NBER WORKING PAPER SERIES

OPPORTUNITY COUNTS:
TEAMS AND THE EFFECTIVENESS OF PRODUCTION INCENTIVES

Brent Boning
Casey Ichniowski
Kathryn Shaw

Working Paper 8306
http://www.nber.org/papers/w8306

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2001

We thank John Engberg, Marty Gaynor, John Ham, James Heckman, Hide Ichimura, Fallaw Sowell and the participants of seminars at NBER, Society of Labor Economists meetings, American Economics Association meetings, and Carnegie Mellon University for their helpful comments. We also thank the Alfred P. Sloan Foundation for their generous support of this research, and thank the many employees of steel minimills for their generosity in sharing their insights and data with us. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Opportunity Counts: Teams and the Effectiveness of Production Incentives
Brent Boning, Casey Ichniowski and Kathryn Shaw
NBER Working Paper No. 8306
May 2001
JEL No. J24, J41, J5, L20, M11

ABSTRACT

This paper investigates the individual and joint effects of group incentive pay and problem-solving teams on productivity. To estimate models of adoption of these work practices and models of the effects of the work practices on productivity, we constructed a data set on the operations of 34 production lines in U.S. steel minimills. Through site visits and interviews, we collected longitudinal data including precise measures on productivity, work practices, and technology of each of these production lines. We find strong support for the proposition that problem-solving teams are an important means for increasing the effectiveness of group incentive pay plans in establishments with complex production processes.

With regard to adoption of work practices, we find that problem-solving teams are adopted only in the presence of incentive pay plans, and that more technologically complex production lines are much more likely to adopt teams. The latter result implies that teams are more valuable in these types of production environments. We also present estimates of the productivity effects of adopting these work practices. Group-based incentive pay, on average, raises productivity, and the adoption of teams in addition to incentive pay leads to a further increase in productivity. The average effect of teams together with group incentives is economically important, corresponding to an annual increase of over 3000 additional tons of steel with a value of over $1.4 million. We also find that the productivity effect of teams is significantly larger in more complex production lines, consistent with the result that more complex production lines are more likely to adopt problem-solving teams. Finally, we show that our estimates of the productivity effects of these work practices are little changed by corrections for possible selectivity bias.

Brent Boning
U.S. Council of
Economic Advisers
Executive Office of
The President
Washington, D.C. 20502

Casey Ichniowski
Graduate School of Business
Columbia Univeristy
New York, NY 10027
and NBER

Kathryn Shaw
G.S.I.A.
Carnegie-Mellon U.
Pittsburgh, PA 115213
and NBER

“Opportunity Counts:
Teams and the Effectiveness of Production Incentives”

Brent Boning*
Casey Ichniowski**
Kathryn Shaw***

March 1, 2001

*   Council of Economic Advisers, Washington, D.C.
**   Columbia University, School of Business, and NBER
*** Carnegie Mellon University, Graduate School of Industrial Administration, and NBER.

We thank John Engberg, Marty Gaynor, John Ham, James Heckman, Hide Ichimura, Fallaw Sowell and the participants of seminars at NBER, Society of Labor Economists meetings, American Economics Association meetings, and Carnegie Mellon University for their helpful comments. We also thank the Alfred P. Sloan Foundation for their generous support of this research, and thank the many employees of steel minimills for their generosity in sharing their insights and data with us.

I. Introduction

This study presents new evidence on the productivity effects of group incentive pay and problem-solving teams. Several recent studies on the economics of firm's internal organization indicate that certain personnel and human resource management (HRM) practices may be complementary inputs in a manufacturing firm's production function. The productivity gain from adopting the complementary practices together would exceed the sum of their individual productivity effects if they were adopted individually.[1] However, theoretical explanations for why certain work practices should be complements are not well developed. As a result, existing research offers conflicting views about which work practices might be profitably combined. Empirical tests of complementarity among HRM practices are also limited. No current study provides evidence on both the adoption of HRM practices and on their productivity effects, even though the argument that HRM practices are complementary has clear empirical implications for both the adoption and productivity effects of the work practices.[2]

To address these limitations in the existing research, we concentrate on two HRM practices in particular – group incentive pay and problem-solving teams. We model why these two specific work practices should be complements. The central hypothesis derived from the model is that group incentive pay and problem-solving teams can be jointly performance enhancing, and that the joint adoption of these two work practices will have their largest effects on productivity in more complex production processes. We test this hypothesis using a unique, personally collected data set on the operations of U.S. steel minimills, and estimate both adoption and productivity equations. This study's data permit particularly convincing empirical tests for several reasons. First, through site visits and interviews at each minimill, we develop a detailed understanding of the production processes of these establishments. These visits allow us to identify and collect accurate data on the work practices, performance, and capital equipment on these lines. Second, the carefully defined sample based on observations from one narrow segment of the U.S. steel industry eliminates many sources of heterogeneity that would confound empirical models of the productivity effects of work practices in more broadly defined cross-industry samples of firms. Third, we obtained longitudinal data on the operations of each minimill in the sample to estimate fixed effects models of the productivity differences that occur before and after the adoption of the work practices.

Empirical results from these panel data on steel minimills consistently support the study's central hypothesis.

- Teams are adopted only in the presence of incentive pay plans, indicating that teams enhance the value of production-based incentives.

[1] Milgrom and Roberts (1995) provide some of the first arguments in favor of this view. See also Kandel and Lazear (1992) and Baker, Gibbons, and Murphy (1994).

[2] Athey and Stern (1998) illustrate limitations on empirical tests for complementarity of observed production inputs when omitted variables affect the productivity of the observed inputs. For empirical work on complementary work practices, see Black and Lynch (1999, and forthcoming), Appelbaum, Bailey, Berg, and Kalleberg (2000), Appelbaum and Batt (1994), Bartel (2001), Batt (1999), Brynjolffson, and Hitt (2001), Dunlop and Weil (2000), Hunter and Lafkas (1998), Huselid and Becker (1996), Ichniowski, Levine, Olson, and Strauss (1996), Kruse (1993), and MacDuffie (1995).

- The complexity of the production lines is an important determinant of the adoption of problem-solving teams. Hazard rate models show that a one standard deviation increase in this study's measure of production process complexity more than doubles the probability that a line adopts teams.
- Group incentive pay raises productivity. We estimate that the adoption of group incentive pay raises productivity of a minimill production line by .49 percentage points.
- Problem-solving teams, when adopted in concert with group incentive pay, produce further productivity gains. However, the added productivity benefit due to problem-solving teams occurs exclusively in more complex production lines. When we allow the productivity effects of teams to vary with line complexity, we find that the added productivity benefit of teams in the most complex production lines is .39 percentage points. In lines of average complexity, the increase due to teams is .21 percentage points. In the sample's least complex production lines, teams do not raise productivity.
- Given the possibility of endogeneity between teams and productivity, we also estimate semi-parametric nonlinear selection models to correct for this potential bias, and find little change in our estimates of the productivity effects of teams and incentives.

In sum, we find that problem-solving teams enhance the effectiveness of group incentive pay in those production lines that have more complex production technologies and products. In these types of production environments, this added opportunity for production workers to use their knowledge to solve problems and make improvements in operations leads to significant productivity gains, while following standard operating procedures appears to suffice in less complex environments.

II. A Model of Incentive Pay and the Adoption of Teams

Before presenting the formal model, it is useful to have a picture of the production process in the steel minimills that we are studying. In a typical steel minimill, scrap steel is brought to the mill, melted, formed into large blocks, cooled and finally reheated and rolled into finished products. Minimills use electric arc furnaces to melt scrap steel to form raw steel bars (billets) and then roll the billets into products for final use, primarily for the construction and automotive industries. Within this overall process, we analyze the production lines that reheat the billet to make it malleable, and pass the reheated billet through equipment, known as stands, that shape and elongate the billet. Finally, these production lines take the shaped and formed steel after it has cooled and cut it to the desired length before it is prepared for shipping.

In this study, we focus on the use and effects of group incentive pay and problem-solving teams in these production lines. The nature of the particular manufacturing setting that we are studying also helps determine the way that these two employment practices might affect an employee's contribution to output. With regard to incentive pay in these minimill production lines, only group incentives are employed. In particular, the total output quantity and the quality from the production line are readily measured, but the contribution or output of individual employees is not. Group incentive pay plans may still be ineffective in raising output because these plans can be subject to free-rider problems, or because production workers might not have the opportunity to control and thus elevate production-line output.

We argue in this section that a second employment practice – problem-solving teams – can overcome this latter problem that can cause group incentive pay to be ineffective. We define a problem-solving team as a group of workers who meet regularly with a goal of solving production problems that managers or workers identify. The team's activities may be governed by formal rules specifying who may join the team, how meetings are conducted, when meetings might be held, and workers' compensation for attending meetings, but the goal of continuously improving production operations is an important part of the definition. Problem-solving teams provide employees in our sample of minimills with opportunities that they would not otherwise have to elevate output. Through their participation in these teams, production workers can affect the performance of these production lines in a number of ways. Examples of the work of problem-solving teams in these lines include: solving a quality control problem by suggesting the insertion of a new gauge on the line, tackling a defects problem by reconfiguring the layout of the line, and improving operations after the introduction of new capital by sharing information on production problems and discussing ways of speeding up the learning curve. Some of the team's suggestions may require limited amounts of new capital equipment and some may require no investments. Production workers play an important role in crafting these technological improvements because of their hands-on knowledge of how the lines operate**.**

Economic analysis of problem-solving teams is not extensive.[3] This limitation is perhaps surprising given the growth of problem-solving teams among contemporary U.S. businesses.[4] Central to this growth in the use of teams is the recognition on the part of managers that production workers possess valuable information about the operation of production lines that engineers and supervisors often lack. Jensen and Meckling (1992) emphasize the value of production workers' input and argue that managers should "co-locate" decision-making authority with employees who have the most relevant information. Furthermore, if knowledgeable line workers are given the authority to make operating decisions, they must also be given the incentives that motivate good decisions (Jensen and Mechling, 1992; Baker, 1992).[5]

The model we present in this section follows in the spirit of these theoretical arguments. It posits that good (performance-improving) decisions do not just happen spontaneously. Jobs must be designed to put decision-making authority in the right hands, and employees must be motivated to exercise that authority in a productive manner. Because problem-solving teams help with the job design issues and incentive pay plans help motivate the employees in the groups, the two work practices serve complementary roles. The model here also extends this proposition by identifying specific conditions under which these two work practices are most likely to exhibit these complementary effects on production.

The Model

To model the role of teams and incentive pay, we propose that there are two ways that production workers can contribute to the plant's performance:

---

[3] Holmstrom (1982) and MacLeod (1988) offer analyses of teams. These examples pertain more to team production than to the kinds of continuous improvement and problem-solving practices that we focus on this study. Aoki's analysis of Japanese manufacturing firms (1988) considers continuous improvement activities of worker teams.

