Preliminary draft
Comments welcome

Insider Econometrics: A Roadmap for Estimating Empirical Models of Organizational Design and Performance

Casey Ichniowski
Columbia University and NBER

and

Kathryn Shaw
Stanford University and NBER

July 10, 2008

"Great advances have been made in theory and in econometric techniques, but these will be wasted unless they are applied to the right data."

-- Zvi Griliches (1994, 2) Presidential Address, American Economic Association

A fundamental objective of the emerging field of organizational economics is an improved understanding of the theory of the firm. As always, theory needs to be grounded in systematic empirical observations and hypothesis testing. Yet it is probably safe to say that Griliches' observation in his presidential address of over a decade ago – that theory and technique often outpace empirical analysis based on the "the right data" – still applies to the economic analysis of organizations. In this paper, we present a roadmap for conducting convincing empirical studies into two fundamental questions about the economics of organizations. First, why do competing organizations – organizations categorized within one industry – make different decisions about practices and policies that impact how they operate? Second, do these different managerial decisions and practices found among otherwise comparable organizations account for differences in key performance indicators?

We term the approach we describe in this paper "insider econometric" research. In any field, an "insider" is someone with "special knowledge and access to confidential information," and both of these characteristics of insiders play important roles in this insider research. While large differences in the managerial practices and performance outcomes of competing organizations pose important economic questions, any number of theoretical models could account for such variation. To gain a better understanding of which theories of the firm best account for the economic diversity of real organizations, insider econometrics research relies on the special knowledge of experienced industry insiders to identify and the most convincing explanations for these differences. It then relies on carefully obtained data on the internal operations of organizations to craft rigorous tests of hypotheses concerning this organizational diversity.

I. What is Insider Econometrics?

A. <u>Defining Characteristics of Insider Econometric Research</u>

Insider econometric (IE) research takes as a starting point the observation that organizations categorized within the same narrowly defined industry classification often make very different decisions and adopt very different practices that affect their operations and performance. Ultimately, the goals of IE research are to understand these differences in firms' decisions about the practices they adopt in their production units and to identify any impacts on economic performance that result from those choices. Thus, three main questions must be addressed in defining IE research: what are *production units* that constitute the samples for IE studies; what kinds of *managerial decisions and organizational practices* do IE studies consider; and how is *performance* measured in IE studies?

Production Unit Observations. By production units, we mean relatively small combinations of economic inputs that constitute a separate and identifiable production process. These production units are the data points in the sample in an IE study, and the output from the production process being studied across the different observations is very comparable. In examples we consider in this paper, a production unit can range in size from anything as small as an individual worker as in the case of windshield installation (Lazear, xxx) or grocery store checkout (Moretti and Mas, xxx), to small work teams consisting of several employees as in the case of apparel manufacturing (Hamiltion etc, xxx), or manufacturing production lines comprised of crews operating similar capital equipment in steel manufacturing (Ichniowski, Shaw, Prennushi, 1997; Boning, Ichniowski, Shaw, 200x). The precise nature of the observations depends on the technology that determines the production process. In general, firms are not the "production units" for an IE study because the activities of an entire firm spans many areas including planning, marketing, production, and research and development and are so diverse that they encompass many different production processes.

Organizational Practices. Firms that own such homogeneous production units can still make many different decisions that affect how these units operate. Practices that IE studies investigate can be any practice or decision that affects the production function of the production units. Several examples we consider in this paper involve differences in human resource practices and personnel policies found in different production units operating in the same narrowly defined industry. Why would one competitor use incentive compensation or team-based work practices, while another does not? Other features that fit under this definition of an organizational design feature could include practices like just-in-time production methods, staffing rules, xxx, and yyy. This definition of organizational design can also include technological differences across organizations in the same industry classification, such as the adoption of new information technologies rather than relying on older technologies. The variation in these decisions and practices may cross-sectional and occur because different firms adopt different practices in their units or because a single firm uses different practices for different units. The variation could also be longitudinal when a single production unit follows different practices at different points in time.

We do not have in mind decisions about the scale of the overall firm or even a plant as measured by differences in the total amount of labor, capital, or other inputs that enter conventional specifications of a firm's production function. Rather, the kinds of design characteristics that IE studies consider are more detailed features that change how a given amount of capital and labor are organized to complete the work involved in the production the firm's goods and services.

Performance Measures. While economic performance is ultimately profits, the level of analysis in IE research that focuses on very micro-level production units and production processes precludes a full analysis of the effects of a given design feature on economic profits. Even an analysis of accounting profits cannot be conducted for analyses of individual workers, work groups, or parts of establishments where profits are not measured. Thus, at this "production unit" level of analysis, IE research will analyze more detailed measures of worksite performance, such as productivity, product quality, choice of product mix, pricing strategy, or unit sales. Studies of workers might alternatively focus on team output, worker output, wages, or incentive payouts that could be a proxy for productivity.

This characteristic of IE research is inevitable given its level of analysis, and this in turn has the important implication that there can be many reasons why some managerial decision or practice that has beneficial effects on some performance outcome might not be adopted by otherwise similar production units. For example, some managerial practice may be shown to promote higher levels of productivity after production units adopt them. It may even be the case that units that did not adopt the practice would enjoy the same kind of productivity gain, but the costs of adopting such a feature may be higher in these units, thus prohibiting adoption in these settings.

Homogeneity and Heterogeneity in IE Samples. As this description implies, IE studies are typically constrained to a single narrowly defined industry classification. The reason for this is that the production units must be homogenous enough that the researcher can use insider information and data to estimate the adoption and/or performance impact equations and minimize the effects of unmeasured variables. Keeping the basic characteristics of the production process constant by focusing on very comparable production units within a single industry helps accomplish this. However, the production units cannot be completely homogeneous. There must heterogeneity in some managerial decision or organizational practice so that the researcher has variation in the data to identify the models and conduct the estimation. The insider information and data then allows the researcher to understand why this type of heterogeneity exists, and then to model the sources of these observed differences within a sample of otherwise very comparable production units.

B. <u>The Econometric Focus of IE Research: Treatment Effects</u>

With this introduction about several defining characteristics of IE research, the econometric structure of these studies can be described in a straight forward way. IE studies seek to identify the impact of a treatment effect on one or more performance outcomes and to understand why that treatment was adopted among only a subset of seemingly homogeneous production unit observations. The treatment is the organizational practice or managerial decision that varies within this otherwise homogeneous sample. Most commonly, the adoption equation would model the selection of the practice. However, the adoption equation could model other types of selectivity. For example, when the observations in the data set are workers within some large organization, such as an entire division or firm, the adoption equation could model the selection of the workers who choose to remain with the business.

In the broader empirical literature that seeks to estimate treatment effects, it is often true that the researcher would prefer to have random assignment of the treatment. For example, in the recent 'natural experiment' literature, researchers have identified data sets in which some accident in nature happens that causes random assignment into different conditions, as when a flu epidemic strikes one area but not another (CITE), when some exogenous occurrence forces identical twins into different environments (CITE), or when policy treatments were purposely administered at random (Lalonde…). The analyst then uses the random assignment into different conditions to study the impact of treatment differences, such as the effect of education or training on wages.

In insider econometrics studies, the aim is not to estimate the effects of random assignment of the treatment. IE studies do not try to "eliminate" the effects of selectivity from the estimated treatment effect because the heterogeneity that remains in these narrowly defined samples in a given industrial classification not only creates variation in the adoption of the treatments, but this remaining heterogeneity also creates the possibility that the optimal decisions for these (slightly different) production units can also be different. For example, firms do not randomly adopt work practices like incentive piece rates. However, just like a firm's managers, the IE analyst could be interested in understanding and modeling why some workers or certain production lines within some large firm would be subject to the piece rate pay while others are not. The econometric models therefore would estimate the "treatment of the treated" effect – the treatment effect for those observations that find it optimal to use the given design feature. While IE research does not focus on the effects of some practice among production unit observations that would not find it optimal to adopt it, it does try to uncover the reasons why adoption is not optimal for the non-adopters and provide data to support those explanations.

Therefore, a guiding principle in conducting IE research in these narrowly drawn intra-industry samples of workers, work groups or production lines is that the differences in practices across the observations are optimal decisions. IE research does not simply assume that some practice that raises productivity for one work team or production line would necessarily raise it in another very comparable one, or that the lack of adoption among some observations is simply a "mistake" or due to a lack of information. Rather, the IE analyst tries to develop a rich and expert understanding of the production setting that allows him or her to uncover the real reasons for the variation in organizational practices – reasons that may not be at all apparent to an outsider – and to understand how this variation could be the result of optimizing behavior by smart managers and workers. To investigate why different decisions and practices can make economic sense even within largely homogeneous samples, IE research thus models the impact of some practice or decision on performance, while jointly trying to understand the reasons why the practice or decision covers only some workers or production lines.

Thus, the two fundamental goals of IE research listed at the start of this section can now be restated. First, IE studies seek to estimate the "treatment of the treated" performance effect for workers or production units that are covered by some managerial decision or organizational practice. Second, IE research simultaneously tries to provide a convincing explanation for why some workers, work groups, or production lines are covered by the given treatment while others are not. To accomplish these goals, the researcher might obtain insider data that can be used to estimate both the selection equation and the equation for the performance effect of the treatment. Often, data sets are not that complete. As the examples we review in this paper will show, it is probably more common for the researcher to estimate models of the effect of the adoption of some treatment on a performance outcome like productivity, and then use his or her insider knowledge to explain why some firms adopt and some do not.

C. Methodological Steps in IE Research

The production units that form the samples in IE studies are producing very comparable goods and are operating and competing in some narrowly defined industry classification. An outsider relying on standard economic models would find differences in organizational practices and managerial decisions puzzling within such a sample. To solve this puzzle and identify the economic causes for this variation and to measure the effects of this variation on performance outcomes, IE research relies on several basic methodological steps that combines econometric analysis using company-based data sets with detailed qualitative investigation of the operations of the production unites within those companies. Ultimately, the aim in these steps is to combine in-depth knowledge about operations within a particular firm or industry from plant visits and interviews with quantitative analysis and statistical tests to reach richer conclusions than could be achieved with either approach alone.

Conducting Preliminary Field Research. Preliminary field research can take many forms from interviews with practitioners to personal inspections of worksites, but the purpose of the field research is to explore the view points of experienced industry insiders on why decision makers have adopted different policies and practices. Here, the special knowledge that defines insiders is important. Their knowledge helps to focus subsequent steps of IE research in important ways. Perhaps most obvious, this initial step can help identify practices or decisions that vary among comparable production units. It can also help explore the managerial thinking behind these different decisions, and thus provide valuable insight into which theoretical models are most relevant and about the kinds of data one would like to collect. Relevant questions for insiders concern the relationship of the decision or practice being investigated to different performance indicators, other determinants of performance besides the organizational practice of interest, and causes for the observed variation in decisions and practices.

