



Insider econometrics: A roadmap with stops along the way[☆]

Kathryn Shaw^{*}

Stanford University, United States
NBER, United States

ARTICLE INFO

Available online 11 September 2009

Keywords:

Personnel economics
Incentive pay
Productivity
Human capital

ABSTRACT

For the past twenty five years, economists have been building theories of the optimal management of firms. For example, economic models suggest that under some conditions, piece rate pay raises performance, and under other conditions, promotions tournaments raise performance. Some of these theories have been tested, others have not. Economists are now using new empirical research tools, that we label “insider econometrics,” to test the impact of management practices on productivity: to model how much productivity changes; to model why management practices raise productivity; and to examine which firms benefit and why from alternative management practices. The methodology we describe is “insider” because it uses inside knowledge and data from within firms. It is “econometrics” because the methodology is often the application of treatment effects methods to modeling changing management practices within firms. However, the methods and challenges of insider econometrics are unique, and we identify several key features that are important in undertaking empirical studies of workers’ productivity. Now that more firms are keeping data on employees, it is time to improve our analysis of the empirical study of the productivity of workers within firms.

© 2009 Elsevier B.V. All rights reserved.

Economists often posit that some form of incentive pay will increase the productivity of workers. But what kind of incentive pay works best to increase productivity? Should firms introduce piece rate pay, versus annual performance bonuses, or higher pay with promotions as in a tournaments model? Should firms adopt more teamwork along with performance pay? If incentive pay raises workers’ performance levels, why does it do so? What are the underlying mechanisms that drive performance improvements?

For the past twenty-five years, economists have been building theories of the optimal management of firms. Some of these theories have been tested; many have not. These theories often have conflicting or ambiguous predictions, and in addition, one wonders, which models are cute theories and which are most relevant models of the theory of the firm?¹

New empirical work also raises new questions. Two empirical regularities stand out. First, researchers have shown that firms that are categorized within very narrowly defined industries have very different productivity levels, and growth rates.² Do these productivity

differences across establishments arise because the firms are making different products, or because the firms are using very different human resource management (HR) practices?³ Second, we see that management practices differ widely across establishments within narrowly defined industries.⁴ Why do firms that appear nearly identical choose very different HR practices?

Anecdotal evidence suggests that part of the explanation is that business products do differ, within very narrowly defined industries, and part is that HR practices differ. Consider Southwest Airlines. They have much higher productivity, as measured by plane turnaround time or by customers per employee, than the traditional carriers, like United Airlines. Southwest Airlines also has very innovative HR practices of teamwork and broad job definitions. Consider Singapore Airlines, which offers high-priced services to business passengers. These different airlines are operating under different product market strategies – some airlines offer low cost, short distance flights for non-business travelers, and others offer high-priced long distance flights for business travelers. Their HR practices should differ to suit these different strategies.

These cross-sectional differences in HR practices have arisen in part because firms have responded differently to the time series “technology shocks in HR practices.”⁵ During these past twenty-five years, there has been a well known and continuous technology shock

[☆] This paper is based on my Adam Smith lecture for the European association of Labour Economists, presented September 2008, Amsterdam. For the comprehensive paper on this topic, see Casey Ichimowski (Columbia University) and Kathryn Shaw, “Insider Econometrics: A Roadmap for Estimating Empirical Models of Organizational Design and Performance,” In Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, forthcoming 2011.

^{*} Stanford University, United States.

E-mail address: Shaw_Kathryn@GSB.Stanford.EDU.

¹ In Lazear’s (1995,1999) President of the Society of Labor Economists speech on the future of personnel economist, he begins by raising exactly these questions.

² In Haltiwanger (2008), the interquartile range in labor productivity across establishments within narrowly defined industries averaged 66 log points; in narrowly defined retail service industry categories, it averaged 57 log points.

³ See Foster, Haltiwanger, and Syverson (2008) showing that product mix affects productivity.

⁴ See Bloom and Van Reenen (2007).

⁵ See Lazear and Shaw (2007) and Shaw (2003a, b, 2002) for these data, and more description and details on the trends in the “technology shocks of HR practices.” For articles emphasizing the productivity consequences of these trends across firms, see Bartel et al. (2007), Bloom and Van Reenen (2007), and Bloom et al. (2009a,b).

in information technologies – the price of computing power has been falling, and as a result firms have invested heavily in new IT technologies, and that some of these technology investments have paid off with higher productivity. Less well known is that managerial practices have seen similar technology shocks. In the 1980s, the Japanese demonstrated that Total Quality Management (TQM) and lean manufacturing could raise productivity and product quality. Manufacturing firms in the U.S. began adopting these practices. Since then, there have been dramatic increases in the use of incentive pay, in teamwork, and in other ‘innovative’ HR practices. About 28% of Fortune 1000 firms used self-directed work teams in 1987; by 1999, about 70% did. About 38% of Fortune 1000 firms used individual incentive pay in 1987; by 1999, about 68% did.⁶ The increased use of incentive pay and teamwork practices is likely to be due in part to greater measurement of workers’ productivity using new information technologies. Productivity has risen over time in industries that are computer-using industries (Jorgenson et al., 2008). But there are new innovative best practices in HR management as well. Some firms have adopted these best practices and some have not. Why have some adopted and some not? What are the productivity differences?

New econometric methods, new data sets, and most importantly, new approaches to empirical research, are being used to address these questions.

In this paper, we describe a new research methodology, that we label “insider econometrics,” that is aimed at testing the impact of human resource practices or other managerial practices on productivity. This methodology is an application of the econometrics of treatment effects to the study of HR practices within firms. The methodology is “insider” based because the insiders within firms, such as the managers or workers, are relied upon heavily for insights or data. It is “econometric” because it relies on obtaining high quality data and using appropriate econometric methods to estimate the impact of practices on performance. Empirical analysis of the determinants of the productivity of firms offers unique challenges, and unique opportunities, that are not found in other applications of treatment effects. This paper describes this methodology.

This paper offers a road map, with stops along the way. The destination is to find a methodology for estimating the impact of human resource practices on productivity. The road to getting there requires stops in which we visit the work of researchers who have very successfully estimated the effects of HR practices on performance. After these visits, the methodology for doing Insider Econometrics takes shape. This methodology is then summarized in a series of Key Features of Insider studies. After describing the Key Features, we outline the steps to take in conducting Insider studies and the potential sources of data.

Insider econometric studies have three elements. First, Insider studies estimate a productivity regression in which the productivity is a function of some “treatment.” The “treatment” is the innovation – and innovations can include new human resource management practices or production technologies or information technologies that a business adopts.⁷ The econometric analysis tests whether or not the given treatment raises the productivity – or some other measure of performance – of workers or establishments in that business.

Second, Insider studies must model the adoption of the treatment; why did the firm adopt the HR practices? It could be that the adoption of the treatment is random – that the researcher finds a ‘natural experiment.’ It could be that the firm conducts an experiment, for example, adopting new incentive pay, in which workers are randomly treated, but the firm who chooses to conduct the experiment is not

random. It could be that the researcher gets data on the ‘Z’ variables that determine why adoption has occurred. Or, it could be that the researcher simply writes about why the adoption of the treatment occurred, based on what the researcher learns from his or her knowledge of the production function and competitive environment, and of what insiders claim, of why adoption has occurred.

Third, the researcher explains why the HR practices increase productivity (if they do). Uncovering why workers behave as they do often introduces new results labeled “behavioral economics.” The growth of empirical behavioral economics has shown that people have interesting “behavioral” responses in the workplace – that is, people care about their status, or people respond to ‘gifts’ of bonuses by working harder even when hard work is not rewarded with higher future pay. Therefore, Insider Econometrics studies very often examine the effectiveness of traditional economic levers, such as incentive pay, but to accurately pinpoint why these practices affect performance, we turn to behavioral economics in the empirical tests and interpretations of results. Behavioral economics also uncovers new economic levers: for example, if status matters, then should firms provide evaluations to employees to clarify the relative rankings of employees?

1. Stops along the way: examples of insider econometrics research

The goal is to estimate the effects of a new human resource practice on productivity. In our stops along the way, we proceed through four examples. Each example displays a different approach to estimating treatment effects within firms (Examples 1 and 2), or across firms within one industry (Examples 3 and 4). Following these examples, the best practices for treatment effects in organizations is provided in the Key Features described in Section 2. For more examples, and in-depth analysis of methods, see Ichniowski and Shaw (2009).

1.1. Example 1: Incentive pay introduced within one firm: Safelite

The human resource management innovation in this study is the adoption of piece rate pay to replace hourly pay within one firm that installs windshields in cars (Lazear, 1999). In this company, each employee has a truck, and drives the truck to the homes of people who have broken car windshields, and the employee installs a new windshield in the car. The production function is individual-specific and measurable: workers do their work alone and install about two to three windshields a day. Because there is no teamwork, and productivity is personal, this is a clean case study of productivity. Lazear collects productivity data provided by the firm. The data is panel data, following about 1100 individual workers for nineteen months before and after the introduction of the incentive pay plan. This is not a balanced sample – workers come and go – and workers’ movement in and out of the firm is key to the results.

