This work is distributed as a Discussion Paper by the

STANFORD INSTITUTE FOR ECONOMIC POLICY RESEARCH

SIEPR Discussion Paper No. 12-001

The Value of Bosses

by Edward P. Lazear, Kathryn L. Shaw, Christopher T. Stanton

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

The Value of Bosses

Edward P. Lazear,
Kathryn L. Shaw,
and
Christopher T. Stanton

Stanford University,
Stanford University,
University of Utah

October 2, 2012

We greatly appreciate the comments of seminar participants at the University of Chicago, Columbia University, Yale University, Stanford University, Harvard Business School, USC, Northwestern University, the AEA meetings, the Society of Labor Economics, the IZA Economics of Leadership Conference, the Utah Winter Business Economics Conference, NBER Personnel Economics and NBER Organizational Economics meetings. We thank our discussants John Abowd, Casey Ichniowski, and Robert Miller for their thoughtful suggestions.

Abstract

How and by how much do supervisors enhance worker productivity? Using a company-based data set on the productivity of technology-based services workers, supervisor effects are estimated and found to be large. Replacing a boss who is in the lower 10% of boss quality with one who is in the upper 10% of boss quality increases a team's total output by about the same amount as would adding one worker to a nine member team. This implies that the average boss is about 1.75 times as productive as the average worker. Additionally, boss's primary activity is teaching skills that persist.

Do bosses have a positive effect on worker output? One extreme view is that bosses are irrelevant and that worker productivity is unaffected by the choice of supervisor. Anyone can do the supervisor's job because the supervisor has little direct impact on worker output. An alternative is that workers are indistinguishable and the output of the firm depends only on how well bosses make use of labor. Using a setting where individual workers frequently switch bosses, the effect of individual bosses on worker productivity is estimated.

Workers depend on their bosses in many ways. First, the hiring decision may rely on input from a worker's prospective supervisor. Second, the supervisor is likely to be important in motivating a worker, which affects worker productivity and the workers' success within the firm. In extreme cases, supervisors discipline and terminate workers. Third, the supervisor acts as mentor or coach, teaching subordinates the techniques that will enhance their productivity. Fourth, supervisors assign tasks to workers and tell them what they must do and may not do on the job.

Despite the potentially important role that supervisors play, the economics literature has largely been silent on the effects that bosses actually have on affecting worker productivity.[1]

[1]The literature has focused on CEOs or managers in detailed occupations. For work on CEOs' productivity, see Bennedsen, et.al (2007a, 2007b), Bertrand and Schoar (2003), Jenter and Lewellen (2010), Kaplan, et.al. (2008), Perez-Gonzalez (2006), Perez-Gonzalez and Wolfenzon (2012), and Schoar and Zuo (2008). The sports sector offers opportunities for strong papers on the effects of coaches on performance (Bridgewater, Kahn, and Goodall, 2011; Dawson, Dobson, Gerrard, 2000; Frick and Simmons, 2008; Goodall, Kahn, and Oswald, 2011; Kahn, 1993; and Porter and Scully, 1982). Recent work in education studies the effects of principals (Branch, Hanushek, and Rivkin, 2012). Regarding hierarchy and managers in law firms see Garicano and Hubbard (2007). Regarding university leaders, see Goodall (2009a, b). Regarding national leaders, see Jones and Olken (2005). Regarding church leaders, see Engelberg, Fisman, Hartzell, and Parsons (2012). Regarding personal traits and leadership, see Kuhn and Weinberger (2005) and Borghans, ter Weel, and Weinberg (2008). Early theoretical work includes Herbert Simon on firm size and compensation (1957) and Rosen on the span of managerial control (1982). For more recent work on leadership, see Hermalin (1998), Rotemberg and Saloner (2000), and Lazear (2012).

Even more to the point, the literature has not been able to speak to the importance of the various mechanisms through which boss effects might operate. Most of this is a data issue, but some of it reflects the fact that the literature has modeled the relationship between boss and worker at an abstract level and has not pushed beyond to examine what is likely to be the most important relationship in the workplace.

The neglect is even more striking when contrasted with the interest in peer effects. There is a large literature, both theoretical and more recently empirical, that has focused on the effects of workers on their peers and team members.[2] Peer effects may be important, but except in a few industries, like academia, where the structure is very flat and workers have much authority over what they do, the relationship with one's boss is likely to be as or more important than that to any other worker. At a minimum, this remains an open question and one that should be investigated.

A significant fraction of resources is devoted to supervision. Among manufacturing workers, front-line supervisors comprised 10 percent of the non-managerial workforce in 2010. Among retail trade workers, front-line supervisor comprised 12 percent of the non-managerial workforce.[3]

By using data from a large service oriented company, it is possible to examine the effects of bosses on their workers' productivity and to compare them to individual worker effects. Daily productivity is measured for 23,878 workers matched to 1,940 bosses over five years from June 2006 through May 2010, resulting in 5,729,508 worker-day measures of productivity. The productivity data are from one production task that we label a TBS job, or "technology-based service" job. The workers are monitored by a computer which provides a measure of productivity. Companies that have TBS jobs like this one include those with retail sales clerks, movie theater concession stand employees, in-house IT specialists, airline gate agents, call center workers, technical repair workers, and a host of other jobs in which an employee is logged into a computer while working. Because of confidentiality restrictions, details about the day-to-day tasks of the workers cannot be revealed for this large company.

[2] For theory, see Kandel and Lazear (1992). For empirical examples, see Mas and Moretti (2009), and Falk and Ichino (2006). For work on teams and complementarities, see Ichniowski and Shaw (2003).

[3] The data is from Bureau of Labor Statistics, Occupational Employment Statistics for 2010. First-line supervisors are an occupational class. For manufacturing, the non-managerial workforce is all those who are not supervisors or managers. For retail, the non-managerial workforce is retail clerks and cashiers.