[4] Osterman (1994, 1997) estimates that the percent of establishments with more than half of their employees participating in teams or quality circles had grown to 62% of U.S. establishments with over 50 employees by the early 1990's and rose modestly in the mid-1990s.

[5] For more on the assignment of responsibility within firms, see Prendergast (1995), Rosen (1982) and Geanakoplos and Milgrom (1991).

1. They can exert "production effort" (e_1) while running the production line. Workers must decide how much production effort to exert – how carefully and quickly they perform routine tasks such as monitoring production and repairing equipment.
2. They can also exert "problem-solving effort" (e_2) in team sessions. When production workers join problem-solving teams, they must decide whether to make some minimal effort on these teams in order to keep their jobs, or whether to expend greater effort.

Problem-solving teams confer decision-making rights on the production employee, thereby providing the employee with the opportunity to put his knowledge of the production process to work. Without these opportunities that teams provide, e_2 will be lower. Incentive pay shares the value of any increased production, giving employees a reason to put forth the extra production and problem-solving effort to increase output.[6]

We use a simple principal-agent framework to model the firm's decision to adopt incentive pay in which the firm balances the cost of compensating employees for working harder and bearing risk with the value of the additional production. Specify the production function as:

$$(1) \qquad \text{output} = f(e_1, e_2) = A + A_1 e_1(I) + A_2 e_2(I,T) + \varepsilon$$

e_1 and e_2 are production and problem-solving effort, I and T represent the firm's implementation of group incentive pay and respectively, and A, A_1 and A_2 are parameters of production.[7] In each of the effort dimensions, employees choose either high or low effort, denoted $e_i = 1$ and $e_i = 0$, respectively, for i=1,2. The effort choice e_1 is a function of the incentive, or $e_1(I)$. The effort choice e_2 is a function of both the incentive and the presence of teams, or $e_2(I,T)$. The random production shock ε is a normally distributed random variable with mean 0 and variance σ^2.

A_1, which parameterizes the effect of production effort, should vary with the manufacturing environment. For example, high values might be associated with very labor-intensive environments, or situations where monitoring is costly. A_2 parameterizes the productivity of problem-solving efforts, and may also vary with the production environment. For example, from field observations and interviews in our sample of minimill production lines, A_2 appears to vary in important ways with the technological complexity of the production process. A very simple production process, involving a relatively small number of steps with

[6] Our model incorporates two types of effort, and thus has a multi-tasking element. However, it differs from the multi-tasking models where the employee chooses between efforts that produce multiple outputs for the firm, such as Holmstrom and Milgrom (1991). In our model, both types of employee effort are directed to the single objective of increasing quality production. Thus, there is not a problem with the employee underproducing one output, and overproducing another output. We also assume that high effort in both production and problem solving is feasible. The discrete effort levels bound total effort: the high effort level is still a limited effort. Constraining high effort to a single dimension is a straightforward extension of the model.

[7] We omit an interaction effect, $A_3 e_1 e_2$, from this model. It is conceivable that a problem-solving improvement might stem from a technological solution that requires a high level of production effort. For example, a production line can be stopped frequently by debris from another stage in the production. Problem solving might identify that debris is causing the breakdowns, and then additional effort can be expended to clean the debris on a regular basis, so that the interaction effect exists. However, it is more likely that the problem-solving team would create a mechanism to shield the equipment from the debris, so that the interaction effect is likely small. Our interest is the interaction effect of the practices rather than the efforts.

established and tested technologies, may have a low value for A_2. In such low complexity lines, extensive meetings among employees to brainstorm for solutions to operating problems are unnecessary. In these lines, following easily specified standard operating procedures will suffice. In contrast, a more complex production process, involving more steps, more sophisticated technology or higher quality standards, will have higher values of A_2. Put differently, A_2 represents the opportunities for worker-directed technological change, and these opportunities should increase with technological complexity of the production process.

The risk-neutral principal pays an output-based linear incentive to the employees, $I=\beta p f(e_1,e_2,T) + \gamma$, where β and γ define the incentive payment and base pay; p is revenue less the (constant) marginal cost per unit of output; and output f depends on efforts e_1 and e_2, and on the presence of teams T.[8] The principal sets the employees' share of revenue, β, to maximize expected profits. We assume that the firm operates in competitive markets for both inputs and outputs.

The principal's expected profits are:

$$\pi(e_1, e_2) = \mathrm{E}[(1-\beta)\, p\, f(e_1, e_2) - k - sT - \gamma]$$

where k is the fixed cost of production, and s is the ongoing cost of problem-solving teams. The basis of the incentive is the increase in revenue less marginal cost. An incentive based on revenues, profits or production does not alter the model's implications.

The agent is an expected utility maximizer with preferences represented by the utility function $u(W,e_1,e_2) = -\exp[-r(W-c(e_1,e_2))]$, where r is a risk-aversion parameter, W is the agent's income, and $c(e_1,e_2)$ is the disutility of effort, which is measured as a monetary cost. We follow Itoh (1994) and model the disutility of effort as $c(e_1,e_2) = (c/2)e_1 + (c/2)e_2 + \delta c e_1,e_2$. If $\delta > 0$, diligence in both dimensions has a higher disutility than just the sum of diligence in each dimension alone, so the model includes an additional cost for performing multiple tasks. We model the employees as a single unit. While this abstracts from potential free-riding problems associated with group incentives,[9] the empirical work to follow tests whether incentive pay is effective and thus implicitly examines whether the free-rider problem is overwhelming.

Implications of the Model

The firm must decide what levels of production and problem-solving efforts to induce by determining whether the cost of inducing high effort levels are justified by the increased revenue generated by high production.[10] The solution of the model is provided in Appendix A1. Here, we summarize hypotheses arising from the model with respect to problem-solving teams and incentive pay. The central proposition of the model

[8] Our motivations for using a linear incentive are that it most closely resembles observed practices in the steel industry, that it makes the problem more tractable, and that it is optimal in many circumstances. See Holmstrom and Milgrom (1987) for a discussion of these points.

[9] Kandel and Lazear (1992) and MacLeod (1988) present models that limit or eliminate the free-rider problem through mechanisms like "positive peer pressure" or other group responses to evidence of individual free riding.

[10] This production function makes explicit the connection between the firm's decision to adopt problem-solving teams and inducing high problem-solving effort; it is clear that the decision to adopt problem-solving teams and the decision to induce high problem-solving effort are actually one decision. Since high problem-solving effort is only effective when a problem-solving team is in place, eliciting high problem-solving effort and instituting problem-solving teams will always occur simultaneously.

is that, all else constant, the joint adoption of problem-solving teams and incentives is a function of the opportunities for incremental technological improvements, or A_2.[11] In our sample, measures of production process complexity are the main proxy of the A_2 parameter. Given this proxy for A_2, we can then specify the following six hypotheses that we test directly with the data from our sample of minimills. With regard to the adoption of the work practices, the above model yields the following propositions:

> *Hypothesis 1*: Teams will be adopted only in the presence of incentive pay.
>
> *Hypothesis 2*: Higher levels of production process complexity promote the adoption of problem-solving teams that complement the adoption of incentive pay.

With regard to the productivity effects of these two work practices, the model implies the following additional hypotheses:

> *Hypothesis 3:* Output will be higher in the presence of incentives, because incentive pay increases workers' production effort by increasing the returns to that effort.
>
> *Hypothesis 4*: Output will be higher when teams are added to incentive pay, because teams provide workers with the opportunity to improve the performance of the line, and incentive pay increases workers' returns to problem solving effort to elevate output. Thus, workers' problem-solving effort rises with teams and incentive pay.
>
> *Hypothesis* 5: The magnitude of the effect of teams plus incentives on productivity will be larger in more complex production lines, if A_2 increases with line complexity.
>
> *Hypothesis 6*: The decision to adopt teams and the performance gains from teams are jointly determined—lines that are more complex (or higher A_2) have a greater output from teams and are therefore more likely to adopt teams. Thus, in modeling the output gains from teams, the decision to adopt teams must be endogenously estimated.

III. Data

To test these hypotheses, we have assembled detailed data on the operations of production lines in steel minimills. We limit our study to a single production process to eliminate many sources of heterogeneity that have confounded previous attempts to estimate the effects of work practices on productivity or to identify determinants of different work practices. The sample does not simply contain minimills that have been reported in the business press as successful adopters of work teams or other work practices. Rather, the sample represents the full breadth of the experience in the minimill segment of the steel industry. These mills utilize a variety of human resource practices, and have a wide range of profitability, from mills that are

[11] Other predictions of the model include that the value of adopting incentives or teams will rise with the price-cost margin. It will fall with the direct disutility of effort *c*, the disutility of multiple tasking δ, the variability of production σ^2 and risk aversion *r*, and with higher transition costs.

unprofitable to those that are highly profitable. The sample pertains to 34 different production lines, owned by 19 different companies, and includes nearly all U.S. minimill production lines of this kind.[12]

We personally visited each minimill from November 1994 through April 1997, conducting extensive interviews and touring each line. We interviewed workers, union officials in organized establishments, plant managers, and HR managers to understand how workers contribute to productivity. These interviews also permit us to identify the actual HRM practices of the lines, and not simply the set of formal policies listed in personnel manuals that may not be used in practice. We also toured the production lines with experienced line supervisors, operators and engineers. These personal inspections allowed us to identify ways in which production processes and technologies still differ from one line to the next, even within a sample drawn from one narrowly defined segment of one manufacturing industry. The visits also allowed us to observe the workings of real problem-solving teams.

Sample

The sample for analysis consists of 2355 monthly observations on the operations of these 34 production lines, or about five years of monthly data per line on average. In the empirical analysis, we estimate models on both the adoption of work practices and on the effects of work practices on productivity. The model of section II identifies three important categories of variables – output measures, the HRM practices, and production line complexity. In addition, the productivity equations that we estimate must also control for other technological features of the lines that affect productivity.

Productivity

Our site visits and interviews allowed us to identify the most convincing measures of productivity and technological features of the production process that can affect the lines' performance. The measure of productivity is YIELD, defined as "good tons" produced (tons that meet industry-established quality standards) divided by the tons that enter the production process. This measure is directly related to amount of steel output, since tons produced is the product of YIELD and "tons charged" (or the amount of steel that enters the production process). Tons charged will vary depending on the size of the mill, its rated capacity, and the products produced. YIELD also captures quality dimensions of output. It is lower when the finished product fails to meet a quality standard, such as failing to conform to the specified dimensions of the product.