Collecting Micro-Level Data on Comparable Production Processes. IE research focuses on some well defined production process that generates homogeneous outputs. Perhaps the most important point to note in this introductory description about the collection of data on these homogeneous operations is that it must anticipate the basic issues in accounting for selectivity in the adoption of the treatment, and the potential effects of unmeasurable variables that exist in any study of a treatment on a performance outcome in a non-experimental setting. IE studies ultimately are most convincing when they provide a persuasive explanation for, and model of, the choice of the specific organizational treatment being studied, and an informed comprehensive model of the performance outcomes.

One very desirable characteristic for the IE study's data set is that it would ideally provide both time-series and cross-sectional variation on the operations of the production units. The time series is valuable because managerial decisions and organizational practices often change over time and would permit the estimation of models that control completely for observation-specific fixed effects. However, cross-sectional variation in observations can also be useful since differences in the adoption of the given treatment may be due to worker- or organization-specific factors that only vary across observations. Or, the researcher may be able to demonstrate how some observation-specific characteristic is responsible for both differences in adoption and the magnitude of the effects of the treatment on performance. The data set ideally would also be comprehensive and cover a "universe" of all production units that meet the criteria for being part of the study.

Of course, no researcher has the option of ordering the idea confidential data set from one or more firms. Instead of confidential data from the firm, researchers may have access to confidential data that is submitted to the government in surveys of firms (in regulated industries or employer-employee matched data); or researchers may themselves survey firms about their production units; or they may simply have some opportunistic access to data on a single firm or several production units through private contacts. Still, the focus on needing to explain both the variation in the treatments and in the performance outcomes at the earliest stages of the design of an IE study can help alert the researcher to opportunities for augmenting the available data.

Conducting Econometric Analysis. The next step is the formulation and estimation of the model. The researcher faces the concern typically faced by standard empirical researchers: how is it possible to test key hypotheses using non-experimental data that was generated as a result of the decisions made by managers and workers rather than from a controlled experiment? Researchers would like to know how firms would have fared if they made different decisions, but of course, they will not have access to the 'unobserved counterfactual." Section II below addresses this issue at length.

Interpreting the Results. At this point, standard empirical researchers would typically conclude their work by interpreting the coefficients of the regression results in light of the proposed theoretical model, testing the results for robustness to alternative explanations, and so on. Insider studies go a step further. Analysts can ask knowledgeable experts to independently verify or interpret the results to see if they fit the informant's views, particularly their knowledge of the production process itself. The researcher may circle back and re-estimate the model. Of course, the results of the model need not conform to the industry expert's priors – the views of industry experts differ widely as a result of circumstances within their firms. Furthermore, the fundamental purpose of insider econometric estimation is to provide real tests of hypotheses rather than informed opinions that one would get in traditional case studies. But this additional round of discussion with industry experts can provide additional insights and refinements not possible in other research settings.

These steps are not always followed in this sequence. For example, if researchers find they have access to an unusual data set, they may start with the potential data and then talk to insiders to develop their ideas and modeling. In examples we review in the paper, we present some examples of publicly available data sets that have been used to do insider analysis.

Ultimately, the empirical patterns documented in insider econometric research should loop back and provide new insights that theorists build into new models, and we suspect the rich detailed empirical work from IE studies with these kinds of features that we consider below in this study have the potential to offer such theoretical insights. The empirical patterns from IE studies can provide important insights into what would be the most promising directions for elaborating on the simpler models of the firm because IE studies will document empirical patterns that new, more elaborate theories would need to take into account.

In the sections that follow, we do not attempt to provide an entirely comprehensive literature review of all studies that have these characteristics and goals. Instead, we highlight several papers as examples of research that rely on insider insights and data to conduct econometric tests of hypotheses relevant to broad field of organizational economics. We take this approach because we suspect that a more detailed treatment of several insider studies can serve as a useful roadmap for future studies by highlighting several important issues empirical researchers face in developing and then testing models about organizations' choices of practices and their impacts on performance outcomes. Since IE researchers must often assemble their own primary source data sets from confidential sources, anticipating these issues before constructing a data set may be particularly important in these settings.

II. Econometric Issues in Estimating Effects of Organizational Treatments

IE research that follows the guidelines presented in Section I is ultimately an investigation of an organizational "treatment" on some performance outcome. The "treatment" refers to a discrete organizational choice variable. It may be the use of piece rate pay or of another human resource management practices, such as the use of teams, the methods for selecting or training workers, the quality of workers, the design of jobs or career ladders, or the degree of information sharing. The treatment can refer to other organizational policies besides work practices: strategies about choice of product mix, the pricing policy, re-organization of worksites under new rules, or an organizational change from deregulation. While these discrete treatments would not encompass adjustments of the quantity of labor or capital, treatments could be the adoption of specific new computer-based technologies for production, a decision to use capital in a just-in-time operations approach, or the use of some new employee recruiting and selection mechanism that affects the quality of the workforce.

While the treatment literature is well known among economic policy analysts, organizational economists are less familiar with the estimation of treatment effects, and moreover, unique issues arise in the application of treatment analysis to studying the impact of organizational practices on firms' performance.

Treatments are, by definition, ones that vary considerably within narrowly defined industries or occupations—for example, piece rate pay would be applied to a group of workers, not an entire firm. Therefore, the level of analysis is organizations or workers or organizational production units within firm. Thus, we see once again why insider data is valuable – both the productivity equation and the selection equation have less measurement error for narrowly defined groups.

A. Performance Outcomes and Treatments

As a way of summarizing the intuition behind the econometric estimation of the treatment effects, consider Figure 1. In Figure 1, we assume we have panel data, or time-series cross-sectional data to estimate the treatment effects. An example helps. The treatment could be the adoption of teamwork, replacing individual work. The cross-sectional data is production units: the units are different plants that adopt teams over time if we have plant-level data; the units are workers within a plant if we have data on workers within one firm that adopts teams. There are two sets of firms: those that will adopt the treatment (like teams) at time t*, so the "treated" group is *T* (T=0); and those that never actually adopt the treatment, labeled the non-treated group *N* (or T=0). Figure 1 assumes we are tracking changes in performance (*Y*) over time for the *T* and *N* type observations. The performance curve slopes upward over time because we assume both groups experience gradual performance increases for reasons such as learning-by-doing or increased human specific capital in production. The treated group experiences a big jump in performance due to treatment, equal to ΔY^T, following treatment at time t*. We do not observe the performance gain for the non-treated group had it been treated: the potential (unobserved) performance gain is called the unobserved counterfactual, and it is the dashed line labeled the "non-treated counterfactual" in Figure 1. In the figure, we display a hypothetical example that assumes that gains from the treatment would be lower among observations where no treatment is observed. We assume that the firms that never adopt the treatment would have lower potential gains to the treatment, equal to ΔY^N.

Figure 1 clearly displays the selection biases. If we estimate the model in OLS, the estimated gain attributed to the treatment will be ΔY^T_{OLS}. The OLS estimate takes the average of the data across only the observed data (or only the bold lines in Figure 1). Is the OLS estimate of the treatment gain biased upwards or downwards? It depends on the researcher's goals. If the researcher wants to know what would happen if the treatment were randomly imposed on units, then the researcher wants to know the average of the ΔY^T and ΔY^N, which is called that *average treatment effect* (ATE). If this is the goal, then the OLS estimate is biased upwards: ΔY^T_{OLS} exceeds the average of the ΔY^T and ΔY^N in the figure. If the researcher wants to know what would happen if the treatment is only adopted by those firms that find it most optimal, then the researcher wants to know only the value of ΔY^T, which is called the *treatment of the treated* (TT). If this is the goal, then the OLS estimate is biased downwards: ΔY^T_{OLS} is less than the ΔY^T gain.[1]

In all cases, the researcher's knowledge about the selection, or adoption equation, is crucial. In the estimation of the average treatment effect, the researcher either has to be sure the treatment is randomly assigned (as it is in some examples given below) or be able to estimate the selection equation as well as the performance equation in order to control for the biases using non-experimental data with non-random assignment. In the estimation of the treated of the treated, the researcher wants to know the selection equation in order to draw conclusions about which types of organizations ought to adopt the treatment in order to obtain the gains of ΔY^T. We discuss these alternative estimation methods more rigorously next.

B. <u>Modeling Organizational Treatment Effects</u>

[1] See Heckman (1990) for a discussion of these different types of treatment effects.

Given the goal of modeling the effects of a treatment (T) on performance (***Y***), one can represent two distinct "switching" regimes for the performance of treated (*T*) and not treated (*N*) observations:

(1) $$Y_{it}^{T} = \alpha^{T} + X^{T}\beta^{T} + u^{T} \text{ for } T = 1, E\left(u^{T}\right) = 0$$

(2) $$Y_{it}^{N} = \alpha^{N} + X^{N}\beta^{N} + u^{N} \text{ for } T = 0, E\left(u^{N}\right) = 0$$

Where the ***X*** is a matrix of control variables that can differ for the treated and non-treated groups (and matrix ***X*** has cross-sectional and time series variation within these groups) and a choice of treatment T equation is

(3) $$I = Z\gamma + v \text{ where T=1 if } I \geq 0\text{; T=0 if } I<0$$

Observed performance for the entire sample is the weighted average of (1) and (2), where the weights are formed with the indicator T variable, so that the treatment group T=1 group weights (1) by T=1 when treatment occurs, and the non-treated group N weights (2) by (1-T)=0 when not treated.

(4) $$\begin{aligned} Y &= Y^{T}T + Y^{N}(1-T) \\ &= \alpha^{T}T + \alpha^{N}(1-T) + X^{T}\beta^{T}T + \left(X^{N}\beta^{N}\right)(1-T) + Tu^{T} + (1-T)u^{N} \end{aligned}$$

Using (4), we want to identify the effect of the treatment, T. However, the estimated treatment effect is likely to be subject to selection bias as implied by equation (3) and displayed in Figure 1. In the next subsections, we discuss the methods of estimating these alternative treatment effects described for Figure 1 above.[2]

C. The Average Treatment Effect

The average treatment effect (ATE) is the effect that researchers often identify as their goal in the broad treatment literature: it is the gain in performance if the treatment were randomly assigned to firms (or individuals in the firm). Differencing (2) and (1) to estimate the gain in performance, it is:

(5) $$E\left(Y^{T} - Y^{N} \middle| X_{it}\right) = \left(\alpha^{T} - \alpha^{N}\right) + \left(\beta^{T}\overline{X}^{T} - \beta^{N}\overline{X}^{N}\right) + E\left(u^{T} - u^{N}\right)T + u^{N}$$

[2] Note also that most researchers assume that $\beta^{T} = \beta^{N}$, so that the effects of the treatment operate only through the intercept, not through interactions with the X variables. Whether this is a reasonable assumption depends on the particular circumstances of the treatment effect, which insiders are better able to address for each individual study.