The primary findings are:

1. Bosses vary greatly in productivity. The difference between the best bosses and worst bosses is significant. Bosses in the top decile increase each worker's output by about 1.3 units per hour more than bosses in the bottom decile. Given that the typical boss supervises about nine workers and the average worker produces about 10.3 units per hour, this amounts to a change in total productivity that is larger than the amount produced by the average worker.
2. Using what we believe is a conservative normalization, the average boss adds about 1.75 times as much output as the average worker, which is in line with the differences in pay received by the two types of employees.
3. The boss's primary job is teaching skills that persist. Contemporaneous motivation of workers is secondary.
4. The worst bosses are unlikely to be retained. Over a one year period, bosses in the lowest 10% of the quality distribution are 67% more likely to leave the firm than bosses in the top 90% of the distribution.
5. The effect of good bosses on high quality workers is greater than the effect of good bosses on lower quality workers, but the effect of sorting is not large.

I. Theoretical Framework

A. Human Capital and Effort

An individual worker's output at time t, q_t, depends on human capital, H_t, which reflects both innate ability and previously learned skills, and on effort, E_t. A natural (although not necessary) specification is multiplicative: harder work results in greater returns to human capital

$$(1) \qquad q_t = H_t * E_t\,.$$

A worker's stock of human capital at time t depends on experiences with current and previous bosses, other variables, the set of which is denoted X_t, and some innate ability, denoted α. Then

$$(2) \qquad H_t = H(X_t, \alpha, b_t)$$

where b_t is the quality-adjusted boss time that a worker has encountered over his or her career up to time t. If the team m to which the worker is assigned contains one boss and N_m workers, then

(3) $b_t = b(d_{jt}/N_{mt}, d_{j\,t-1} / N_{m\,t-1}, ..., d_{j\,0} / N_{m\,0})$

where d_{jt} is the quality of boss j with whom the worker is paired at time t and N_{mt} is the size of the team at time t. Past bosses can affect the worker's output at time t because some of the knowledge and work habits acquired from those bosses may be retained.

Analogously, effort is

(4) $E_t = Z(X_t, \alpha, b_t)$.

Substituting (2), (3) and (4) into (1) yields

$$q_t = H(X_t, \alpha, b(d_{jt}/N_{mt}, d_{j\,t-1} / N_{m\,t-1}, ..., d_{j0} / N_{m0})) \times Z(X_t, \alpha, b(d_{jt}/N_{mt}, d_{j\,t-1}/N_{m\,t-1}, ..., d_{j0}/N_{m0})$$

or

(5) $q_t = f(X_t, \alpha, b(d_{jt}/N_{mt}, d_{j\,t-1} / N_{m\,t-1}, ..., d_{j\,0} / N_{m\,0}))$.

Bosses, in the context in which we study, are most important in their ability to teach and motivate workers. For the most part, they do not engage in task assignment, hiring, or other aspects of the supervisor job, although they may play some role in firing and in promotion. One might expect that the motivation effect of bosses works primarily through effort, E_t, and that the teaching role of bosses works primarily through skill level, H_t, but there is nothing in the specification that requires this.

Bosses also have some endowment of skills and these skills need not be uni-dimensional. For example, it may be that nature endows boss skills such that good teachers are also good motivators. Or the endowments may be negatively correlated: Good drill sergeants may make terrible psychotherapists. There may be some ability to trade these skills off. A boss with any given set of endowed skills might be able to turn one into another by spending a larger fraction of time focused on teaching or motivating.

This framework suggests the following empirical questions:

E1: Do bosses matter? Specifically, do they raise workers' output? If so, by how much?

E2: Do bosses vary in their quality or are they homogeneous?

E3: Do bosses matter because they teach or because they motivate? Which dominates?

B. Sorting Bosses to Workers

Sorting is key. Should good bosses be matched with good workers or with bad workers? It is conceivable that good bosses are more valuable to less able workers because the most able workers can learn by themselves and are innately highly motivated. The reverse is also possible.

The best workers may be able to take better advantage of the knowledge and motivation that a good boss passes on.

The empirical question is:

E4: Comparative Advantage: To which workers should the best bosses be assigned? Do good bosses improve productivity more for the best workers (stars) or more for the worst workers (laggards)?

C. Workers are Additive; Bosses are Multiplicative

Why is a research scientist who makes a great breakthrough so valuable to a firm? It is because the innovation enhances the productivity of a large number of workers. The effect is multiplied by the number of workers that the innovation affects.

The same is true of bosses. A good claims processor can process a larger number of claims than a poorer one, but the effects are limited to the claims that the worker processes himself. Supervisor quality affects output primarily through the work of subordinates and an increase in supervisor quality is multiplied by the number of individuals who are touched by that supervisor. Thus, the effects are multiplicative for supervisors and additive for workers.

Formally, the total output of the firm, Q, is the sum of the individual worker's outputs, q_i,

$$Q = \sum_i q_i$$

so $$\frac{\partial Q}{\partial d_j} = \sum_i \frac{\partial q}{\partial d_j}$$

If the boss effect is the same on each worker, then

$$\frac{\partial Q}{\partial d_j} = N_M \frac{\partial q}{\partial d_j}.$$

The worker effect on total output is simply

$$\frac{\partial Q}{\partial \alpha_i} = \frac{\partial q}{\partial \alpha_i}.$$

Thus, the effect of worker talent on output is just the effect itself, whereas the effect of the boss's talent on output is multiplied by the number of individuals that she supervises.

If peer effects are operative then, like bosses, peers could have multiplicative effects on productivity. This generates the empirical question:

E5: Are boss effects larger than peer effects?

II. Data

The data are from an extremely large service company that has daily records on worker output, linked to the boss to which the worker was assigned on each day.[4] The period covered is June, 2006 to May, 2010. There are 23,878 workers and 1,940 bosses. The unit of analysis is a worker-day and there are about 5.7 million worker-days over the entire period.

The company has multiple service functions, but the data used come from one task classification where workers are involved in general customer transactions. The task is one that is repeated, but each experience has some differences. The choice of one task for analysis ensures that all workers in the sample are engaged in approximately the same activity.