Employees can affect YIELD in a number of ways. Workers improve yield through any activities or decisions that prevent "wrecks" on the line when material jams in the equipment or by limiting the number of electrical or mechanical failures in the equipment. These activities help keep lines running more continuously. Workers can also correct quality problems before a significant amount of unacceptable quality steel is produced. According to managers, YIELD is the most useful measure of line performance. It captures both the output and quality dimensions of productivity. We were able to obtain yield data for 2251 of the 2355 observations. The mean value of the YIELD variable is 93.3, so that 6.7% of the output of this

[12] This sample includes the vast majority of U.S. minimills that make the kinds of steel products that we consider in our sample. We were unable to visit only eight such U.S. minimills. Based on our conversations with managers at these mills and at the mills in our sample, our sample provides an accurate picture of the range of performance outcomes and HRM practices in this industry, and accurately reflects the nature of the available technology for producing these steel products.

sample's mills is lost in the production process or does not meet the required quality specifications. YIELD has a standard deviation of 3.20, and a minimum and maximum of 73.05 and 100 respectively.

HRM Practices

Section II models how incentive pay and problem-solving teams can increase productivity. The dummy variable INCENTIVE indicates the presence of an incentive pay scheme at the mill. As indicated above, given the relative ease of measuring production line and mill-level output, and the difficulty of measuring individual employee contributions, incentive pay plans in this sample are always group incentives. We define an incentive pay scheme as being present when pay is a function of production, product quality, profits or some combination of these.[13] The use of some incentive plan is nearly universal by the end of our sample period, with 91% of the lines having an incentive plan by the end of our monthly data for each line.

While incentive pay is widespread in the sample, the use of problem-solving teams is more varied. At the beginning of our data, 12 percent of lines have teams and by the end 42 percent have teams. The dummy variable TEAMS represents the presence of a formal problem-solving structure at the mill. We classify lines as having TEAMS only if they had a formal structure designed to involve production workers in identifying ways to improve production. Thus, a line that only had crew meetings to share information and organize a shift was not classified as having teams. Further, lines that would informally gather to correct an immediate problem were not classified as having teams, because the focus was not on improving the production process, but simply repairing an immediate breakdown. Given the widespread adoption of the INCENTIVE measure, it is also the case that teams exist together with incentive pay. Teams without incentives exist in only 1.5% of all monthly observations and this situation lasts only a few months in any mill. By the end of the sample period, no mill has teams without an incentive. When we estimate productivity equations, we are therefore essentially estimating the effect of adopting group incentives and the effect of adopting group incentives plus teams relative to productivity levels observed in workplaces with the neither incentives or teams.[14]

Complexity of Production

Differences in the technological complexity of production processes are central to the model of Section II. While the sample of minimill production lines eliminates many types of heterogeneity across the sample's observations, production lines in the sample are not identical in terms of the complexity of their operations. In our sample of minimills, the complexity of steel products and the technological complexity of

[13] The most common incentives in the steel industry make pay a function of production by the crew or of profits of the mill. Profits are based on the mill's performance, not company performance. A few mills include multiple criteria in their incentive scheme; these criteria are combined linearly.

[14] While our model focuses on the complementarity between group incentives and problem-solving teams, recent research also suggests that complementarities may exist among a broader set of HRM practices (Milgrom and Roberts, 1995; Ichniowski, Shaw and Prennushi, 1997). We therefore collected data on nine other HRM practices measuring employment security and no layoff policies, training and education programs, extensive pre-hire screening of job applicants, employee visits to customers, and combined operator-maintenance job definitions. We do find that lines with teams and incentives use 5.9 of these supporting practices on average, while incentive-only lines use 3.6 of these other practices. While lines with teams are much more likely to do more extensive pre-hire screening and are much more likely to have workers visit customers, the precise set of additional practices that exist in lines with teams varies from mill to mill. Still, these data suggest that it may be appropriate to think of the estimated productivity effects of the INCENTIVE variable and the TEAM PLUS INCENTIVES interaction variable in the empirical work to follow as measures of the effects of an "Incentive Without Teams System" and an "Incentives with Teams System." However, we do not have sufficient data to test for separate or added productivity effects from these individual practices, given the multicolinearity between practices.

the production process go hand in hand. More intricate minimill products with more demanding tolerances are made on lines with more elaborate production technology.

We construct two related measures of this complexity. First, based on our plant visits, we define four product classes. Product 1 involves the least complex production in which billets are simply reduced to smaller sizes, and product 4 involves the most complex production process that produces intricate steel shapes with very tight tolerances in product specifications. The four variables, PROD1 through PROD4, measure the percentage of the mills' production in each of these four categories. Each mill has a clear majority of its production in one of the four categories. Furthermore, with only one exception, 100% of each mill's production is accounted for by either one or two production categories. When production is split between two categories, it is split across relatively similar groups (e.g., 80% in PROD1 and 20% in PROD2). Second, we also measure production process and product complexity with the variable COMP, which is a weighted sum of the four production PROD1-PROD4 variables, with weights equal to a representative market price for each product group.

Other Control Variables

Econometric specifications of the adoption model and the productivity model are given in the next section. We describe there any additional control variables beyond the HR practice and complexity variables that we include in those models. For example, the productivity equation includes controls for vintage of the capital, experience of the managers and the work force, and a time trend for learning curve effects. While we describe these other controls in more detail in the next section, Table 1 provides a reference list of the entire set of dependent and independent variables that we use in the adoption and productivity models.

IV. Estimation

The theoretical model in Section II concludes with six predictions about the adoption of problem-solving teams and incentives and about the impact of these HRM practices on productivity. We take up the hypotheses concerning adoption and productivity in turn.

Adoption Equations

Hypothesis 1 of section II – that teams are adopted only in the presence of incentives – is borne out by the data on the distribution of these two HRM practices. Incentives are widespread in our sample, and are universal across all production lines by the end of our sample period. Problem-solving teams always follow the introduction of incentives. Because group incentives are so widespread, our multivariate analysis of the adoption of HRM practices focuses on problem-solving teams, which are present in 42% of these production lines by the end of our sample.[15]

Hypothesis 2 proposes that a more complex production process increases the value of adopting teams—production lines that are more complex provide workers with more opportunities to identify ways of elevating output by solving problems. Represent the decision to adopt teams with the equation:

[15] Since all mills eventually adopt incentives, incentive pay may be productivity-enhancing for all. Late adoption may be governed by transition costs (such as the need to negotiate with the union) rather than differences in the performance gains from incentives. In addition, the theoretical model does not generate specific predictions about the impact of complexity on incentives.

(2) $$T_{it} = 1 \qquad \text{if and only if } \Gamma Z_{it} + u_{it} > 0$$

where $\Gamma Z_{it} + u_{it}$ is the profitability of adopting teams. Lines adopt teams when the net gains are positive. The *Z* vector contains variables influencing adoption, as specified below.

We estimate hazard rate models of the probability of adopting problem-solving teams, conditional on the mill not having already adopted teams and on the features of the production and managerial environment. We estimate these hazard rate models of the adoption of teams using data from 31 mills that had the opportunity to adopt teams. Three of the 34 lines in our data set enter our sample period already having adopted teams. The sample for the hazard rate models is 1,686 monthly observations from the 31 lines that began the sample period without teams. We estimate both exponential and Weibull hazard models, where the Weibull model permits duration dependency. For our estimation, we assume that January 1984 was the first date when teams could have been adopted.[16]

Table 2 presents the results from the hazard rate models. The Weibull and exponential specifications yield qualitatively similar results. The shape parameter, which indicates the duration dependence in the Weibull models, is approximately .8 and not statistically significantly different from 1 in either model. The point estimates indicate negative duration dependence—the longer a mill operates without teams, the less likely the mill will adopt problem-solving teams. However, the insignificance indicates that the exponential model is a legitimate specification.

The results in Table 2 provide strong support for Hypothesis 2: *the complexity of the production process is a significant determinant of the adoption of problem-solving teams*. Mills with the two most complex production processes, those concentrating on PROD3 and PROD4, are much more likely to adopt problem-solving teams than mills with the less complex processes. Lines with higher production concentrations in PROD1 products are the least likely to adopt teams (Table 2, columns 1 and 3). The composite measure of complexity, COMP, also has a positive and significant effect on adoption, again indicating that the higher the complexity, the more likely the adoption of teams (Table 2 columns 2 and 4). Based on the column 2 exponential hazard model, a one standard deviation increase in the COMP measure (or 103.9 units) increases the probability of adopting teams by 210%. In a given month, the probability that a line adopts teams is, on average, .0041. A one standard deviation increase in COMP implies an increase in this probability to .0086. Over a five-year period, the mean probability of adopting teams would be .13, and this figure would increase to .40 with a one standard deviation increase in COMP.

Two other factors included in these adoption equations are consistently significant determinants of teams adoption. The significance of these variables may reflect the effects of risk aversion or other factors suggested in the Section II model. For example, OPAGE, the average age of the employees, has a negative effect, suggesting that older workers could be more risk averse, have higher costs of effort, or are more

[16] Cole (2000) presents evidence that the quality movement was transferred from Japan to the United States in the early 1980s and the associated teams were introduced into American manufacturing firms around that time. Thus, we interpret teams as a managerial innovation or "shock" that entered the U.S. in the mid-1980s. Empirical results are not sensitive to the choice of a specific start date. Using start dates before 1984 does not change the results significantly. Three mills were built after 1984, so they could adopt no earlier than they began operations, and the start date is set accordingly. For most mills, the data we obtain begins after 1986; we estimate the hazard rate models conditional on the mill having not adopted by the first date that we observe the mill.

resistant to change.[17] The positive and significant coefficient on CEONEW may suggest the importance of transition costs, as managers who have used traditional managerial styles and have a long association with a given set of practices may be less likely to accept new practices such as problem-solving teams.[18]

Productivity Equations

The adoption models provide one way to investigate the notion that problem-solving teams (in conjunction with incentives) are more effective in more complex production environments. We also have collected data to provide direct estimates of the productivity effects of teams and incentives in different production environments.

Begin by writing the productivity model as:

$$(3) \qquad Y_{it} = \alpha_0 + \alpha_1 I_{it} + \alpha_2 I_{it}T_{it} + \alpha_3 C_{it}I_{it}T_{it} + \theta X_{it} + u_{it}$$

Y_{it} is our measure of productivity, YIELD, I_{it} represents the adoption of incentives (with or without teams), $I_{it}T_{it}$ the joint adoption of teams in addition to incentives, $\alpha_3 C_{it}I_{it}T_{it}$ is the interaction between complexity and teams with incentives, and X_{it} a set of control variables for production that are not affected by the presence of incentives or teams.