In Figure 1, this would equal the average of the ΔY^T and ΔY^N if treatment were randomly assigned so we observed both value.[3]

The problem is that the treatment is often not assigned randomly, either across organizations or within them. Rewrite the expected residual in (5) as

(6) $$E\left[\left(u^T - u^T\right)T + u^N\right] = E\left[\left(\left(u^T - u^N\right)\middle| T = 1\right)\cdot P_r\left(T = 1\right)\right]$$

If there is selection bias, and we estimate (5) using OLS, the value of the error term in (6), will be correlated with the expected gains, $E\left(\alpha^T \middle| T = 1\right) - \left(\alpha^N \middle| T = 0\right)$ for the treatment (if we assume $\beta^T = \beta^N$). A particular form of this selection bias was displayed in Figure 1: the estimated gain in performance is larger for those plants that actually adopt the treatment compared to those that don't, so the selection bias is positive.

There are a large number of approaches available for addressing the selection bias that arises when aiming to estimate the ATE. Researchers who have worked with panel data may wish that simply estimating the model with fixed effects for the cross-sectional units (or, fixed effects for each plant), but fixed effects estimation will not solve the selection bias problem is your goal is to estimate the ATE. The reason it will not solve it is evident in Figure 1 – fixed effects will control for differences in intercept for adoptors and non-adopters, but if there remain unobserved reasons why actual adopts gain more than the non-adoptors would gain, fixed effects are still biased. Fixed effects would be unbiased for Figure 1 if we had information on performance the unobserved counterfactual (dotted line).[4]

The only way to estimate the ATE is to add information on the nature of the selection bias, or on the Z variables in (3), and then use instrumental variables or selection correction methods that aim to condition the performance outcome on selection into treatment to remove the correlation between the residual and the treatment. The best

---

3 In Figure 1, $\left(\alpha^T - \alpha^N\right)$ differs for the adoptors and non-adoptors: with random assignment, it would not differ.

4 The 'natural experiment' literature in economics looks for instances in which a policy was adopted randomly, and then the researcher can use fixed effects estimation. This is called difference-in-differences estimation because the researcher is differencing over time within units to difference out plant-specific omitted variables, and differencing across cross-sectional units to differences out any unobserved shocks time series shocks to these plants.

methods for doing this depends on what we can assume about (3).[5] In one example discussed below, this approach was executed is to use semi-parametric estimation within fixed effects models for changes in performance that difference out the selection bias variable.

D. The Treatment of the Treated

While estimating the "average treatment effect" is often the goal in many studies, it is often not a relevant goal for researchers studying organizational performance. For example, consider the treatment of introducing piece rate pay as a replacement for hourly pay. To estimate the average treatment effect, one would take a sample of companies and randomly impose piece rate pay on a subsample and estimate the outcome. Clearly, such an experiment would be hard to imagine. What both researchers and the managers of organizations want to know is the answer to the question: under what production conditions is it optimal to use piece rate pay and given that it is optimally introduced, what are the expected performance gains from it?

Answering these questions requires an estimate of the treatment of the treated. Here, the relevant treatment effect is the expected gains conditional on treatment.

(7)

$$E\left(Y^T - Y^N \middle| x, z, T=1\right) = \left(\alpha^T - \alpha^N \middle| T=1\right) + \left(\beta^T \overline{x}^T - \beta^N \overline{x}^{NIT} \middle| T=1 + E\left(\left(u^T - u^N\right) \middle| T=1\right)\right.$$

which is the gain from choosing the treatment (T=1) given the values of X=x and Z=z, where $\overline{X}^{NIT}$ is the value of the X's for the organization (or worker) that was not treated but could have been $\left(\overline{X}^{NIT} = E\left(X^N \middle| T=1\right)\right)$.[6]

In the example concerning piece rate pay, we need to consider two points. First, among a sample of firms adopting piece rate pay, what was the mean performance gain?

---

[5] See Imbens (2004) for a discussion of alternative methods of estimating the ATE. See Imbens and Lemieux (2008) for a discussion of regression discontinuity analysis for estimating the ATE when the researcher has an indicator function that specifies that selection occurs when a linear variable crosses a threshold value (so this is rarely true for organizational policy changes).

[6] Note that the α^T and α^N and the β^T and β^N are different for the treated group than for the non-treated group (as shown in Figure 1), and for simplicity in writing (7) we assume we care only about the treated group coefficients and do not make that distinction.

This "treatment of the treated" effect can be estimated with first differences (before and after treatment) or instrumental variables. Second, it is important consider that the estimated treatment affect is conditional on the decision to choose piece rates in (3), whether optimal or not. If we can assume that all firms choosing piece rates did so optimally, then we have a clear model of the treatment of the treated for optimal adoption. If firms are not adopting piece rates "optimally," then the interpretation of the estimated effect from (7) is unknown. To estimate the treatment of the treated effect in such examples where panel data contain observations before and after treatment, one can estimate the top performance line in Figure 1, dropping from the sample the non-treated group. Assuming that you care about the treatment of the treated group, this can be a reasonable approach (as long as no other variables are correlated with the treatment).

[need to condense this to a footnote: In essence, we are estimating the expected mean treatment effect for those individuals or firms who are at the margin of adopting or not adopting the treatment (such as incentive pay or some other organizational treatment). This parameter – the mean effect for those at the margin – has been labeled the Mean Treatment Effect (Bjorklund and Moffitt, 1987, Heckman and Vytlacil, 2001, and references therein) or the Local Average Treatment Effect (LATE, Imbens and Angrist, 1994, and other forms of it in Heckman, 1997, Angrist, et. al, 2000). Since this parameter pertains to those at the margin of adoption, we are considering the treatment of the treated. Given the objective of estimating this parameter, there are nevertheless a large number of alternative estimation methods that depend on several practical issues: the nature of the treatment being applied, the observed variables that are available for identifying the selection equation, what one is willing to assume about the value of the observed variables in predicting selection, and the assumptions about the distributions of the unobserved variables. A consideration of these important issues is available in the references above. But the overall point for this discussion of insider studies of the effects of organizational changes on performance outcomes is that there is tremendous heterogeneity in the nature of the parameter to be estimated and in the actual estimate of the parameter given alternative assumptions.

E. <u>Methods of Estimating Treatment Effects for Organizational Practices</u>

[this section is too long – I haven't edited it in light of what we say in the examples after it]

In sum, our goal in this section is discuss the selection issue, often called the selection bias. If we assume that the (conditional) mean productivity of firms who adopt teams is the same as those who never adopt teams, and this assumption is false because the expected productivity affects the adoption decision, then invoking this assumption will create a selection bias in the estimated value of the treatment (where the estimated value of treatment is Y – Y in this simple example). A very extensive literature has developed around methods for reducing this selection bias. In applying that the selection bias literature to organizational changes, the primary way in which organizational economics diverges from the economic policy literature is that we often care as much about *why* the policy was chosen as we do what the productivity gains are from that practice. Two advantages of IE style research for assessing these treatment effects should be noted. First, in some cases, studies with confidential data on detailed worker or organization characteristics will provide better data for constructing the selection equations in the model.

Secondly, because there are many assumptions we can make to estimate the unobserved counterfactual, in-depth knowledge of the organization that arises from the insider visits will often provide the researcher with a great deal of guidance on what assumptions can reasonably be made to estimate the performance outcome for the unobserved counterfactual, or to estimate the expected outcome if the firm had not adopted the practice. Even when confidential sources do not provide quantitative data for modeling the selection process, the insiders' perspectives often provide an understanding of the selection process and guide the analyst toward sensitivity analysis that tests whether returns to the treatment vary within various identifiable subsamples of observations. The essential point is that inherent heterogeneity in people and production processes in turn produces heterogeneity in how a given treatment impacts performance outcomes.

While economic performance is ultimately profits, the level of analysis of IE studies will likely preclude a full analysis of the effects of the given treatment on economic profits. Even an analysis of accounting profits cannot be conducted for analyses of individual workers, work groups, or parts of establishments where profits are not measured. Thus, at these very micro-levels of analysis, research will study more detailed measures of worksite performance, such as productivity, product quality, choice of product mix, pricing strategy, or unit sales. Studies of workers might alternatively focus on team output, worker output, wages, incentive payouts or other compensation measure. This is an inevitable characteristic of IE research given its micro-level analysis, and it has the important implication that there can be many reasons why a practice with beneficial effects on some performance outcome might not be adopted by otherwise similar organizations. For example, some organizational design decision may promote higher levels of productivity, but the costs of adopting such features may be prohibitive for in certain settings.

[could add something like the following, but may go after the empirical, or in it – haven't edited that yet. It is about the tradoffs of getting homogeneous versus heterogeneity that is mentioned in the intro and I don't know where else we are discussion this. Steel has homogeneous production units but exog heterogeneity in the adoption of the HR practice as a function of profits. Mas and Moretti have heterogeneity in the production units themselves (i.e., the quality of the workers) but homogeneity in the adoption – that is, randomness in the assignment of workers to work shift teams.]

The treatment analysis described above assumes the researcher observes changes in an organizational practice in their data. That need not be the case: the researcher may compare the effects of different practices across plants (i.e., steel). If the plants are very similar, and we have controls for the differences, then the treatment may be considered exogenous or random because it wasn't the production process that caused the treatment to be adopted. In this case, the variation could come from variation in the profitability of adoption, due to transition costs, not variation in productivity due to adoption.]

III. Empirical Applications

This section considers several examples of IE style of research as illustrations of how insider knowledge and confidential data about a specific production process can be combined with a thoughtful application of econometric estimation to produce a rich understanding of the ways the various organizational policies and practices affect performance outcomes. We consider two main categories of studies – single firm studies in where worker selectivity must be carefully considered when estimating effects of the organizational policy, and multi-establishment studies where the central questions center on which worksites did and did not adopt the given policy.

A. Single Firm Studies: Estimation of Treatment Effects with Worker Selection

Many of the most compelling papers on the effects of organizational treatments on performance outcomes have been based on data on workers in a large firm before and after the adoption some new organizational practice. In some cases, it can even be argued that the firm's decision about which workers, work groups, or work sites to the treatment is random. As we review studies that illustrate IE research in such single firm settings, we will find that selection issues are still important even when the policy treatments have features of random assignment.

1. Insider Knowledge of Worker Selection after Different Treatments

Lazear (2000) tackles the fundamental economic question of the effects of monetary incentives on employee performance, specifically the effect of piece rate compensation (the treatment) on worker-specific output (the performance outcome). The sample he analyzes is drawn narrowly and pertains to all workers performing one specific job in the same firm, the Safelite Company, before and after the implementation of a piece rate compensation plan. Lazear describes the production process that generates the data for this firm-specific setting. A worker drives a truck to a customer's location and installs a windshield in a customer's car that has a broken windshield. Thus, the production function is worker-specific in this case. Lazear estimates performance equation (1) with number of windshields as the dependent variable, piece rate pay (versus hourly pay) as the treatment, and controls for the workers' tenure. Because each worker forms an individual production unit, he is able to estimate the production function for the firm. The average value of the worker productivity measure is 44% higher after the firm implements piece rate pay.

