Mas and Moretti (2009) gather data on where checkout clerks stand (relative to each other in the store) to show that the large gain in the speed of grocery checkout clerks when a high-performing clerk joins the clerk's shift occurs when the new high-performing co-worker watches the lower performer. Gant, Ichniowski, and Shaw (2002) gather data on the communications social networks of all employees to show that inter-worker communication and problem-solving activities are much more pronounced among steel production lines with new HRM practices and higher productivity than they are in other lines. Many recent studies in behavioral economics are aimed precisely at showing how people behave in response to incentives or other HR practices, and some of the most convincing evidence in insider econometric research describes the behavior of workers in treated and untreated conditions.

In sum, one advantage of doing insider research is that data are often available on why the treatment is effective, and these data allow the researcher to identify specific behavioral mechanisms that explain the connection between the management policy and the productivity outcomes. Insider studies are more convincing if they go beyond the estimation of a treatment effect. If additional data are collected about why the treatment is effective, the empirical work is more persuasive.

B. Detailed Steps to Take When Faced with Data

When faced with potential data from one or more firms, what should the researcher do? The following list offers some guidance.

1. Interview managers, workers, and others to learn about the mode of production and the management issues faced by and addressed by the firm. Ask them to tell you, what determines productivity? Learn the production process.

2. Identify a research question based on these interviews. Most researchers would not know the questions or potential hypotheses without the plant visits. Rule out some potential hypotheses.
3. Identify why this firm (or firms) have adopted these management practices and other firms have not.
4. Gather additional data needed for better tests of the hypotheses. Determine your econometric methods and how you will identify the appropriate treatment effect. If data describe the adoption of one new treatment within a single firm, can you collect data on factors that explain heterogeneity in observed treatment effects? If data describe the adoption of the treatment across multiple firms or locations, can you collect data on the reasons for adoption versus non-adoption? Decide what is endogenous versus exogenous, and where there might be measurement error or omitted variable bias.
5. Evaluate and interpret your results. Empirical researchers typically test the robustness of their models by introducing functional form changes. Insider econometrics studies go a step further by asking the insider informants if the results make sense.
6. Collect additional data for more comprehensive and convincing tests. For example, though the primary model is a production function, it could be very persuasive to interview workers to gather other evidence to interpret or retest the results.

Insider studies identify what is most likely to affect productivity, and also identify what can be ruled out as a determinant of productivity. Ruling out potential hypotheses can be as important as formulating new testable hypotheses. Across all firms, productivity may grow due to learning curves, teamwork, new product development, or new management practices. In cross-industry studies of firm-level data, the list of possible productivity determinants is endless. In insider studies, many possible reasons for productivity growth can be ruled out (or controlled for in the regressions) given the narrow focus on a single production process, so the economist can focus on testing the hypotheses that are most relevant to that firm or industry or theoretical model.

Initial discussions with experienced insiders about how the production operations work can lead naturally to other steps. For example, the researcher may end up exploring managers' opinions about why some workers, work groups, or establishments are covered by the management treatment, while elsewhere the treatment is absent. He must then translate the real world experience of the industry insiders into a research design that tests hypotheses about the fundamental questions that define insider econometric research reviewed at the outset of this chapter. Why do some employees or organizations that appear to be competing in the same narrowly defined industry work under one set of managerial practices while others work under other practices? When new practices are adopted, did productivity change? How do changes in productivity vary across workers or establishments? Are there any changes in employee behavior that can help explain the changes in productivity?

Section VIII. Economic Lessons Learned from Insider Econometric Research

Each insider study provides its own insights for firm- and industry-specific settings, and each tests different economic models. At the same time, many of the empirical findings from one study to the next echo common themes. These broader economic themes emerging from insider studies can, we believe, provide guidance about the most promising avenues for developing richer theories of competing firms and their employees. In this section, we review several broader themes emerging from multiple insider studies.

1. *Management matters. Effects of management practices on productivity can be big.*

The fundamental conclusion emerging from insider econometric research reviewed in this chapter is that management matters, and that it matters in several ways. First, insider studies routinely document substantial productivity effects from the adoption of new management practices. As examples, the magnitudes of the overall productivity effects among windshield installers in Lazear (2000) or fruit harvesters in Barankay, Bandiera, Barankay, and Rasul (2005) are substantial – 44% and 58% respectively. Studies in many other industry settings reviewed above also show dramatic productivity increases after adoption of other new practices. If, as seems likely, the productivity benefits of the new practices among the adopters exceed the costs of adopting to them, then managers' search for optimal practices is critical for firm performance. While it is beyond the scope of this study, one can speculate that the large magnitudes of the "treatment of the treated" productivity effects reviewed in this chapter imply that the economy-wide impact of management practices on productivity is also large.

Second, the large productivity increases that occur after the adoption of some new management practice are still not universal within a given industry setting. Insider studies also routinely document heterogeneity in the magnitude of productivity benefits among workers and plants using the same basic production process. Management therefore matters in another way. Managers need to understand subtle within-industry differences to decide if a new management practice will increase productivity for their specific operations.

Third, a wide array of management practices impact productivity of workers, work groups, and plants in different industries. The number of management practices considered in insider studies, as reviewed in the list of studies in Table 1, is very large.[35] Thus, managers must evaluate a very wide range of policy areas when deciding what practices would be the optimal ones for their workers, their technologies, and their industries.

2. *The productivity of individual workers is determined by his or her work group.*

Several insider studies demonstrate that work groups matter. An individual's productivity is a function of his co-workers' productivity. Two work group effects stand out. First, productivity spillovers due to peer effects matter. When an employee is working next to a high performer, the employee works harder. This is a sense of peer pressure, not teamwork. We discussed this above regarding the Mas and Moretti (2009) paper showing that when a faster check-out clerk replaces a slower clerk on a shift, the shift becomes more productivity not simply because the new clerk is faster than the old one, but because the other clerks also become faster (Mas and Moretti, 2009).

The second work group effect is that employees are complements in production when they work as teams, and teams can be much more productive than individuals. Teamwork implies that workers do not have identical skills and thus workers can complement one another. In a team-based production function, output is not the sum of individual output, but instead is a function of the multiplicative interaction between

[35] Even within the area of compensation policies, there are many other types of compensation plans that we have not reviewed. For example, tournaments are readily modeled in laboratory experiments, so there is extensive experimental evidence showing that tournaments can motivate individuals better than other kinds of incentive pay in some settings (see Bull, Schotter, and Weigelt, 1987). See the insider studies of Knoeber and Thurman (1994) for producers of broiler chickens; Drago and Garvey (1998) for Australian firms, Ehrenberg and Bognanno (1990) for professional golfers. Different methods for determining CEO or other top executive pay would be yet another set of management practices that impact firm performance.

employees. Several insider papers reviewed above delve into the details of what makes work teams productive.[36]

3. Employees respond to management practices at both the extensive margin of whether to work for the firm and the intensive margin in deciding how hard to work at their jobs.

When firms introduce a new incentive pay plan, we expect workers to work harder and to target their efforts towards the output that is rewarded. This increase in effort operates on the worker's "intensive margin" – on workers effort "within" the firm. Other management practices, like teamwork or innovative job design, also aim to increase effort at the intensive margin. However, management practices also have a big effect on workers at the "extensive margin" – on workers' decision to join the firm or the work group that offers the particular practice. Workers that don't like incentive pay avoid firms that pay for performance (Lazear, 2000). In other cases, within the firm, management can put a policy such as teamwork in place, but then leave it to the employees to join teams. Workers can remain employed even when they choose not to be covered by the new practice. In these cases, insider studies again find there is selectivity in who joins teams (Hamilton, Nickerson, and Owan, 2003). The key result is that when a firm chooses its management practice, it will cause worker sorting. Thus, the productivity gains that firms experience from management practices result from a combination of effort (the intensive margin) and optimal sorting (the extensive margin).

4. Firms compete using different product market strategies, and these different strategies require internal organizational structures to match the strategies.

The idea that businesses in an industry compete in different product niches has long been recognized in the business strategy literature. Low cost mass producers co-exist in industries alongside other businesses that serve customers who prefer higher cost

---

[36] In steel finishing lines, problem-solving teams raise productivity because worker interactions become richer and more frequent. Problem solving requires interaction among the co-workers (Gant, Ichniowski, and Shaw, 2002). In apparel manufacturing, the productivity of work teams depends on the mix of skills and productivity levels of the team's members (Hamilton, Nickerson, and Owan, 2003) In insider studies of orchards, the existence or absence of friendships among co-workers determines the extent to which effort of the fruit pickers changes after various kinds of incentive pay plans are adopted (Bandiera, Barankay, and Rasul, 2005). In pharmaceuticals, the productivity of research on one project depends on knowledge spillovers from other related projects (Henderson and Cockburn, 1996). In many settings then, the true production function specifies not just an aggregate amount of workers but must capture the interactions among the individual workers.

specialty versions of the products. Several insider econometric studies document this link between product market strategy and internal management practices.[37] This link is an important result. It explains, in part, why some firms within an industry adopt certain human resource management practices and other similar firms don't. It also explains why traditional output-per-manhour productivity measures will be higher for some firms than others in the same industry.