After the move to piece rate pay, where workers are paid for the number of windshields installed, the average productivity of workers in the firm rises by 44%. However, when individual workers are followed over time (by controlling for person fixed effects), before and after the piece rate pay, the average worker’s personal productivity gain rises by a lesser 30%. The conclusion is that piece rate pay is especially effective because the pay practice itself produces worker selection: more productive workers (with higher average levels of productivity) are more likely to work for Safelite when they are paid piece rate pay than hourly pay. That is, the workers who arrive after the new incentive pay are more productive, on average, than those who are hired before the incentive pay plan (and those who may quit after it is introduced).

This effect is displayed in Fig. 1. There are three types of workers. Worker Type 1, shown by the bold line, is the average worker whose productivity rises after the incentive pay. Worker Type 2 is the low ability worker, who was hired before incentive pay and then quits when it is introduced. Worker Type 3 is the high-ability worker who is

⁶ The increased use of innovative HR practices, such as teamwork and careful hiring, has also occurred within some very narrowly defined industries, such as valve manufacturing (Bartel et al., 2007).

⁷ We therefore will use the terms organizational innovation and organizational treatment interchangeably for the remainder of the paper.

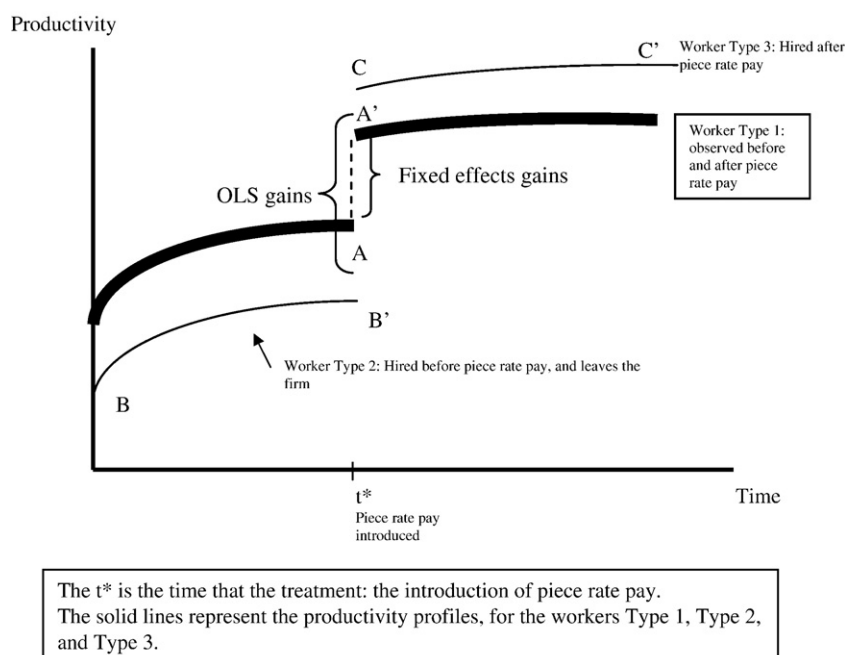


Fig. 1. Example 1 – incentive pay introduced within one firm (Safelite) (Lazear, 2000).

hired after the pay plan. The total performance gain is the sum of the within person gain for Type 1 worker, and the across person gain in hiring Type 3 and losing Type 2.

In equation form, productivity is as follows:

$$y_{it} = \beta I_{it} + \gamma_i x_{it} + \alpha_i + u_{it} \quad (1)$$

where y_{it} is productivity, i is the person, t is time, I_{it} is a dummy variable equal to one when the treatment is in place in the plant, x_{it} is the set of production function control variables, α_i is the person fixed effect, and u_{it} is the error term. It is important that we control for person fixed effects to measure the within-person productivity gains from incentive pay. The key is that people sort after the pay plan is introduced. Those with low α_i sort away from Safelite (Shaw and Lazear, 2008).

The fundamental feature of this paper, and all papers that successfully look inside just one firm, is that there are *differences in workers' responses* to the incentive pay. That is, there is *heterogeneity* in the response of workers to the treatment, or to the incentive pay: the workers who stay at the firm after the adoption of incentive pay, or who are hired after incentive pay, have higher average levels of productivity, α_i . It is this heterogeneity that makes the paper interesting – heterogeneity in workers' abilities and therefore in workers' responses to the pay plan – that give the paper its clever result that workers sort.

Only one firm is adopting a new practice – incentive pay – so why is this particular firm interesting? Readers of the paper are not interested in windshield installation. Readers are also not overwhelmingly interested in whether piece rate pay raises productivity in this one firm for this type of job, though the big size of the productivity gain is remarkable. What makes this paper really interesting is that the author shows *why* productivity goes up, and the reason for the performance gain is theoretically important. Productivity goes up because after the firm adopts incentive pay, workers raise their effort, but in addition, the workers who are the most productive are those who self select to work in the company. This result feels *intuitively generalizable because the paper makes a fundamental economic point*: HR practices are signals that cause sorting. The key point for economic theory is that the firm's selection of its optimal HR practices acts as a signaling device to sort workers into the firm. This is likely to be true for many different HR practices, not just incentive pay, and for many different firms, not just Safelite.

Is the treatment exogenous or random? It is exogenous to the individual worker. It is not exogenous to the firm: the firm chooses incentive pay because they hoped it would raise productivity. In later examples, we look at the why some firms adopt new HR practices and some don't, and the impact of adoption across firms.⁸

1.2. Example 2: Incentive pay introduced within one firm: fruit pickers

The human resource management experiment here is also the change in an incentive pay plan (Bandiera et al., 2005a,b). In this case, the job is picking fruit, which is done by a workforce that is hired only during the summer fruit season. The firm undertakes a series of experiments of changes in payment schemes.⁹ In one experiment, the firm changes the incentive pay of the fruit pickers, introducing piece rate pay, as in Safelite. In the first half of the picking season, the workers are paid based on relative output: 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 (but keeping the mean pay fixed). In the second half of the season, the firm switches to piece rate pay, in which the payment per unit of output is fixed ex ante and does not vary as a function of how co-workers perform. The data is panel data on 142 workers covered for 108 days.

After the move to piece rate pay, the average productivity of workers in the firm rises by about 58%. Nearly all workers increase their effort, but also, the variance of output rises markedly – there is a jump in the number of high performers. Once again, the readers of this paper are not particularly intrigued by either the occupation of fruit picking or this specific case of piece rate pay. The reason *why* productivity rises is interesting. Under the relative payment scheme, workers are holding back their effort because they don't wish to harm their friends. This evidence comes from data the authors link to the

⁸ Even though Lazear does not address the treatment by the firm, there are clear conclusions about which firms should adopt this treatment, of piece rate pay. The firms that should adopt the treatment are those that would benefit the most from the combination of the effort effect and the selection effect that the HR practice induces. There is also complementarity between HR practices and information technology: the firm adopts this treatment because they purchase new IT that makes measurement of productivity cheaper – in that sense, adoption is due to a technology shock.

⁹ For papers on their other experiments, on managerial pay or on teamwork, see Bandiera et al. (forthcoming, 2009a, 2009b, 2007).

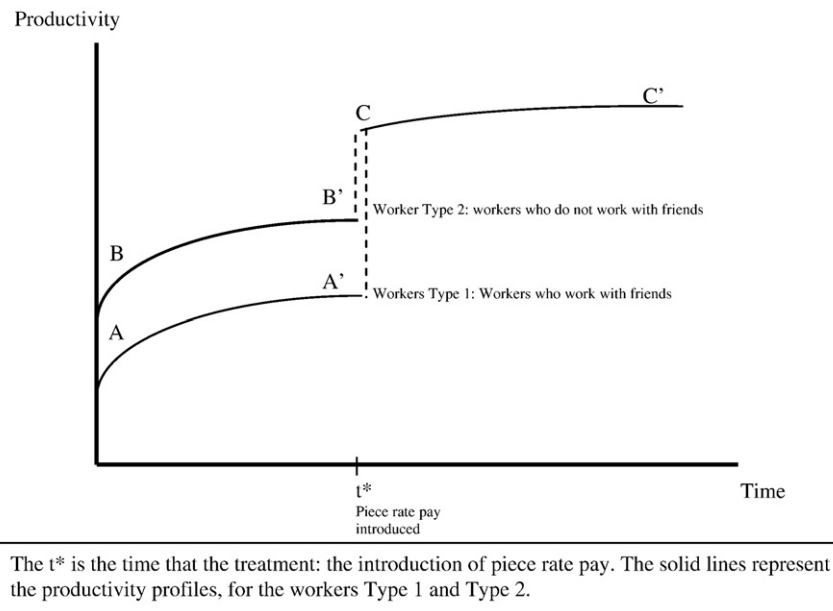


Fig. 2. Example 2 – incentive pay introduced within one firm (fruit pickers) (Bandiera et al., 2005a,b).

performance data. The authors survey the workers and ask about who their friends are. Under the original relative payment scheme, workers withhold their effort the most when friends are present in their work group, because the worker's personal high effort lowers the pay of their friends. (Under the relative scheme, average pay for the day is fixed, so that if you earn more, your friends earn less.) However, after piece rate pay is introduced, this effect goes away, because the worker's personal effort does not affect the pay of their friends. So, everyone works harder with piece rate pay.