Lazear then re-estimates the effect of piece rate incentives controlling for fixed (worker) effects. From personnel records, Lazear identifies employees who worked for the firm in the period before the treatment but left before piece rates were implemented and employees who joined the firm after the piece rate plan began. In Figure 2, outcomes for the former group are represented by line BB while outcomes for the latter group are shown in line CC. As is clear from the figure, the fixed effects model that examines outcome changes only within the treatment group produces the gain of AA', while the OLS estimation across groups produces the gain of B'C. The fixed effects estimate of the effects of piece rates on worker productivity falls to 22%. The fall for the fixed effects is due to worker selection.

Specific features of this study allow us to make several key points about selection. First, from insider knowledge about the treatment, the firm randomly assigned workers by region to the piece rate policy. It would be tempting and natural to assume that the estimated treatment effect is the average treatment effect since workers (who are the production units) are randomly assigned to piece rate pay. However, selection also occurs through worker turnover which he can identify in the confidential personnel data. Thus, Lazear estimates a treatment of the treated effect. He shows that when production function (1) is estimated with worker fixed effects, the gains from piece rate pay among a fixed set of workers before and after the adoption of piece rate pay is 22%, while the overall change in output across all workers in the firm before and after the policy treatment is much larger, or 44%, a difference which Lazear attributes higher quality workers self selecting into the firm after incentive pay is adopted.

2.Selection Bias as an Advantage in Organizational Research

Lazear uses endogenous selection as a key advantage for learning more about how incentive pay affects worker performance in (certain) real organizations. We learn that when a firm introduces a new incentive compensation scheme, the firm gains for two reasons. Effort for a given worker rises in response to the incentive pay. But in addition, more productive workers self sort into the firm in response to the incentive pay scheme. In terms of Figure 2, workers of low quality BB leave and workers of high quality CC are hired after piece rates. Often in the treatment literature, analysts try to eliminate the effects of any selection issues when estimating the treatment effect since understanding the nature of the selection will not be the aim of the research. In organizational studies, we want to understand why the selection occurs, and the inside data and analyses often provide that understanding.

Still, Lazear's estimate of a 22% increase in worker productivity due to piece rates is a mean treatment effect conditional on selection. The paper estimates mean treatment effects that attempt to measure the return to incentive pay. But in organizations, the effect of one treatment is likely to be conditional on other treatments. In the case of the windshield repair study, the estimated mean effect would change if the firm also introduced a new method of hiring workers that changed the quality of new workers beyond what occurs under self-sorting, or of lowering turnover that retained more of the lower quality workers than was actually the case. Again, insider knowledge about the context is important for understanding how to interpret the empirical patterns in the study. The simplicity of the production function excludes some potential complicating effects – for example, there is no teamwork. Thus, by picking the homogeneous production environment and understanding it well, he is able to focus on two key changes – pay practices and worker selection – and not worry that certain classes of omitted variables such as other practices are responsible for the estimated results.

The above discussion considers one of the two key selection issues that arise in all models attempting to estimate the effects of some policy treatment on organizational performance – the selection of production units receiving the treatment. In the case of Safelite where workers are the production units, this is the selection of workers who do and do not work under the new piece rate treatment. The second selection issue concerns the treatment itself – why did Safelite adopt piece rate pay when it did? According to the study, Safelite adopted piece rate pay in part because they purchased new information technology that kept immediate records of individual output. One can consider adoption to be a result of an exogenous technology shock. One can thus argue that piece rates would have had similar effects in earlier time periods but the costs of the necessary performance measurement would have made piece rate pay unprofitable. While the study also reports that incentive compensation remained uncommon for managers and workers in other occupations at Safelite whose output is still difficult to measure, the important point to note here for IE research more generally is that a single treatment applied at one point in time across a broad group of employees does not permit a model of the selection of treatment itself.

3. Insider Knowledge and Data to Assess Causality Behind Treatment Effects

Persuasive studies not only measure differences in outcomes for a given set of subjects under different treatment conditions but they also follow their subjects closely in a way that provides richer insights. One can imagine a medical study comparing those treated by a new drug versus those receiving a placebo. If the drug has desirable effects, one could also imagine the benefits from collecting additional data that tracked the treated subjects that offered insights into how the new drug produced the effect that it did. IE studies can also address the analogous issue for organizations and provide insight into mechanisms through which some treatment had a measurable effect on performance.

Bandiera, Barankay, and Rasul (2005) also use firm-specific data to study the effects of piece rate incentive pay, but in the context of a fruit picking farm. The farm begins with a relative payment scheme in which each worker earns more per unit of fruit if he picks more than his peers, but as overall productivity goes up, the farm lowers the wage per unit of fruit picked. The farm switches to a flat piece rate with a fixed fee per piece picked. Productivity of workers under the plain piece rate plan is over 50% higher compared to the 'relative' payment scheme. The paper is interesting and informative not just because they follow the same workers in the same work environment under two different compensation "treatments," but Bandiera et al. go on to identify why the flat piece rate pay is more productive. By using data on who is friends with whom among the pickers, they show that under relative pay workers hold back on their effort when they are concerned about hurting their friends: if a picker works too hard it lowers the pay of their friends. The key to the study of organizations is not just that piece rate pay is more effective, but that using inside knowledge and data allowed the analysts to document why workers behave differently under different treatments, and thus establish why one treatment leads to a different outcome than another. A richer theory of an organization's productivity, according to these results, needs to acknowledge a role for relationships, in this case friendships, among co-workers.

Mas and Moretti (forthcoming) also analyze worker-specific productivity data from a single firm and conclude that relationships among workers are an important determinant of how the workers perform. Here, the data come from six stores of a single supermarket and provide measures of how fast checkout clerks process customers' items. The treatment that Mas and Moretti consider is the introduction of a new worker onto a shift and they test how the other clerks' performance responds when the new worker's typical checkout speed is relatively fast or slow.

They first document that the introduction of a worker whose typical level of productivity has historically been high increases the speed of the co-workers. They go on to document theoretically interesting variations in the magnitude of this kind of productivity spillover among co-workers. In particular, they find that the biggest increases in checkout speeds of co-workers are concentrated among the slower clerks. Thus, a store maximizes overall checkout speed by maximizing the diversity of typical productivity levels of its clerks. They then use inside information on the physical layout of the checkout stations to further document that increases in speed of a slower clerk occurs only when the faster co-workers face him. Of the various theoretical models that could explain such productivity spillovers, this more specific empirical pattern uncovered in this IE study indicates that monitoring among co-workers is important.

4. Insider Knowledge about the Selectivity Behind Which Workers are Treated

In an example like Lazear's study of Safelite, we noted the inherent limitation in the researcher's ability to examine certain aspects of selectivity in the organizational treatment. For Safelite, in regions where the piece rate system was implemented, why was it rolled out when it was? In other studies, however, when there is endogenity of the organizational treatment itself, analysts can develop insider data document important economic aspects about that selectivity.

One example that illustrates this aspect of IE research in a single-firm context is Hamilton, Nickerson, and Owan's study (2003) of teamwork in apparel production. The use of teams in many forms is becoming increasingly common in many firms. Theoretical models of teams posit that there are gains to teamwork when individuals on the team can work in a complementary fashion – when the multiplicative interaction between team members raises output more than the sum of the individual parts. At the same time, productivity can decline in moving from individual work to team production when free rider issues are important. Since some employees in a given occupation or industry work in teams while others do not, the question of the effects of teams on performance is one that IE style research could usefully address. Hamilton, Nickerson, and Owan address this increasingly important question by analyzing data on worker and work group productivity in apparel making. Their empirical analysis begins by documenting an increase in productivity after workers shift to teamwork. This finding suggests that the gains from worker skill complementarities or peer monitoring more than offset the potential losses of output from a free-rider effect in these teams.

Their analysis goes further and uses rich data on the workers to provide a deeper understanding of the team selection process. In particular, their study provides an example of the endogenous selection of production units that undertake the teamwork "treatment" – the decision by workers to work as a team is endogenous. Since the gains to teamwork are hypothesized to be a result of skill heterogeneity, selection of team members is critical. The composition of the team should affect the estimated returns to teamwork. However, there could also be adverse selection in this case, where the best sewers should avoid teams and stick with individual piece rate pay. How do these competing forces play themselves out?

Using insiders' knowledge about the organization and detailed confidential data on the history of workers' performance before and after teams, Hamilton et al are able to test these important selection hypotheses by looking at the quality of the individual performers on the teams and the pre-team performance of workers who form the teams. Their analysis documents different ways that worker selection on the teams is important to the estimated differences in productivity before and after team formation. First, they document that better sewers are more likely to form teams earlier. Thus their results follow the general pattern laid out in basic Figure 1 illustration (which when adapted to the specifics of this study, can be represented by Figure 3). Those who choose teams first have higher gains than those who choose teams later because the early choosers are more productive in terms of both their level of output and in the changes in output after moving to teamwork. These authors identify still other selection issues that are important for understanding the variation in the performance effects after the formation of teams. Team productivity grows more when the team is comprised by individuals who exhibited greater heterogeneity in individual performance prior to the use of teams, a result that is similar in spirit to the conclusion above about productivity spillovers among supermarket clerks (Mas and Moretti, forthcoming).

In all of the subsections above, data from a single firm are used to understand the size of the productivity gains of a new organizational treatment in a way that accounts carefully for issues related to the selection of workers.[7] Estimated productivity differences are affected by workers' individual quality (Lazear, 2000), the workers' quality prior to forming teams (Hamilton et al., 2003; Mas and Moretti, forthcoming), and detailed characteristics related to social relations among workers (Bandiera et al., 2005). There is no one 'treatment effect.' The magnitude of the estimated effect of the treatment varies as a function of the selection of workers with different characteristics. Organizational research does not just try to difference out the effects of such selectivity. Instead, it must consider differences in the estimated size of organizational treatment effects under different econometric models to gain insights about the nature of the selectivity and about how these treatments affect the performance of firms.

B. Effects of Treatments Applied in Different Organizations

The above studies focus on the estimated effects of a change in an organizational practice within one firm. These within-firm models seek to understand the impact of the organizational change by comparing workers who are and are not subject to the organizational change in ways that carefully consider the endogeneity of which workers are covered by the given change. The estimation in these studies ultimately rely on a difference-in-difference model, for workers who do and don't adopt.

Of course, by looking only within one firm, the models cannot make the availability of the practice itself endogenous. In Hamilton et al (2003), insider data help identify which workers for teams once the firm decided to adopt their teamwork policy. The single firm study however cannot identify why this particular apparel making firm adopted the overall firm-wide policy about teamwork while another apparel maker did

[7] In other studies in which the person is the production unit, worker selection affects the impact of the treatment. For example, Landers, Rebitzer, and Taylor (1996) model the hours of lawyers, showing that large firms using tournaments produce longer work hours. Here selection is also important. Given shared compensation under the partnership structure of law firms, adverse self selection of low quality lawyers into high hours firms is a concern. These firms must screen workers by using their record of billable hours in the promotions decision.

not,[8] or how much some other firms might gain. In terms of the econometric estimation, using within-firm data in which there is only one firm-wide decision to adopt the overall firm policy, the analyst cannot estimate difference-in-differences with the treatment decision as the unit of analysis rather than the worker.