5. Management innovations often require the adoption of multiple practices that are complements with each other.

Complementarities among management practices are pervasive. That is, the productivity gain from one management practice, like problem solving teams, interacts with the gain from other practices, such as training, selection or compensation practices. Therefore, managers who are considering adopting one new management innovation cannot calculate the marginal gain from that practice independent of the set of other practices (Milgrom and Roberts, 1990). There is extensive evidence from insider papers that choosing the right system of complementary practices increases productivity more than choosing individual practices in isolation.[38]

This review of broader themes emerging from the empirical findings of insider econometric research highlights how diverse and complex the determinants of productivity really are. Managers who want to achieve high levels of productivity from their workforces must consider a wide array of factors. Which employees are working together? What are the skill differences across co-workers and what are their personal

---

[37] Minimills that make complex steel products adopt problem-solving teams more often than minimills making commodity products (Boning, Ichniowski, and Shaw, 2007). Valve manufacturers making more customized valve products adopt more innovative information technologies within their capital equipment and adopt more innovative training and teamwork (Bartel, Ichniowski and Shaw, 2007). Software companies operating in market segments that have big upside gains to innovative software development pay higher wages and offer more performance pay than software companies operating in more traditional product market niches (Andersson, Freedman, Haltiwanger, Lane, and Shaw, 2009). Southwest Airlines, targeting low-price airline customers, uses different hiring and training and pay plans than does Singapore Airlines that targets high-price customers.

[38] Ichniowski, Shaw and Prennushi, (1997) document this result in the steel industry. In auto assembly plants, innovative HRM systems are typically adopted along with lean manufacturing methods and just-in-time inventory procedures (MacDuffie, 1995). In apparel manufacturing, new HRM systems are adopted when new information technologies that tie the manufacturer directly to large retailers are adopted (Dunlop and Weil, 2000; Appelbaum, Bailey Berg, and Kalleberg, 2000). Decisions to adopt new information technologies are also linked together with decisions to adopt a series of new HRM policies in valve manufacturing (Bartel, Ichnowski, and Shaw, 2007). Bresnahan, Brynjolfsson, and Hitt, (2001) also find that HR practices and new information technologies are complementary.

realationships like? When I adopt a new policy, am I likely to attract the kind of employee I really want, or will the better performers leave my firm? Is the new practice I am considering consistent with other practices I have in place, or do I need to change multiple practices? If I change my competitive business strategy, will I need to reconfigure how I manage my workforce? Insider econometric research paints a picture of the manager's job as one that is extremely challenging. As the first theme above states, management certainly matters.

IX. Data Sources for Insider Econometric Studies

In the insider studies reviewed above, researchers worked with firms to obtain data. That is not always necessary. As described next, researchers can use publically available data and augment it with data from other sources or from their own survey. Data sets may contain information on productivity or on management practices, but not both: researchers must combine data from several sources to get both. Many insider studies are now being conducted in education and health care, and in firms in developing economies. As researchers obtain data in these or other sectors, new insider studies will be possible. In this list of possible data sources, we include lengthy footnotes with references to papers that have used data from these kinds of sources. However, authors of these papers have also been creative in their use of data that spans these categories.

1. Insider data from within one firm. Here, the firm provides the data.[39]
2. Insider data from multiple firms within the same industry. Researchers obtain data from the firms (through visits) or conduct surveys of the firms to learn HR practices.[40] For surveys, the researchers use insider guidance to develop

[39] Autor, Levy, and Murnane (2003). Bandiera, Barankay and Rasul (2009). Bandiera ,Barankay and Rasul (2007). Bandiera ,Barankay and Rasul (2005). Bartel (2004). Batt (1999). Bhattacherjee (2005). Encinosa, Gaynor and Rebitzer (2007). Freeman and Kleiner (2005). Griffith and Neely (2007). Hamilton Nickerson and Owan (2003). Jones and Takao. (2007), Kalmi and Kauhanen (2006). Knoeber and Thurman (1994), Lazear (2000). Lo,Ghosh and Lafontaine (2006).

[40] Bartel, Ichniowski and Shaw (2007). Bloom, Eifert, Mahajan, McKenzie and Roberts (2009). Bloom, Genakos, Sadun and Van Reenen (2009). Bloom and Van Reenen (2007). Boning, Ichniowski, and Shaw (2007). Dunlop,and Weil (2000). Fernie and Metcalf (2003). Huselid and Becker (1996). Ichniowski and Shaw (2003). Ichniowski and Shaw (1995). Kruse (1993). Kruse, Blasi and Park (2006). MacDuffie (1995). Oyer (1998). VanReenen and Bloom (2007). Van Biesebroeck (2006), Berg, Appelbaum, Bailey, and Kalleberg (1996), Dunlop and Weil (2000).

and administer their own survey, and the challenge is to get productivity data from the survey (or match it from a different data set, like census data).

3. Insider analysis using data from regulated industries.[41] Regulated and government-funded industries – such as education, health care, trucking, and electricity generation – often have data available. The researcher may need to conduct surveys to obtain the HR information to match to the productivity data.

4. Insider analysis is also done using Census data or data sets from consulting firms or industry associations.[42] Some industries are followed by the U.S. Census Department, such as trucking, retail trade, and manufacturing industries. Some industries have consulting firms that follow the industry – as in venture capital. And some industries have industry associations or the desire to publicize their firm's attributes – as in franchising and law. These data sets might follow HR practices, or productivity, or both.

5. Insider analysis using employer-employee matched data sets that are available for entire countries. In Europe, the U.S., and some developing countries, the census departments are matching data on individual workers to data from the firm. These data sets span all workers in all industries. But it may be wise to take subsets of the data sets, for certain industries or occupations, and delve deeper into the production function. Though the data sets do not match HR information to these data, researchers can use the wages and mobility of all the workers in the firm to infer the sorts of HR practices that firms are adopting.[43]

6. Insider analysis using data from developing countries is now emerging. There are now studies where firms are allowing researchers to randomly assign new management practices to plants in a single industry (e.g., Bloom, Eifert, Mahajan, McKenzie and Roberts, 2009). Such studies that measure average treatment effects offer some unique opportunities to identify observations that enjoy relatively large or small productivity impacts, which in turn could help identify the causes for any observed heterogeneity in the treatment effects.

[41] Asch (1990). Baker and Hubbard (2003). Dee and Keys (2004). Dranove, Kessler, McClellan and Satterthwaite (2003). Eberts, Hollenbeck, and Stone (2000), Graciano and Heaton (2006). Hoxby (2002)., Jacob and Levitt (2003). Kreuger (1999). Lavy (2008). Lavy (2002). Lim (2007). Rivkin, Hanushek and Kain (2005). Hall, VanReenen, and Propper (2008). Rosenthal, Song and Landon (2004). Imberman, Kugler and Sacerdote (2009). Duflo, Hanna and Ryan (2007). Kalnins and Wilbur (2006).

[42] Autor and Scarborough (2008). Black and Lynch(2004). Black and Lynch (2001). Baker and Hubbard (2003), Hubbard (2000), Garicano and Hubbard (forthcoming). Lafontaine and Shaw (1999). Lafontaine and Shaw (2005). Landers, Rebitzer and Taylor (1996). Oyer and Schaefer(2007).

[43] See Andersson, Freedman, Haltiwanger, Lane and Shaw (2009) for work using employer-employee matched data for the software industry. See Lazear and Shaw (2007) for examples of other papers and the range of possible data sets.

In sum, to obtain data on both productivity outcomes and management practices, the researcher will often need to merge several datasets. Helper (2000) discusses additional issues and ideas on working with firms and their data. When working with firms, it is important for the researchers to show what they can do for the firm, while acknowledging the importance of the firm's generosity in these efforts aimed at furthering our knowledge of how businesses really operate.

X. Conclusion

Insider econometric research has goals at two different levels. First, insider econometric research analyzes information from inside the firm with modern econometric tools to estimate the impact of organizational practices on performance. Second, the broader goal of insider research is to use these estimated performance results to improve our understanding and theories of the firm, and where possible, to offer suggestions to managers for improved organizational practices within their firms.

We describe insider econometric studies as research that focuses on three basic questions:

1. *Does a new management practice raise productivity*?
2. *Why does the new management practice raise productivity?*, and
3. *Why is the new management practice adopted?*

Ultimately, designing and executing a study that develops persuasive answers to these questions in part art and part science. Judgments about which management practices to study and how to measure them, how to balance homogeneity and heterogeneity in the design of the sample and data set, and about the many aspects of modeling the determinants of productivity and adoption are critical parts of the research.