This effect is displayed in Fig. 2. There are two types of workers. Worker Type 1 has productivity jump of $A'C$ – they work with friends and have the biggest gain due to piece rates. Worker Type 2 has productivity jump of $B'C$ – they do not work with friends and thus have a lower productivity gain after incentives. As in Safelite, there is *heterogeneity* in responses, of Eq. (1), and the researcher estimates the average response, and estimates the variance in the responses to explain why the incentive pay plan works.

In equation form, there is now heterogeneity in the response to the incentive pay, or the production function is

$$y_{it} = \beta_i I_{it} + \gamma_i X_{it} + \alpha_i + u_{it} \quad (2)$$

where β_i now varies across workers. Worker group Type 1 has a large productivity response to incentive pay; worker group Type 2 has a smaller response.¹⁰

This paper is interesting because the point of the paper feels *intuitively generalizable: the paper aims to make a fundamental economic point*. The conclusion is that workers internalize the social externality – the externality that their effort harms others – during the relative payment scheme. Social connections shape productivity. Social connections shape the response to the incentive payment scheme.

1.3. Example 3: Human resource innovation introduced within one industry: steel integrated mills

The human resource management innovation in this study is the adoption of bundles of new human resource practices (Ichniowski and Shaw, forthcoming). Within integrated steel mills, one area of the mill

is devoted to steel finishing lines that take a coil of very thin steel sheet and unwind the coil and treat it or coat it, for use in products like auto body exteriors. Workers can improve the productivity of this mill, by watching for defects, such as bubbles in the surface of the steel coating, and correcting the problem before the line must be shut down when low quality steel is produced. Shutting down the line reduces productivity. The panel data the authors collect is monthly productivity data for the teams that run the finishing line area of the mill, for 36 finishing mills owned by 17 companies, followed over five years.

The HR innovation is the adoption of new high performance HR practices that are increasingly typical of manufacturing. The researchers identify four “systems” of HR practices, ranging from the traditional HR system to the most innovative HR system. The most innovative high performance system has high marks on all seven dimensions of HR – new incentive pay, careful hiring, teams, training, information sharing, broad job design, and employment security. In the progression of HR practices, the “traditional system” has none of these practices (System 4), the next system has more information sharing (System 3), the next one has more teamwork (System 2), and the top innovative-HR system adds better incentive pay and careful hiring on top of these (System 1).

Note that this paper models team-based productivity; the two papers on Safelite and fruit pickers above modeled individual worker's productivity. In steel mills, individual productivity cannot be measured objectively (though it can be measured subjectively by supervisors). Also, the authors are measuring productivity for one type of team within the mills – the team that operates the finishing line – not the entire firm.

As the mills move from the traditional HR system to the most innovative HR system, productivity rises substantially. What makes this Insider study interesting? Once again, readers of the paper are not interested in this particular product, steel production. The authors show that high performance HR innovations can produce very big productivity gains in manufacturing. However, what makes this result interesting is that the author shows *why* productivity goes up. It goes up because when a firm adopts a set of complementary HR practices, all these practices together significantly change workers' behavior. Firms that adopt a single HR practice in isolation, such as only teamwork, will see very weak gains (if any) compared to adopting the set of HR practices.

The authors also provide, in separate papers, further evidence on *why* productivity rises. Productivity rises because the steel mill operators become problem solvers. In the innovative-HR mills, the workers form social networks in which they communicate much more heavily with

¹⁰ The heterogeneity in the response of the workers to incentive pay, or variation in β_i , may have also been present in the Safelite data and contributed to workers' sorting.

each other to solve operating problems. The firm is thus investing in HR practices that build “connective capital” – practices that build connections among workers to enhance workers’ problem solving (Gant et al., 2002; Ichniowski and Shaw, 2007). Introducing incentive pay alone does not change productivity – it gives workers the incentive to increase output, but other HR practices give the workers the opportunity (through teamwork and information sharing) and the ability (with training and highly-skilled new workers) to increase productivity.

Thus, the conclusion feels *intuitively generalizable because it aims to make a fundamental economic point*: firms need to think about the right bundle of HR practices that fits the firm’s strategy, not just the individual HR practices. Given the innovative-HR bundle, production line operators become problem solvers, replacing the engineers as sole problem solvers.

Is the treatment exogenous or random? In the two examples above – Safelite and fruit pickers – the treatment was exogenous to the worker, but not exogenous to the firm. In the case of this steel study, the authors claim the treatment is exogenous to the productivity regression (1), though not to the firm’s profits. The logic is as follows. Starting in the 1980s, informed people within steel firms realized that steel mills abroad (primarily in Japan) were raising productivity by having workers undertake problem solving. Some mills adopted the new practices; some didn’t, or made more gradual changes. The mills that adopted the new practices were the greenfield (newly built) mills, or the reconstituted mills (with new owners but old capital), that could hire new workers while adopting new HR practices. The impediment to change in old mills was the transition costs: all mills have expected higher productivity from new HR, but some mills didn’t adopt because their higher transition costs lowered the profitability of adopting.¹¹ In these very homogeneous production lines, the adoption was exogenous to productivity gains. There is homogeneity in the production function across mills, but heterogeneity in the adoption, I_{it} , because mills that could not change their workforce and hire new employees had larger transition costs.

1.4. Example 4: Teams introduced within one industry; steel minimills

The HR innovation here is the adoption of problem solving teams in steel minimills (Boning et al., 2007). Steel minimills produce steel bar products, like re-bar used in construction and highways and large steel beams used in constructing buildings. The production process is very similar across mills – the mills melt steel scrap (like auto bodies) and cast it into the bar products. It is important to visualize the production process: hot steel bars are processed through rolling mills stands that shape the bars into different forms (round or I-bar, etc.) and elongate the bars. In these mills, the workers use their 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 productivity data on this one type of product line within the minimills. The data is panel data; monthly productivity data for 36 steel mills owned by 20 companies followed over five years. In the beginning of the data, about 10% of the mills had teams; by the end, about half of the mills had teams.¹²

After the move to teamwork, in which operators join together to solve production problems, productivity rises by about 6%.¹³ Why does

productivity go up? As in the fruit picker studies, in this study, there is heterogeneity in the effect of the innovation on productivity: the β_i varies across the mills, as in Eq. (2). This is displayed in Fig. 3. The mills are categorized into two groups: mill Type 1, produces the most complex products; mill Type 2 produces less complex products. The adoption occurs at time t^* . The Type 1 firms that adopt teams see their productivity jump by ΔY^T when teams are adopted: the top bold line is the time-path of productivity for the “treated group.” The Type 2 plants that do not adopt teams see their productivity grow over time due to normal learning: the lower bold line is the productivity growth as a function time t for the non-adoptors or “non-treated group.” For these mills, their potential gains to adoption (i.e., the unobserved counterfactual) are lower (as displayed by the dashed line), so they rarely adopt.

The empirical results reflect the productivity gains displayed in Fig. 3. First, mills are sorting. The mills that are expected to have the highest return to teamwork, or the Type 1 mills making complex products, are much more likely to adopt teams. The complex mills adopt 78% of the time, and the not-complex mills (producing commodity products) adopt 28% of the time. Second, the mills producing the complex product have higher productivity gains from teamwork than those producing the commodity product.¹⁴

So what makes this Insider study interesting? Readers of the paper are not especially interested in steel production. The aim of the paper is to make a fundamental economic point: the optimal HR practices follow the strategy of the firm. Moreover, the gains to teamwork are selective: complex production lines gain the most; commodity lines do not benefit from problem solving. Even though teamwork is considered a new “best practice” for manufacturing, not all firms gain equally; not all firms should adopt this best practice. *Even though these points apply to steel, they feel intuitively generalizable across industries.*¹⁵

This paper aims to estimate the “treatment of the treated” effect of on productivity. For Insider methodology, when studying the performance of firms, it is often more important to estimate the treatment of the treated effect than the average treatment effect. In the econometrics literature, the average treatment effect is the effect obtained from random adoption. In firms, we do not want firms to adopt HR practices randomly; we want them to adopt them optimally. Learning about the process of non-random adoption is as important as learning about the productivity gains. This is true for Safelite – where individual workers responded differently to HR – and true for minimills – where mills respond differently.