Hypotheses 3 and 4 state that output will be higher when incentive pay elicits greater production effort, or $\alpha_1 > 0$, and output will be higher when problem-solving teams elicit problem solving effort, or $\alpha_2 > 0$. The productivity regression results support these hypotheses: incentive pay and problem-solving teams raise performance.[19] Column 1 of Table 3 presents the initial OLS estimates of the productivity equation (3), assuming that the error term u_{it} is distributed i.i.d. with mean 0. Both practices, incentives and teams, raise performance. In Column 3, we relax the assumption that u_{it} is distributed i.i.d., and allow for the possibility of line-specific fixed productivity effects and line-specific auto-regressive errors:

$$(4) \qquad Y_{it} = \alpha_0 + \alpha_1 I_{it} + \alpha_2 I_{it}T_{it} + \alpha_3 C_{it}I_{it}T_{it} + \theta X_{it} + \gamma_i + u_{it}$$

where $u_{it} = \rho_i u_{it-1} + \varepsilon_{it}$. Though our data are specific to one type of production process and we have extensive controls, some line-specific fixed effects might exist because some key variables, such as managerial quality, may be omitted from (3) and could be correlated with line performance. In estimating our model and collecting the data, we attempted to proxy managerial quality with managerial tenure, but such measures serve

[17] Tenure has a significant positive effect in some specifications. OPLONG is a dummy equal to 1 if the average operator tenure at the mill is over 10 years, and in each of the models, the coefficient on OPLONG is positive and significant. Assuming tenure is an indicator of employee problem-solving ability, the estimation may provide additional support that the potential productivity of the teams affects adoption. Mills that are part of a single-mill company (SOLO) are less likely to adopt teams, possibly because they cannot spread transition costs, such as costs of training programs and learning how to implement teams, across mills.

[18] We also considered the possible effects of unionization. Unions, by providing a framework for communication and management-employee interaction, might contribute to team formation. The data indicate that the unionization effect on adoption is insignificant. When the model is expanded to include a variable for unionization rates, the coefficient on a union dummy variable is 0.585, with a p-value of .539.

[19] The coefficients on the control variables are in the expected direction and, in many cases, significant. The variables LINEAGE, LINEAGE2 and EDATE have a high degree of colinearity, so interpretation is difficult, but EDATE is consistently positive and significant, indicating the existence of learning curves in the industry. Performance falls with operators' average age and rises with average tenure. This pattern mirrors the effects of these variables in the equations for the adoption of teams shown in Table 2, reinforcing a conclusion that longer tenure raises performance on the job and therefore raises the value of teams, but age controlling for tenure does not.

only as proxies for managerial technical and leadership skills.[20] The fixed effects estimates reflect the increased performance within mills from introducing the practices, rather than a comparison between mills that do and do not have incentives. While the magnitudes of the estimated coefficients change after allowing for these line-specific effects in columns (3) and (4), the pattern of results is the same. The results in column (3) show that adopting incentive pay (with or without teams) raises productivity by .493 percentage points, and that the joint adoption of teams and incentive pay increases the yield by another .248 percentage points.

The effect of problem-solving teams on output is amplified in Hypothesis 5, which predicts that problem-solving effort will be more valuable with more complex production processes (C_{it}), because opportunities to use teams to develop improvements in production operations are greater in complex environments, thus $\alpha_3 > 0$. In the OLS results of column (2) of Table 3, the coefficient on the complexity times teams plus incentives (i.e., the coefficient α_3 on COMP•TEAMS) is positive and highly significant, and in the fixed effects model of column (4) it remains significant, though less sizable. Note that when COMP•TEAMS is added in (2) and (4), the coefficient on TEAMS alone turns negative (though insignificant in the fixed effects model.[21]

To interpret these productivity effects, we evaluate the models for different values of the complexity variable. Because the results in column (4) indicate the existence of significant line-specific productivity effects, we focus on the point estimates from the column (4) specification. The point estimates of the coefficients on the TEAMS (which includes incentives) dummy and on the COMP•TEAMS interaction term indicate that teams have a net positive effect in more complex production environments. At the maximum value (585.2) and average value (431.3) of the complexity variable, the additional increase in productivity due to the adoption teams would be .39 and .13 percentage points, respectively. At the minimum value of the complexity (284.7), the estimated coefficients on the TEAMS and COMP•TEAMS variables imply a relative reduction in productivity of -.12 percentage points once teams are adopted. Perhaps accordingly, teams are not common in production lines with the lowest levels of production process complexity. The estimated coefficients on the TEAMS and COMP•TEAMS variables from the column 4 specification imply that the net effect from adopting teams would be zero when the complexity variable is equal to 356 – a level of complexity well below the average value of complexity in our sample.

Productivity effects of these magnitudes are economically important. We can interpret the magnitudes of these productivity effects further by translating them into effects on tons of steel and on revenues and profits. Based on the fixed effects specification in column 3 of Table 3, the adoption of incentives and teams would increase the annual output of a typical mill (producing about 400,000 tons of finished product) by over 3,000 tons of steel. This increase translates into additional profits of approximately $1.4 million at 1995 prices.[22] Based on the fixed effects specification in column 4 of Table 3 that allows the effects of teams to vary with production process complexity, the increase in profitability from adopting incentives plus teams would be even larger. For example, in a line with the highest value of the complexity variable, the increase in output would be larger and worth approximately $2.4 million per year.

[20] The misspecification stemming from the difficulty of modeling managerial skills has long been recognized. An early treatment is Griliches (1957).

[21] The terms TEAMS and COMP•TEAMS are jointly significant, with a p-value of 0.026.

[22] An increase in yield does not increase variable costs, so we simply convert the increased production into profits, using the lowest-priced product. This omits the costs of maintaining teams, and any additional incentive pay. It also omits any cost savings stemming from the teams or improved operating practices, such as lower replacement costs for damaged equipment.

We conclude the tests of Hypotheses 3 through 5 by trying an alternative functional form for the interactions between incentives, teams, and complexity. We re-estimate the productivity models and permitting the effect of incentives (as well as teams) to vary between low and high complexity lines, where complexity is reduced to a (0,1) dummy variable rather than its linear effect. We re-estimate the productivity equation, including the following four dummy interaction variables: incentive pay only in low complexity lines; incentive pay plus teams in low complexity lines; incentive pay only in high complexity lines; and incentive pay plus teams in high complexity lines. Productivity for these groups is measured relative to the group of observations that have neither incentives nor teams in models that control for line complexity and the other controls in the Table 3 models. For the purposes of defining these four dummy interaction variables, we consider low complexity lines to be those lines that concentrate on PROD1 and PROD2 steel products.[23]

OLS and fixed effects models yield qualitatively similar results to those presented earlier with the linear complexity variable. Estimated coefficients (and standard errors) from the model that allows for both line-specific productivity effects and line-specific autoregressive errors are: Incentives and No Teams in Low Complexity Lines, .419 (.141); Incentives with Teams in Low Complexity Lines, .100 (.222); Incentives and No Teams in High Complexity Lines, .935 (.301); and Incentives with Teams in High Complexity Lines, 1.358 (.299).

These coefficients really highlight the detailed patterns emerging from the effects of incentives and teams. First, teams do not add significantly to the productivity levels that low complexity lines achieve with only group incentives in place. In particular, while the point estimate of the incentives plus teams coefficient (.100) is less than the point estimate of the incentives without teams coefficient (.419) for low complexity lines, an F-test indicates that the hypothesis that these two coefficients are equal cannot be rejected at the .10-level. Second, teams do improve productivity in high complexity lines above what they achieve with only incentives. For high complexity lines, the coefficient for incentives plus teams (1.358) is significantly different from the coefficient for incentives without teams (.935) at the .01-level. Even when we allow the effect of incentives to vary with line complexity, we find that teams provide a productivity advantage in high complexity lines above and beyond what group incentives on their own produce in these lines. This particular result provides direct support for Hypothesis 5. Third, incentives on their own appear to matter more in high complexity lines than in low complexity lines (i.e., .935 > .419), but the difference in the two coefficients that measure the effects of incentives in high and low complexity lines is significant only at the .12-level.

In sum, the results from the productivity models, like those from the adoption equations, support the central propositions of the study's model. Incentive pay and teams improve productivity (Hypotheses 3 and 4), and moreover, when highly complex lines adopt teams and incentives, performance rises more (Hypothesis 5).

IV. Selection Bias in the Productivity Equation: Focus on the Teams•Incentives Effect

In this section, we address the possibility of selection biases that arise in our estimation of the productivity effects of the work practice variables, particularly the TEAMS variable. Recall that the effectiveness of problem-solving teams is represented by A_2 in the production function, and our proxy for A_2

[23] Specifically, if 80 percent or more of the line's production was in PROD1 and PROD2 combined, we classified the mills as low complexity. By this standard, 60 percent of the observations are low complexity, and 40 percent high complexity.

rises with production-line complexity. In the empirical productivity models above, ($\alpha_2 + \alpha_3 C_{it}$) provides our estimate of A_2. If there is an unobserved component to C_{it} (and therefore to A_2), the endogeneity of teams and complexity can produce a selection bias when we estimate equation (3). Thus, Hypothesis 6 from our theoretical model points out that A_2 measures the impact of teams on output, and therefore A_2 also influences the return to adopting teams, so this endogeneity of team adoption must be addressed in the estimation of the productivity regression. Recall also that the TEAMS variable represents all those lines that have both teams and incentive pay, though the variable is labeled TEAMS.

To address this possible selection bias, it is important to be explicit about the nature of the performance gains from teams that we seek to determine. Do we want to know the expected gains for only the subset of lines for which it is optimal to adopt teams, or do we want to know the expected gains if any random line adopts teams? This question can be rewritten in equation form. The "treatment" here is the introduction of teams. The expected gains from the random adoption of teams is labelled the "average treatment effect," because these gains would be expected if firms randomly adopted teams. This is the expected value $E(Y^T_{it} - Y^N_{it} | X_{it})$, where Y^T_{it} is the output from the use of teams, Y^N_{it} is the output without teams for line *i*, and *X* is the set of control variables that govern production. Knowledge of the size of the average treatment effect would be valuable if, for example, the government were to advocate that all firms adopt teams. Alternatively, another measure of the gain from teams is the expected gain for those who find it optimal, or the "treatment of the treated." This is the expected value $E(Y^T_{it} - Y^N_{it} | X_{it}, T_{it}=1)$. This would be the relevant gain for those firms whose characteristics match the characteristics of the adopters. Thus, we consider two possible mean counterfactuals, the "average treatment effect" and the "treatment of the treated."[24]

To define these counterfactuals, separate the teams and non-teams effects in a switching regression framework:

$$(5) \quad Y_{it}^T = \alpha_0 + (\alpha_2 + \alpha_3 C_{it})T_{it} + \theta^T \tilde{X}_{it} + \gamma_i^C T_{it} + \gamma_i + \varepsilon_{it}^T$$

$$(6) \quad Y_{it}^N = \alpha_0 + \theta^N \tilde{X}_{it} + \gamma_i + \varepsilon_{it}^N$$

where the superscripts *T* and *N* refer to the presence of Teams and No Teams, respectively, the vector $\tilde{X}$ contains the control variables and the dummy for an incentive, γ_i is a mill-specific effect that does not depend on teams, and for simplicity we omit the serial correlation in the residuals. For the remainder of this section, we will refer to the T*I, or teams and incentive interaction, as the "teams effect," and use the short-hand notation of T to stand for this interaction. While we use the term "teams effect" to simplify the exposition, teams co-exist with incentives in our data, so we are implicitly describing the teams effect conditional on the joint use of incentives.