The approach that insider econometric research takes to address these questions combines qualitative field research and quantitative econometric research, with the insights from interviews and field research informing the econometric analysis in important ways. Thus, the insider econometrician must be an expert translator. The researcher must translate the real world experience of the industry insiders into a data set and research design that addresses these fundamental questions about worker and firm behavior. The research designs of insider econometric research rely on rich micro-level data on workers or work groups from a single industry and production process. While the empirical tests pertain to one industry setting, it is precisely this kind of detailed information about new management practices, worker characteristics, the production process, product characteristics, and other aspects of the industry setting that allows the researcher to test new theories about what makes workers and firms more productive.

References

Akerlof, George A. and Rachel Kranton, (2000) "Economics and Identity," *Quarterly Journal of Economics,* 15(3): 715-753.

Akerlof, George A. and Rachel Kranton, (2005) "Identity and the Economics of Organizations," *Journal of Economic Perspectives,* 19(1): 9-32.

Andersson, Fredrik, Clair Brown, Benjamin Campbell, Hyowook Chiang and Yooki Park, (2006), "The Effect of HRM Practices and R&D Investment on Worker Productivity," Working paper.

Andersson, Fredrik, Matthew Freedman, John C. Haltiwanger, Julia Lane and Kathryn L. Shaw, (2009), "Reaching for the Stars: Who Pays for Talent in Innovative Industries?," Economic Journal, Royal Economic Society, 119(538) :F308-F332.

Appelbaum, Eileen, Thomas Bailey, Peter Berg, and Arne Kalleberg (2000), *Manufacturing Advantage: Why High-Performance Work Systems Pay Off* Ithaca, NY: Cornell University Press.

Appelbaum, Eileen, and Rosemary Batt (1994), "*The New American Workplace: Transforming Work Systems in the United States*," Ithaca, NY: Cornell University Press.

Arthur, Jeffrey (1992), "The Link between Business Strategy and Industrial Relations Systems in American Steel Minimills," *Industrial and Labor Relations Review*, 45: 488-506.

Arruñada, Benito, Luis Garicano and Luis Vázquez (2001), "Contractual Allocation of Decision Rights and Incentives: The Case of Automobile Distribution." *Journal of Law, Economics, and Organization,* 17(1):256-283.

Asch, Beth J. (1990), "Do Incentives Matter—The Case of Navy Recruiters." *Industrial and Labor Relations Review*, 43(3): 89-106.

Athey, Susan and Scott Stern (1998), "An Empirical Framework for Testing Theories about Complementarity in Organizational Design," *NBER* working paper 6600. Revise & Resubmit, *Management Science*.

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

Autor, David (2003), "Why Do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics* November 116(4): 1409-1448.

Autor, David, Frank Levy, and Richard Murname (2003), "The Skill Content of Recent Technological Change: An Empirical Investigation." *Quarterly Journal of Economics* November 118(4): 1279-1333.

Autor, David, Frank Levy, and Richard Murname (2002), "Upstairs Downstairs: Computers and Skills in Two Floors of a Large Bank." *Industrial and Labor Relations Review*, April, 55(3): 432-447.

Azoulay, Pierre (2004), "Capturing Knowledge within and across Firm Boundaries: Evidence from Clinical Development." *American Economic Review*, December, 94(5): 1591-1612.

Azoulay, Pierre and Joshua Graff Zivin (2005), "Peer Effects in the Workplace: Evidence from Professional Transitions for the Superstars of Medicine." Working paper.

Baker, George P. and Thomas N Hubbard (2003), "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information," *American Economic Review*, 93(3): 551-572.

Bandiera, Oriana, Iwan Barankay and Imran Rasul (2009), "Social Connections and Incentives in the Workplace: Evidence from Personnel Data" *Econometrica,* Econometric Society, 77(4): 1047-1094.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2007), "Incentives for Managers and Inequality among Workers: Evidence from a Firm Level Experiment" *Quarterly Journal of Economics* May, 122(2): 729-773.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, (2005), "The Evolution of Cooperative Norms: Evidence from a Natural Field Experiment," CEPR Discussion Paper No. 5358.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, (2005), "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *Quarterly Journal of Economics*, 120(3): 917-962.

Bartel, Ann, Ciaran Phibbs, Patricia Stone, and Nancy Beaulieu, (2009), "Human Capital and Productivity: The Case of Nursing Teams," Working paper.

Bartel, Ann (2004), "Human Resource Management and Organizational Performance: Evidence from Retail Banking" *Industrial and Labor Relations Review*, 57(2): 181-203.

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw (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, Casey Ichniowski, Kathryn Shaw, and Ricardo Correa (2009), "International Differences in the Adoption and Impact of new Information Technologies and New HR Practices: The Valve Making Industry in the U.S. and U.K." in *International Differences in the Business Practices and Productivity of Firms*, eds. Richard Freeman and Kathryn Shaw, University of Chicago, National Bureau of Economic Research.

Baron, James and Michael Hannan (2002), "Organizational Blueprints for Success in High-Tech Start-Ups: Lessons From The Stanford Project On Emerging Companies." *California Management Review*, 44(3): 8-36.

Batt, Rosemary (1999), "Work Organization, Technology, and Performance in Customer Services and Sales," *Industrial and Labor Relations Review*, July 52: 539-64.

Bayo-Moriones, Alberto, José Enrique Galdón Sánchez and Maia Güell (2004), "Is Seniority-Based Pay Used as a Motivation Device? Evidence from Plant Level Data." No 4606, CEPR Discussion Papers.

Berg, Peter, Eileen Appelbaum, Thomas Bailey, and Arne L. Kalleberg (1996) "The Performance Effects of Modular Production in the Apparel Industry." Industrial Relations, 35(3): 356-74.

Bertrand, Marianne, Esther Duflo,and Sendhil Mullainathan (2004), "How Much Should We Trust Difference-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249-75.

Bhattacherjee, Debashish (2005), "The Effects of Group Incentives in an Indian Firm: Evidence from Payroll Data," *Labour*, 19(1): 147-173.

Black, Sandra and Lisa Lynch (2004), "What's Driving the New Economy?: The Benefits of Workplace Innovation", *The Economic Journal,* 114(493): 97-116.

Black, Sandra and Lisa Lynch (2001), "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity", *Review of Economics and Statistics,* August 83(3):434-445.

Bloom, Nick, Mirko Draca, and John Van Reenen (2008), "Trade Induced Technical Change? The Impact of Chinese Imports on IT and Innovation," November.

Bloom, Nick, Benn Eifert, Aprajit Mahajan, David McKenzie and John Roberts (2009), "Management as a Technology: Evidence from Indian firms," working paper.

Bloom, Nick, Christos Genakos, Raffaella Sadun and John Van Reenen (2009), "Management Practices Across Asia, Europe and the US," Working paper.

Bloom, Nick and John Van Reenen (2007), "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, November 122(4): 1351-1408.

Blundell, Richard and Monica Costa Dias (2002), "Alternative Approaches to Evaluation in Empirical Microeconomics," *Portuguese Economic Journal*, 1, 91-115.

Bollinger, Bryan, Phillip Leslie and Alan Sorensen (2009) "Calorie Posting in Chain Restaurants" Working paper.

Boning, Brent, Casey Ichniowski, and Kathryn Shaw (2007), "Opportunity Counts: Teams and the Effectiveness of Production Incentives," *Journal of Labor Economics,* 25(4): 613-650.

Breshnahan, Timothy, Erik Brynjolfsson and Loren Hitt (2002), "Information Technology, Work Organization and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics*, 117(1): 339-376.

Cappelli, Peter, and David Neumark (2001), **"**Do High-Performance Work Practices Improve Establishment-Level Outcomes?" *Industrial and Labor Relations Review*, July 54(4): 737-775.

Caroli, Eve and John Van Reenen (2001), "Skill biased Organisational Change? Evidence from British and French Establishments" *Quarterly Journal of Economics,* 171(4): 1449-1492.

Chan, Tat, Jia Li and Lamar Pierce (2009), "Compensation and Peer Effects in Competing Sales Teams," Working paper.

Chevalier, Judy and Glenn Ellison (1999), "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, May 114(2): 389-432.

Chu, Sean, Phillip Leslie and Alan Sorensen (2009) "Bundle-Size Pricing as an Approximation to Mixed Bundling" forthcoming in *American Economic Review*.

Clark, Kim B. (1980), "The Impact of Unionization on Productivity: A Case Study," *Industrial and Labor Relations Review*, 33(4): 451-469.

Clark, Kim B. (1980), "Unionization and Productivity: Micro-Econometric Evidence," *Quarterly Journal of Economics*, 95(4): 613-639.

Clark, Kim B. (1984), "Unionization and Firm Performance: The Impact on Profits, Growth, and Productivity," *The American Economic Review*, 74(5): 893-919.

Dee, Thomas and Benjamin Keys (2004), "Does Merit Pay Reward Good Teachers? Evidence form a Randomized Experiment," *Journal of Policy Analysis and Management*, 23(3): 471-488.

Doyle, Joseph J., Ewer, Steven and Wagner, Todd, (2008), "Returns to Physician Human Capital: Analyzing Patients Randomized to Physician Teams" NBER Working Paper No. 14174.

Drago, Robert and Gerald T. Garvey (1998), "Incentives for Helping on the Job: Theory and Evidence" *Journal of Labor Economics*, 16(1): 1-24.