For example, to develop even richer estimates of the effects of piece rate pay on performance than are developed in the studies by Lazear (2000) or Bandiera et al. (2005), we would like to have observations on the introduction in piece rate pay in organizations that are at the margin in introducing it, find a matched sample of firms that are also at the margin but did not introduce it (for exogenously determined reasons), and then estimate difference-in-difference models to compare the effects of piece rate across the two settings. No study using existing firm-level data can achieve this. However, this objective is an important benchmark to keep in mind when considering existing studies that do examine the effects of different organizational treatments across organizations. We now consider the estimation of the effects of organizational treatments on performance outcomes using insider data from many organizations.

1. Insider Data on Practices Across Organizations and Treatment of the Treated Estimates

Consider again the effects of teams on productivity analyzed in Hamilton et al. (2003), but now the unit of observation is the plant, not the individual. In terms of the basic comparisons illustrated in Figure 1, there are two sets of plants: those that have adopted teams, and those that have not. As before, the performance of both groups will change over time, with Figure 1 assuming gradual performance increases for both groups. The team-treated group experiences a big gain in performance following treatment. We do not observe the performance gain for the non-treated group, but the figure assumes that the plants that did not adopt teams would experience smaller gains from adopting teams since they did not adopt them.

Figure 1 suggests the possibility of differences-in-differences estimation with fixed effects models. Clearly full sample OLS estimates without controls for plant fixed effects are biased. If we estimate such an OLS model, the gains to teams will appear bigger, because the treated group has a higher level of performance that would be

[8] Cross-establishment research on HRM practices including teamwork among apparel makers does offer information on this kind of selectivity. See Dunlop and Weil (1996) and Berg, Appelbaum, Bailey, and Kalleberg (1996)

erroneously attributed to treatment. However, difference-in-differences assumes that the treatment sample is the same as the untreated conditional on observed X's. To be specific, the difference-in-differences estimation is appropriate for Figure 1 if we assume that the estimated effect of time can be estimated parametrically so that the non-treated group has a lower growth of productivity over time ensure than the treated group. That is, as in equation (4) above, the βX varies across the treated and untreated subsamples, and the parametric estimation of βX captures these differences. If instead, the slope of the non-treated is lower than the slope of the treated (as in Figure 1), and this is not measured well with the observed X's, then the gains to teams will be overestimated because the higher learning curve of the group that selects teams will be attributed to the team effect, not to differences in selection. If there are sufficient data, a reasonable choice for estimating the treatment of the treated without this bias is to limit the sample to only those who select teams.[9]

Using longitudinal data on the sample of U.S. steel minimills, Boning, Ichniowski, and Shaw (2007) estimate the effects of teams on measures of productivity in this industry according to this procedure that focuses on productivity changes within the group that adopted teams. Comparing estimates of the effects of teams in models that do and do not restrict the sample to only those mills that introduced teams, they find that estimated treatment effects are very similar. Apparently, in this case, there is little difference in the growth rates of productivity for the non-adopters and the adopters prior to adoption.

*Shor*ten this paragraph:

Insider knowledge here is used to rule out certain explanations for differences in adoption of practices like teams across plants. The "special knowledge" that insiders offer can also be used to identify factors that do help explain why some firms were 'treated' and some were not. In the case of minimills, site visits and interviews revealed that teamwork and problem solving was more important in mills that made more complex

[9] One of the first "organizational treatments" to receive considerable attention in empirical econometric studies was union status, and this literature offers examples of treatment of the treated estimation within samples of plants that became unionized. Clark's analysis (1980) of cement plants is an early exemplar of an IE approach that generates compelling treatment of the treated estimates of the positive effects that unions can have on productivity. This research also reflects the use of interview evidence with cement plant managers to uncover mechanisms that explain the estimated productivity effects.

products. Without site visits, one might simply have assumed that homogeneous production facilities like these, all within the exact same narrow standard industrial classification, is also homogenous. However, Boning et al present insider data that reveal that, for the four different classes of minimill products observed in these plants from the least complex product type to the most, the frequency of teams was: 23% of all plants in product class one (plants making the simplest minimill steel products); 33% of plants making product class 2; 67% of plants making product class 3; and 100% of the plants making the most intricate type of steel products. These patterns then lead to additional theoretical analyses to explain why problem solving teams are more commonly adopted when organizations make more complex products and empirical analyses that examine changes in the productivity metrics within plants making different kinds of products. The econometric productivity models go on to demonstrate that the introduction of teams in minimills making more (less) complex products leads to larger (smaller) changes in mill productivity, just as the team adoption patterns would predict.

The example here offers an interesting point about the value of IE research for addressing the two basic questions about the adoption and subsequent performance effects of organizational practices. Broad-based surveys of establishments in many industries may be able to identify that an establishment in one industry like minimills is more or less likely to have teamwork policies than an establishment in some other industry. Such research may also be able to identify certain establishment-level characteristics correlated with team policies. However, research following an IE approach was able to identify a very powerful predictor of the adoption of teamwork policies based on field work that also revealed that problem solving activity by teams was particularly useful for mills making more intricate steel products. IE research did not simply rule certain classes of explanations for the adoption of teams, but instead presents concrete data on what led certain minimills and not others to be the adopters of teams. This insider insight is in turn permitted the estimation of specific models of minimill productivity that document how the productivity impacts of teams varies across these organizations.

2. Insider Data on Practices Across Organizations with Exogenously Determined Treatments

While policies and practices that vary across plants in these intra-industry samples are not randomly assigned by plant managers, some IE research that considers the effects organization-specific practices on performance concludes that estimated changes in productivity observed among policy adopters might well be a reasonable estimate of the change in productivity that non-adopters would experience. Ichniowski, Shaw, and Prennushi (1997) study the effects of different systems of HRM polices on productivity in production lines in integrated steel mills. This analysis also includes estimates from econometric models with controls for plant-specific fixed effects and finds significant productivity benefits associated with adopting new HRM systems.

Their conclusion for differential adoption of these productive HRM systems across similar plants is that certain plants have higher transition costs of adopting these practices. In particular, the most productive HRM systems were more likely to be observed in "greenfield" lines starting up production and in older "reconstituted" lines that started up operations by new owners after being closed temporarily than they were in older lines that operated continually since inception (Ichniowski and Shaw, 1999). While the desired counterfactual on the productivity effects among non-adopters can never be observed, the authors conclude in this case that the most compelling conclusion for this setting is that adoption is governed not by the expected productivity gains from the new HRM practices, but by differences in transition costs across sites. Insider insight was again essential for pinpointing the specific empricial pattern that all of the lines adopting the most prodcuctive HRM systems were new startup line or old lines that had been closed and then reopened by new owners. The authors conclude that the new HRM practices could increase productivity in other continuously operating lines, but it would be more costly to implement these new work systems in an old, but continuously operating, worksite.

Perhaps the most important point to make for this analysis of IE methods is that insider insights are critical for indentifying fine-grained distinctions that differentiate adopters of organizational policies from very comparable non-adopters. Even survey questions about age of plants would not capture the distinctions in this industrial setting. Furthermore, the IE research here directs theoretical analysts to consider transition costs related to social relations and attitudes in the workplace rather than the age of capital equipment as the more important determinant of HRM adoption. Finally, the example here illustrates the importance of remembering that performance metrics in micro-level studies do not measure overall profitability to the firm. In this case, the study argues that HRM adoption would raise steel-making productivity among non-adopters, but transition costs would make this decision an unprofitable one.

Mas (2005) provides an illustration of IE style research from a very different setting than manufacturing in which exogenous determination of the treatment may be important in interpreting observed empirical patterns. Mas looks inside public sector organizations and examines how well New Jersey police perform in solving crimes. The outcomes in this study are police pay and the treatment is the use of interest arbitration to determine the police union contracts. Mas finds that when there is a large loss for the union workers in arbitration – measured by larger differences between the union demand and the arbitrator award – workers are subsequently less productive. In interpreting these patterns, Mas argues that, while arbitration may be endogenous to various firm-specific factors, the wage outcome of arbitration can be considered exogenously determined and this is what drives his organizational change results. Ultimately, he concludes that it is not the level of the wage that affects policemen's performance it is the wage loss relative to expected wages. Like other IE research, the study pertains to the very specific group of unionized police and those subject to state law permitting contract arbitration. But the patterns observed for this setting suggest a specific conclusion that workers are motivated to perform in part by the difference between wage outcomes and their wage expectations that may apply to other settings.

C. Addressing <u>Omitted Variable Bias in IE Estimation</u>

In any econometric study of the effect of some treatment condition on a performance outcome based on non-experimental data, one can always imagine reasons why estimated effects of the treatment could be attributable to some factor correlated with both the treatment and the performance outcome that is omitted from the analysis. The selection model presented in section II and represented in Figure 1 encapsulates this concern. In particular, one may still argue that treatment estimates in studies reviewed above (e.g., fixed effects estimates of the effect of teams on productivity in before/after comparisons in a sample of team adopters) are still overestimates of the treatment-of-the-treated effect. One need only argue that some other productivity enhancing factor was adopted around the same time as teams in these plants.

Certain advantages of IE research in considering omitted variable issues are clear. Extensive interviews or field research, brings a richer understanding of the determinants of the organizational treatment and of the performance outcome. Furthermore, omitted variables are less likely in the narrowly drawn samples of homogenous production that characterize IE studies.

1. Measuring Unmeasurables

IE research provides examples of a direct approach for considering whether estimated effects of organizational treatments might be due to omitted (time-varying) variables. In the studies considered in section III.2 that analyze data from multiple organizations, one could reasonably be concerned that more productive organizational practices simply capture effects really attributable to better managers. Better managers adopt different practices than do other managers, such as teamwork policies, but superior performance in the organizations is due to the managers and not the practices. In section II presentation, management quality enters the adoption equation (3) as a Z variable, and also is an element of the X vector in the performance equations.

Insider access to organizational information data can allow researchers ways of addressing such potential sources of bias in estimates. For example, in the case of HR practices adopted in integrated steel mills, Ichniowski, Shaw, and Prennushi (1997) create explicit measures for the time periods when different managers were in charge of the steel making operations, and their productivity models are thus able to include direct controls for a factor like managerial quality that is often an "unobservable variable." The combination of narrowly drawn samples that is sometimes coupled with unique opportunities to identify and incorporate measures of factors that would remain unmeasured without the special knowledge that an IE approach can often allow IE research reasonably convincing ways for addressing omitted variable concerns.

2. Was There More than One Organizational Treatment?

Milgrom and Roberts (1990) offer a simple description of operations of contemporary manufacturing operations as an illustration of a potentially important point that would affect IE research seeking to estimate the effect of an organizational treatment on performance outcomes – many organizational practices may adopted as systems of practices because of complementarities among the practices. For example, if a plant decides that introducing teams is optimal, it may also institute additional training, more careful selection of workers for attitudes toward teamwork, information sharing among employees and other practices that support effective problem solving by teams. If these organizational practices are complements, then all should enter the production function multiplicatively. That is, if you are looking at the effects of teamwork on productivity, but omit a complementary work practice, the return to teamwork as a lone policy treatment will be biased upward.