2. Key features to consider when doing insider econometrics

“Great advances have been made in theory and in econometric techniques, but these will be wasted unless they are applied to the right data.” (Griliches, 1994)

The aim, or destination, for Insider Econometrics is to combine the “right data” with new econometric techniques to model when, why, and how much HR practices raise productivity. Firms choose the HR practices. Researchers choose their research questions, choose their data sets, and choose their econometric methods for evaluating the effects of the HR practices. Therefore, the Insider Econometric methodology must guide researchers in making these decisions, or guide readers as they evaluate the insider work that has been done.

¹¹ See Ichniowski and Shaw (1997) for the full analysis of why some mills adopt the innovative HR practices and some do not.

¹² The teams are team-based systems – other supporting practices, like enhanced training and careful hiring, support the teams. This data is quite different from that used in the example above – the finishing lines in example 3 are located within integrated steel mills (that make steel from raw materials); these bar-product rolling mills are located within minimills (that make steel from melted cars). The integrated mills produce the flat-rolled product for one type of consumer; the minimill bar mills produce their products for a range of heterogeneous consumers. Both studies obtain data for just that area of the mill in which one team runs one production process.

¹³ The dollar value of the gain is large because a percentage point gain implies a large dollar return in a very large fixed capital plant, such as a steel mill.

¹⁴ Note that the commodity mills made have made a ‘mistake’ in adopting, and thus we have the information on the productivity outcome. If they followed strategy perfectly, they would never adopt and no gains could be compared. Alternatively, if the transition costs of adopting teams are smaller among these commodity mills, it may be profitable to adopt teams.

¹⁵ 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. across all firms (Bartel, Ichniowski, and Shaw, 2007).

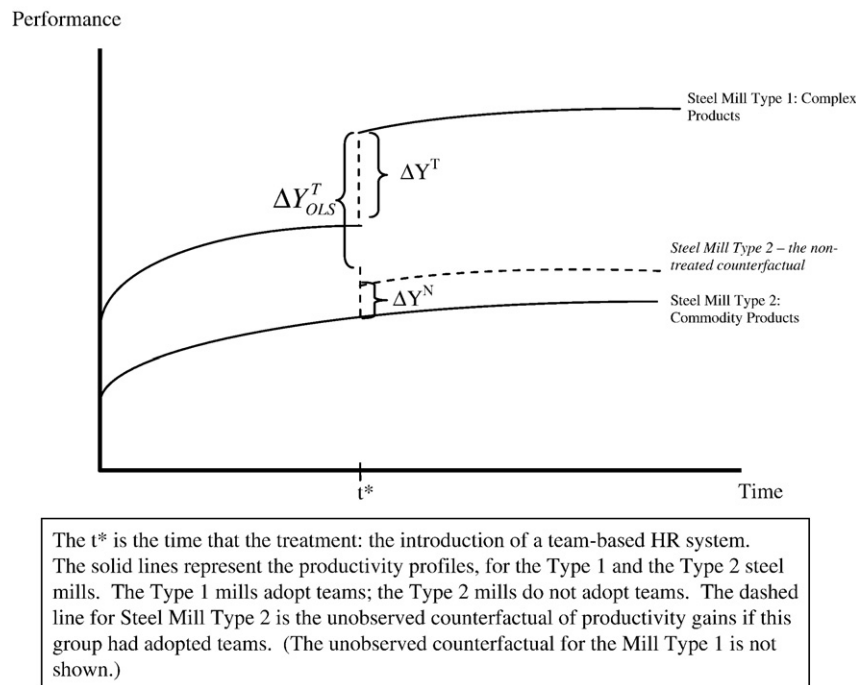


Fig. 3. Example 4 – teams introduced within one industry (steel minimills) (Boning et al., 2007).

The four steps along the way were chosen to illuminate the Key Features. These four Key Features are the basic rules of thumb for conducting an Insider study.

2.1. Key Feature 1: Find a treatment

Exceptionally well run companies, that appear to be doing everything right, may not be useful to the researcher doing Insider Econometrics. Business school cases can be written about these companies, describing what the insiders feel is working well relative to their experiences in other companies. But often there is no variation in the treatment – or in the HR practices – within the firm, because the firm does not change its HR practices over time. There must be variance in the treatment I_{it} either across individual i , where i is firms or workers, or across time t . The best studies are panel studies that have variation across people and time to do difference-in-differences analysis. Many studies also have variation in β_i , or in the response to the treatment. The advice, “find a treatment,” seems obvious. But the points below show that treatments can come from many sources. There are not natural experiments in the firm; firms rarely randomly adopt new HR practices.¹⁶ Insider Econometrics aims to learn about why firms adopt, and the impact of adoption, from a wide variety of different types of treatments, or HR practices. Different types of treatments are described further in Key Feature 3 below.

2.2. Key Feature 2: Test a hypothesis that is generalizable: model fundamental economic behavior

As a result of the rapid rise in the use of computerized performance-measurement software in companies, there are a large number of companies that possess worker productivity data, but not nearly as many companies ought to be the focus of an Insider study. For

example, thousands of firms have introduced problem solving teams since 1980, but very few of these firms would be valuable source of data for new studies.¹⁷ What makes an Insider study valuable research?

An Insider study is most successful when the reader believes the study is *generalizable* because it models *some fundamental economic behavior*.¹⁸ The research papers displayed above aim to make generalizable points: worker productivity rises in Safelite because workers sort themselves into the firm in response to the new incentive pay; worker productivity rises for fruit pickers because the workers stop withholding their effort when new piece rate pay is introduced; worker productivity rises in steel mills because steel mills adopt innovative HR practices that combine incentive pay with complementary practices, such as teamwork and training; worker productivity rises in minimills when teamwork practices are introduced to complement the firm's product market strategy, aimed at producing complex products. These are fundamental economic points: HR practices serve as important signaling devices to potential workers; social influences affect responses to incentives; HR practices are complements; product market strategy drives optimal HR strategy. In these papers, the researcher shows *why* the HR practice is effective. The logic of the model then persuades the reader that these points may be relevant in many firms, and thus are generalizable points. Evidence that the points are generalizable is not offered. These micro studies cannot offer such evidence.

2.3. Key Feature 3: Design your ‘experiment’ to reflect the tradeoff between homogeneity and heterogeneity

Insider Econometric studies have a fundamental flaw – the researcher and the reader never know whether the conclusion from each study is generalizable beyond the specific company or industry.

¹⁶ Firms will increasingly experiment in the future, as they have more data on productivity of employees. The fruit-pickers studies show examples of such experiments (Bandiera et al., 2005, 2007, 2009a,b, forthcoming). Still, these experiments are not random. There may be random shocks that induce HR changes. There are also instances of random sorting of peers (Mas and Moretti (2009)), and of union strike or bargaining outcomes that alter productivity (Mas, 2004, 2006, 2008).

¹⁷ Ideally, researchers would model the introduction of teams in every firm, and then take the average team effect to be the economy wide effect that could be used by policymakers to determine why productivity rises over time in the U.S. economy. However, in practice, so many studies are too costly to undertake.

¹⁸ Within economics, the field of experimental economics follows this practice. Experiments are done using students in lab settings, assuming that all people would behave as students do if encountering similar settings in the workplace.

It would be much nicer to estimate production functions across all industries, and therefore conclude how much and when incentive pay increases performance across many different industries. Such cross-industry studies have two drawbacks. First, the production function is likely to suffer from omitted variable bias. When estimating a production function across industries, the treatment or HR practice is likely to be correlated with many possible omitted variables. By estimating a production function within a firm or within an industry, omitted variable bias is reduced. The researcher can include the relevant variables that influence productivity. And, the researcher can explain *why* the key HR practice should increase performance, and why other possible omitted variables are irrelevant or uncorrelated with the HR practice. Second, the adoption of the treatment, or the HR practice, is hard to model across quite different industries. Across industries, there are many possible influences affecting the adoption of something like incentive pay. Thus, the homogeneity of the production process has advantages.

A general rule of thumb is that the more homogeneous the data set, the more persuasive is the researcher's claim that he is estimating one common production function across all the 'i' units in the data set. The researcher is better able to obtain control variables that are specific to that production function, and to reduce the likelihood of omitted variable bias.

Striving for homogeneity in production units is, however, also a drawback in Insider studies, because there must also be heterogeneity in the treatment variable, I_{it} , across the production units to estimate the β coefficient on I_{it} in regression (1). If every worker or plant were truly identical to every other worker or plant in the data set, then the adoption of the treatment would also be identical across the units. There would be no variation in the data. In the four examples above, there is heterogeneity in the treatment, (in I_{it}), or heterogeneity in the response (in β_i), across the i -units that are workers or teams or plants. People or plants cannot be exactly identical in every way.

Thus, in designing an Insider study, the researcher faces a tradeoff between homogeneity and heterogeneity, because the researcher seeks homogeneity in the production units for the estimation of the production function, but heterogeneity in the treatment variable, I_{it} , or in the response to the treatment. Given a very homogeneous production function, the researcher must ask, why are some observations covered by the treatment and other observations are not? Overall, the goal is to strike a purposeful balance between homogeneity in the production function, and heterogeneity in the treatment effect.¹⁹

2.4. Key Feature 4: Model why the treatment was adopted

In Insider econometrics, we are modeling the introduction of a new HR practice. That's the treatment. Therefore, we must ask, what changed? Why is the firm introducing the HR practice now and not previously? There are three basic reasons why firms are adopting new HR practices.