Dranove, David, Daniel Kessler, Mark McClellan and Mark Satterthwaite (2003), "Is More Information Better? The Effects of "Report Cards" on Health Care Providers," *Journal of Political Economy*, 111(3): 555-588.

Duflo, Esther, Rema Hanna and Stephen Ryan (2007), “Monitoring Works: Getting Teachers to Come to School,” *BREAD* Working Paper No. 103.

Dunlop, John T. and David Weil (2000), “Diffusion and Performance of Modular Production in the U.S. Apparel Industry,” in *The American Workplace: Skills, Compensation and Employee Involvement*, eds. Casey Ichniowski, David Levine, Craig Olson, and George Strauss (Cambridge University Press).

Eberts, Randall, Kevin Hollenbeck, and Joe Stone (2000), “Teacher Performance Incentives and Student Outcomes,” *The Journal of Human Resources*, 37(4): 913-927.

Ehrenberg, Ronald G. and Michael L. Bognanno (1990), "The incentive effects of tournaments revisited: Evidence from the European PGA tour," *Industrial and Labor Relations Review*, 43(3): 74-88.

Elfenbein, Daniel W and Josh Lerner (2003), “Ownership and Control Rights in Internet Portal Alliances, 1995-1999.” *RAND Journal of Economics*, 34(2): 356-69.

Encinosa, William E. III; Martin Gaynor and James B. Rebitzer (2007), “The Sociology of Groups and the Economics of Incentives: Theory and Evidence on Compensation Systems,” *Journal of Economic Behavior & Organization*, Elsevier, 62(2): 187-214.

Fernie, Sue and David Metcalf (2003), “Its Not What You Pay It’s the Way That You Pay It and that’s What Gets Results: Jockeys’ Pay and Performance.” *Review of Labour Economics and Industrial Relations*, 13(2): 385-411.

Forbes, Silke J. and Mara Lederman (2009), “Does Vertical Integration Affect Firm Performance? Evidence from the Airline Industry.” Working paper.

Freeman, Richard B. and Morris M. Kleiner (2005), “The Last American Shoe Manufacturers: Decreasing Productivity and Increasing Profits in the Shift from Piece Rates to Continuous Flow Production,” *Industrial Relations*, 44(2): 307-330.

Galizzi, Monica, Petra Miesmaa, and Craig Slatin (2006), “Occupational Injuries, Workers’ Reporting and Firms Policies in the Health Care Industry: The Challenges and Rewards of Combining Qualitative and Quantitative Research Methodologies,” Working paper.

Gant, Jon, Casey Ichniowski and Kathryn Shaw (2002), “Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations,” *Journal of Economics and Management Strategy*, 11(2): 289-328.

Garicano, Luis and Tom Hubbard (forthcoming), “Specialization, Firms and Markets: The Division of Labor Within and Between Law Firms.” *Journal of Law, Economics and Organization*.

Garicano, Luis and Paul Heaton (2009), “Information Technology, Organization, and Productivity in the Public Sector: Evidence from Police Departments.” *Journal of Labor Economics* forthcoming.

Garicano, Luis and Tom Hubbard (2007), "The Returns to Knowledge Hierarchies." US Census Bureau Center for Economic Studies Paper no CES-WP-07-01.

Garicano, Luis and Paul Heaton (2006), "Computing Crime: Information Technology, Police Effectiveness, and the Organization of Policing," *CEPR* Discussion paper no 5837.

Garicano, Luis and Thomas N. Hubbard (2005), "Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyers' Human Capital," *CEPR* Discussion Papers 4924.

Gaynor, Martin, James B Rebitzer and Lowell J. Taylor (2001), "Incentives in HMOs" *Macroeconomics* 0111001, EconWPA.

Gibbons, Robert and J.Roberts (forthcoming), *Handbook of Organizational Economics* Princeton University Press, forthcoming 2010.

Gil, Ricard (2009), "The Interplay of Formal and Relational Contracts: Evidence from Movies." Working paper.

Griffith, Rachel and Andy D. Neely (2007), "Incentives and Managerial Experience in Multi-Task Teams: Evidence from within a Firm" CEPR Discussion Paper No. DP6107.

Griliches, Zvi.(1994), "Productivity, R&D, and the Data Constraint," *American Economic Review,* 84(1): 1-23.

Haltiwanger, John, Fredrik Andersson, Matthew Freedman, Julia Lane and Kathryn Shaw (2009), "Reaching for the Stars: Who Pays for Talent in Innovative Industries?" *Economic Journal Royal Economic Society* 119(538): F308-F332.

Haltiwanger, John (2008), "Top Ten Questions for Understanding Firm Dynamics and Productivity Growth" Keynote address at CAED.

Hall, Emma, John Van Reenen and Carol Propper (2008), "Can Pay Regulation Kill? The Impact of Labor Markets and Skills on Hospital Productivity" CEPR Discussion Paper 843.

Hamilton, Barton, Jack Nickerson, and Hideo Owan (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.

Heckman, James J. (1990), "Varieties of Selection Bias," *American Economic Review*, 80(2): 313-18.

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith, and Petra Todd (1999), "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5): 1017-1098.

Heckman, James J., Robert Lalonde, and J. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*, vol. 3, eds. Orley Ashenfelter and David Card (Elsevier Science: Amsterdam), chapter 31: 1865-2097.

Helper, Susan (2000), "Economists and Field Research: 'You Can Observe a Lot by Just Watching." *American Economic Review*. May. 90(2): 228-232.

Henderson, Rebecca, and Iain Cockburn, (1996) "Scale, Scope, and Spillovers: Determinants of Research Productivity in the Pharmaceutical Industry." RAND Journal of Economics, Spring, 27(1), pp. 32-59.

Hoxby, Caroline (2002), "Would School Choice Change the Teaching Profession?" *Journal of Human Resources*, 37(4): 846-891.

Hubbard, Thomas N. (2003), "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking," *American Economic Review*, 93(4): 1328-1353.

Hubbard, Thomas N. (2000), "The Demand for Monitoring Technologies: The Case of Trucking." *The Quarterly Journal of Economics*, May, 115(2): 533-560.

Huselid, Mark and Brian Becker (1996), "High Performance Work Systems and Firm Performance: Cross-Sectional Versus Panel Estimates," *Industrial Relations*, 35(3): 635-672.

Ichino, Andrea and Giovanni Maggi (2000), **"**Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm" *The Quarterly Journal of Economics*, 115(3): 1057-1090.

Ichniowski, Casey and Kathryn Shaw (2003), "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices" *Journal of Economic Perspectives,* winter, 17(1): 155-180.

Ichniowski, Casey, Kathryn Shaw and Giovanna Prennushi (1997), "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," *American Economic Review*, 86: 291-313.

Ichniowski, Casey and Shaw, Kathryn (1995), "Old Dogs and New Tricks: Determinants of the Adoption of Productivity-Enhancing Work Practices" *Brookings Papers on Economic Activity: Microeconomics*: 1-65.

Ichniowski, Casey and Shaw, Kathryn (1999), "The Effects of Human Resource Systems on Productivity: An International Comparison of U.S. and Japanese Plants," *Management Science*, 45: 704-722.

Imbens, Guido (2004), "NonParametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *The Review of Economics and Statistics*, 86(1): 4-29.

Imbens, Guido and Thomas Lemieux (2007), "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, February 142(2): 615-635.

Imberman, Scott, Adriana Kugler, and Bruce Sacerdote (2009) "Katrina's Children: A Natural Experiment in Peer Effects from Hurricane Evacuees," *Mimeo*, April.

Ingram, Paul and Peter Roberts (2000) "Friendships Among Competitors in the Sydney Hotel Industry," American Journal of Sociology, 106(2): 387-423.

Jacob, B. and S. Levitt (2003), "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *The Quarterly Journal of Economics*, 118(3): 843-877.

Jones, Derek C. and Kato, Takao (2007), "The Impact of Teams on Output, Quality and Downtime: An Empirical Analysis Using Individual Panel Data," *IZA* Discussion Papers 2917.

Jones, Derek C., Panu Kalmi, and Antti Kauhanen (2006), "Teams, Performance-Related Pay, Profit Sharing and Productive Efficiency: Evidence from a Food-Processing Plant," Working paper.

Jorgenson, Dale W., Mun C. Ho and Kevin Stiroh (2003) "Growth of U.S. Industries and Investments in Information Technology and Higher Education," in NBER/CRIW Conference on Measurement of Capital in the New Economy.

Jorgenson , Dale W. and Kevin J. Stiroh (2000), "U.S. Economic Growth at the Industry Level," *American Economic Review*, American Economic Association, 90(2) :161-167.

Joskow, Paul (1985), "Vertical Integration and Long-Term Contracts: The Case of Coal-Burning Electric Generation Plants." *Journal of Law, Economics, and Organization*, 1: 33-80.

Joskow, Paul (1987), "Contract Duration and Relationship-Specific Investment: Empirical Evidence from Coal Markets." *American Economic Review*, 77: 168-85.