---

[24] We must choose between these different mean counterfactuals because alternative concurrent states are never observed. That is, for every line *i* we would like to know the gains from teams, or to know ($Y^T_{it} - Y^N_{it}$), where Y^T_{it} is the output from TEAM use and Y^N_{it} is the output without teams for line i. However, each team's non-chosen state is never observed—for line i that chooses teams, the concurrent alternative of no-teams is not observed. Thus, we cannot calculate individual gains, and must instead turn to the calculation of expected population gains, with assumptions about

Equation (5) contains a complexity variable that may have observed and unobserved components for the econometrician, though we assume the manager always knows the value of complexity. Assume that true complexity is equal to $C_{it} + \gamma_i^C$, where C_{it} is measured complexity and γ_i^C is line-specific unobserved complexity. The interaction between complexity and teams adoption in the production equation produces observed and unobserved components, or $C_{it}T_{it} + \gamma_i^C T_{it}$, where the unobserved $\gamma_i^C T_{it}$ will complicate our estimation.

In defining the teams subsample in equation (5), we consider two alternatives. As defined, (5) refers to the subsample of lines that *currently have teams*. However, it may be more appropriate to separate the subsample of lines that will *eventually have teams*. Rewrite (5) as:

$$(5') \quad Y_{it}^{\tilde{T}} = \alpha_0 + (\alpha_2 + \alpha_3 C_{it})T_{it} + \theta^{\tilde{T}} \tilde{X}_{it} + \gamma_i + \gamma_i^C T_{it} + \varepsilon_{it}^{\tilde{T}}$$

where superscript $\tilde{T}$ refers to the set of lines that will eventually adopt teams (and N in (6) would become the set of lines that never adopts teams). The primary advantage of the subsample in (5') is that we can estimate the teams coefficient within (5') by using the before-after data on performance. If mills that eventually adopt teams are inherently different from mills that never adopt teams (due to perhaps a mill-specific complexity difference), then the bias function using (5') is the more appropriate model. We will present results for both (5) and (5').

Treatment of the Treated

From the perspective of managerial decisions, the "treatment of the treated" may be the most relevant counterfactual to estimate. A manager typically wishes to know the expected gain from teams given the characteristics of his production line. Thus, if the manager knows the characteristics of his line (such as its complexity), he wishes to know both when adoption is valuable (from the hazard rate model of adoption) and the expected gains conditional on adoption, or $E\left[Y_{it}^T - Y_{it}^N \mid \tilde{X}_{it}, T_{it} = 1\right]$. This is the treatment of the treated.

To estimate the treatment of the treated effect, rewrite (5') by differencing it within mills. Assume that a mill adopts teams in the interval between times *t-k* and *t*, and that prior to the adoption of teams, productivity evolves according to:

$$(7) \qquad Y_{it}^{N|T} = \alpha_o + \theta^{N|T} \tilde{X}_{it} + \gamma_i + \varepsilon_{it}^{N|T}$$

which rewrites equation (6), with the symbol of N|T representing the non-teams group that eventually adopts teams. If we then estimate the productivity regression by differencing between periods *t-k* and *t*,

$$(8) \qquad Y_{it}^{\tilde{T}} - Y_{i,t-k}^{N|T} = (\alpha_2 + \alpha_3 C_{it})T_{it} + \theta^{\tilde{T}} \tilde{X}_{it} - \theta^{N|T} \tilde{X}_{i,t-k} + \gamma_i^C T_{it} + \varepsilon_{it}^{\tilde{T}} - \varepsilon_{i,t-k}^{N|T}$$

---

the alternative states, or the counterfactuals. To organize our discussion of these alternatives, we use the comprehensive review by Heckman, LaLonde, and Smith (1999).

the coefficient on T_{it} is the estimated treatment of the treated effect. Note also that the gains to treatment include the observed effects, $\alpha_2 + \alpha_3 C_{it}$, as well as the unobserved effects, γ_i^C. To estimate this effect, we are making a key parametric assumption about the unobserved counterfactual. We are assuming that $Y^{N|T}$ (which is unobserved after teams are adopted) would have evolved according to the parametric form $\theta^{\tilde{T}} \widetilde{X}_{it}$ in (7). Or, we are estimating the teams effect assuming that ($\theta^{\tilde{T}} \widetilde{X}_{it} - \theta^{N|T} \widetilde{X}_{it}$) accurately measures the evolution of the change in Y due to non-teams influences. [25]

The results of estimating (8) are in column 1 of Table 4, where (8) is estimated with fixed effects. The results are little changed relative to those for the full sample in Table 3, column 3 which keeps the non-teams subsample in the analysis. There is a slight increase in the coefficient on TEAMS and a bigger increase in the coefficient on INCENTIVE. The results here, which drop the non-teams subsample from the analysis, suggest that the coefficients in the non-teams sample are similar to those for the TEAMS subsample. In column (2), we interact complexity with teams and find a significant positive interaction: more complex lines have higher gains to teams. This result surfaces despite the fact that the sample of lines that choose to adopt teams are already themselves more complex lines.

Average Treatment Effect

For public policy purposes, we may wish to know the expected gains to teams if all lines were encouraged to (randomly) adopt teams. This would also be the relevant counterfactual if a manager decided to adopt teams because he saw others doing so, but did not evaluate whether teams were optimal for him. This is also the effect that is typically estimated in much of the existing literature on productivity effects of HRM practices that typically rely on cross-industry samples of establishments. Since the effects of HRM practices (such as teams) vary with intra-industry factors (such as production process complexity), cross-industry studies that do not account for factors that mediate the effects of teams will produce estimates of "average treatment effects."

To find the expected gains to teams for a randomly chosen production line, we estimate the average treatment effect, $E\left[Y_{it}^T - Y_{it}^N \mid X_{it}\right]$. To consider alternative approaches to estimating this unobserved expected value, begin by combining (5') and (6) in one equation and add the assumption that the control variables have equivalent coefficients, or $\theta^T = \theta^N$:

$$\text{(9)} \qquad Y_{it} = \alpha_o + \left(\alpha_2 + \alpha_3 C_{it}\right) T_{it} + \theta \widetilde{X}_{it} + \gamma_i + \left(\gamma_i^C + \varepsilon_{it}^T - \varepsilon_{it}^N\right) T_{it} + \varepsilon_{it}^N$$

Using (9), the average treatment effect is

[25] Note that we are effectively using a matching estimator, in which the performance data for the years prior to adoption are used as the matched comparison group. We are also assuming that there are no differences in the distribution of the residuals for mills before and after treatment. For example, we assume that there is no Ashenfelter (1978) dip prior to treatment, in which performance falls just prior to team adoption, and that there is no change in the variance of the

$$(10) \qquad E\left(Y_{it}^{\tilde{T}} - Y_{it}^{N} \mid X_{it}\right) = \left(\alpha_2 + \alpha_3 C_{it}\right) + E\left[\left(\gamma_i^{C} + \varepsilon_{it}^{\tilde{T}} - \varepsilon_{it}^{N}\right)T_{it} + \varepsilon_{it}^{N}\right]$$

The expected value of the residual on the right-hand-side is not equal to zero, but equals

$$(11) \qquad E\left[\left(\gamma_i^{C} + \varepsilon_{it}^{\tilde{T}} - \varepsilon_{it}^{N}\right)T_{it} + \varepsilon_{it}^{N}\right] = E\left[\left(\left(\gamma_i^{C} + \varepsilon_{it}^{\tilde{T}} - \varepsilon_{it}^{N}\right)\mid T_{it} = 1\right)\Pr\left(T_{it} = 1\right)\right]$$

Because the choice of $T_{it} = 1$ is determined in part by γ_i^{C}, $E\left[\left(\gamma_i^{C} + \varepsilon_{it}^{\tilde{T}} - \varepsilon_{it}^{N}\right)T_{it} + \varepsilon_{it}^{N}\right] \neq 0$ (and most likely E[•]>0). This is the nature of the selection bias. We have introduced the unobserved complexity term, γ_i^{C}, to emphasize that mills are likely to adopt teams as a function of this unobserved variable, and it in turn enters the production function and thus biases the results. To facilitate our discussion below, we focus on this element of the residual in (11), with the assumption that $E\left[\varepsilon_{it}^{\tilde{T}} - \varepsilon_{it}^{N}\right] = 0$.

There are a number of alternative approaches to estimating the average treatment effect, but we focus on one, semiparametric estimation with fixed effects. To examine the robustness of this estimator relative to the typical alternatives, the results for several alternatives are reported in Appendix A2. Relative to these alternatives, the advantage of semiparametric estimation is a gain in robustness, arising from the less restrictive assumptions regarding the behavior of firms and the error terms (dropping normality). The approach that we follow is adapted from Heckman, Ichimura, Smith, and Todd (1998) and we use Andrews and Schafgans (1996) (which modifies Heckman, 1990) to recover the teams coefficient. For details on estimation, see the Appendix A3.

Temporarily disregarding fixed effects, rewrite the productivity equations (5') and (6) as

$$(12) \qquad Y_{it}^{\tilde{T}} = \theta\, \tilde{X}_{it} + g_{T}\left(p_{it}\right) + \varepsilon_{it}^{\tilde{T}}$$
$$(13) \qquad Y_{it}^{N} = \theta\, \tilde{X}_{it} + g_{N}\left(p_{it}\right) + \varepsilon_{it}^{N}$$

where $\tilde{X}$ includes TEAMS, INCENTIVE, and the control variables, p is the probability of adopting teams, and superscript $\tilde{T}$ refers to those lines that eventually adopt teams, and superscript N to those lines that never adopt teams. We assume that the nonparametric bias correction functions, $g_k(p_{it})$ for $k = \tilde{T}, N$, capture any bias in the model. The conditional expectation of (12) and (13) given p_{it} is

$$(14) \qquad E\left[Y_{it}^{k} \mid p_{it}\right] = \theta\, E\left[\tilde{X}_{it} \mid p_{it}\right] + g_k\left(p_{it}\right)$$

for $k = \tilde{T}, N$. Taking the difference of the productivity equation from the conditional expected equation (differencing (12)-(14) and (13)-(14)) produces

$$(15) \qquad Y_{it}^{k} - E\left[Y_{it}^{k} \mid p_{it}\right] = \theta\left(X_{it} - E\left[\tilde{X}_{it} \mid p_{it}\right]\right) + \varepsilon_{it}^{k}$$

---

residual. These assumptions are confirmed for our data—mills that adopt are not those that see a drop in performance which is rewritten as

(16) $\quad y_{it}^{k} = \theta \, \tilde{x}_{it} + \varepsilon_{it}^{k}$

where the bias correction term, $g_k(p_{it})$, is differenced out of the equation. Thus, the coefficient estimates from (16) are the unbiased average treatment effects for TEAMS. To estimate (16), begin by estimating the expectations, $E\left[Y_{it}^{k} | p_{it}\right]$ and $E\left[X_{it} | p_{it}\right]$, using Kernel estimation given the p_{it} from the hazard model results reported in Table 3. Given these expected values for each line *i* at time *t*, we create the variables in (16) and estimate with OLS. However, because our productivity regressions are likely to contain important line-specific fixed effects (as shown in Table 3), we estimate these in fixed effects form (see the Appendix A3 for details).