First, firms introduce new HR practices because there are exogenous technology shocks to the economy. Falling computer prices over time are a clear 'technology shock' that has caused HR practices to change. As computer prices fall, firms purchase new software that stores data used for managing people and machines. Therefore, the firms' HR practices change. Firms are more likely to introduce pay for performance when the firm can cheaply measure and store data on performance of workers.

There are also "HR technology shocks" over the last thirty years. In the early 1980s, most manufacturing firms did not use incentive pay, did not use teamwork, and did not involve their production line employees in problem solving. Engineers solved production line

problems, and workers came to work to operate the machinery. This changed when the Japanese showed that involving workers in problem solving, via Kaizen processes, could raise productivity and product quality. Manufacturing firms in the U.S. saw this 'shock' to management practices and began to change their HR practices (These were described above in the Introduction.). As described in Bloom and Van Reenen (2007), there are new 'best practices' in manufacturing and elsewhere. Firms are gradually adopting these new best practices.

In fact, there are a range of different types of technology shocks that result in the adoption of new management technology. Falling computer prices from scientific innovations induce information technology treatments; falling trade barriers across countries produce falling relative prices and import competition²⁰; new managerial technologies arose from the Japanese and others introduction of quality improvement processes, lean management, teamwork, or performance pay.²¹ These factors then interact: greater IT investment means more measurement of productivity and greater use of incentive pay.²²

The second reason why firms are adopting new HR practices is that firms change their product market strategies. For example, a firm may change from producing a commodity-type product to producing a customized product. In some sense, this is a cross-sectional change: firms look around at other firms like them in their industry, and make changes in their product market strategy to adjust. These are cross-sectional within industry adjustments, as firms react to others in their industry. The technology shocks above are time series adjustments: there are time series changes in technology that cause firms to change their HR practices. Of course, even within industries, firms are changing their product market strategy because there are, again, shocks to their environment: new competitors enter (as in competition from China); or new products are developed; or new production processes are developed.²³ When firms change their product market strategy, their optimal HR strategy changes.

Third, firms may decide to experiment with new HR practices, because they have not yet found their optimal practices, or because their internal conditions are changing. For example, stock options may no longer be paying off, or an aging workforce may need different incentives. Often, however, the desire to experiment is motivated by outside shocks – by the change in best practices that are summarized in our first two reasons.

Given this backdrop, let's revisit the sources of data displayed in the examples above.

Data across workers within one firm is displayed above, and in many more papers referenced below. The most widely cited studies using data from one firm estimate the effects of an HR practice on productivity, and introduce a key second element in their models – these studies always focus on *heterogeneity in the workers' responses to the treatment*. In several studies reviewed here, researchers model how and why workers' responses differ when new incentive pay or new teams are added in the firm. In response to new incentive pay: some workers leave the firm and some stay and become more productive (Lazear, 2000); some workers begin working harder in response to incentives when their friends are not in their work group, etc. (Bandiera et al., 2005a,b, 2007). In the field of behavioral economics, the researcher may care less about the specific HR practice, but more about documenting interesting human behavior

²⁰ See Bloom, Draco, and Van Reenen (2009) for empirical evidence that reduced trade barriers increased the adoption of innovative HR practices.

²¹ See Bloom and Van Reenen (2007) for models of the adoption and effects of lean management and similar practices on organizational performance in manufacturing. See Lazear and Shaw (2007), Shaw (2003a,b, 2002) for the descriptive evidence that HR managerial innovations are technology shocks to management practices.

²² For emphasis on the IT and HR interactions, especially incentive pay, see Acemoglu (2003), Lemieux et al. (2009); Lazear and Shaw (2007).

²³ See Bloom, Schankerman and Van Reenen (2007) for evidence on spillover effects between firms as a result of new technologies versus new competitors.

¹⁹ This point has been made for years by scientists as they design their lab experiments.

that impacts performance and happens to interact with HR policies.²⁴

In these studies, the treatment is exogenous to the individual workers within the firm. Thus, the treatment estimation reflects an exogenous natural experiment across workers. The treatment is not exogenous to the firm; firms introduce a new HR practice, like incentive pay, in response to either the new technology shocks, or in response to what their competitors are doing. In the case of Safelite, the firm shifted to piece rate pay after they bought new computers that kept track of workers' daily productivity.²⁵

In some cases, there is no new HR treatment within the firm, but identification of the treatment effect comes from the random movement of workers across different work areas within the firm. Mas and Moretti (2009) follow workers as they change shifts randomly within a grocery store. They show that when a worker joins a new peer group with more top performer, the worker's own productivity rises. That is, there is strong peer group influence on productivity. It is the worker's random movement that identifies the peer group effect, not the firm's change in HR policies.

With the decline in unionization in the U.S., we tend not to model union effects on productivity. However, strikes, or unanticipated contract negotiations outcomes, can have significant productivity effects and have implications for other HR practices (Mas, 2004, 2006).²⁶

Data across work areas within one firm models the productivity of workers across teams (or work areas) within one firm, such as different retail stores owned by one firm (Lafontaine and Srinivasan, 2009; Griffith and Neely, 2007), or different teams within one firm (Hamilton et al., 2003). In these studies, individual productivity is not measured; the team is the relevant work unit. The estimation issues are the same as those using data across firms. However, these studies may also have worker-level data available to augment the team-level data (Hamilton et al., 2003).

Data across establishments or firms is a primary source of the identification of the effect of an organizational innovation on productivity. Usually, the researcher has panel data, or data across establishments over time, and the treatment is identified from the introduction of new HR practices over time across firms. Then, the researcher must also model why the HR practices were adopted – it could be due to an exogenous shock such as falling computer prices. Falling computer prices or HR technology shocks can produce a quasi-natural experiment across firms, adoption is not econometrically modeled. In the case of integrated steel mills above, the adoption is a function of the costs of transition which are exogenous to the productivity regression.

These papers tend to do before-after analysis, using the time series element of the data to estimate the productivity regressions using some form of fixed effects regressions. It is also possible to identify a treatment effect of HR on productivity using cross-sectional data, across firms or establishments. In the integrated steel finishing lines (Stop 3 above), the treatment is identified from the introduction of new HR practices among some firms and not among other firms. If the production functions are extremely homogeneous across plants in the cross-section, the treatment effect can be estimated from the production function equation.²⁷

²⁴ In these studies, the treatment is not exogenous to the firm or worker, but the behavioral response is modeled. The behavioral economics literature that uses data from firms is very much in the spirit of insider econometrics, but is too voluminous to cite here.

²⁵ A famous case of changes in HR practices due to new IT use is Harrah's ties workers' pay to customer evaluations that they got through their marketing and investments (HBS 9-403-008).

²⁶ The original insider econometrics work was done by Kim Clark (1980a,b, 1984) to model the effects of unionization on productivity.

²⁷ The reason for using cross-sectional data is that there is often more variation in HR practices across plants than within plants over time: firms don't change their HR practices very often over time. If HR practices vary across plants due to exogenous historical accident or transition costs and production functions are very similar, the cross-sectional variation is valuable. Most often, researchers display both their cross-sectional regression results and panel fixed effect regression results, to display the different identification strategies (Ichniowski et al., 1997).

3. Three additional corollaries

The four Key Features above are the most striking features of Insider studies. However, there are several corollaries that describe useful, but not required, features of Insider studies.

3.1. Corollary 1: It is always nice to get more data to prove your point

Increasingly, empirical papers use more than one data set. This is especially true for Insider Econometric studies. The first goal of an Insider study is to determine how much an innovative HR practice improves productivity. After that, the researchers wish to show why the HR practice improves productivity. For fruit pickers and steel mills, the researchers surveyed the workers to display the nature of their social networks as the cause of the productivity gains.²⁸ Alternatively, the researcher can link multiple data sets from the firm – such as linking the firm's data set on the productivity of the workers with the firm's dataset on the characteristics of the workers.²⁹

Next, researchers want to know why some firms adopt innovative practices and some don't. In the steel mill studies above, the researchers gathered data on the characteristics that would affect adoption. Many other papers do this as well.³⁰ Alternatively, a separate survey may be done by the researchers. Bloom and Van Reenen (2007) survey medium sized manufacturing firms to obtain very extensive data on the practices that firms have and the reasons they may have adopted these practices. Their own survey data is then linked to the government data on productivity.³¹

3.2. Corollary 2: Add descriptive evidence from insiders

Interviews with insiders are used in two ways in these studies. First, insiders help develop the key testable hypotheses. Insiders often have views on whether teams increase performance, but they don't have the perspective or the data to test their hypothesis. Second, insiders should be used to interpret the results, and then re-estimate the hypotheses if needed. This qualitative information should be added to the research paper. After the empirical results are presented, the reader wants to understand if there is likely to be an omitted variable bias or an endogeneity bias. A good story can go a long way in reassuring the reader that the estimated model is a good way of interpreting the reality of the firm. Has the researcher found "truth?"