Kalnins, Arturs and Wilbur Chung (2006), "Social Capital, Geography, and Survival: Gujarati Immigrant Entrepreneurs in the U.S. Lodging Industry," *Management Science*, 52(2): 233-247.

Kandel, Eugene and Edward Lazear, (1992), "Peer Pressure and Partnerships," *Journal of Political Economy* , 100(4): 801-17.

Knoeber, Charles R. and Walter N. Thurman (1994), "Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production," *Journal of Labor Economics*, 12(2):155-79.

Kortum, Samuel S. and Josh Lerner (2000), "Assessing the Contribution of Venture Capital to Innovation." *RAND Journal of Economics*, 31(4): 674-692.

Kreuger, Alan (1999), "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics,* 114(2): 497-532.

Kreuger, Alan and Alex Mas (2004), "Strikes, Scabs and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires," *Journal of Political Economy* 112(2): 253-289.

Kruse, Douglas (1993), *Profit Sharing: Does it Make a Difference?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Kruse, Douglas, Joseph Blasi, and Rhokeum Park (2008), "Shared Capitalism in the U.S. Economy: Prevalence, Characteristics, and Employee Views of Financial Participation in Enterprises"Working paper, presented at NBER-Russell Sage conference on shared capitalism October.

Lafontaine, Francine and Margaret Slade (2008), "Inter-Firm Contracts: Evidence", *Journal of Economic Literature*.

Lafontaine, Francine (1992), "Agency Theory and Franchising: Some Empirical Results" *Rand Journal of Economics*, 23: 263-83.

Lafontaine, Francine and Kathryn Shaw (1999), "The Dynamics of Franchise Contracting: Evidence from Panel Data," *The Journal of Political Economy*, October, 107(5): 1041-1080.

Lafontaine, Francine and Kathryn Shaw (2005), "Targeting Managerial Control: Evidence from Franchising," *Rand Journal of Economics,* Spring 36(1): 131-150.

Landers, Renee M., James B Rebitzer, and Lowell J. Taylor (1996), "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms,"*American Economic Review*, 86(3): 329-348.

Lavy, Victor (2008), "Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance Based Pay Tournaments among Teachers," NBER Working paper.

Lavy, Victor, and Joshua Angrist (2009), "The Effects of High Stakes High School Achievement Awards: Evidence from a Group-Randomized Trial," *American Economic Review*, 99(4): 1384-1414.

Lavy, Victor (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Students Achievements," *Journal of Political Economy*, December, 110(6): 1286-1317.

Lazear, Edward and Kathryn Shaw (2007), "Personnel Economics: The Economist's View of Human Resources" *Journal of Economic Perspectives* 21(4): 91-114.

Lazear, Edward (2000), "Performance Pay and Productivity," *American Economic Review*,90(5):1346-1361.

Lazear, Edward and Paul Oyer (2007), "Personnel Economics," NBER Working Papers 13480.

Lazear, Edward (1995), *Personnel Economics*, Cambridge, MA: MIT Press.

Lazear, Edward (1999), "Personnel Economics: Past Lessons and Future Directions," *Journal of Labor Economics*, 17(2): 199-236.

Lazear, Edward (1999), “Globalization and the Market for Teammates,” *The Economic Journal*, 109(454): 15-40.

Ledford, G. E. Jr., Lawler, E. E., III, and Mohrman, S. A., (1995), “Reward Innovation in Fortune 1000 Companies.” *Compensation and Benefits Review*, 76-80.

Lemieux, Thomas, Bentley MacLeod, and Daniel Parent (2009), “Performance Pay and Wage Inequality,” *Quarterly Journal of Economics*, February, 124(1): 1-49.

Lerner, Josh and Julie Wulf (Forthcoming), “Innovation and Incentives: Evidence from Corporate R&D.” *Review of Economics and Statistics*.

Lerner, Josh and Robert P. Merges (1998), “The Control of Technology Alliances: An Empirical Analysis of the Biotechnology Industry.” *Journal of Industrial Economics*, June, 46(2): 125-156.

Leslie, Phillip (2004), "Price Discrimination in Broadway Theatre," *RAND Journal of Economics*, 35(3): 520-41.

Levitt, Steven D. and John List (2008) “Field Experiments in Economics: The Past, the Present, and the Future,” NBER Working Paper no. 14356.

Lim, Claire (2007), “Turnover and Accountability of Appointed and Elected Judges” (Ph.D. Dissertation), University of Pennsylvania.

List, John, Sally Sadoff and Mathis Wagner (2008) “So You Want to Run an Experiment, Now What? An Introduction to Optimal Sample Arrangements,” September, working paper.

Lo, Desmond, Mrinal Ghosh and Lafontaine, Francine (2006), “The Role of Risk, Incentives, and Selection in Salesforce Compensation Contracts,” Working paper.

MacDuffie, John Paul (1995), “Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry,” *Industrial and Labor Relations Review*, 48: 197-221.

McMillan, John and Christopher Woodruff (1999), “Interfirm Relationships and Informal Credit in Vietnam.” *The Quarterly Journal of* Economics, March, 114(4): 1285-1320.

Mas, Alexandre (2006), “Pay, Reference Points, and Police Performance,” *The Quarterly Journal of Economics*, 121(3): 783-821.

Mas, Alexandre and Enrico Moretti (2009), “Peers at Work,” *American Economic Review*, September, 99(1): 112-145.

Mas, Alexandre (2008), “Labor Unrest and the Quality of Production: Evidence from the Construction Equipment Resale Market,” *Review of Economic Studies* 75(1): 229-258.

Mas, Alexandre (2004), "Do Disappointing Pay Raises Lower Productivity? Final-Offer Arbitration and the Performance of New Jersey Police Officers". Princeton University.

Masten, Scott E (1984), "The Organization of Production: Evidence from the Aerospace Industry", *Journal of Law and Economics*, 27: 403-17.

Meyer, Bruce (1995) "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, April, 13 (2): 151-161.

Milgrom, Paul and John Roberts (1995), "Complementarities and Fit: Strategy, Structure and Organizational Change in Manufacturing," *Journal of Accounting and Economics*, 19(2-3): 179-208.

Monteverde, Kirk, and David Teece (1982), "Supplier Switching Costs and Vertical Integration in the Automobile Industry." *Bell Journal of Economics*, 13: 206-12.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor (2002), "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment," *American Economic Review*, 92(4): 850-873.

Misra, Sanjog and Harikesh Nair (2008), "A Structural Model of Sales-Force Compensation" Working paper, Stanford University.

Neumark, David, Roy J. Bank and Kyle D. Van Nort (1996), "Sex Discrimination in Restaurant Hiring: An Audit Study," *The Quarterly Journal of Economics*, August, 111(3): 915-941.

Nickerson, Jack A. and Brian S. Silverman (2003), "Why aren't all Truck Drivers Owner-Operators? Asset Ownership and the Employment Relation in Interstate for-Hire Trucking." *Journal of Economics and Management Strategy*, 12(1): 91-118.

Oliner, Stephen D. and Daniel E. Sichel (2000). "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives*, 14 (4):3-22.

Osterman, Paul (1994), "How Common is Workplace Transformation and Who Adopts It?" *Industrial and Labor Relations Review*, 47(2): 173-188.

Osterman, Paul (2000), "Work Reorganization in an Era of Restructuring: Trends in Diffusion and Effects on Employee Welfare" *Industrial and Labor Relations Review*, 53(2): 179-196.

Oyer, Paul and Scott Schaefer (2007), "Personnel-Economic Geography: Evidence from Large US Law Firms," *2nd Annual Conference on Empirical Legal Studies Paper.*

Oyer, Paul (1998), "Fiscal Year Ends and Non-Linear Incentive Contracts: The Effect on Business Seasonality," *The Quarterly Journal of Economics*, February 113: 149-185.

Perez-Gonzalez, Francisco (2006), “Inherited Control and Firm Performance,” *American Economic Review*, December, 96(5):1559-1588.

Philippon, Thomas and Ariell Reshef (2007), “Skill Biased Financial Development: Education, Wages and Occupations in the U.S. Financial Sector,” NYU Working Paper.

Prendergast, Canice (1999), “The Provision of Incentives in Firms,” *Journal of Economic Literature*, 37(1): 7-63.

Rivkin, Steven, Eric Hanushek, and John Kain (2005), “Teachers, Schools, and Academic Achievement,” *Econometrica* 73(2): 417-458.

Rosenthal, Meredith, Rushika Fernandopulle, HyunSook Ryu Song, and Bruce Landon (2004), “Paying for Quality: Providers’ Incentive for Quality Improvement,” *Health Affairs* 23(2): 127-141.

Shaw, Kathryn Adam Jaffe, Josh Lerner and Scott Stern (2003), “The Human Resources Revolution: Is it a Productivity Driver?” *Innovation, Policy and the Economy,* eds Chicago: University of Chicago, National Bureau of Economic Research 69-114.