The results of the semiparametric estimation are in columns 3 and 4 of Table 4. The estimated TEAMS effect rises modestly for the estimation of equation (16) (column 3). In the regression of column 4, we change our definition of teams status. In column 3, teams status is defined as the line eventually having teams, so the teams coefficient can be estimated from the lines that move from non-teams to teams status. In column 4, we define teams status as currently having teams. In this case, the regressions are divided into teams/non-teams, so the TEAMS variable is differenced out of the regression, and thus its coefficient is not estimated. We therefore introduce the econometric model of Andrews and Schafgans (1998) that estimates the mean values for mills with teams and without, and uses these to produce an implied mean return to teams. The estimates in column 4 are very similar to those in 3—the estimated TEAMS effect falls to .396.

Summary of Treatment Effects

For both the average treatment effect and the treatment of the treated, we find that the selection-corrected coefficient on TEAMS rises modestly, rather than declining due to positive selection bias. One very likely reason for the increase is the way in which complexity enters the productivity regression.

The complexity variable enters the performance equation in two opposing ways. The TEAMS variable is positively correlated with complexity—more complex lines are more likely to adopt teams (recall Table 2)—and TEAMS raise performance. On the other hand, more complex lines have lower levels of yield. It is harder to get high levels of yield when you are producing complex products with more involved equipment. Accordingly, in our regression results for equation (3), the complexity variable C_{it} has a significant negative effect on yield (see Table 3). Therefore, if there is an unobserved omitted complexity term in the performance regression, the coefficient on TEAMS will be biased downward, because TEAMS is positively correlated with an omitted complexity variable that has a negative effect on performance. Indeed, when the observed complexity variable C_{it} is omitted from the regression, the coefficient on TEAMS declines (for OLS and fixed effects estimation). When selection corrections are added, the coefficient on TEAMS rises, because the various selection corrections purposely reduce the correlation between the TEAMS variable and the unobserved complexity error term.

---

prior to adoption.

The main conclusion of the analysis of this section is that we find little bias in the estimated TEAMS effect. This conclusion is perhaps not so surprising given our research methodology, since the careful construction of our sample should reduce this selection bias. That is, we have limited our study to very similar steel rolling mills because we wish to reduce the possible correlation between teams use and other variables that influence productivity. This limitation enhances the probability of getting unbiased treatment effects.[26] In fact, we find that the characteristics of our treated group are very similar to our nontreated group. But the careful selection of our sample also means that we are not choosing a control group from the general population of all firms, and therefore do not know whether these results from the steel industry would transfer to all firms.

The Adoption of Two Endogenous Practices and Bias in the Test for Complementarity in Production

Thus far, we have addressed in this section the selection bias that arises from endogeneity of the decision to adopt problem-solving teams. But we have not addressed the selection bias that can arise if incentives are endogeneous as well. Athey and Stern (1998) stress that this endogeneity bias can be particularly problematic when there are two or more practices and the researcher is testing for complementarity between the practices. In particular, correlated unobservables in the returns to practices will lead to biased tests for the complementarity of the practices. This section now addresses this concern.

Consider a simple example of selection bias in our two-practice model. Rewrite the production function, (3), dropping the direct and interactive effects of the complexity variable, C_{it}, to simplify the exposition.

(17) $$Y_{it} = \alpha_0 + \alpha_1 I_{it} + \alpha_2 T_{it} + \alpha_3 I_{it} T_{it} + \theta X_{it} + u_{it}$$

Assume, however, that the effects of *I* and of *T* on performance depend upon an unmeasured variable, such as the presence of a culture of a positive work ethic in the mill. Mills with a "positive culture" may be more likely to adopt teams and incentives and the returns to these practices may rise with culture. To develop implications for modeling complementarity, assume that the joint effect of $I_{it}T_{it}$ on productivity, or the α_3, does not depend upon the culture, but that α_1 and α_2 do. In this case:

(18) $$\alpha_1 = \alpha_{11} + \alpha_{12}\text{Culture}_{it}$$

(19) $$\alpha_2 = \alpha_{21} + \alpha_{22}\text{Culture}_{it}$$

so that a positive culture increases the value of both incentives and teams, and thus increases the probability of adopting incentive pay and teams. Then insert (18) and (19) into (17) and rewrite as:

(20) $$Y_{it} = \alpha_0 + (\alpha_{11} + \alpha_{12}\text{Culture}_{it})I_{it} + (\alpha_{21} + \alpha_{22}\text{Culture}_{it})T_{it} + \alpha_3 I_{it} T_{it} + u_{it}$$

or

(21) $$Y_{it} = \alpha_0 + \alpha_{11} I_{it} + \alpha_{21} T_{it} + \alpha_3 I_{it} T_{it} + (\alpha_{12} I_{it} + \alpha_{22} T_{it})\text{Culture}_{it} + u_{it}$$

[26] Heckman, Lalonde and Smith (1999) point out that when "analysts.. compare comparable people…much of the bias in using nonexperimental methods is attentuated." (page 3). We attempt to reduce this bias by comparing comparable lines.

Athey and Stern (1998) show that when α_3 is zero (so there is no true complementarity between practices I_{it} and T_{it}), econometric results will still give the appearance of complementarity when culture is unobserved and when the returns to *I* and *T* are a function of the unobserved culture as in (18) and (19). The estimate of α_3 will be biased upwards and the correlation between teams and incentives will be positive.

This potential problem is not likely to be a source of significant bias in our estimates of the joint effect of teams and incentive pay. In particular, with the rare exception of a few months in one mill, mills do not adopt teams without incentive pay. This is important in that it suggests that α_2 is effectively close to zero, because the value of teams alone are so low as to never be adopted without incentives. If we assume that α_2 is zero, it is evident in (21) that there is no upward bias in the complementarity effect, because the coefficient on Culture in that equation reduces to (α_{12} I_{it}) and the only potential bias is in the incentive effect, as the productivity equation (17) becomes:

(22) $$Y_{it} = \alpha_0 + \alpha_{11} I_{it} + \alpha_3 I_{it} T_{it} + (\alpha_{12} I_{it})Culture_{it} + u_{it}$$

With α_2 close to zero, we therefore have focused the econometric analysis of Section IV on the potential endogeneity of Teams and the resulting selection bias in the Teams coefficient (a bias that is not shown in the equation above). While there would be efficiency gains to estimating a full structural model as described in Athey and Stern (1998), we also see no convincing instruments for the adoption of Incentives, in part because they are nearly universally adopted. Finally, note that the Athey and Stern model uses cross-sectional data. With our panel data set, the fixed effects estimates control to some degree for this type of selection bias. For example, if culture is specific to a mill and does not change with the adoption of new practices, then first differencing the data when estimating the productivity equation (22) as we have done in the fixed effects model will eliminate the bias term. In fact, this may be the explanation for the reduction in the magnitude of the coefficient on the incentive variable shown in Table 3 after fixed effects are incorporated in the productivity equation. But to account for the significant effects of teams and incentives in the fixed effects models in columns 3 and 4 of Table 3, this omitted variable (which we proxy by Culture in the above analysis) must be a factor that varies over the five year time periods for the lines in our sample.

VI. Conclusion

The main theoretical hypothesis of this paper is that, in more technologically complex production environments, problem-solving teams are an important means for elevating the effectiveness of group-based incentive pay. Econometric evidence using a unique data base on the HRM practices and performance of U.S. steel minimills provides consistently strong support for this proposition. In models of the adoption of problem-solving teams, we find that a one standard deviation increase in this study's measure of the complexity of minimill production more than doubles the likelihood that mills adopt teams. This empirical pattern implies that teams are more valuable in more complex production environments.

We also provide direct evidence on the productivity effects of group incentives and problem-solving teams by estimating models of the productivity of minimill production lines. These estimates show that the productivity increase that lines achieve when they combine teams with group-based incentive pay is larger in more complex production lines. Group-based incentive pay on its own raises line productivity by some .51 percentage points. The adoption of teams in addition to group incentives leads to a further productivity increase of .13 percentage points in production lines of "average complexity" and a larger increase of .39

percentage points in the most complex production lines. The average productivity effect of teams and incentives that we estimate is economically important, corresponding to an annual increase of some 3000 additional tons of steel valued at over $1.4 million. We also show that our estimates of the productivity effects of these HRM practices are little changed by corrections for possible selectivity bias. Furthermore, the conclusion that teams and group incentives are complementary HRM practices in complex production lines is not likely to be the result of some unmeasured input in the production function that affects the productivity of teams and incentives.

Our analysis underscores a point that is not well emphasized in existing studies of incentive pay. Most theoretical models of the firm emphasize the need to give workers the *incentive* to increase output (by designing an optimal incentive pay plan) and the *ability* to increase output (by investing in the necessary skills and human capital).[27] We show here that *opportunity* to increase output is also important. The returns to incentive pay are greater when problem-solving teams provide production workers with the time and opportunity to diagnose problems and to design and implement technological changes on their lines. The prevailing assumption in existing economic analyses may have been that workers would find ways of increasing output given the incentive to do so, or that incentive pay should not be offered when workers do not have an opportunity to influence output. Our model and results emphasize that *when job design is endogenous, incentive pay can be made more effective by structuring jobs to increase workers' opportunities to respond to the incentives.* Holmstrom and Milgrom (1991, 25-26) make a similar point when they argue that "job design is an important instrument for the control of incentives … [and that] job design can enhance the value of incentive pay, by allocating people to jobs in which incentives are most efficient." Our analysis helps make explicit an important mechanism through which job design enhances the value of incentive pay. Team-based job design provides workers with a greater opportunity to react to incentives.