To provide some context, consider an example of a technology-based service job: workers doing computer-based test grading. In most states, students take standardized tests, such as the "Star" tests in California. The students' handwritten essays (in subjects from science to English) are scanned into a computer, and then the graders of these tests sit in large rooms, where they grade each essay on a computer. The graders' work is timed and checked for quality. Graders must be at their desk a certain percent of the day (defined as 'uptime' below), which is recorded, and have modest amounts of incentive pay. They are given frequent feedback on their performance. Their bosses sit with them to teach them grading skills and to motivate the workers. While this may seem like an unusual example, we made a number of plant visits to companies like this and all visits shared this typical scenario.

These jobs are labeled technology-based service jobs because the company uses some form of advanced IT system to record the beginning and ending time for each transaction, or to record the daily volume of transactions, for each worker. As described above, many production processes in services now fit this description. The technology that is used to measure performance may be a new computer-based monitoring system (as in the standardized test grading above), an ERP (Enterprise Resource Planning) system that records a worker's

[4] In reality, the boss is recorded as the regular boss for that day. If there was a substitute boss, say because the usual boss was absent, this would not be picked up in the record.

productivity each day (such as the number of windshield repair visits done by each Safelite worker (Lazear, 1999; Shaw and Lazear, 2008)), cash registers that record each transaction under an employee ID number, call centers, or computer-monitored data entry. These technology-based service jobs are likely to be widespread and represent a major IT-based shift in computerization and measurement of worker productivity. Although some of these jobs are outsourced to firms outside the U.S., many remain in the U.S., particularly when the customer interaction is face-to-face or the work is idiosyncratic and skilled (as in test grading).

The technology-based service workers studied herein are constantly learning. New products or processes are introduced over time so there is learning by workers and the potential for teaching by bosses on the job.

Work takes place in "areas" and the group of workers associated with a given area is labeled a "team." The average daily team size is 9.04 workers and each team is managed by one boss. The team is identified through the worker's link to a boss identification number; all workers with the same boss that day are said to be part of the team. Workers switch bosses about four times per year.[5] It is these switches to different bosses that permits estimation of the effect of bosses on workers' productivity.

There are two measures of output: output-per-hour and uptime as reported in Table 1. Each worker handles about 10.3 transactions per hour. In any hour at work, workers miss some time for breaks and personal time, leaving their work areas and thereby slowing the entire system. This is rare. The mean uptime is 96.3%.

Most of the variation is in output-per-hour rather than in uptime. The standard deviation of output-per-hour is 30.8% of its mean; the standard deviation of uptime is 2.8% of its mean. Consequently, the initial discussion and results focus on variation in output-per-hour. Later, the basic analysis is repeated for uptime.

III. The Basic Results

The empirical specification starts with equation (5) and assumes the structure

$$(6) \quad q_{it} = X_{it}\beta + \alpha_i + \left(\frac{d_j}{N_{m_{it}}}\right) + g\left(\left(\frac{d_{jt-1}}{N_{m_{it-1}}}\right), \ldots, \left(\frac{d_{j0}}{N_{m_{i0}}}\right)\right) + \varphi_{ij} + \varepsilon_{it.}$$

---

[5] The worker-boss pair is defined by the usual worker-boss pairing. If a boss were absent on any given day, the usual boss would be the one of record.

where X_{it} is a set of control variables, described below, and α_i is the worker fixed effect. Each boss j's quality is assumed to be invariant over time, but workers change bosses so in the most complete specification, it is necessary to keep track of the past bosses to whom the worker was assigned. The effect of the boss on all workers taken together is denoted d_{jt} for boss j at time t. The subscript t is dropped for the contemporaneous effect so d_j refers to the boss on the current worker day, t. Recall that $N_{m_{it}}$ is the size of the team for worker i at time t. The parameter φ_{ij} captures concurrent interactive effects, or match effects, in worker-boss pairs. It is conceivable that worker i is better suited to boss j than to boss k and φ_{ij} allows that generality.

A. Brief Discussion of Estimation Issues

Before discussing the estimates, it is important to flag a couple of potential problems that may arise in estimation. First, there may be non-random assignment of workers to bosses. There is some evidence of non-randomness, but it seems less pronounced than might be expected. Much of the technical analysis that follows in section VI below addresses this issue.

If worker quality can be measured well and the functional form of the estimating equation is properly specified, then non-random assignment is not a problem. More specifically, even if good workers are more frequently assigned to good bosses, there is no bias in the estimates of the boss effect so long as the model controls for worker quality and so long as the model allows a given boss to affect workers differently.

At the most basic level, the inclusion of worker effects controls for worker quality, but there are more sophisticated and more comprehensive ways both to test for the extent of the problem and to treat it. A variety of methods is used and described in more detail in the subsequent analysis. They include a general form of mixed estimation, which allows for interaction effects between workers and bosses. All approaches yield the same qualitative conclusions. Bosses are both important to worker productivity and vary in their effectiveness.

Second, because boss (and worker) fixed effects are estimated with error, there can be more variation in the estimated effects than exist in reality. Here, too, a number of approaches are used to correct estimates for this problem. A typical boss effect is estimated with about 3,000 worker-days of data, but a boss effect that is estimated from a small number of worker-days will be estimated less precisely than one estimated with a large number of worker-days. The most basic approach to deal with this issue is to weight calculations by worker-days to downplay the effect of noise on the estimated boss effects. Other more sophisticated approaches are also employed, including one that uses a likelihood-based approach to adjust for sampling error. Again, these are explained in more detail below. The choice of technique affects the estimated effects, but the least sophisticated correction of merely weighting by worker-days and the most sophisticated correction provide very similar results, with both boss effects varying significantly, even after the corrections are made.

B. Regression Results

The most straightforward way to analyze boss effects on worker productivity is to estimate a simple variant of (6) that includes both boss and worker fixed effects. Since most of what follows reinforces the conclusions of this analysis, this is a good place to start because it provides the main intuition behind the primary findings of the study.

The simple fixed-effect specification follows directly from (6), given a number of simplifying assumptions. If it is assumed that all teams are of the same size equal to N_m, that there are no match effects, $\varphi_{ij}=0$, and that there are no effects of past bosses on current output, then (6) reduces to

(7) $$q_{it} = X_{it}\beta + a_i + \delta_j + \varepsilon_{it}.$$

Note that δ_j replaces the d_j because d_j is the effect of the boss on the entire team so $\delta_j = d_j / N_m$.