Shaw, Kathryn Donna Ginther, and Madeline Zavodny (2003), “Technology Shocks and Problem-solving Capacity,” in *Technology, Growth, and the Labor Market,* eds Boston: Kluwer Academic Publishers 235-258.

Shaw, Kathryn Nathalie Greenan, Yannick L’Horty, and Jacques Mairesse (2002), “By What Means Does Information Technology Affect Employment and Wages?” *Productivity, Inequality, and the Digital Economy: A Transatlantic Perspective,* eds Cambridge: MIT Press.

Schotter C. Bull, A. and K. Weigelt. (1987), “Tournaments and piece rates: An Experimental Study,” *Journal of Political Economy* . 95 :1-33.

Syverson, Chad, Lucia Foster and John Haltiwanger (2008), “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, March.

Syverson, Chad (2004), “Market Structure and Productivity: A Concrete Example.” *Journal of Political Economy*, 112(6): 1181-1222.

Syverson, Chad, Austan Goolsbee and Steven Levitt (2009), “What Determines Productivity at the Micro Level?” Microeconomics. Working Paper.

Van Biesebroeck, Johannes (2006), “Complementarities in Automobile Production,” *Journal of Applied Econometrics* 22(7): 1315-1345.

Blundell, Richard and Monica Costa Dias (2002), “Alternative Approaches to Evaluation in Empirical Microeconomics,” *Portuguese Economic Journal*, 1, 91-115.

List, John, Sally Sadoff and Mathis Wagner (2008) "So You Want to Run an Experiment, Now What? An Introduction to Optimal Sample Arrangements," September, working paper.

Meyer, Bruce (1995) "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, April, 13 (2): 151-161.

TABLE 1

A Partial List of Additional Insider Econometric Studies,
By Management Practices Analyzed

Sets of Complementary Management Practices	**Performance Outcome**	**Why?**
Ichniowski, Shaw and Prenusshi (1997) Data: Team Monthly productivity, team-level five years, 36 steel production lines (17 companies).	Productivity rises by 10% due to combined use of incentives, teamwork, careful hiring and other practices.	When firms implement multiple human resource management practices, the complementarities between the practices raises performance.
Griffith and Neely (forthcoming) Data: retail stores, monthly sales, months, one company.	Productivity rises when the multi-dimensional. Balanced Scorecard performance evaluation is introduced.	When Balanced Scorecard incentive pay is adopted, productivity rises only if managers are experienced.
Tat Chan, Jia Li and Lamar Pierce Data: daily sales for 85 cosmetics sales workers, 791 days, eleven companies.	Individual incentive pay raises individual productivity, but lowers within- team cooperation and lowers group productivity.	Individual incentive pay raises personal productivity, reduces cooperation and reduces overall productivity. Team-based incentive pay increases cooperation within team and increases competition across teams (or firms).
Baron and Hannon (2002) Data: 200 technology startup companies in Silicon Valley, annual data, 1995 and 1997.	Sets of innovative "star" HR practices raise growth 140%, and reduce the probability of firm failure 210%, relative to autocratic HR practices.	The initial set of human resources practices that the startup firm chooses has a large affect on all measures of firm success: "star" firms "recruit only top talent, pay them top wages, and give them resources and autonomy. In "autocratic" firms "you work, you get paid."
MacDuffie (1995) Data: 62 auto assembly plants.	Productivity and quality outcomes are higher in plants that have systems of innovative HR practices. Effects are larger if plant also has management practices for flexible production systems.	The "organizational logic" of lean manufacturing production processes is no supported by individual HR practices but by multiple HR practices that are elements of a system.

Sets of Complementary Management Practices	Performance Outcome	Why?
Dunlop and Weil (2000) Data: 42 business units in the men's suit, shirt and pants sector and in the women's jeans and undergarment sectors of the apparel making industry.	Lead times are substantially shorter when team-based systems of HR practices in modular assembly production processes are used.	The adoption of team-based HR systems in modular production processes is driven by the adoption of new complementary sales-tracking information technologies by large retailers who required apparel makers to fill smaller orders more rapidly. The effects of new HR practices on performance outcomes are not independent of the effects of the use of new information technologies.
Berg, Appelbaum, Bailey, and Kalleberg (1996) Data: Production records and worker surveys from over 100 workers in 4 plants of 2 apparel making companies.	Systems of HR practices associated with the "module system of manufacturing" exhibit lower production costs, reduced throughput production times, reduced work-in-process inventories, and increased variety of clothing styles than HR practices under the "bundle system."	Team-based HR practices under module production allows greater coordination of workers' activities that eliminate production bottlenecks and solve production problems.

Incentive Pay	**Performance Outcome**	**Why?**
Lazear (2000) Data: monthly measures of average daily output for 3000 workers who install car windshields, 19 months, one company.	The Productivity of individual workers rises 44% due to a switch to piece rate pay, from hourly pay.	Higher effort by workers due to incentive pay raises productivity by 22%. Better sorting or selection of workers to the job raises productivity by 22% more.
Bandiera, Barankay and Rasul (2005) Data: daily output for 142 workers, 108 days, picking fruit, one company.	Productivity of workers rises 58% when switch to piece rate pay from a relative pay scheme.	Higher Effort of workers rises due to piece rate pay, because workers were withholding effort when friends were on their work teams.
Nagin, Rebitzer, Sanders and Taylor (2002) Data: daily data, 16 call centers, one company	Daily productivity of workers rises by about 3% when pay is tied to call monitoring.	Monitoring reduces instances of workers reporting unsuccessful sales calls as successes. Perceptions of a "fair work culture" also reduce this unproductive behavior.
Misra and Nair (2009) Data: monthly sales, 87 sales agents for contractors for 38 months, one company.	Total revenues rose 9% following an incentive pay plan that has a linear incentive (no quotas) and that pays out more frequently.	A pay plan that is highly non-linear and pays out over long intervals (quarters rather than months) results in gaming the system and lower productivity for current employees.
Lavy (2008) Data: student test scores, one year, 629 teachers in Isreali school system	Productivity of students rose 3% after new tournament bonus pay for teachers, and rose 1 to 6% in other similar plans: see Angrist and Lavy 2002; Lavy, 2002.	Students' test scores in English and Math rose due to a change in teaching methods and increase in teachers' effort in response to incentives.
Lerner and Wulf (2007) Data: 140 firms, 700 firm years, compensation for R&D heads, matched to patent data.	R&D group productivity rises by 26%, as measured by citations or patents, when long term incentives (measured as a percent of pay) rise by 50%.	Use of long-term incentives for corporate heads of R&D departments rose 50% from 1988 to 1998 and is likely to have improved R&D project selection.

Peer Effects	**Performance Outcome**	**Why?**
Azoulay and Zivin (2005) Data: Total publications by 4764 star scientists and their co-workers,	Productivity rises 17% when a scientist is co-located near a star scientist.	Collaboration with a star raises personal productivity, but when a star is geographically nearby it raises productivity further.
Ichino and Maggi (2000) Data: retail bank employees, 28642 employees, 442 branches, 20 years, one company.	Productivity (measured by absenteeism) rises by over 15% when employees are located with peer employees who work harder.	Worker productivity is strongly influenced by personal background and culture, but these influences are reduced when workers are paired with more productive peers.
Mas and Moretti (2009) Data: all supermarket transactions, 6 stores, 2003-2006. One company for the checkout clerks,	Within a 10 minute work interval, personal productivity rises 1.7% when working near a peer who is 10% more productive than average.	The productivity gains from peer effects show that low productivity workers work harder when a high productivity worker watches them work.

Teamwork	**Performance Outcome**	**Why?**
Hamilton, Nickerson and Owan (2003) Data: weekly productivity, 132 workers, 156 weeks, apparel sewing, one company.	Team-based pay increased productivity 14% compared to individual piece rate pay.	Team-based pay increases team output through collaboration among workers with complementary skills. Highly productive workers raise team productivity more.
Bartel, Phibbs, Beaulieu and Stone (2009) Data: 431 hospital intensive case units, monthly productivity, 48 months one company.	Hospital units with more "unit-specific" human capital have a 1.5% reduction in patients' length of stay.	Teamwork within nurses' work area combines nurses' human capital which raises productivity.
Boning, Ichniowski and Shaw(2007) Data: 2355 monthly productivity, team-level, observations for 34 steel minimills, five years,19 companies.	There is a 11 to 20% reduction in defective output when mills move to team-based problem solving.	Team-based problem solving is more valuable to the firm when the steel mill produces complex products than when it produces commodities.
Bandiera, Barankay and Rasul (2009) Data: 407 fruit picking teams, 15 fields, 109 days	Productivity rises 24% when a team based tournament reward system is introduced.	Workers of similar ability form teams in response to incentives; teams are not formed on the basis of friendship.