The possibility of complementarities among organizational practices counsels the analyst to investigate the nature of the organizational treatment carefully. This task is perhaps more manageable within narrowly drawn IE-style samples since specific production settings like those considered in this study, like windshield replacement, apparel making teams, fruit picking, minimill steel lines, and others, may have a more limited set of relevant organizational policies. Existing IE research suggests that complementarity among organizational practices is an important concern with regard to complementarities among HRM practices (Ichniowski, Shaw and Prennushi, 1997); among new HRM practices and just-in-time production practices (MacDuffie, 1995); and among new information technologies and new HRM practices (Breshnahan, Brynjolfson, and Hitt, 2002; Bartel, Ichniowski, and Shaw, 2007).

IV. Obtaining Insider Data

Several observations can also be made based on this review about alternative methods researchers have used to access insider data.

A. <u>Insider Data from Within Firms</u>

First, the seemingly most common kind of insider data used in IE studies is data from a single firm. Several of the studies highlighted above were based on datasets the researchers had obtained from a firm.[10] Within the firm, IE research focuses on distinct production processes and within any firm there can be many different output generating "production functions." The production function could be individual workers, work groups within an establishment, different establishments within a division, and so on. With increasingly sophisticated information collecting and storage technologies now common place in many businesses and firms, and the greater likelihood of cultivating insider access through one or two key officers of a firm, we see this as a particularly promising area for growth of new IE research.

Second, the number of studies about performance differences across different plants or other micro-level observations that share a very specific production process based on personally assembled industry-specific data sets has also grown in recent years. This approach requires data from many firms operating in the given industry. While this approach obviously requires significant time and effort, it can result in exactly the kind of rich understanding of an entire industry context that IE research strives for. Methods for managing this approach to data collection and its costs are therefore worth considering. One development here is the commitment to network building in well-established industry councils supported by the Alfred P. Sloan Foundation. The number of industry-specific study councils supported by the Sloan Foundation has now grown to twenty-six, some of which have been in operation for many years. From a practical standpoint, this approach to research can benefit from the support of a small board of industry advisers and the existing councils could be beneficial in facilitating data collection.

B. <u>Insider Data Available Through Custom Industry-Specific Surveys</u>

Data collection for industry-specific research across multiple organizations and firms that encompasses many production units does not have be gathered from site visits to each production unit. Several alternatives are possible for managing the costs of IE

[10] For tips on obtaining data from a firm, see Helper (19xx) and Freeman and Shaw (2008). Our primary advice is: find a personal contact; make some part of it valuable to the firm; emphasize the teaching and educational value; and consider the legal issues of data confidentiality early in the project. Working with large firms can be the riskiest; smaller firms or privately held firms are more likely to approve publication of your results, though there are certainly exceptions to this.

data collection. In our own work, we conducted five plant tours in one industry with the aim of developing an industry-specific survey to study the effects of new information technologies and new HRM practices on productivity and worker skill requirements (Bartel, Ichniowski, and Shaw, 2007). The field visits and interviews did not generate the confidential IE data, but it did provide the "special insider knowledge" that we used to develop a survey in the language specific to the technologies and operations of that industry. Contacts for the survey sample were obtained through public data sources on business establishments in the U.S., and the survey was conducted by a telephone by a survey research firm. This approach again involves data collected personally by the analyst, but in a way that economizes on the costs of data collection by reducing the number of site visits.

Public data sets often lack information on organizational practices, so when possible, researchers may survey firms for their organizational practices information and then match that to existing production data from the firm. Bloom and Van Reenen () provide an excellent example of this.[11]

Other studies that fit the section II criteria for IE style research have made effective use of existing data on workers or establishments within certain industries. The Census Centers in the U.S. provide access to surveys with detailed information on specific industries. Perhaps the best known exemplars here are studies of the effects of the adoption of new information technologies on trucker productivity (Hubbard, 2003; Baker and Hubbard, 2003; 2004) and occupation-specific research on lawyers (Garicano and Hubbard, 2003).

Our data primarily come from the web pages of law firms.3 We downloaded information from
lawyer bios for all lawyers at fifty-eight large American law firms. The information we collected
included the lawyer's name, which law school she graduated from, the date of graduation, and
the person's rank at the law firm.

[11] Bloom and Van Reenen (2008) estimate models of organizational performance at the firm-level and thus are not technically an insider study, but their careful analysis of the evolution of manufacturing to lean "modern manufacturing analysis" is conducted with a survey approach that is an exempler of very carefully insider knowledge. See also B and Lynch () for a firm-level survey. See Hoxby (2002) for a firm-level survey, but conducted with schools that are homogeneous at the firm level.

C. Using Data from the Internet or Industry Publications

The Internet is proving to be a valuable source of data for industrial organizational economists, and in the future may well be of increasing value to organizational economists. Most firms do not post employee records on the Internet, but there are exceptions when firms want to display their human capital to attract business or new hires.[12] Most individuals whose personal information is on the Internet generally do not list their firms, but such data links may be available in the future. In the past, some industries have provided data on firms within their industry to attract business, as in the case of the franchising industry.[13]

C. Using Existing General Survey Data to Create Insider Studies

Finally, agencies of the U.S. government, as well as government agencies of many European countries, have begun to make available what have been labeled employer-employee matched data sets, often on a limited basis to researchers inside the government. These data sets contain administrative data from firms and on every worker in the firm, including data on wages, demographic features of the employees, and at times their occupations and promotions. In U.S data, workers are tracked over time between firms. These data sets are very broad in scope and may not contain the kind of rich institutional detail on many organizational characteristics an IE researcher would desire. However, in some cases, these data have been linked to specific production data of firms or to surveys of organizational practices of firms. Because these data sets are so large (often covering all individuals in the country for large sectors), subsets of the data can be used for industry analysis. For example, in the U.S., data have been collected on the performance and internal structure of firms for trucking, semiconductors, and software industries. (Abowd, Haltiwanger, Lane, 2006; Andersson, Freedman, Haltiwanger, Lane, and Shaw, 2006).[14]

[12] See Oyer and Schaeffer (2007) and Oyer (2008) for examples of law firms that post their lawyers biographies, and the use of vitas posted by academic economic departments.
[13] See Lafontaine and Shaw (1999, 2005).
[14] See Baker and Hubbard (), Hubbard ().

D. Using Data from Regulated Industries, the Government Sector, or Developing Countries

Manufacturing has often been the source of data for IE studies because output is readily measured and data has been obtained more easily, but service industries are increasingly a large part of the economy and data is increasingly available. If we combine the government sector and the regulated sector (like health care) of the economy may make up to as much as 40 percent of the economy, and policymakers are inceasingly concerned with making these sectors more productive. We focus on education and health care as examples of increasingly important regulated industries. Examples of insider work that has been done on public data include: Bushnell and Wolfram's study (2006) of the effects of variation in individual performance on organizational productivity in electricity generation; in the legal sector for judges (Lim, 2007); in the private sector setting of the hotel industry, studies by Kalinins (2006) and Ingram and Roberts (2000) analyze the effects of social networks on performance of hotels by augmenting publicly available data on hotels with their own personally collected data on networks.

The education industry is currently undergoing external shocks that is forcing the industry to become more productive, just as manufacturing experienced external shocks twenty-five years ago.[15] The productivity of the education sector, in terms of output per dollar input, has declined over time in the U.S., so the external shocks to the industry are in the form of competition from abroad – other countries have caught up with our educational levels or excellence – and from the rising educational needs of our knowledge-based economy. Policymakers are responding by demanding that education be more competitive: they are imposing the use of vouchers or charter schools or accountability standards. Organizational practices are changing slowly in response, as schools introduce more performance-based pay, or other changes in organizational practices. Some organizational changes are imposed by regulators, as in the form of mandatory class size reductions.

Experiments in organizational change are thus imposed on schools, and data is available from public schools to test the effects of the organizational change on the

[15] Hoxby makes this point () and discusses the education industry's organizational changes in response to these shocks.

productivity of teachers. Due to the immense quantity of this literature, we will not summarize the results or the types of data. Unlike most organizational datasets, in education there are true experiments, such as the imposition of a teacher incentive pay, or reductions in class size, with output often measured as changes in test scores.[16] Given the increasing availability of data on teachers, we can also assess whether hiring or training teachers to be higher quality would increase their performance.[17] There is also a large literature that does not use experimental data, but that uses either natural experiments, or more standard non-experimental data.[18] Education is an important area for future research, as more experiments are likely and more public datasets with information on teachers' pay and performance will be increasingly made available.

The health care industry is also an industry which has been external shocks that have resulted in differential adoption of new organizational practices, but for which we still often lack the data to directly test whether changes were effective in raising productivity. Here the unit of analysis is likely to be a disease-specific outcome that is measured at either the hospital level or the patient level. The shocks are new technology, or new 'report cards' legislated to improve patient information, or demand shocks arising from the payment schemes of health care insurers. The empirical literature can often measure the shock and the outcome from publicly available data, thus forming a natural experiment, but lacks data on the changes in organizational practices that facilitate the hospital's change in performance. Therefore, the researchers are likely to use interview evidence or their insider knowledge of the diseases or hospitals to provide their interpretation of the likely changes in organizational practices. These explanations can be persuasive.[19] Thus, this literature has the advantage of providing many natural experiments that eliminate the selection problems, but must use their insider knowledge to interpret the results. In some instances, researchers have obtained high quality data when medical providers have introduced a change, such as incentive pay or new

[16] For work on incentive pay, see Lavy (2004), Eberts, Hollenbeck and Stone (. For class size studies, see, for example, Krueger (1999), Angrist and Lavy (1999), and Hanushek ().

[17] For teacher quality effects, using the public Texas data on teachers, see Hanushek ().

[18] For natural experiments on organizational practices and student outcomes, see Hoxby (2002) and for class size, see (). For non-experimental data, see the reviews by Hanushek (0 and Hoxby (

[19] We make no attempt to summarize this extensive literature, but for examples of insurance payment schemes see Acemoglu and Finkelstein (2006) and see Dranove, Kessler, McClellan, and Satterthwaite (2003) for evidence on heath care report cards.

information technologies, and have estimated the impact.[20] The change in organizational practices is very pervasive.[21] Therefore, as more hospitals install more high quality internal databases that measure output, researchers will gain new opportunities for analysis.

Health care and education share a common feature: regulators of these industries have enacted regulations to make them more productive – to monitor the quality of their output. Thus far, much of the empirical literature in these industries has tested whether these changes have made the industries more productive, as in testing, for example, whether vouchers raise teachers' performance. It is undoubtedly the case that these industries must make organizational changes to become productive: teachers are not likely to respond to regulations if they are not given the incentive to do so. The empirical literature has often not yet obtained the data to test whether organizational changes have been made to raise performance, or the sectors have not made the needed changes to increase performance. As both these changes occur in the future, there will be more data to test the impact of change.