Table 2 reports the results.[6] All models contain tenure controls given by

$$g(ten) = 1\{ten < 365\} * f(ten) + \mathcal{L}1\{ten \geq 365\} + 1\{ten \geq 365\} * f(365),$$

where f is a fifth order polynomial over the first year.[7] After the first year, f is evaluated at 365, and $\mathcal{L}$ captures this plateauing of the tenure profile. This functional form is chosen because month fixed effects and a worker's tenure profile are nearly colinear after including worker fixed effects.[8] Not surprisingly, and consistent with prior work in other industries,[9] worker output is increasing and concave in tenure.

---

[6] Throughout the remainder of the paper, details about the estimation procedures are contained in the notes accompanying each table.

[7] For these jobs, a portion of the learning is firm-specific and a portion is occupation-specific, and the regressions do not hold constant the latter because the data contains only the start date with the current firm, not general occupational experience. Therefore, the tenure coefficients combine firm-specific learning with occupational learning for those who did not arrive with previous occupational experience, but estimate firm-specific learning for those who arrived with previous experience.

[8] Estimates of $\mathcal{L}$ suggest that the discrete jump at day 365 is less than 3% of the total effect of tenure. Alternative assumptions about the tenure profile do not change the magnitude of worker and boss effects.

The columns of Table 2 display several different specifications: OLS without any fixed effects, worker effects, boss effects, and worker plus boss effects. The results of Table 2 make the most important points, which survive various later specifications changes. Column 2 shows that worker effects are clearly important. When worker effects are included, the R-squared rises from 0.061 to 0.237 with an F(23877, 5705571)=55.6 (p-value = 0), rejecting the null hypothesis that the set of individual fixed effects are zero. Column 3 shows that boss effects contribute to worker productivity, but absent worker fixed effects, these boss effects include worker sorting to bosses. The inclusion of boss effects in addition to worker fixed effects, in Column 4, yields an F statistic of F(1939, 5727509)=102.1 (p-value = 0), rejecting the null hypothesis that the boss fixed effects are jointly zero.

Rows (1) and (4) report the standard deviations of the worker and boss effects when the fixed effects are weighted by worker-days. This approach corrects in the simplest way for excess volatility in estimated boss effects δ_j, volatility that can result from sampling error when there are few workers per boss. To interpret these estimates, see, for example, the 1.34 coefficient in row (1). This states that when the estimation is done with worker fixed-effects included and with boss fixed-effects excluded, the standard deviation in worker fixed effects is 1.34 transactions or "units" per hour.

More important, the standard deviation of the boss effects, shown in row (4), is 3.44, which is about 2 ½ times larger than the standard deviation of the worker fixed effects themselves. There is significant variation in the quality of bosses that is reflected in the amount by which they can affect the productivity of the teams that they supervise. The magnitude of this effect is interpreted in section D below.

Alternative approaches to weighting the boss fixed effects, reported in rows (2)-(3) and (5)-(6), leave the basic conclusions unchanged. The results in rows (2) and (5) use a more sophisticated approach following Rockoff (2004) to reweight the fixed effects. The estimated boss effects are assumed to be drawn from a normal distribution with unknown variance σ_δ^2. The variance of the estimated boss effects is then composed of the sum of sampling error and the true variance of the boss effect. An estimate of the parameter σ_δ^2 can be obtained by maximizing the likelihood of the individual estimates, which is just the normal likelihood with mean zero and

[9] See Lazear (2000), Shaw and Lazear (2008) for examples of estimated productivity-tenure profiles.

variance equal to σ_δ^2 plus sampling error. With an estimate of the variance of the sampling error in hand, it is straightforward to recover an estimate of σ_δ^2. An alternative set of baseline results, reported in rows (3) and (6), are the unweighted standard deviations of the boss effects. The high values of these boss effects reflect sampling error. Some bosses are estimated to be very good or very bad simply because of noise associated with not appearing frequently in the data. Given this undesirable sampling error, these unweighted baseline estimates are of little interest.

C. A More Refined Specification: Mixed Effects

The main story, that there is significant variation in boss quality and that bosses affect worker productivity, is robust to various corrections. One approach is the mixed effects model. The estimates of σ_δ^2 and σ_α^2 are obtained by treating α_i and δ_j as random effects with a mixed estimation that allows an arbitrary correlation between the random effects design matrix [A D] and X, where A and D are matrices of worker and boss indicators and X is the matrix of other right hand side variables. The identifying assumptions are:

E(α|X) = E(δ|X)=0 and $Cov\left(\begin{matrix}\alpha\\ \delta\\ \varepsilon\end{matrix}\middle| X\right) = \begin{bmatrix}\sigma_\alpha^2 I_{\#W} & 0 & 0\\ 0 & \sigma_\delta^2 I_{\#B} & 0\\ 0 & 0 & R\end{bmatrix}$ where $I_{\#W}$ and $I_{\#B}$ are identity matrices with sizes corresponding to the numbers of workers and bosses, respectively. The matrix R is the covariance matrix of the errors. For the purposes of estimating equation (7), R is assumed to be $\sigma_\varepsilon^2 I_N$. The covariance parameters σ_δ^2 and σ_α^2 are estimated via restricted maximum likelihood, using the procedure detailed in Abowd, Kramarz, and Woodcock (2006).

Mixed effects estimation has one advantage over the prior fixed effects estimation. When fixed effects are estimated with a short panel, the estimated fixed effects are not consistent. The results in Table 2 weight the estimated fixed effects by worker-day frequency or by maximum likelihood to reduce the role that some short panel data plays in calculating the standard deviation of the boss effects. In mixed effects, no weighting is needed because the standard deviation of boss effects is calculated. Therefore, for most regressions below, the mixed effects model is the one that is reported, not the fixed effects model. In a few regressions below, only the fixed effects model can be estimated and those results are reported and methods noted.

The results of the mixed effects model are comparable to the results of the fixed effects model. A standard deviation change in boss quality increases total output by about 4.61 units per hour (Panel B of Table 2). This is larger than the 3.44 estimate using fixed effects weighted by worker days and larger than the maximum likelihood estimate of 3.53.