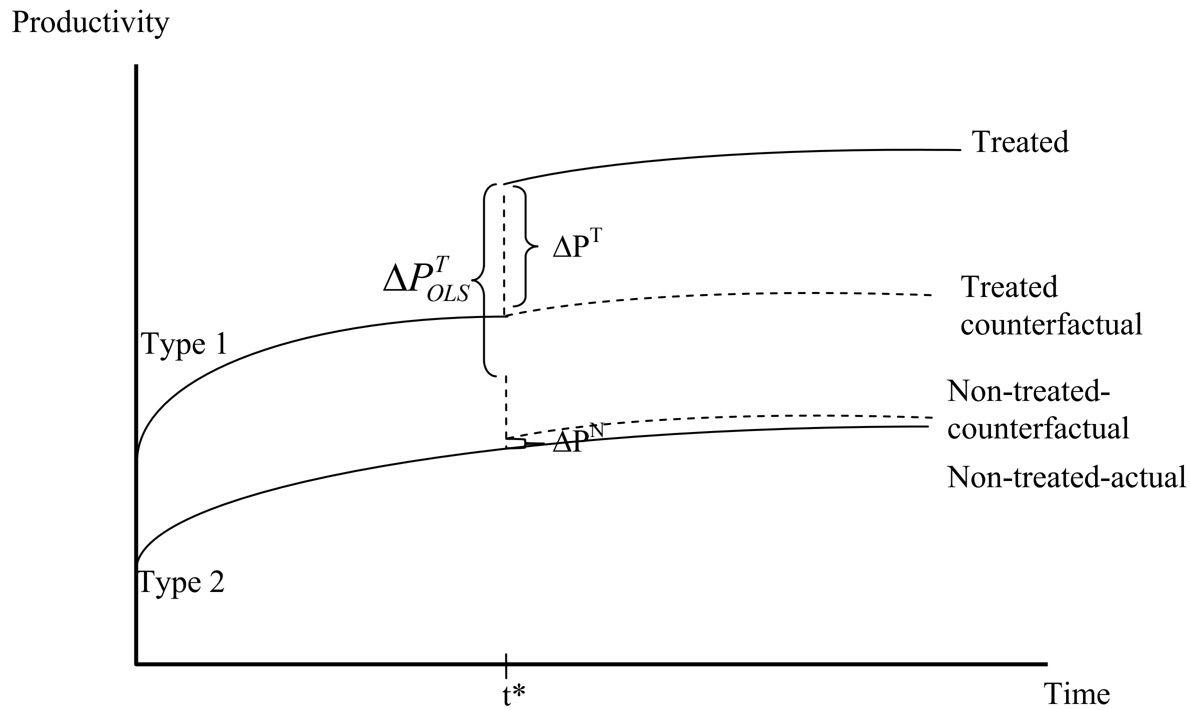
Social Networks and Information Sharing	Performance Outcome	Why?
Kalnias and Chung (2006) Data: 2800 hotels, annual data 1991-1999 on revenues, exit, and entry	Close social ties from same demographic group lowers the failure rate of hotels.	Employees from the same demographic group (Gujarati immigrants) share more information about management practices that work.
Bandiera, Barankay and Rasul (2007) Data: daily worker output, 94 days, picking fruit, one company.	Productivity rises by about 9% when workers are socially connected to managers who are paid incentives.	When the firm switches to incentive pay for managers, overall firm productivity rises, as managers allocate jobs to the most efficient workers rather than to their friends.
Gant, Ichniowski and Shaw (2002) Data: Daily communication ties among workers, 642 workers, 7 steel production lines.	Direct communications for problem-solving is five times greater for workers on production lines with a set of innovative management practices.	Problem-solving activity, as measured by amount of direct communications on operational issues, rises when workers are employed by firms using a set of management practices, like teamwork, incentive pay, and careful hiring.

Hiring Practices	Performance Outcome	Why?
Autor (2001) Data: Workers employed by 1033 temporary help firms in 1994.	The match quality of workers and firms is higher when workers self-sort in response to the firms' offers of general training.	Providing general human capital training includes worker to self sort so the most able are hired by firms offering the greatest amount of general training. (The temporary help firms then sell this worker-quality information to firms).
Autor and Scarborough (2008) Data: 1363 retail stores, one company.	Productivity rises due to switch from informal screening to test-based screening.	Testing yields more productive hires resulting in a 10% increase in workers' average tenure.
Goldin and Rouse (2000) Data: orchestra auditions for 7065 individuals, 508 audition rounds, 14,121 person-rounds	Promotion of women at each round of audtions increases by 50% when auditions are blind, and hiring of women increases substantially.	Better hiring techniques, such as reviewing orchestral auditions with a blind shield, substantially raises the percent of women employed.
Lazear (2000) Data: monthly measures of average daily output for 3000 workers who install car windshields, 19 months, one company.	Productivity rises 22% when the most productive workers are hired and stay at the auto windshield installation firm.	Human resources management practices like incentive pay, cause workers to self select to the firms at which the workers are most productive.

Boundaries of the Firm	Performance Outcome	Why?
Forbes and Lederman (2009) Data: 2 million flights from 72 U.S. airports over 260 days.	Flight delays at the same airport on the same day are significantly shorter for airlines that are vertically integrated – those who have an owned, rather than independent, regional partner with which it connects.	Direct ownership of partner airlines promote more efficient decision making when airlines need to adapt to non-routine conditions.

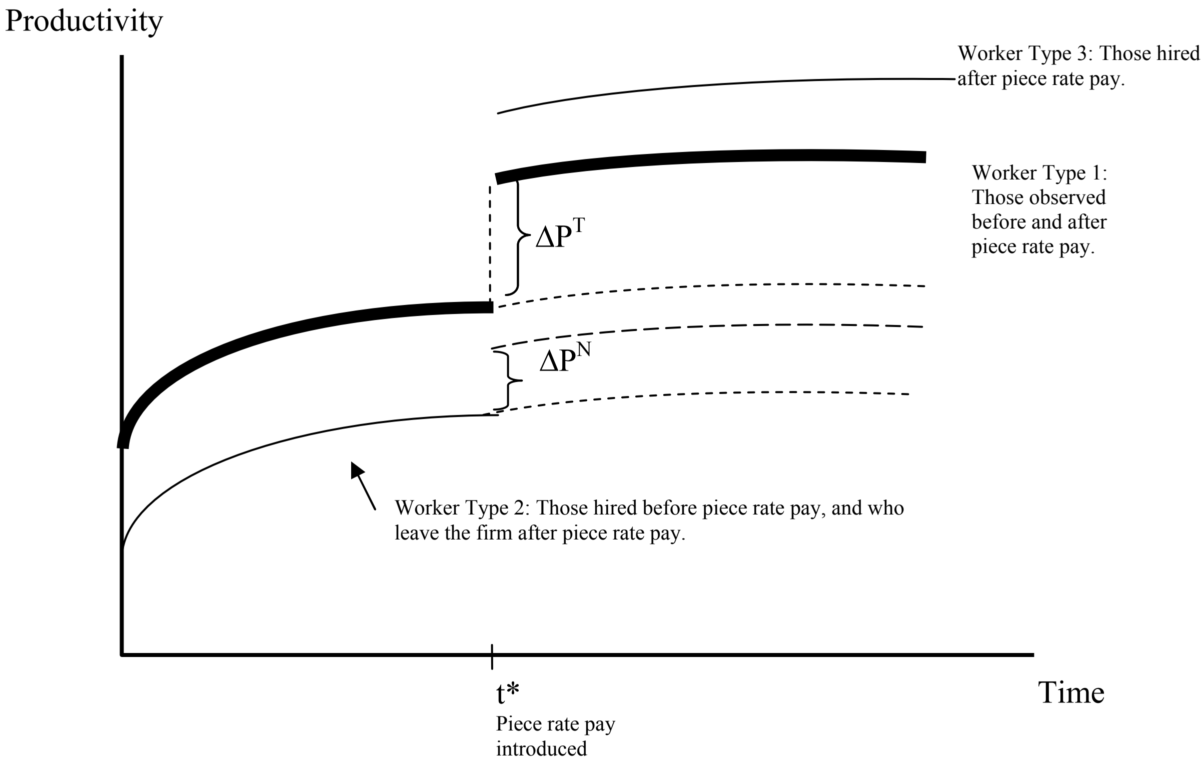
Figure 1:

The Productivity Effects of Adopting a New Management Practice
Hypothetical Age-Productivity Profile for Treated and Non-treated Groups