3.3. Corollary 3: Report likely complementarity between HR practices

It is often true that HR practices are complements, or are adopted in bundles. In the case of information technology purchases, more data in-house often means more detailed incentive plans. While the researcher may be focusing on one HR practice, it is useful to report other practices that were adopted in unison or prior to the HR practice of interest.³²

²⁸ For Safelite, the reason why productivity rises is evident in the firm's data set, containing data not just on workers' productivity, but on the hiring and exit of workers that produces the results. For steel, see Gant, Ichniowski, and Shaw (2002).

²⁹ Griffith and Neely (2007) follow a firm that randomly imposed Balanced Scorecard incentive pay on all retail store managers for a large firm of retail plumbing supplies. But the stores that see productivity rise from this treatment are the stores that have more experienced managers who are likely to know how to respond to get the most out of the incentive pay through improved store performance.

³⁰ See Hubbard (2003), Baker and Hubbard (2003), and Bartel et al. (2007).

³¹ The underlying hypothesis of their paper is that the HR technology shocks over time have introduced new best practices for firms, and all firms would gain from adding the new best practices, such as pay for performance rather than pay for tenure. This same hypothesis underlies the integrated steel mills study of Ichniowski et al. (1997).

³² See Roberts (2004) and Milgrom and Roberts (1995) for discussions of complementarity. If two HR practices are complements, then firms adopt them jointly and the multicollinearity between these practices makes it impossible to readily separate their effects in the production function. Specifically, if the practices are truly complements across all firms in the data, then firms will *always* adopt the practices together and there will be no data to identify the separate effects of the individual practices. For empirical work, see Ichniowski et al., 1997, and Breshnahan et al., 2001.

4. Steps to take

The Key Features above describe what we aim to achieve in developing hypotheses and seeking data. How is an Insider study really conducted? It is essential that insiders be involved – often to obtain the data, but more importantly, to guide and interpret the research. The steps reflect this.

1. Interview managers, workers, and others to learn about the mode of production and the management issues faced by and addressed by the firm.
2. Identify a research question based on these interviews. Most researchers would not know the questions or potential hypotheses without the plant visits.
3. Gather data and test the hypotheses. Determine your econometric methods to identify the appropriate treatment effect. Decide what is endogenous versus exogenous, and where there might be measurement error or omitted variable bias.
4. Evaluate and interpret your results. Empirical researchers typically test the robustness of their models by introducing functional form changes in the production function. Insider Econometrics studies go a step further by asking the insider informants if the results make sense. However, the informants may have missed the key points – they could be too close to the phenomenon being studied to independently evaluate what they've done and why it has succeeded or failed. By combining insider information and an outsider's perspective, the researcher may be closer to the truth.
5. Get more data if needed for 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.

Overall, what Insider Econometric studies can do very successfully is present a *preponderance* of the evidence to convincingly evaluate the hypothesis. It is the merging of quantitative and qualitative data that makes these studies convincing. Increasingly, journal editors wish to see all the steps above, and are pleased to have qualitative evidence added to quantitative models.

5. Obtaining data

In the studies highlighted above, the researchers worked with firms to obtain data. That is not always necessary. For a more complete description of alternative data sources, see [Ichniowski and Shaw \(2009\)](#). Data sets often contain information on productivity or HR practices, but not both: researchers must combine data from several sources to get both. The alternatives data source are:

1. Insider data from within one firm. Here, the firm provides the data.³³
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.³⁴ For surveys, the researchers use insider guidance to develop 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.³⁵ A good many industries – such as education, health care, trucking, and

electricity generation – are regulated and 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 consulting firms' data sets or industry associations' data sets.³⁶ Some industries are followed by the U.S. Census Department, such as trucking, or 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 and 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.³⁷

In sum, to obtain both productivity data, and HR practices data, the researcher will often need to merge several datasets. For more details on working with firms, see [Helper \(2000\)](#). To work with firms, it is important for the researchers to show what they can do for the firm, or to ask for the firm's generosity in working to further our knowledge of how businesses operate.

6. Econometric methods

The Insider Econometrician is implementing econometric treatment analysis. The “treatment” is the innovative management change, such as a change in human resource practices or any other change. The “treatment effect” is the estimated impact of the treatment innovation on productivity, or on any other performance measure for the firm or worker. There is a very voluminous literature on the estimation of treatment effects.³⁸ The majority of the treatment literature has traditionally focused on the estimation of the effects of a policy change – such as an increase in taxes or unemployment insurance – on an outcome variable – such as labor supply. For Insider studies, we aim to estimate the impact of managerial practices on performance, and thus there are special issues that arise.

For Insider analysis, there are many ways in which the econometrics of the treatment effect is focused on specific issues not emphasized in the broader econometrics literature. First, the Insider Econometric methods revolve around the estimation of basic fixed effects productivity regression models – where workers or firms are the fixed effects in the production function – with different methods added to reflect the possible endogeneity of the HR practice in the regression. The methods for modeling adoption depend upon the type of data and type of experiment: alternative methods were displayed above in Key Feature 4 (“model why the treatment was adopted”). Second, the Insider research often wants to permit heterogeneity in the treatment effect across workers: we expect workers will differ in their performance responses to changes in a management practice. Third, the Insider researcher often wants to know why the treatment had an effect on performance. They

³³ Autor et al. (2003). Bandiera et al., 2005, 2007, 2009a, b, forthcoming. Bartel (2004). Batt (1999). Bhattacharjee (2005). Encinosa, Gaynor and Rebitzer (2005). Freeman and Kleiner (2005). Griffith and Neely (2007). Hamilton et al. (2003). Jones and Kato (2007). Jones, Kalmi and Kauhanen (2006). Knoeber and Thurman (1994). Lazear (2000). Lo et al. (2006).

³⁴ Bartel et al. (2007). Bloom et al., (2009a,b). Boning et al. (2007). Dunlop and Weil (2000). Fernie and Metcalf (2003). Ichniowski and Shaw (2003). Ichniowski and Shaw (1995). Kruse (1993). Kruse, Blasi, and Park (2008). MadDuffie (1995). Van Biesebroeck (2006).

³⁵ Asch (1990). Baker and Hubbard (2003). Dee and Keys (2004). Dranove et al. (2003). Eberts et al. (2000). Garicano and Heaton (2006). Gaynor et al. (2001). Hall et al. (2008). Hoxby (2002). Hubbard (2003). Jacob and Levitt (2003). Kreuger (1999). Lavy (2008). Lavy (2002). Lim (2007). Rivkin et al. (2005). Reese, Hall, and Propper. Rosenthal et al. (2004). Imberman et al. (2009). Duflo and Hanna (2005). Kalnins and Chung (2006).

³⁶ Autor and Scarborough (2008). Garicano and Hubbard (2005). Lafontaine and Shaw (1999). Kreuger and Mas (2004). Lafontaine and Shaw (2005). Landers et al. (1996). Oyer and Schaefer (2007).

³⁷ See Anderson, 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 the range of possible data sets.

³⁸ See Imbens and Wooldridge (forthcoming) for a review, and List (forthcoming) for a review of the social experiments literature, which has similar features to our Insider Econometrics methods.

aim to show why either in the productivity regressions or in auxiliary regressions with different data.

Several typical approaches to modeling the treatment adoption are shown in the examples above. When one firm is adopting one practice, like incentive pay, estimation of the production function with worker fixed effects is sufficient (see Example 1, Lazear, 2000; Example 2, Bandiera et al., 2005a,b). When one firm is adopting different practices in different parts of the firm, like a controlled experiment, then differences-in-differences or matching estimators are used. For example, in Griffith and Neely (2007), some stores within a retail plumbing supply firm adopt the Balanced Scorecard incentive pay plan, and some don't. In some cases, researchers have a natural experiment, or a very good instrument for the treatment. Unlike in Griffith and Neely, in most studies, whether different firms are adopting new practices, or different areas within a firm are adopting practices, the HR practices are endogenous to the firm. In this case, the goal is to model adoption, or explain to the reader that the research has two features (see Example 4, Boning et al., 2007). First, the production function estimates display the "treatment of the treated effect;" the coefficient on teams or incentive pay in the production function estimates their effect on productivity conditional on adoption of these HR variables by these firms. Second, the researcher then adds descriptive evidence or a model of why some firms adopt and some don't.