D. The Impact of Bosses

How much do bosses matter? There are two notions of the impact of bosses. One is the increase in productivity that a typical worker would achieve by moving from a poor boss to a good boss. The other is the increase in the productivity of all team members resulting from more time with the average boss.

Using the mixed effects results, the boss who is at the 90th percentile of boss quality distribution increases productivity by 5.9 units/hour more than the boss at the 50th percentile. Comparably, replacing a boss who is in the lower 10% of boss quality with one who is at the 90th percentile increases a team's total output by about the same amount as would adding one worker to a nine member team.

It is important to remember that the estimates of boss effects are lower bounds of the variance in boss effects because of selection. The worst conceivable boss is not likely to be in our sample of bosses. Consequently, the observed distribution is likely to be a truncated version of the underlying potential distribution of boss effects. Even if the distribution of boss effects had no variance, this would not mean that bosses did not matter. It would merely imply that all bosses affected worker output to the same extent. The conclusion is that even among the selected sample of those who are employed as bosses, there is large variation in the effect of bosses on worker output.

The fact that bosses vary significantly in their productivity-enhancement effects implies by necessity that bosses must matter. It can only be the case that a good boss affects productivity by much more than a bad boss when bosses affect productivity in the first place. If bosses were mere decorations, one would expect no variation in boss effects beyond sampling error. The fact there is wide variation in boss effects implies that there is a substantial productivity effect that bosses confer on their teams.

These numbers are not the productivity of the boss, but merely the difference in productivity between very good and very poor bosses. To obtain an estimate of the productivity effect, it is necessary to normalize the estimates. Boss fixed-effects, like all fixed-effects, are relative. They do not provide an estimate of the level of productivity enhancement, but only of the difference between one boss's effect and that of another.

One normalization that may be reasonable and on which evidence is provided below, is based on the idea that those who are bosses are promoted and hired into that position are superior to the best workers. Bosses obtain and retain their jobs only by being more productive as bosses than they would be as workers. Otherwise, comparative advantage would dictate that they operate as workers rather than bosses. It is also reasonable to expect that those who are promoted to boss are identified as being more able than the average worker because they were exceptional producers when they were workers themselves. Of course, promotion mistakes can be and are made, but they tend to be weeded out (as shown later).[10] Therefore, let us assume that the poorest bosses have productivity that is equal to that of the better workers. Specifically, assume that the boss who is at the 10th percentile of the boss quality distribution is as productive as a worker who is at the 90th percentile of the worker quality distribution. The 10th percentile boss is then worth about 12.1 transactions per hour, which is the number of units that the 90th percentile worker produces in a typical hour. Given this benchmark and knowledge of the parameters of the distributions governing worker and boss effects, it is possible to calculate the level of productivity for every boss.

This normalization implies that the average boss produces about 18 transactions per hour in enhanced productivity of that boss's subordinates. Were no bosses present, the typical team of nine workers would handle 18 fewer transactions per hour on a mean of about 100 transactions. This is consistent with our discussions with the firm on levels of compensation. No compensation data are available to us, but we were told that bosses, who are almost twice as productive as workers by this measure, earn between 1.5 to 2 times as much as workers.

E. Demand Variation

Because the firm is large, it should be able to cope with demand shocks by changing the number of workers that it has working at any point in time so as to keep the load constant. Still, unanticipated changes in demand may catch the firm off-guard and may result in changes in output-per-worker. Thus, the question: Are the results sensitive to demand shocks? Perhaps some workers and even bosses look exceptionally good not because they are inherently more productive, but because they encounter more periods of high demand and have lower slack time as a result. Although the measure of output is average transaction time (from which output-per-

[10] See Lazear (2004) for a theoretical exposition of promotion decisions under uncertainty and the effects of error.

hour is inferred), it is possible that workers might speed up when there is a long queue of customers waiting for service.

To control for demand shocks, the regression contains day of the week and monthly dummies.[11] But there is an additional control for daily demand shocks: The mean contemporaneous (daily) output for all other team members, excluding the worker, will pick up demand variations that are specific to the team at any given time.

The regression results are easily summarized. The inclusion of a peer mean output reduces the standard deviation of the boss effect from 3.44 to 2.90, when weighted by worker-days, comparable to row 4 of Table 2.[12] The drawback of including the peer mean output variable is over-controlling: If a good boss raises the mean of every worker on the team, then including peer mean output will inappropriately lower the estimated boss effect. Therefore, the peer mean output variable is excluded from future analyses. But the conclusion is that even with the most powerful proxy for demand, the standard deviation of boss effects remains large.

F. Variation in Team Size

Bosses should have more impact on workers' productivity in smaller teams. Reintroduce team size from (6) in the productivity regression (7) as

$$(8) \qquad q_{it} = X_{it}\beta + a_i + \left(\frac{d_j}{(N_{m_{it}})^{\theta}}\right) + \varepsilon_{it}$$

where θ permits boss attention for any individual to decline at a nonlinear rate with team size. Equation (8) is estimated as a fixed effects model with nonlinear least squares. There is evidence that rising team size lowers the impact of the boss on workers' productivity: The estimated θ is .197. However, the estimated standard deviation of boss effects is modestly larger. The standard deviation of boss effects is 5.32 when multiplied by the observed team size in the data and weighted by worker-days and is 4.41 using mean team size of 9.04 (these estimates are comparable to the functional form of row 4, Table 2).[13]

---

[11] Whether market conditions affect output depends on how good the firm is at adjusting the number of employees at work so as to keep the transaction arrival rate close to constant for any given worker, despite varying demand conditions. In our discussions with the firm, we know that the firm attempts to adjust the number of hours worked so as to minimize slack. Still, there is variation in part because the firm must observe slack persisting for a long enough period of time before it makes sense to send some workers home.