Developing countries are making changes in how they provide health care and education, as well as other sectors, and experiments are more likely. Once researchers set up connections in developing countries, by working with NGO's or foundations, the cost of an experiment in organizational change is quite cheap. There is therefore an exponentially growing literature on how we can provide incentives for the success of the provision of, for example, malaria drugs or teacher attendance.[22] While we might not typically think of these examples as tests or organizational change, they undoubtedly are, and as organizational economists, there is no doubt that learning about the effectiveness of practices in developing countries is of great social value. Organizational issues are not limited to the developed world.

---

[20] See Nagin, Rebitzer, Sanders, and Taylor (2002), Gaynor, Rebitzer, and Taylor (2004), and Encinosa, Gaynor and Rbitzer (2007) for incentive pay and see Athey and Stern (2002) for information technology.

[21] See Rosenthal, et.al. (2004) for survey of the prevalence of physician bonus schemes for patient outcomes.

[22] See Kremer (2003), Duflo, Hanna, and Ryan (2005), and see Duflto, Glennerster, and Kremer (2006) for a review of methods and literature in this area.

V. Conclusion

We introduced our summary of Insider Econometric studies of the effects of organization-specific policy treatments on performance outcomes with the Griliches' observation about the need for empirical research using the "right data" to catch up with advances in theory and technique. The discussion here does not offer a boiler plate check list of the characteristics that several informative IE style studies all share in common. But the discussion does offer a road map with a destination – reasoned and rigorous tests of hypotheses about the effects of organizational policies and practices on performance outcomes. Since true random assignment of policies to organizations and of workers to organizations does not occur in real organizations, and since important economic insights about organizations are gained by evaluating the precise ways in which selectivity in these treatments does take place, it is important for the analyst to anticipate these concerns in developing and collecting the quantitative data used in IE research.

IE research analyzing narrowly drawn samples by its very nature cannot produce findings that can be broadly generalized across many different industrial settings with their own distinctive production processes. But, several important advantages of IE style research for assessing treatment effects should be noted. First, studies with confidential data on detailed worker or organization characteristics will provide better data for constructing the selection equations in the model. Second, even when confidential sources do not provide quantitative data for modeling the selection process, the insiders' perspectives often provide an understanding of the selection process and guide the analyst toward sensitivity analysis that tests whether returns to the treatment vary within various identifiable subsamples of observations. Third, models of the performance outcomes can be more thoroughly understood and specified. Finally and perhaps most importantly, it is precisely these kinds of rich data that are required to test many recent theories in the field of organizational economics. The empirical results of this style of insider research can prove to be especially persuasive and lead to a better understanding of the economic diversity of organizational practices and performance and ultimately a richer and more accurate theory of the firm.

Duflo, Esther, Rema Hanna, and Stephen Ryan, (2005), "Monitoring Works: Getting Teachers to Come to School," (with Rema Hanna and Stephen Ryan), NBER Working Paper No. 11880, 2005.

Duflo, Esther, Rachel Glennerster, and Michael Kremer (December 2006). "Using Randomization in Development Economics Research: A Toolkit,". NBER Technical Working Paper No. 333.

Kremer, M. (2003): \Randomized Evaluations of Educational Programs in Developing Coun-
tries: Some Lessons," *American Economic Review*, 93(2), 102{106).

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

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

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

Eberts, Randall, and Kevin Hollenbeck, and Joe Stone, "Teacher Performance Incentives and Student Outcomes," *Journal of Human Resources*, 57 (4): 913-927.

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.

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

Caroline M. Hoxby, 2000. "The Effects Of Class Size On Student Achievement: New Evidence From Population Variation," *The Quarterly Journal of Economics*, MIT Press, vol. 115(4), pages 1239-1285, November.

Caroline Hoxby, 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working Papers 7867, National Bureau of Economic Research, Inc.

Lavy, Victor, (2002). "Evaluating the Effect of Teachers' Group Performance Incentives on Students Achievements," *Journal of Political Economy*. December. 10(6): 1286-1318.

Hoxby, Caroline (2000), "The Effect of Class Size on Student Achievements: new Evidence from Population Variation," *Quarterly Journal of Economics*. 115(4): 1239-1285.

Angrist, Joshua and Victor Lavy, (1999) "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*. 114(2): 533-575.

Jacob, B and S. Levitt (2003). "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*. 118(3): 848-878.

Kreuger, Alan (1999), "Experimental Estimate of Education Production Functions," *Quarterly Journal of Economics*. 114: 497-552.

Lim, Claire (2007) "Turnover and Accountability of Appointed and Elected Judges" (Ph.D. Dissertation), 2007

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

Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector
NBER Working Paper 12254, May (2006). Revised and Resubmitted, Journal of Political Economy.

Rosenthal, Fernandopulle, Ryu Song, and Landon (2004). "Paying for Quality: Providers' Incentive for Quality Improvement," *Health Affairs* 23(2): 127-141.

Oyer, Paul (2008) Ability and Employer Learning: Evidence from the Economists Labor Market, *Journal of the Japanese and International Economies,* 22, June , 268-289.

____________ and Scott Schaefer (June 2007) Personnel-Economic Geography: Evidence from Large US Law Firms. Working paper.

References

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

Andrews, Donald W. K. and Marcia Schagfans (1996) "Semiparametric Estimation of the Intercept of a Sample Selection Model," *Review of Economic Studies,* vol. 65, pp 497-517.

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*, vol. 45, pp. 488-506.

Asch, Beth J. (1989) "Do Incentives Matter—The Case of Navy Recruiters." *Industrial and Labor Relations Review*, vol. 43 (Supplement 1), pp. s89-s106.

Ashenfelter, Orley (1978) "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, vol. 67, pp. 648-660.

Ashenfelter, Orley and Hosken, Daniel "The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies," *Working paper*, 2006.

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

Autor, David, Frank Levy, and Richard Murname (March 2001) "The Skill Content of Recent Technological Change: An Empirical Investigation."

Autor, David, Frank Levy, and Richard Murname (March 2001) "Upstairs Downstairs: Computers and Skills in Two Floors of a Large Bank."

Autor, David H. (2001) "Why do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics*, vol. 116, pp. 1409-1448.

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

Bandiera, Oriana, Barankay, Iwan, and Rasul, Imran "Incentives and Favoritism in the Workplace: Evidence from Personnel Data," *Working paper*, 2005.

Bandiera, Oriana, Barankay, Iwan, and Rasul, Imran (2005) "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *Quarterly Journal of Economics*, vol. 120, pp. 917-692.

Bandiera, Oriana, Barankay, Iwan, and Rasul, Imran "Social Connections and Incentives in the Workplace: Evidence from Personnel Data," *Working Paper*, 2006.

Bandiera, Oriana, Barankay, Iwan, and Rasul, Imran (2006) "The Evolution of Cooperative Norms: Evidence from a Natural Field Experiment," *Advances in Economic Analysis & Policy*, Vol. 6: No. 2.

Bartel, Ann (2003) "Human Resource Management and Organizational Performance: Evidence from Retail Banking" *Industrial and Labor Relations Review*, vol. 57, no. 2.

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw (2007) "How Deos Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills" Quarterly Journal of Economics, 122(4), pp. 1721-1758.

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

Berg, Peter, Eileen Appelbaum, Thomas Bailey, and Arne L. Kalleberg (2000) "The Performance Effects of Modular Production" in *The American Workplace: Skills, Compensation and Employee Involvement*, eds. Casey Ichniowski, David Levine, Craig Olson, and George Strauss (Cambridge University Press).

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

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

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

Black Sandra, Lisa Lynch and Anya Krivelyova (2004) "How Workers Fare When Employers Innovate", *Industrial Relations,* vol. 43(1), (January)

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).

Boning, Brent, Casey Ichniowski, and Kathryn Shaw (2007) "Opportunity Counts: Teams and the Effectiveness of Production Incentives," *Journal of Labor Economics, vol. 25, no. 4, pp. 613-650.*.

Breshnahan, Timothy, Erik Brynjolfsson and Loren Hitt (2001) "Information Technology, Work Organization and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics*.

Cappelli, Peter, and David Neumark (2001) **"**Do High-Performance Work Practices Improve Establishment-Level Outcomes?" *Industrial and Labor Relations Review*, Vol. 54, No. 4 (July) pp. 737-775

Chevalier, Judith A. and Morton, Fiona Scott "State Casket Sales Restrictions: A pointless Undertaking?" W*orking paper*, 2006.

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

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

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

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

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

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).

Encinosa , William E. III; Gaynor, Martin and Rebitzer, James B. (1997) "The Sociology of Groups and the Economics of Incentives: Theory and Evidence on Compensation Systems," Agency for Health Care Policy and Research.

Fernie, Sue and Metcalf, David (1999) "It's 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*, vol. 13, pp. 385-412.

Freeman, Richard B. and 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*, vol. 44, pp. 307-330.

Galizzi, Monica, Miesmaa, Petra, and Slatin, Craig "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*, 2006.

Garicano, Luis and Heaton, Paul "Computing Crime: Information Technology, Police Effectiveness, and the Organization of Policing," *Working paper*, 2006.

Garicano, Luis and Hubbard, Thomas N. "Learning about the Nature of Production from Equilibrium Assignment Patterns," *Working paper*, 2005.

Garicano, Luis and Hubbard, Thomas N. "Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyers' Human Capital," *Working paper*, 2005.

Gaynor, Martin, Rebitzer, James B., Taylor, Lowell J. (2001) "Incentives in HMOs" *NBER Working Paper* #8522.

Griliches, Zvi. (1994) "Productivity, R&D, and the Data Constraint" *American Economic Review,* vol. 84, no. 1, pp. 1-23.

Guertzgen, Nicole "Rent-Sharing and Collective Bargaining Coverage – Evidence from Linked Employer-Employee Data," *Working paper*, 2005.

Hamilton, Barton, Nickerson, Jack and Owan,, Hideo (2003) "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, vol. 111, pp. 465-497.

Heckman, James J. (1990) "Varieties of Selection Bias," *American Economic Review*, vol. 80, pp. 313-8.

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith, and Petra Todd (1998) "Characterizing Selection Bias Using Experimental Data," *Econometrica*, vol. 66, pp. 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).

Holmstrom, Bengt (1982) "Moral Hazard in Teams," *Bell Journal of Economics*, vol. 13, pp. 324-340.

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

Holmstrom, Bengt and Paul Milgrom (1993) "The Firm as an Incentive System," *American Economic Review*, vol. 84, pp. 972-991.

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

Hunter, Larry and John Lafkas (1998) "Information Technology, Work Practices, and Wages," *Proceedings of the 50th Annual Meeting of the Industrial Relations Research Association.*

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

Ichniowski, Casey, David I. Levine, Craig Olson and George Strauss (Fall 1996) "What Works at Work: Overview and Assessment" *Industrial and Labor Relations Review*, vol. 35, pp.299-334.