For more details on how to think about alternative estimation strategies for production functions focused on HR practices, see Ichniowski and Shaw (2009). Occasionally, researchers have data in which there is a natural experiment in HR, but that is rare.³⁹

7. Conclusion

There are three mechanisms firms have for raising workers' productivity through optimal HR practices. The firm can add an individual HR practice, like incentive pay plan, aimed at increasing the workers' effort on the job. The incentive pay plan operates on the workers' intrinsic margin – it alters the workers' behavior within the current firm. Second, the firm can add an individual pay practice, like incentive pay, to induce optimal sorting of workers. Firms and workers are matched: the firms that have the biggest gains to incentive pay, in their production environment, will attract the workers who are most productive. The incentive pay plan operates on the workers' extrinsic margin of behavior – it causes workers to sort across firms. Third, the firm can package together different HR practices to form a complementary systems of HR practices that together raise productivity. Adopting a new incentive pay is more effective when the firm also adds more on the job training or team problem solving. All HR practices, like incentive pay or teams or training or careful hiring, affect productivity in these three ways. For example training doesn't just raise productivity on the job, it also causes workers who value training highly to sort to the firm that offers it. The four 'stops along the way,' in Section 1 above, are aimed at showing how, when, and why HR practices are effective in raising productivity in the three ways of increasing workers' effort, sorting, or bundling HR practices.

HR practices differ across firms because firms have different production functions and operate in different labor and product markets. Why do firms ever change their HR practices? There are three sources of disequilibrium. First, there are time series 'shocks' to optimal best practices. Our knowledge of how to use HR practices evolves and improves over time. Thirty years ago, hourly pay was common; today, variable pay has been added. Second, there are cross-sectional shocks – a competing firm takes an action that causes another firm within its industry to change its strategy and thus its HR

practices. These within-industry shocks arise from new product development or new research. Third, firms may decide to experiment with new HR practices, because they have not yet found their optimal practices, or because their internal conditions are changing.

Ultimately, implementing Insider Econometrics is part art and part science. Insider econometrics papers are aimed at illuminating why HR practices are adopted, and how effective they are, given the shocks that cause changes in adoption rates. Insider econometrics methodology is about formulating the experimental design to achieve this goal. The four stops were chosen to offer different "views" of how one does Insider Econometrics. On any road trip, the sightseer takes photos along the way. No road trip should stop at the same views; we want photos of different sites. Insider Econometric researchers are taking snapshots of what some firms do, to learn from these snapshots. The destination of Insider Econometrics research is to produce beautiful snapshots that deepen our understanding of the theory of the firm. Just as photography has rules of thumb – for composition and lighting to produce great pictures – we have rules of thumb laid out as the Key Features of Insider Econometrics. The Key Features are aimed at addressing the researcher's question, what do I do when faced with an idea or data from a firm. Successful studies offer a preponderance of scientific evidence that is pieced together and selected artfully. When done, the researcher has beautiful, and convincing, composition of evidence.

References

- Acemoglu, Daron, 2003. Patterns of Skill Premia. *Review of Economic Studies* 70 (2), 199–230.
- Andersson, Fredrik, Freedman, Matthew, Haltiwanger, John, Lane, Julia, Shaw, Kathryn, 2009. Reaching for the Stars: Who Pays for Talent in Innovative Industries? *Economic Journal* 119 (538), 308–332.
- Asch, Bethl., 1990. Do incentives matter—the case of Navy recruiters. *Industrial and Labor Relations Review* 43 (3), 89–106.
- Autor, David, Levy, Frank, Murnane, Richard, 2003. The skill content of recent technological change: an empirical investigation. *Quarterly Journal of Economics* 118 (4), 1279–1333 November.
- Autor, David H., Scarborough, David, 2008. Does job testing harm minority workers? Evidence from retail establishments. *Quarterly Journal of Economics* 123 (1), 219–277 February.
- Baker, George P., Hubbard, Thomas N., 2003. Make versus buy in trucking: asset ownership, job design, and information. *American Economic Review* 93 (3), 551–572.
- Bandiera, Oriana, Barankay, Iwan, Rasul, Imran, 2005a. The Evolution of Cooperative Norms: Evidence from a Natural Field Experiment. CEPR Discussion Paper No. 5358.
- Bandiera, Oriana, Barankay, Iwan, Rasul, Imran, 2005b. Social preferences and the response to incentives: evidence from personnel data. *Quarterly Journal of Economics* 120 (3), 917–962 August.
- Bandiera, Oriana, Barankay, Iwan, Rasul, Imran, 2007. Incentives for managers and inequality among workers: evidence from a firm level experiment. *Quarterly Journal of Economics* 122 (2), 729–773 May.
- Bandiera, Oriana, Barankay, Iwan, Rasul, Imran, 2009a. Social Connections and Incentives in the Workplace: Evidence from Personnel Data. *Econometrica* 77 (4), 1047–1094 July.
- Bandiera, Oriana, Barankay, Iwan, Rasul, Imran, 2009b. Team Incentives: Evidence from a Field Experiment, working paper.
- Bandiera, Oriana, Barankay, Iwan, Rasul, Rasul, Imran, in press. Social Incentives in the Workplace. forthcoming in the *Review of Economic Studies*.
- Bartel, Ann, 2004. Human resource management and organizational performance: evidence from retail banking. *Industrial and Labor Relations Review* 57 (2), 181–203.
- Bartel, Ann, Ichniowski, Casey, Shaw, Kathryn, 2007. How does information technology affect productivity? Plant-level comparisons of product innovation, process improvement, and worker skills. *Quarterly Journal of Economics* 122 (4), 1721–1758.
- Batt, Rosemary, 1999. Work organization, technology, and performance in customer services and sales. *Industrial and Labor Relations Review* 52, 539–564 July.
- Bhattacharjee, Debashish, 2005. The effects of group incentives in an Indian firm: evidence from payroll data. *Labour* 19 (1), 147–173.
- Bloom, Nicholas, Schankerman, Mark, Van Reenen, John, 2007. Identifying Technology Spillovers and Product Market Rivalry. NBER Working Papers 13060, National Bureau of Economic Research, Inc.
- Bloom, Nicholas, Draco, Mirka, Van Reenen, John, 2009. Trade induced technical change? The impact of Chinese imports on technology, jobs and plant survival – evidence from a panel of plants. Working paper.
- Bloom, Nick, Van Reenen, John, 2007. Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics* 122 (4) November.
- Bloom, Nick, Eifert, Benn, Mahajan, Aprajit, McKenzie, David, Roberts, John, 2009a. Management as a Technology: Evidence from Indian firms, working paper.
- Bloom, Nick, Genakos, Christos, Sadun, Raffaella, Van Reenen, John, 2009b. Management Practices Across Asia, Europe and the US, working paper.

³⁹ Perez-Gonzalez (2006) estimates the effect of CEO quality on firm productivity using a natural experiment: in family firms, passing the control of the firm to a child is more likely when the child is a boy, and passing control of the firm to a family member lowers the productivity of the firm. Other examples include Mas (2006) and Mas and Moretti (2009).