[12] It falls similarly from 3.53 to 3.07 when using the likelihood weighting estimates (comparable to row 5 in Table 2), and falls from 4.61 to 2.73 when using the mixed effects estimates

[13] Team size may not be exogenous. Better bosses may be assigned to larger teams. This would reduce the variance of the estimated boss effects. We have been unsuccessful in finding an appropriate instrument to deal with this

G. The Boss Effects are Identified

The estimates have presumed identification of boss effects. How are they identified? Holding constant the worker's quality, α_i, the boss effect, δ_j, is identified by those workers who switch bosses. The boss effects are estimated off of "changers." In order to estimate the effect of a boss on workers' productivity, the same boss must work with different workers, whose abilities are known through the worker fixed effects. If a given worker switches from boss A to boss B and his productivity rises, then the change in productivity is attributed to the change in bosses. For any given boss, the boss effect is therefore estimated as the average increase across all workers who work for that boss when they switch to that boss (or average decrease when they switch from that boss). More precisely, the boss effects are estimated within "groups" of connected workers in the graph-theoretic sense.[14] If a separate group of bosses and workers is not connected, no worker nor boss ever interacts with any other worker or boss in the non-connected group. Within each group, there must be one normalization of the boss effects and one normalization of the worker effects.

The data are sufficient to estimate the boss effects within each connected group. For each worker, there is an average of 240 days of daily productivity data (or about a calendar year of data). Each worker changes bosses about 4 times during this interval. Therefore, when the boss is the unit of analysis, his team members have, on average, touched 4.7 other bosses. Given the average number of workers per boss, the number of worker changers per boss is 49 (or 80 if weighted by the number of observations per boss). These are sizable numbers. As a result, 99.99% of the daily data is in the largest connected group, with only 560 of the 5.7 million observations and 11 of the 1,940 bosses outside of the largest connected group.

IV. Teaching and Motivating by Bosses

The most general specifications allow prior bosses to affect current period productivity. In the productivity regression, let us call that which persists "teaching" and that which is only contemporaneous "motivation." Teaching is simply defined as that part of what bosses do that has some persistence in its effect on output. It might involve skill transfer or providing the

---

problem.

[14] Paraphrasing, "When a group of [workers] and [bosses] is connected, the group contains all the [workers] who ever worked for any of the [bosses] in the group and all the [bosses] to which any of the workers were ever assigned" (Abowd, Kramarz, and Woodcock, 2006).

worker with a good work ethic and good work habits. As long as it is persistent, we will think of this as a skill that was taught to the worker. Motivation is defined as that which affects performance today, but dies out immediately. A kind word that makes a worker push harder for an hour or two might be included in this kind of effect. Its persistence is limited to the day on which the boss inspires the worker to improve productivity.

Teaching re-introduces the lag structure of (6), where the persistent (teaching) portion of past boss effects is λ. Assume the past boss's effect is independent of the length of the spell with that boss, so that

$$(9) \qquad q_{it} = X_{it}\beta + \alpha_i + \delta_j + \sum_k^j 1\{t > \tau_{ik}\}\gamma^{t-\tau_{ik}}\lambda\delta_k D_{ik,t-\tau} + \varepsilon_{ijt},$$

where the term τ_{ik} captures the last calendar day that worker i works with past boss k and the matrix $D_{ik,t-\tau}$ indicates past boss assignments in period t-τ. Assume also that past skills depreciate at rate $\gamma^{t-\tau_{ik}}$, so that the estimated γ reported below is the monthly rate of decay, using (t-τ_{ik})/30. In sum, equation (9) contains worker and boss fixed effects and a lagged boss effect that is permitted to depreciate by $\gamma^{t-\tau_{ik}}$ after the worker moves to a new boss. Estimation is via non-linear least squares with fixed effects.

Table 3 contains the results. Teaching accounts for 67 percent of the effect of bosses on workers' productivity. That is, the amount that the past boss effect persists to the present is estimated to be λ=.67. However, the skills learned from past bosses also depreciate; the monthly rate of decay, γ, is estimated to be 0.75. Therefore, after 6 months, about 18% of a boss's teaching remains. Past skills might depreciate if workers learn new products or processes over time, as they do in most TBS companies. But the bottom line is that bosses are mostly providing knowledge that does not depreciate instantaneously.

V. Heterogeneity in Boss Effects

A. Match Effects

Does the treatment effect of boss quality on worker productivity vary with the quality of the worker? There may be heterogeneity in the treatment effects: the effect of a boss on a worker's productivity may depend on the quality of the worker. Good bosses, especially those with teaching skills, may be most useful for those workers who have the toughest time learning or for those who have the most to learn. But it is possible that the reverse is true: our most distinguished academics teach Ph.D. students, not kindergarteners, because the basic skills learned when young are easily taught by less skilled individuals.

It is unclear, a priori, whether a new boss has a comparative advantage with a high human capital or a low human capital worker. From (1), (2) and (4), note that

$$\frac{\partial q}{\partial b} = H\,\frac{\partial E}{\partial B} + E\,\frac{\partial H}{\partial B}.$$

Even if ∂E/ ∂B and ∂H/∂B were greater for the high H than for the low H workers, because high H workers have greater stocks of human capital, the sign is indeterminate. As such, it is important to estimate this to determine how bosses should be sorted so as to make the most of comparative advantage.

To introduce match effects, the estimating equation is an amended version of (7) that includes the match effects, φ_{ij} , and is based on the general specification in (6). The estimating equation is

(10) $q_{it} = X_{it}\beta + \alpha + \delta_j + \varphi_{ij} + \varepsilon_{it}$.

where φ_{ij} is the match quality between person i and boss j as previously defined. The match effects are estimated in Table 4 using mixed effects estimation. The identifying assumptions are now : E(α|X) = E(δ|X)= E(φ|X)=0 and $Cov\left(\begin{matrix}\alpha\\ \delta\\ \varphi\\ \varepsilon\end{matrix}\middle| X\right) = \begin{bmatrix}\sigma_\alpha^2 I_{\#W} & 0 & 0 & 0\\ 0 & \sigma_\delta^2 I_{\#B} & 0 & 0\\ 0 & 0 & \sigma_\varphi^2 I_{\#M} & 0\\ 0 & 0 & 0 & R\end{bmatrix}$ where #M is the number of distinct matches in the data and R is again assumed to be $\sigma_\varepsilon^2 I_N$. Note that the assumption of a diagonal covariance matrix does not mean there is zero systematic correlation between output and high ability workers and bosses, but only that these terms are captured in the realized values of φ.