The empirical results also provide additional understanding of the causes for limited adoption of innovative work practices like teams that promote employee participation. Cross-industry studies of firms with and without teams or other innovative HRM practices often purport to estimate effects of these work practices on measures of firm productivity or profitability. Yet, when positive performance effects are documented in these studies, the question naturally arises as to why all firms would not adopt these new HRM practices. There are transition costs associated with changing HRM practices—costs that can be significant, because workers must invest in new skills and forge new communications relationships.[28] Our analysis here goes beyond transition costs—to uncover fundamental differences in the expected value of innovative HRM practices across plants that should limit their adoption. Not all plants should adopt teams—the benefits appear to be high for plants with complex production processes, but nonexistent for plants without complex processes. Thus, complex plants should invest in problem-solving teams and incentive pay, and plants producing simple commodity products should use standard operating procedures for production. Finally, our results suggest that, as U.S. manufacturers invest more and more in new computer and information technologies, the use of new work practices like problem-solving teams will become increasingly common:

[27] For reviews of the adoption of incentive pay and the optimal structure of incentive pay, see Brown (1990), MacLeod and Parent (1999), Gibbons (1996), Gibbons and Waldman (1998), Lazear (1995), and Prendergast (1995). There are also other incentive structures that can work well with teams—see Baker, Gibbons, and Murphy (forthcoming) on relational contracts and Kreps (1997) on extrinsic versus intrinsic incentives.

[28] For more on transition costs, see Appelbaum, et.al. (2000), Ichniowski and Shaw (1995), and Gant, Ichniowski and Shaw (2001). For more on the need to elevate the skill level of workers, see Autor, Levy, and Murname (2001a) and (2001b), and Breshnahan, Brynjolfsson and Hitt (2001).

information technology puts information in the hands of production workers and innovative work practices give them the authority to use that information.[29]

[29] See Breshnahan, Brynjolfsson and Hitt (2001) and Brynjolfsson and Hitt (2000).

References

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.

Aoki, Masahiko (1988) *Information, Incentives and Bargaining in the Japanese Economy* (New York: Cambridge University Press).

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.

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

Athey, Susan and Scott Stern (1998) "An Empirical Framework for Testing Theories About Complementarity in Organizational Design," MIT Sloan School Working Paper no. 4022BPS-98.

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."

Baker, George (1992) "Incentive Contracts and Performance Measurement," *Journal of Political Economy*, vol. 100, pp. 598-614.

Baker, George, Robert Gibbons and Kevin Murphy (1994) "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, vol. 108, pp. 1125-1156.

Baker, George, Robert Gibbons and Kevin Murphy (forthcoming) "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*.

Bartel, Ann, "Human Resource Management and Performance in the Service Sector: The Case of Bank Branches," Columbia University working paper, 2001.

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

Black, Sandra and Lisa Lynch (forthcoming) "How to Compete: The Impact of Workplace Practices and IT on Productivity," *Review of Economics and Statistics*.

Black, Sandra and Lisa Lynch (May 1996) "Human Capital Investments and Productivity," *American Economic Review Papers and Proceedings*: vol. 86: 263-267.

Black, Sandra and Lisa Lynch (1999) "The New Workplace: What Does it Mean for Employers," *Industrial and Labor Relations Association Papers and Proceedings.*

Brynjolfsson, Erik, and Loren Hitt (Fall 2000) "Beyond Computation: Information Technology, Organizational Transformation, and Business Performance," *Journal of Economic Perspectives*.

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

Cole, Robert E. (2000) "Market Pressures and Institutional Forces: The Early Years of the Quality Movement," in *The Quality Movement and Organization Theory*, eds. Robert E. Cole and W. Richard Scott (Sage Publications, Inc.: Thousand Oaks, CA), pp. 67-88.

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.

Freeman, Richard B. and Joel Rogers (1999) *What Workers Want*, Ithaca, NY: Cornell University Press.

Gant, Jon, Casey Ichniowski, and Kathryn Shaw (2001 forthcoming) "Social Capital and Organizational Change in High Involvement and Traditional Work Organizations," *Journal of Economics and Management Strategy*.

Geanakoplos, John and Paul Milgrom (1991), "A Theory of Hierarchies Based on Limited Attention Span" *Journal of the Japanese and International Economics*, vol. 5, 205-225.

Gibbons, Robert (1996) "Incentives and Careers in Organizations," MIT Working Paper.

Griliches, Zvi (1957) "Specification Bias in Estimates of Production Functions," *Journal of Farm Economics*, vol. 39, pp. 8-20.

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 (1987) "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, vol. 55, pp. 303-328.

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.

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.

Itoh, Hideshi (1994) "Job Design, Delegation, and Cooperation: A Principal-Agent Analysis" *European Economic Review*, vol. 38, pp. 691-700.

Jensen, Michael C. and William H. Mechling (1992) "Specific and General Knowledge and Organizational Structure," in *Contract Economics*, eds. Lars Werin and Hans Wijkander (Blackwell: Oxford), pp. 251-74.

Kandel, Eugene and Edward Lazear (1992) "Peer Pressure and Partnerships," *Journal of Political Economy*, vol. 100, pp. 801-817.

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

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

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

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

Lazear, Edward (1999) "Globalization and the Market for Teammates," *Economic Journal*.

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.

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.

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.

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

Wand, M.P. (1994) “Fast Computation of Multivariate Kernel Estimators,” *Journal of Computational and Graphical Statistics*, vol. 3, pp. 433-445.

Appendix A1: Solution of the Section II Model

In the first stage of the model solution, the firm finds the incentive that will induce its employee to exert different effort levels, and the expected profits from those incentive and efforts. The employee's preferences are represented by the utility function $u(W,e_1,e_2) = - \exp[-r(W-c(e_1,e_2))]$, where r is a risk-aversion parameter, W is the agent's income, and $c(e_1,e_2)$ is the disutility of effort, which is measured as a monetary cost. We follow Itoh (1994) and model the disutility of effort as $c(e_1,e_2) = (c/2)e_1 + (c/2)e_2 + \delta ce_1,e_2$. If $\delta > 0$, diligence in both dimensions has a higher disutility than just the sum of diligence in each dimension alone, so the model includes an additional cost for performing multiple tasks. We model the employees as a single unit. This abstracts from potential free-riding problems associated with group incentives.[30] Of course, in our empirical work we test whether incentive pay is effective and thus implicitly test whether the free-rider problem is overwhelming. The risk-neutral principal (employer) pays an output-based linear incentive to the employees, $W=\beta p f(e_1,e_2,T)+\gamma$, where β and γ define the incentive payment and base pay; p is the revenue less the (constant) marginal cost per unit of output, and f is the production function.[31] The principal sets the employees' share of revenue, β, to maximize expected profits. We assume that the firm operates in competitive markets for both inputs and outputs. The firm's expected profits are $\pi(e_1,e_2) = \mathrm{E}[(1-\beta)p f(e_1,e_2) - k - sT - \gamma]$ where k is the fixed cost of production, and s is the ongoing cost of problem-solving teams. Table A1 presents the incentive structure and profits arising from each of the effort levels.

Table A1: Effort combinations, incentives and expected profits

Effort (e_1,e_2)	Incentive	Expected Profit
(0,0)	$\beta=0$	$pA-k-W$
(1,0)	$\beta=c/(2pA_1)$	$p(A+A_1)-k-W-c/2-[r\sigma^2c^2/(8A_1^2)]$
(0,1)	$\beta=c/(2pA_2)$ (Note: only feasible when $A_2>A_1$)	$p(A+A_2)-k-s-W-c/2-[r\sigma^2c^2/(8A_2^2)]$
(1,1)	if $A_2>A_1$, $\beta=[(1+2\delta)c]/(2pA_1)$ if $A_1>A_2$, $\beta=[(1+2\delta)c]/(2pA_2)$	$p(A+A_1+A_2)-k-s-W-c-\delta c-[r\sigma^2c^2(1+2\delta)^2/(8A_1^2)]$ $p(A+A_1+A_2)-k-s-W-c-\delta c-[r\sigma^2c^2(1+2\delta)^2/(8A_2^2)]$

[30] Kandel and Lazear (1992) and MacLeod (1988) present models that limit or eliminate the free-rider problem. This problem is that each employee's incentive to contribute to production falls as the number of employees N rises, because the incentive pay from the additional production is shared evenly among all N employees while the burden of the effort is not shared, so that for large N, the employee will become a free rider on the effort of his peers. If everyone acts in this manner, there is no increase in effort and no incentive pay. Kandel and Lazear (1992) develop a model in which this drawback can be overcome, by encouraging managerial practices that build positive peer pressure in the plant, causing employees to monitor the behavior of their peers so that each produces his share and all gain the incentive pay. MacLeod suggests that the group will enforce high effort, because any drop in production will be attributed to free-riding by some member, and the group will respond with all individuals reverting to low effort, so that free rider does not receive the incentive. Thus the group response enforces common high effort.

In the second stage of the model solution, the firm chooses the effort combination that maximizes expected profits, and sets β equal to the value that achieves that effort combination. In accordance with intuition, inducing high effort is optimal when the productivity of that effort is high. The profit functions in Table A1 show that a number of factors affect the decision to adopt incentives or teams. Further details on the solution of the model are available from the authors on request.[32]

Appendix A2: Alternative Estimators of the Average Treatment Effect

There are several alternative estimators that are commonly used in this literature, so the results of these alternatives are reported herein to evaluate the robustness of the results reported thus far

Fixed Effects Estimation. The fixed effects estimator can be considered equivalent to a difference-in-differences estimator, in which we compare productivity gains for teams adopters to those for non-teams adopters (thereby ruling out any economy-wide or life-cycle shocks that may be occurring as teams are adopted). The assumptions underlying this estimator are the same as those underlying the fixed effects estimator used for the treatment of the treated model above (concerning the parametric functional form and distribution of the residuals). If we assume that γ^T=0, so adoption is a function of only observed complexity, the fixed effects model produces an unbiased estimate of the average gain to teams. These results were shown previously in columns 3 and 4 of Table 3.

Instrumental Variables Estimation with Fixed Effects. Assuming now that the parametric forms of (5) and (6) are correct, then the instrumental variables estimator provides an unbiased estimate of the average treatment effect if we find instruments for the TEAMS variable that are uncorrelated with the unobserved effect, γ^T_{it}, but that capture the random aspects of teams selection (assuming that some lines adopt teams even when they do not have the higher product complexity). Based on our preceding adoption model, the instruments that we have are proxies for transition costs. The results of this estimation is that both coefficients on the INCENTIVE and TEAMS variables become highly insignificant, though larger in value. These results suggest that our instruments are inadequate-- the R-squared for our TEAMS variable as a function of transition costs is a low .01.

Heckman Correction with Fixed Effects. In the Heckman selection model, the average treatment effect can be estimated consistently if the non-zero portion of the expected value of the residual is controlled for in the regression by introducing the inverse of the Mills ratio. We assume that mills adopt teams when the payoff to adoption crosses a minimum threshold of zero (equation (2)), and assuming joint normality between the error in this selection equation and the production equation.In our model, the vector of variables identifying adoption are the same as those used in the hazard models of Table 2 (product complexity, new

---

[31] Our motivations for using a linear incentive are that it most closely resembles observed practices in the steel industry, that it makes the problem more tractable, and that it is optimal in many circumstances. See Holmstrom and Milgrom (1987) for a discussion of these points.