- Boning, Brent, Ichniowski, Casey, Shaw, Kathryn, 2007. Opportunity counts: teams and the effectiveness of production incentives. *Journal of Labor Economics* 25 (4), 613–650.
- Breshnahan, Timothy, Brynjolfsson, Erik, Hitt, Loren, 2001. Information technology, work organization and the demand for skilled labor: firm-level evidence. *Quarterly Journal of Economics* 117 (1), 339–376.
- Clark, Kim B., 1980a. The impact of unionization on productivity: a case study. *Industrial and Labor Relations Review* 33 (4), 451–469.
- Clark, Kim B., 1980b. 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. *American Economic Review* 74 (5), 893–919.
- Dee, Thomas, Keys, Benjamin, 2004. Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management* 23 (3), 471–488.
- Dranove, David, Kessler, Daniel, McClellan, Mark, Satterthwaite, Mark, 2003. Is more information better? The effects of “report cards” on health care providers. *Journal of Political Economy* 111 (3), 555–588.
- Duflo, Esther, Hanna, Rema, 2005. Monitoring Works: Getting Teachers to Come to School. *NBER Working Paper* No. 11880.
- Dunlop, John T., Weil, David, 2000. Diffusion and Performance of Modular Production in the U.S. Apparel Industry. In: Ichniowski, Casey, Levine, David, Olson, Craig, Strauss, George (Eds.), *The American Workplace: Skills, Compensation and Employee Involvement*. Cambridge University Press.
- Eberts, Randall, Hollenbeck, Kevin, Stone, Joe, 2000. Teacher performance incentives and student outcomes. *Journal of Human Resources* 37 (4), 913–927.
- Encinosa III, William E., Gaynor, Martin, Rebitzer, James B., 2005. The Sociology of Groups and the Economics of Incentives: Theory and Evidence on Compensation Systems. *IZA Discussion Papers* No. 1851.
- Fernie, Sue, Metcalf, David, 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.
- Foster, Lucia, Haltiwanger, John, Syverson, Chad, 2008. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review* 98 (1), 394–425 March.
- Freeman, Richard B., Kleiner, Morris M., 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.
- Gant, Jon, Ichniowski, Casey, Shaw, Kathryn, 2002. Social capital and organizational change in high-involvement and traditional work organizations. *Journal of Economics and Management Strategy* 11 (2), 289–328 Summer.
- Garicano, Luis, Heaton, Paul, 2006. Computing Crime: Information Technology, Police Effectiveness, and the Organization of Policing. *CEPR Discussion paper* no 5837.
- Garicano, Luis, Hubbard, Thomas N., 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, Rebitzer, James B., Taylor, Lowell J., 2001. “Incentives in HMOs” *Macroeconomics* 0111001, EconWPA.
- Griffith, Rachel, Neely, Andy D., 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.
- Hall, Emma, Van Reenen, John, Propper, Carol, 2008. Can Pay Regulation kill? The Impact of Labor Markets and Skills on Hospital Productivity. *CEP Discussion Paper* 843.
- Haltiwanger, John, 2008. Top Ten Questions for Understanding Firm Dynamics and Productivity Growth” Keynote address at CAED.
- Hamilton, Barton, Nickerson, Jack, Owan, Hideo, 2003. Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy* 111 (3), 465–497.
- Helper, Susan, 2000. Economists and Field Research: You Can Observe a Lot Just by Watching. *American Economic Review*. *American Economic Association* 90 (2), 228–232 May.
- Hoxby, Caroline, 2002. Would school choice change the teaching profession? *Journal of Human Resources* 37 (4), 846–891 Autumn.
- Hubbard, Thomas N., 2003. Information, decisions, and productivity: on-board computers and capacity utilization in trucking. *American Economic Review* 93 (4), 1328–1353.
- Ichniowski, Casey, Shaw, Kathryn, 1995. Old Dogs and New Tricks: Determinants of the Adoption of Productivity-Enhancing Work Practices. *Brookings Papers on Economic Activity: Microeconomics*, pp. 1–65.
- Ichniowski, Casey, Shaw, Kathryn, 2003. Beyond incentive pay: insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives* 17 (1), 155–180 Winter.
- Ichniowski, Casey, Shaw Kathryn, 2007. “Connective Capital: Building Problem-Solving Networks Within Firms”, working paper
- Ichniowski, Casey and Kathryn Shaw, forthcoming. “Insider Econometrics: A Roadmap for Estimating Empirical Models of Organizational Design and Performance,” In Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, forthcoming 2011.
- Ichniowski, Casey, Shaw, Kathryn, Prennushi, Giovanna, 1997. The effects of human resource management practices on productivity: a study of steel finishing lines. *American Economic Review* 86, 291–313.
- Imbens, G., and J. Wooldridge, forthcoming. “Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*.”
- Imberman, Scott, Adriana Kugler, and Bruce Sacerdote, (2009) “Katrina's Children: A Natural Experiment in Peer Effects from Hurricane Evacuees,” *Mimeo*, April.
- Jacob, B., Levitt, S., 2003. Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118 (3), 843–877.
- Jones, Derek C., 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., Kalmi, Panu, Kauhanen, Antti, 2006. Teams, Performance-Related Pay, Profit Sharing and Productive Efficiency: Evidence from a Food-Processing Plant, Working paper.
- Jorgenson, Dale W., Ho, Mun S., Stiroh, Kevin J., 2008. A Retrospective Look at the U.S. Productivity Growth Resurgence. *Journal of Economic Perspectives* 22 (1), 3–24.
- Kalnins, Arturs, Chung, Wilbur, 2006. Social capital, geography, and survival: Gujarati immigrant entrepreneurs in the U.S. lodging industry. *Management Science* 52 (2), 233–247.
- Knoeber, Charles R., Thurman, Walter N., 1994. Testing the theory of tournaments: an empirical analysis of broiler production. *Journal of Labor Economics* 12 (2), 155–179.
- Kreuger, Alan, 1999. Experimental estimate of education production functions. *Quarterly Journal of Economics* 114 (2), 497–532.
- Kreuger, Alan B., Mas, Alexander, 2004. Strikes, Scabs, And Tread Separations: Labor Strife And The Production Of Defective Bridgestone/Firestone Tires. *Journal of Political Economy* 112 (2), 253–289 April.
- Kruse, Douglas, 1993. Profit Sharing: Does it Make a Difference? W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.
- Kruse, Douglas, Blasi, Joseph, Park, Rhokeyum, 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, Shaw, Kathryn, 1999. The dynamics of franchise contracting: evidence from panel data. *Journal of Political Economy* 107 (5), 1041–1080 October.
- Lafontaine, Francine, Shaw, Kathryn, 2005. Targeting managerial control: evidence from franchising. *Rand Journal of Economics* 36 (1), 131–150 Spring.
- Lafontaine, Francine, Srinivasan, Jagadeesh, 2009. “Within-firm Labor Productivity across Countries: A Case Study” In: Freeman, Richard, Shaw, Kathryn (Eds.), *International Differences in the Business Practices and Productivity of Multinational Firms: A Case-Based Approach*, University of Chicago Press.
- Landers, Renee M., Rebitzer, James B., Taylor, Lowell J., 1996. Rat race redux: adverse selection in the determination of work hours in law firms. *American Economic Review* 86 (3), 329–348.
- Lavy, Victor, 2002. Evaluating the effect of teachers' group performance incentives on students achievements. *Journal of Political Economy* 110 (6), 1286–1317 December.
- Lavy, Victor, 2008. Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance Based Pay Tournaments among Teachers NBER working paper.
- Lazear, Edward, 1995. *Personnel Economics*. MIT Press, Cambridge, MA.
- Lazear, Edward, 1999. Personnel economics: past lessons and future directions. *Journal of Labor Economics* 17 (2), 199–236.
- Lazear, Edward P., 2000. Performance pay and productivity. *American Economic Review* 90 (5), 1346–1361.
- Lazear, Edward, Shaw, Kathryn, 2007. Personnel economics: the economist's view of human resources. *Journal of Economic Perspectives* 21 (4), 91–114 Fall.
- Lemieux, Thomas, MacLeod, Bentley, Parent, Daniel, 2009. Performance pay and wage inequality. *Quarterly Journal of Economics* 124 (1), 1–49 February.
- Lim, Claire, 2007. “Turnover and Accountability of Appointed and Elected Judges” (Ph.D. Dissertation). University of Pennsylvania.
- List, John A., forthcoming. “Introduction to field experiments in economics,” *Journal of Economic Behavior and Organization*.
- Lo, Desmond, Ghosh, Mrinal, Lafontaine, Francine, 2006. The Role of Risk, Incentives, and Selection in Salesforce Compensation Contracts,” Working paper.
- Mas, Alexandre, 2004. Do Disappointing Pay Raises Lower Productivity? Final-Offer Arbitration and the Performance of New Jersey Police Officers.
- Mas, Alexandre, 2006. Pay, reference points, and police performance. *Quarterly Journal of Economics* 121 (3), 783–821.
- Mas, Alexandre, 2008. Labor Unrest and the Quality of Production: Evidence from the Construction Equipment Resale Market. *Review of Economic Studies* January.
- Mas, Alexandre, Moretti, Enrico, 2009. Peers at work. *American Economic Review* 99 (1), 112–145 September.
- Milgrom, Paul, Roberts, John, 1995. Complementarities and fit: strategy, structure and organizational change in manufacturing. *Journal of Accounting and Economics* 19 (2–3), 179–208.
- Oyer, Paul, Schaefer, Scott, 2007. Personnel-Economic Geography: Evidence from Large US Law Firms. 2nd Annual Conference on Empirical Legal Studies Paper.
- Perez-Gonzalez, Francisco, 2006. Inherited control and firm performance. *American Economic Review* 96 (5), 1559–1588 December.
- Rivkin, Steven, Hanushek, Eric, Kain, John, 2005. Teachers, schools, and academic achievement. *Econometrica*. 73 (2), 417–458.
- Roberts, John M., 2004. *The Modern Firm: Organizational Design for Performance and Growth*. Oxford: Oxford University Press.
- Rosenthal, Meredith, Fernandopulle, Rushika, Song, HyunSook Ryu, Landon, Bruce, 2004. Paying for quality: providers' incentive for quality improvement. *Health Affairs* 23 (2), 127–141.
- Shaw, Kathryn, 2002. “By What Means Does Information Technology Affect Employment and Wages?” In: Greenan, Nathalie, L'Horty, Yannick, Mairesse, Jacques (Eds.), *Productivity, Inequality, and the Digital Economy: A Transatlantic Perspective*, Cambridge: MIT Press.
- Shaw, Kathryn, 2003b. “The Human Resources Revolution: Is it a Productivity Driver?” In: Jaffe, Adam, Lerner, Josh, Stern, Scott (Eds.), *Innovation, Policy and the Economy*, Chicago: University of Chicago, National Bureau of Economic Research 69–114.
- Shaw, Kathryn, 2003a. “Technology Shocks and Problem-solving Capacity,” In: Ginther, Donna, Zavadny, Madeline (Eds.), *Technology, Growth, and the Labor Market*, Boston: Kluwer Academic Publishers 235–258.
- Van Biesebroeck, Johannes, 2006. Complementarities in Automobile Production. *NBER Working paper* #12131.