When match effects are introduced, the standard deviation of the match effect is .68 (Table 4). This is small relative to the standard deviation of the boss effect at 2.97.[15]

It is possible to compute an overall variance in boss effects, taking into account that the

[15] The match effects cannot be estimated for the full sample of 5.7 million observations, so the data is divided into regional subsamples and the results reported in Table 4 are the averages across these subsamples. The results of column (1) are comparable to column (4) Panel B results of Table 2, but with different normalizations of the omitted boss in each subsample. The standard deviation of the boss effects falls from 4.61 (Table 2) to 3.75 (Table 4) when the regressions weight the mean of each boss and worker distribution to be 0 within each random subsample.

individual boss effect now depends on the worker with which the boss is combined. For each boss, there are as many boss effects as there are different workers that boss j has supervised. The standard deviation in boss effects is now calculated across boss-worker pairs, as the standard deviation of $\delta_j + \varphi_{ij}$. The result of this calculation is a standard deviation of boss effects that is 3.05.

B. Assignment

With estimates of φ_{ij} in hand, it is possible to calculate whether good bosses should be matched to good workers or to bad workers. Bosses are classified as "good" or "bad" according to whether their estimated boss effect δ_j is above or below the median.[16] Workers are also classified as "good" or "bad" according to whether their estimated worker effect α_i is above or below the median. The aim is to model heterogeneity in treatment effects by categorizing match effects by the quality of the boss and quality of the worker in the boss-worker pairs.

The designations of good/bad bosses and good/bad workers are formed from the distribution of the effects holding constant the match effect. The designations are unbiased. There are four cells of (good-boss, good-worker), (good-boss, bad worker), (bad-boss, good-worker), and (bad-boss, bad-worker). As long as each of these cells is populated with data, unbiased estimates can be obtained. All that is required is that some good bosses are matched with good workers, some bad bosses are matched with good workers, and some good bosses are matched with bad workers, and some bad bosses are matched with some bad workers. Each of our four cells of boss/worker pairs for the good/bad combinations will measure the mean outcome for the quality groups designated.

The results are contained in Table 4. The issue here is one of comparative advantage: how best to allocate the bosses. The results in Table 4 provide a clear answer. The mean match effect of good bosses with good workers ("stars") is .112, whereas the mean match effect of bad bosses with stars is .032. As a result, moving a star worker from a bad boss to a good one increases the workers output, on average, by .08 units per hour. But pairing bad workers ("laggards") with good bosses yields a match effect of -.063, whereas pairing those workers with bad bosses yields a match effect of -.060. The numbers are virtually identical (actually, bad workers do slightly better with bad bosses than they do with good bosses). The prescription is

[16] The best linear unbiased predictors of $a_i, \delta_j,$ and φ_{ij} are obtained by solving Henderson's mixed model equations. For details, see Abowd, Kramarz, and Woodcock (2006).

that maximizing the value of bosses requires that the better bosses be assigned to the better workers. Workers and bosses should be matched positively because good bosses (defined as good for the average worker) increase the output of stars by more than they do of laggards. Still, the effects are not large. The gain from switching a laggard who was paired with a good boss and a star who was paired with a bad boss increases output by about .083 units per hour on a mean output of 10.26 units.

VI. Non-random Assignment of Workers to Bosses

There may be non-random assignment of experienced workers to bosses. Interview evidence from visits to the company revealed that for the first assignment after being hired, the worker is randomly assigned to bosses, filling in on teams for workers who have departed. Because this is a high turnover job, much of the assignment in the firm is random, reflecting the fact that new workers are randomly assigned to fill open slots. There are two sources of non-random assignment with subsequent worker movements between teams. Older workers may be assigned to older bosses because both groups get their preferred shift choices. Star workers may be assigned to star bosses when stars are given their preferred boss or shift as a reward for their success.

The following section assesses the sensitivity of the estimates of the boss effects given the possibility of non-random assignment. A series of tests suggests that non-random assignment, in this context, is unlikely to be a significant problem for the estimates of boss effects.

A. Time-Varying Individual Ability

The boss effects are unbiased if the model controls for the worker quality that induces the sorting of workers to bosses. The simplest way to control for worker quality is to introduce worker fixed effects, as was done in Table 2. However, consider the case in which worker quality is time varying and thus the worker fixed effects do not control fully for worker quality. If workers are matched to bosses based on each worker's time-varying output, then there will be some correlation between the residual worker quality and the design matrix of boss effects. Suppose that worker i's performance is trending up. At some point, worker i gets assigned to good boss m, moving from inferior boss s. The trend in i's productivity continues upward after moving to boss m, which makes the boss m effect appear larger than it would be had boss m been assigned to worker i when i was new. Alternatively, if workers' productivity shocks are transitory, but assignment is based on these idiosyncratic shocks, the boss effects may be understated. After the worker is assigned to the good boss, the positive transitory component vanishes, making it appear that the good boss reduced the worker's productivity.

To incorporate time-varying ability, lagged worker productivity, output-per-hour$_{i,t-1}$, is added to productivity regression (7). Two different specifications are reported, using 1 and 2 days of lagged productivity.[17] The results in Table 5 demonstrate that the standard deviation of the boss fixed effect declines slightly when lagged worker productivity is added. The relevant comparison is the net present value (NPV) of the standard deviation of the boss effect, because the boss effect feeds into the worker's lagged output.[18] The NPV of the standard deviation of the boss effects is either 3.39 or 3.17, depending on the assumed lag structure. The variance in boss effects is largely unchanged from the 3.44 value of the baseline model of Table 2 (row 4).

B. Heterogeneity in the Treatment Effect

If there is no heterogeneity in the response of workers to bosses, then any non-random assignment of workers to bosses is irrelevant for the estimation of boss effects. The estimation of the match effects, introducing φ_{ij}, is the method of permitting workers to differ in their responses to bosses. While there is heterogeneity in workers' responses to bosses, it is modest relative to the primary, uniform effect that bosses have on all workers.