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*, vol. 87, pp. 291-313.

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

Imbens, Guido and Thomas Lemieux (2008), "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, (February) 142: 615-635.

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

Jones, Derek C. and Kato, Takao "The Impact of Teams on Output, Quality and Downtime: An Empirical Analysis Using Individual Panel Data," *Working paper*, 2005.

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

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

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

Kreps, David (1997) "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, vol. 87, pp. 359-365.

Krueger, Alan B. (2003) "Strikes, Scabs and Tread Separations: Labor Strife and the Production of Defective Bridgeston/Fireston Tires," *NBER Working paper* #9524.

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 (2006) "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).

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

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

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

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

Lazear, Edward (1999b) "Globalization and the Market for Teammates," *Economic Journal*, vol 109, pp. 15-40.

Levenson, Alec "Knowledge Work and Teams: Analyzing Labor Productivity when Tasks are Interdependent," *Working paper*, 2006.

Lo, Desmond, Ghosh, Mrinal, and Lafontaine, Francine "The Role of Risk, Incentives, and Selection in Salesforce Compensation Contracts," *Working paper*, 2006.

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*, vol. 48, pp. 197-221.

MacLeod, Bentley (1988) "Equity, Efficiency and Incentives in Cooperative Teams," *Advanced Economic Analysis of Participatory Labor-Managed Firms*, vol. 3, pp. 5-23.

MacLeod, Bentley and Daniel Parent (1999) "Job Characteristics and the Form of Compensation," *Research in Labor Economics*, vol. 18: 177-242.

Mas, Alexandre "Pay, Reference Points, and Police Performance," W*orking paper*, 2005.

Mas, Alexandre and Moretti, Enrico (forthcoming) "Peers at Work," *American Economic Review*.

Mas, Alexandre "Labor Unrest and the Quality of Production: Evidence from the Construction Equipment Resale Market," *Working paper*, 2006.

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

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

David Neumark, Roy J. Bank; Kyle D. Van Nort, "Sex Discrimination in Restaurant Hiring: An Audit Study," *The Quarterly Journal of Economics*, Vol. 111, No. 3. (Aug., 1996), pp. 915-941.

Osterman, Paul (1994) "How Common is Workplace Transformation and Who Adopts It?" *Industrial and Labor Relations Review*, vol. 47, pp. 173-187.

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

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

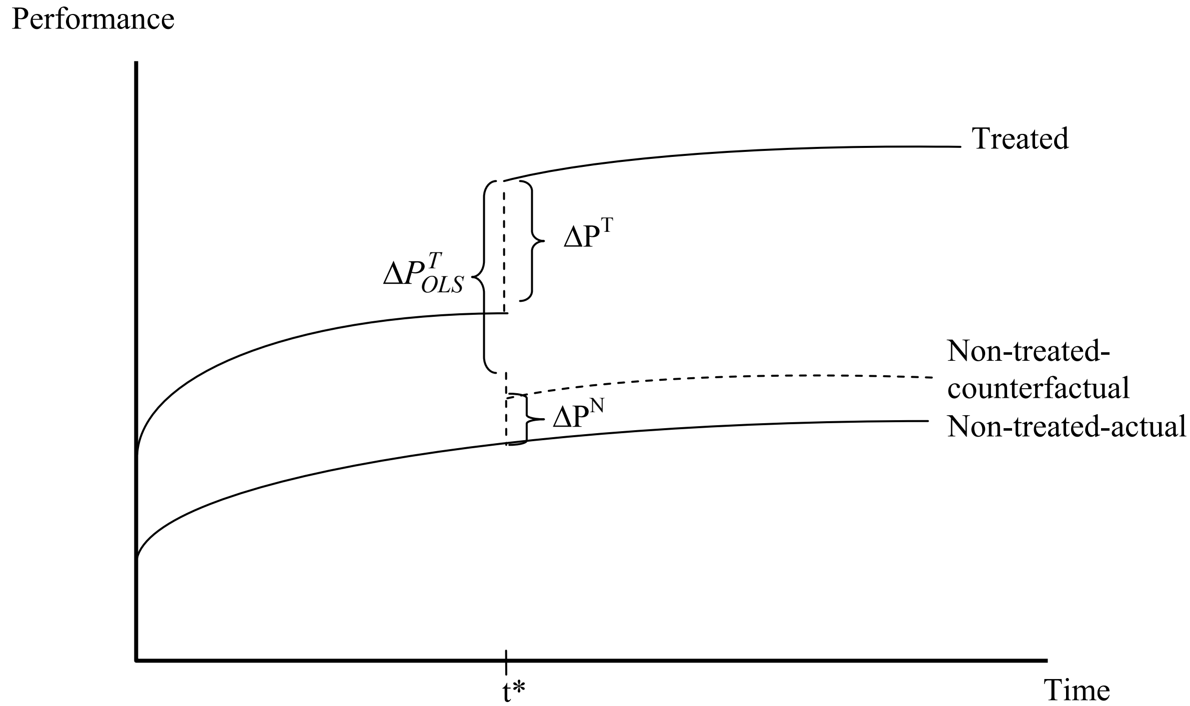
Prendergast, Canice (July 1995) "A Theory of Responsibility in Organizations," *Journal of Labor Economics*, vol 13, no. 3, pp 387-400.

Prendergast, Canice (1999) "The Provision of Incentives in Firms," *Journal of Economic Literature*, vol. 37, pp. 7-63.

Rosen, Sherwin (October 1982) "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, vol. 13;pp. 311-323.

Van Biesebroeck, Johannes (2006), "Complementarities in Automobile Production," *NBER Working paper* #12131.

Figure 1



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

Figure 2: Using Selection as an Advantage in Estimating Treatment Effects
(Lazear, 2000)

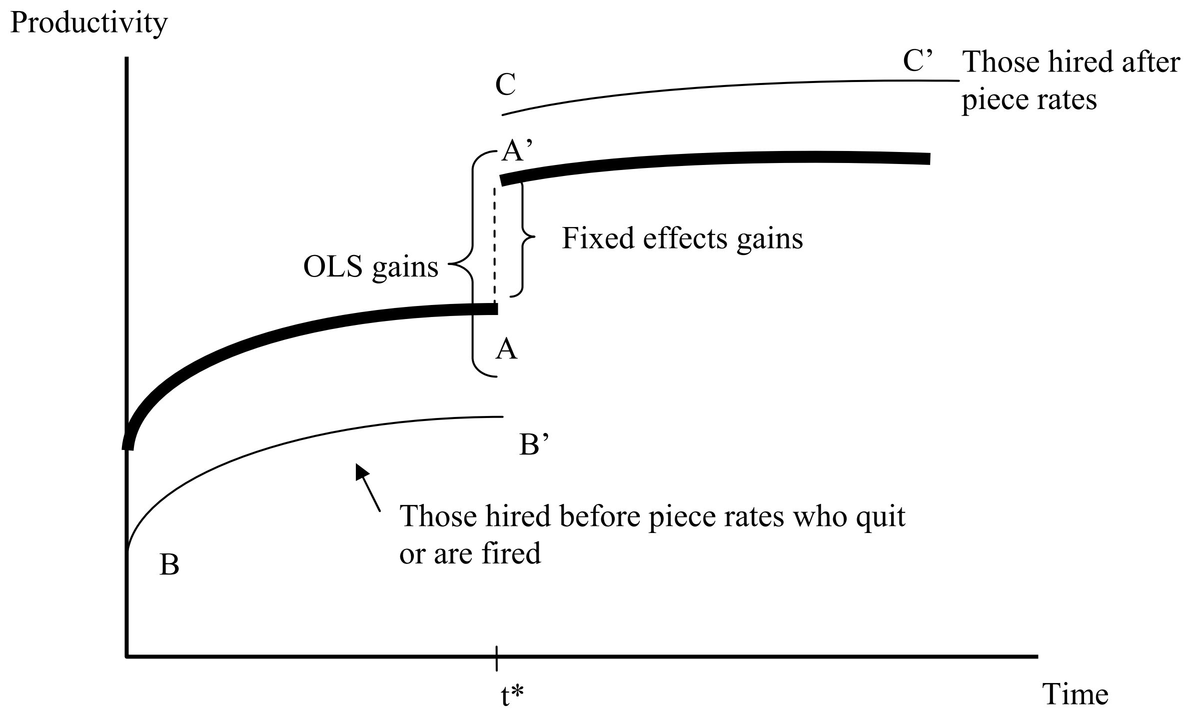


Figure 3: Heterogeneity in the Selection of Production Units
(Hamilton, Nickerson, and Owan, 2003)

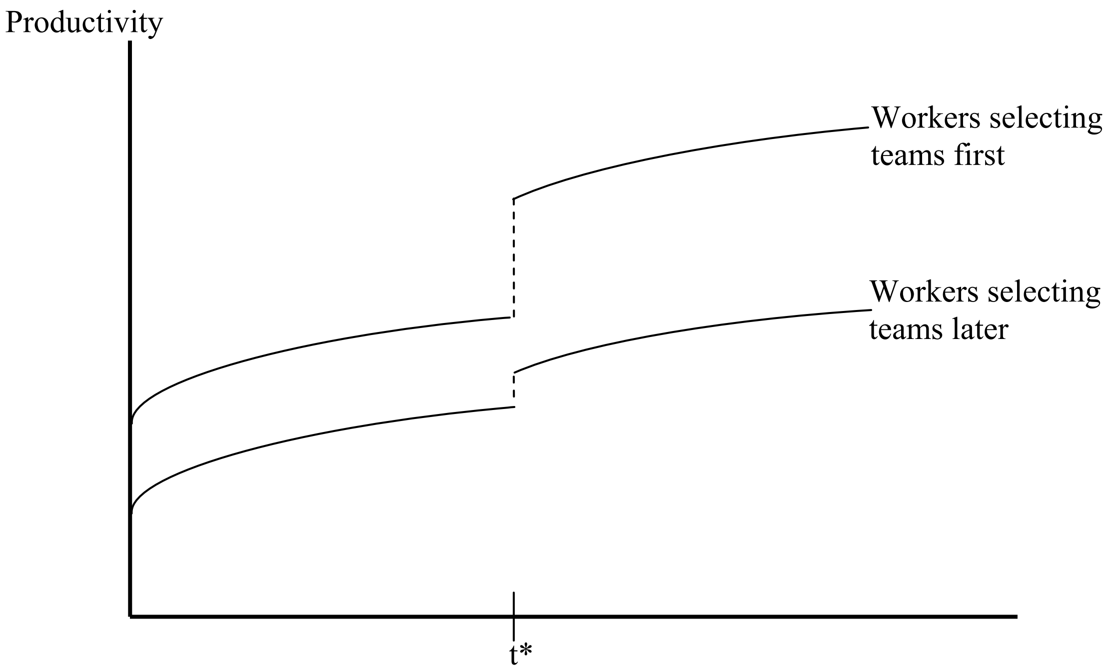


Figure 4: Using Organizational Data to Assess Causality
(Bandiera, Barankay, and Rasul, 2005)