[32] Note the asymmetry introduced as a result of the requirement that problem-solving effort be accompanied by an enabling mechanism to be effective. The firm can restrict problem-solving effort, and choose to induce production effort even when problem-solving effort is more effective than production effort. However, if production effort has a higher payoff, then the firm cannot induce only problem solving effort. The incentive designed to bring about problem-solving effort will instead bring about production effort. (The employees would be making a decision that is profit maximizing; the firm would be wasting the cost of teams.)

managers, employee experience, and transition costs). These selection corrected results are comparable to those of Tables 3 and 4: .337(.0124) INCENTIVE; .366 (.009) TEAM. The coefficient on the inverse of the Mills ratio is significantly positive, as a model of positive selection bias would suggest. The drawback of the Heckman model is that if the assumption of joint normality of the errors is violated, the estimates will not be consistent. For this reason, we focus on the semiparametric selection correction.

Appendix A3: Semi-parametric estimation process.

In this appendix, we provide details on the non-parametric estimation process. The steps involved in the estimation are:

1. Select a bandwidth and a kernel for the nonparametric estimation. We use the biweight kernel, and a bandwidth of .5. Most analysis of nonparametric estimation indicates that the choice of kernel does not affect the results, but the bandwidth can. We also used bandwidths of .3 and .7, with very little change in the estimated coefficients.

2. Obtain the index parameter. These are the estimated probabilities of adopting teams, obtained from the hazard rate model of adoption with the dependent variable TEAM.

3. Trim the data. Given the kernel and bandwidth, kernel density estimates are obtained, and low frequency observations are trimmed. We trimmed 2.5% of the data. Trimming was done separately for the two groups of data (the teams adopting and the non-adopting groups).

4. Obtain nonparametric regression estimates. Using the trimmed data, the Nadaraya-Watson local constant estimates are obtained. To reduce the computational burden, local linear binning (see Wand (1994)) was used.

5. Transform the data, and conduct OLS estimation on the resulting variables (stacking the data for both groups, imposing the constraint of equal coefficients for both groups).

6. For the model where the bias is determined by the current status of teams, the mean must be obtained. The coefficient estimates are used to determine the residuals. Then the following formula from Andrews and Schafgans (1996) can be applied:

$$\mu_n = \frac{\Sigma_{i=1}^{n} (Y_i - X_i'\theta) D_i \, s\left(Z_i\hat{\beta} - r_n\right)}{\Sigma_{i=1}^{n} D_i \, s\left(Z_i\hat{\beta} - r_n\right)}$$

where $Z_i\hat{\beta}$ are the index values from the hazard rate model estimation of adoption, $s(\cdot)$ is a smoothing function, and r_n is a smoothing parameter that goes to infinity as sample size gets large. (In the estimation, we replace the index with the probability of adoption, and assume the smoothing parameter goes to one with sample size.) The results were sensitive to the smoothing parameter, with higher estimated values of teams for larger smoothing parameters. The reported values were obtained with a smoothing parameter of .3. Conducting this estimation for the samples with TEAM=1 and TEAM=0 provides estimates of the mean for the team and non-teams subsamples. The difference of the estimates is the return to teams for a randomly selected line.

Table 1
Variable Definitions and Means

YIELD	Output tons as percent of input tons	93.27
INCENTIVE	Incentive in use (dummy)	.81
TEAMS	Formal problem-solving teams in use (dummy)	.27
COMP	Scale for the complexity of the mill's production	431.3
PROD1	% of Production in low complexity group 1	.43
PROD2	% of Production in intermediate complexity group 2	.25
PROD3	% of Production in intermediate complexity group 3	.13
PROD4	% of Production in high complexity group 4	.20
GMTEN	Tenure of general manager at mill (years)	5.81
RMTEN	Tenure of rolling mill manager at mill (years)	7.24
OPAGE	Average age of operator at mill (years)	6.03
OPLONG	Average tenure of operators over 10 years (dummy)	.45
EDATE	Number of days since January 1, 1960	11724
LINEAGE	Age of the equipment (years)	9.32
EDATE	Time scale (days since January 1, 1960)	11,724
MILLPR	Weighted average of price for mill's products ($/ton)	420.69
CEONEW	Company CEO in position less than 2 years (dummy)	.14
GMNEW	Mill GM in position less than 2 years (dummy)	.18
RMNEW	Rolling Mill Manager in position less than 2 years (dummy)	.11
TURN	Annual turnover of the mill (% of workforce)	5.45
SOLO	Corporation contains only 1 mini-mill	.38

Table 2
Adoption of Problem-Solving Teams
Hazard Rate Models

	Exponential	Exponential	Weibull	Weibull
PROD2	1.102	---	1.256	---
	(1.156)		(1.405)	
PROD3	4.096***	---	4.185***	---
	(0.991)		(1.173)	
PROD4	3.213***	---	3.121***	---
	(1.152)		(1.209)	
COMP	---	0.007**	---	0.007**
		(0.003)		(0.003)
OPLONG	3.045***	2.624***	3.170***	2.672***
	(0.815)	(0.969)	(1.041)	(1.030)
CEONEW	1.772***	1.965***	1.737**	1.955***
	(0.654)	(0.725)	(0.690)	(0.735)
GMNEW	-0.249	-0.227	-0.254	-0.229
	(0.884)	(0.956)	(0.896)	(0.966)
RMNEW	1.336	1.343	1.299	1.324
	(0.817)	(0.943)	(0.875)	(0.969)
OPAGE	-0.212**	-0.175	-0.215**	-0.179
	(0.093)	(0.088)	(0.088)	(0.092)
SOLO	-1.312**	-0.031	-1.271**	-0.033
	(0.621)	(0.764)	(0.642)	(0.777)
LINEAGE	0.228	0.203	0.225	0.205
	(0.238)	(0.247)	(0.241)	(0.253)
LINEAGE2	-0.010	-0.009	-0.010	-0.010
	(0.009)	(0.010)	(0.010)	(0.010)
TURN	-0.06	-0.044*	-0.059	-0.041
	(0.059)	(0.026)	(0.057)	(0.029)
Shape parameter[a]			0.827	0.903
			(3.417)	(4.323)
Number of Mills	31	31	31	31
Time at "risk" (months)	1,703	1,703	1,703	1,703
X^2	56.44	28.24	52.45	26.10

Figures in parenthesis are standard errors.
*** means p-values $-value$<$.01, ** means p-values$<$.05, * means p-values$<$.1

a – A shape parameter less than 1 indicates negative duration dependence.

Table 3:
Productivity Regressions
Dependent Variable: Yield

	1 OLS	2 OLS	3 FE/AR1	4 FE/AR1
INCENTIVE	1.111***	1.127***	0.493***	0.511***
	(0.109)	(0.108)	(0.128)	(0.127)
TEAMS	0.186*	-2.106***	0.248***	-0.610
	0.107	(0.554)	(0.092)	(0.509)
COMP•TEAMS	---	5.000E-03***	---	1.712E-03*
		(1.00E-03)		(9.84E-04)
EDATE	3.19E-04***	3.19E-04***	2.08E-04***	2.21E-04***
	(4.12E-05)	(4.10E-05)	(5.34E-05)	(5.42E-05)
LINEAGE	-0.011	0.002	0.022	0.018
	(0.024)	(0.024)	(0.025)	(0.026)
LINEAGE2	-3.28E-03***	-3.57E-03***	-1.10E-03	-9.28E-04
	(8.41E-04)	(8.40E-04)	(8.65E-04)	(8.84E-04)
PROD2	-4.668***	-4.701***	-1.306***	-1.312***
	(0.152)	(0.151)	(0.434)	(0.434)
PROD3	-2.929***	-2.972***	38.544	44.499
	(0.188)	(0.188)	(107.022)	(107.117)
PROD4	-6.325***	-6.627***	39.107	45.026
	(0.160)	(0.175)	(107.015)	(107.110)
GMTEN	0.114***	0.112***	-0.036***	-0.037***
	(0.012)	(0.012)	(0.012)	(0.012)
RMTEN	-0.058***	-0.050***	0.020*	0.019*
	(0.012)	(0.013)	(0.012)	(0.012)
OPAGE	0.042 ***	0.030**	-0.190***	-0.189***
	(0.015)	(0.015)	(0.066)	(0.066)
OPLONG	-0.339**	-0.203	1.092***	1.118***
	(0.132)	(0.136)	(0.165)	(0.163)
N	2250	2250	2250	2250
R^2	0.751	0.753		
X^2			10196.500	9981.730

Other control variables in the regressions include nine technology variables, turnover, crew size, and shift length.

Figures in parentheses are standard errors.
*** means p-values $-value$<$.01, ** means p-values$<$.05, * means p-values$<$.1

Table 4

Productivity Regressions Adjusting for Selection Bias
Dependent Variable: Yield

	Treatment of the treated (Subsample of adopters)		Semiparametric Estimation	
	FE/AR1	FE/AR1	Bias by current or future teams	Bias by current teams
INCENTIVE	0.615***	0.627***	0.3666***	0.2309**
	(0.148)	(0.147)	(0.1245)	(0.1305)
TEAMS	0.288***	-0.806	0.3964***	---
	(0.094)	(0.507)	(0.0996)	-
COMP•TEAMS	---	2.176E-03**	---	---
	-	(9.80E-04)	-	-
PROD2	-1.353**	-1.376**	---	---
	(0.610)	(0.610)	-	-
PROD3	2.416	1.502	---	---
	(4.410)	(4.399)	-	-
PROD4	-0.214	-0.723	---	---
	(1.798)	(1.799)	-	-
OPAGE	-0.219**	-0.227**	-0.2353***	-0.2225***
	(0.100)	(0.100)	(0.088)	(0.0877)
OPLONG	1.006***	0.996***	1.0709***	1.1511***
	(0.178)	(0.178)	(0.1671)	(0.1647)
EDATE	1.35E-04	1.64E-04*	0.0401	0.0413**
	(9.67E-05)	(9.70E-05)	(0.0219)	(0.0218)
LINEAGE	0.001	-0.016	0.0595***	0.0725***
	(0.037)	(0.038)	(0.0271)	(0.027)
LINEAGE2	2.42E-03	3.32E-03**	-0.0029***	-0.0033***
	(1.58E-03)	(1.66E-03)	(0.0009)	(0.0009)
Implied TEAMS				.5799***
				(9.98e-04)
N	1372	1372	2249	2249
R^2			.1722	.1633
X^2	3047.090	3010.610		

Other control variables in the regressions include nine technology variables, turnover, crew size, and shift length.
Figures in parentheses are standard errors.
*** means p-values$<$.01, ** means p-values$<$.05, * means p-values$<$.1

Note: EDATE measured in days in columns 1 and 2, and months in columns 3 and 4