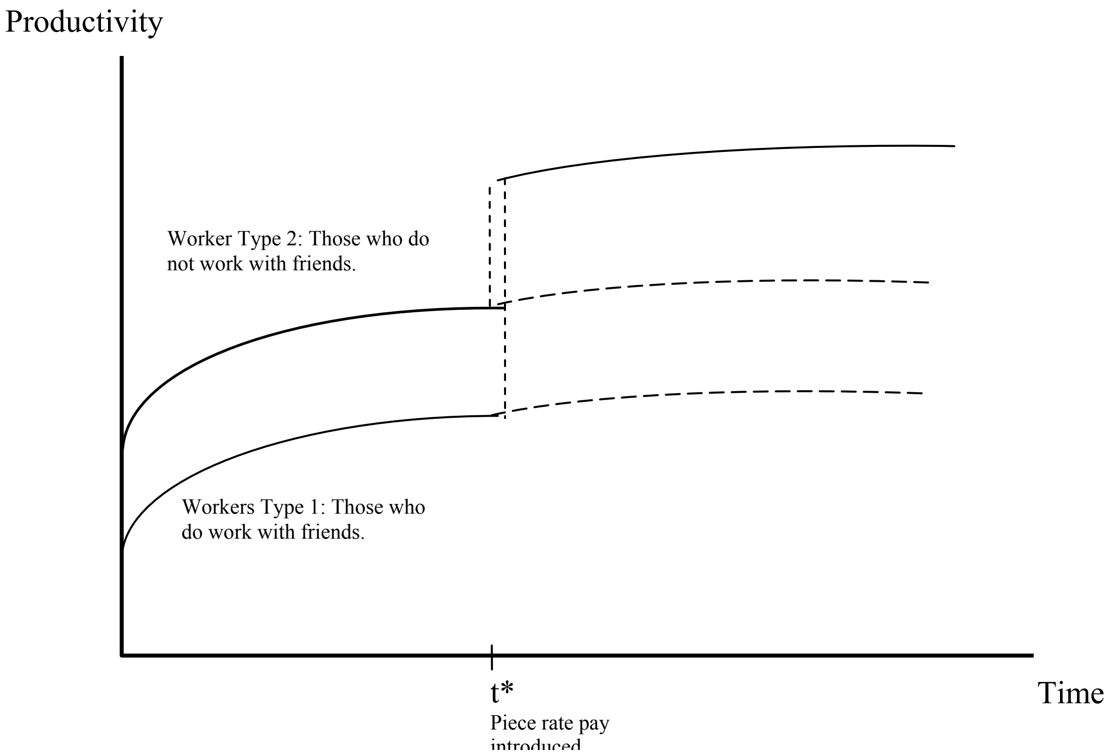
Where t* is the time that the treatment, such as an organizational change, occurs.

Figure 2a:
Incentive Pay Introduced Within One Firm (Safelite)
(Lazear, 2000)



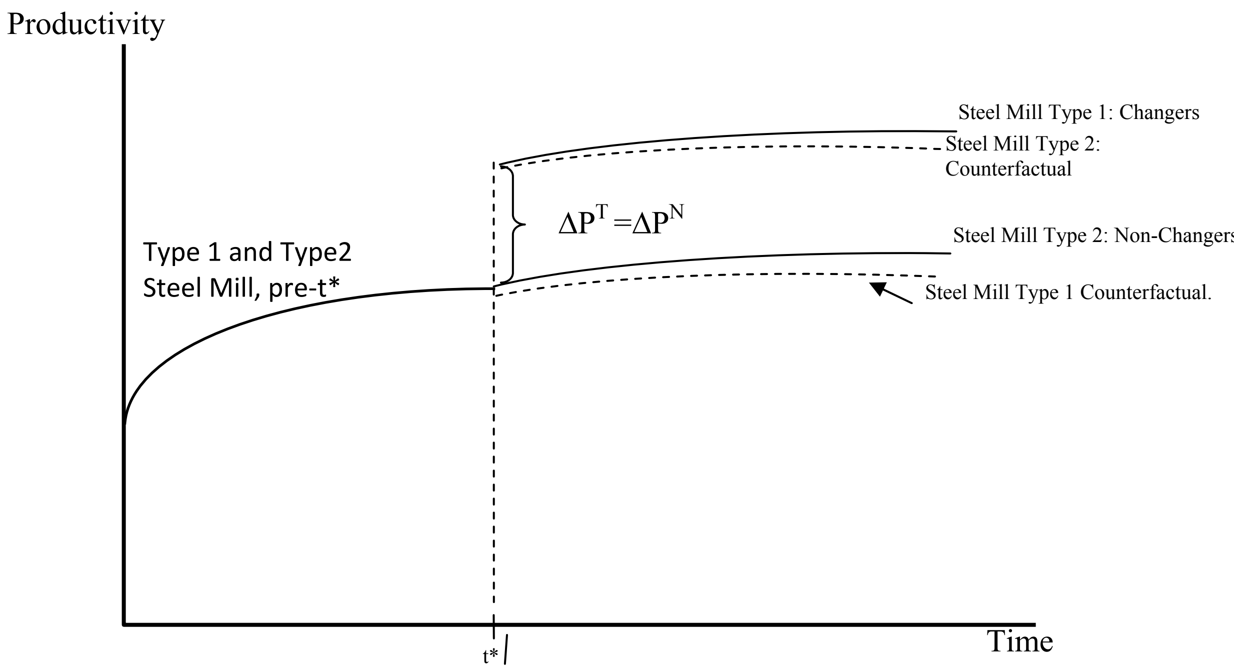
t* is the time of the treatment: the introduction of piece rate pay. The solid lines represent the observed productivity profiles, for the workers Type 1, Type 2, and Type 3. The dotted lines represent unobserved productivity profiles for different worker types.

Figure 2b:
Incentive Pay Introduced Within One Firm (Fruit-pickers)
(Bandiera, Barankay, and Rasul, 2005)



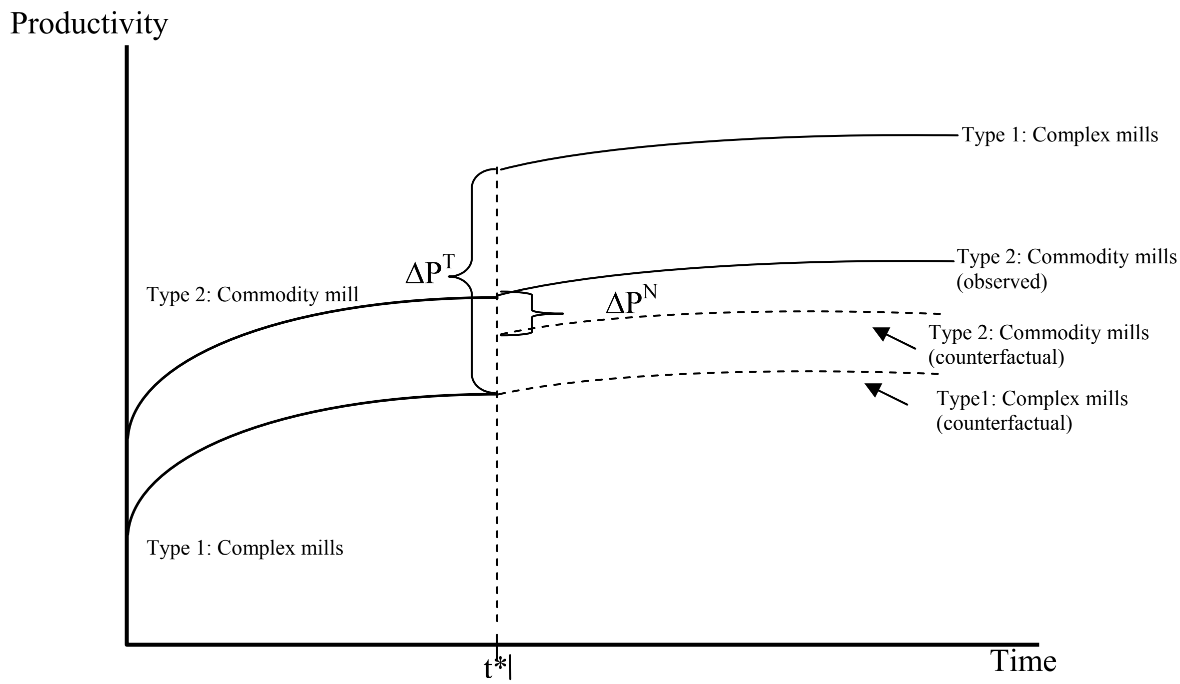
The t* is the time of the treatment: incentive pay plan changes from a relative pay plan before t* to a per unit piece rate plan after t*. Insider information identifies Type 1 workers as those with friends as co-workers; Type 2 are workers without friends as co-workers.
The solid lines represent the actual productivity profiles for workers Type 1 and Type 2; the dashed lines represent the unobserved counterfactuals for these worker types.

Figure 3a:
Innovative HRM Systems Introduced in Steel Finishing Lines
(Ichniowski, Shaw, and Prennushi, 1997)



Note: t* is the date of the treatment: the introduction of new HRM systems. Type 1 lines are changers; Type 2 lines are non-changers. Type 1 and Type 2 lines are identical in their underlying production processes: there is one learning curve for both types. Because these lines are identical, the counterfactual for the changers is equal to the observed age-productivity profiles for the non-changers; the counterfactual for the non-changers is equal to the observed post-t* profile for the changers.

Figure 3b:
Problem Solving Teams Introduced in Steel Minimills
(Boning, Ichniowski and Shaw, 2007)



Note: t* is the date of treatment: the adoption of teams. Type 1 lines make more complex products that therefore have lower levels of productivity pre-t* relative to Type 2 lines that make simpler products. Complex lines are more likely to adopt teams since complex lines have large gains from teamwork and problem solving.