The mixed effects estimator also provides a specification test to assess whether bosses and workers are sorted based on their idiosyncratic match effects. To understand the logic behind the test, consider an alternative method to estimate the match effects. Jackson (2012) calls this alternative method the orthogonal match fixed effects estimator, in which the match effects are calculated as the mean of the residual for each boss-worker spell after fixed effect estimation. The orthogonal match fixed effects estimator imposes that the mean of the match effects for each worker and boss is zero by construction. In contrast, the mixed effects estimator allows the observed match effects to deviate from zero for each boss and worker. The mixed effects estimator instead imposes that the *potential* match effects are zero (Jackson, 2012). This means that if a boss and worker were paired at random, the expected match effect would be zero,

[17] Consistency of these estimates relies on large T asymptotic theory.
[18] The NPV of the boss effect is the standard deviation of the boss effect divided by (1- $\eta_{i,t-1}$), where $\eta_{i,t-1}$ is the coefficient on the lagged dependent variable.

but there is nothing that restricts the match effects to be mean zero for the actual subset of matches that do occur.

The implication is that the mean of the match effects for each worker and boss will be zero in the mixed effects estimation if the assignment of workers to bosses is nearly random. If the assignment of workers to bosses is not random, then the estimated match effects from Table 4 may deviate from zero because the workers are being assigned to bosses to reflect match specific gains. When using individual workers or bosses as the unit of analysis and weighting by the number of assignments that each worker or boss has in the data, the mean of the workers' average match effect across workers is -0.0004 with a standard deviation of 0.124 and the mean of the bosses' average match effect across bosses is -0.0004 with a standard deviation of 0.044.[19][20] These results are consistent with the identifying assumptions. The zero means of the observed match effects within workers and bosses suggests that there is little sorting of workers to bosses on the basis of the expected match-specific component of productivity.

C. Using Randomly Assigned Workers as a Robustness Check

The next test examines whether the estimated boss effects predict well out of sample. The experiment is as follows. Because new workers were nearly randomly allocated to their first boss due to the stochastic nature of quits and vacancies, we estimate boss quality using data from the older workers and see if those boss quality measures can predict the productivity of the new workers when there is known random assignment. If the estimated boss effects for older workers can predict the productivity of new workers, then there is likely to be little non-random assignment of workers to bosses after controlling for the worker fixed effects.

In step one, boss fixed effects were estimated on a subset of matched worker-boss data for older workers who have experienced at least two previous boss switches. The estimated boss fixed effects were then saved. In step two, the quartiles of the initial boss distribution were

[19] Due to the limited number of observed assignments, some workers or bosses with a sequence of lucky pairings are likely to have match effects that deviate from zero. However, under the null hypothesis that assignment is independent of the latent match effect between bosses and workers, as the number of boss assignments increases for worker i, the mean match effect for worker i should converge to 0. The same logic applies to the mean match effect for boss j. Because individual bosses tend to have more observed matches in the data, the distribution of the mean match effect for bosses should have a smaller variance than the distribution of the mean for workers.

[20] This exercise gives unequal weight to workers and bosses depending on the frequency of their observed assignments. Using the worker as the unit of analysis and calculating each worker's mean match effect from the specification in Table 4 yields an equally-weighted, overall mean across workers of 0.0027 with a standard deviation of 0.187. The same calculation for bosses yields an equally-weighted mean match effect across bosses of -0.011 with a standard deviation of 0.128.

calculated from the estimated boss effects in the experienced worker sample.[21] If the boss fixed effects predict well out of the sample (i.e., beyond the older group on which they were estimated), then new workers who are randomly assigned to a good boss should have higher output than a new worker who is randomly assigned to a bad boss.

To test this, a regression is estimated. Among new hires, the dependent variable is the mean output during the 30 day period prior to the first boss switch. The results in Table 6, Column 1, suggest that the estimated boss fixed effects predict well out of sample. The coefficient on the highest quartile is 0.22 and is highly statistically significant. The excluded category is the bottom quartile. For comparison purposes, the raw interquartile range of the estimated boss fixed effects for this sample is 0.49. The coefficient on the second highest quartile is 0.07. Thus, boss effects that are estimated using the older sample of workers do a reasonable job of predicting the productivity of the new hires who are assigned randomly. On average, a good boss is always a good boss.

Finally, it is possible to estimate the distribution of boss effects using only the sample of new workers. Using only new workers, the estimated standard deviation of boss mixed effects is 4.43 (excluding worker effects because there is only 1 observation per worker). This is very close to the estimate of 4.61 in Table 2.

D. Testing for Non-random Boss Transitions

Column 2 of Table 6 tests for non-random sorting given the ex-ante classification of bosses. Consistent estimation of the individual boss effects requires orthogonality between the design matrix representing current boss assignments and the concurrent and lagged residuals. While a test cannot be carried out using concurrent residuals, it is possible to test whether residuals from the initial boss assignment predict the quality of future bosses. The test is implemented by regressing the residuals from Table 6, Column 1, on indicators for the quartiles of the distribution of the subsequent boss. Under the null hypothesis of random assignment, the quartiles of future boss fixed effects should be unrelated to the lagged residual. An $F(4, 5946)$ test that all parameters are jointly zero does not reject the null (p-value = 0.17).[22] While this test

---

[21] Quartiles are used for this calculation as a partial solution to problems arising due to measurement error in the individual boss fixed effects. Even with this measurement error, there is a statistically significant and positive coefficient when regressing individual productivity for new workers on the boss effects recovered for the old workers (results not shown).

[22] The quartiles of the future boss distribution are treated as fixed for the purposes of this test, which biases the test

cannot speak to non-random allocation of bosses and workers that occurs later in a worker's career, when coupled with the external validation of the estimated boss effects on a separate sample of workers, the results suggest that non-random sorting is unlikely to be a major problem for estimation.

in favor of rejecting the null hypothesis of random boss transitions.

VII. Boss Attrition

It seems reasonable to suppose that the boss selection process is such that the observed bosses are the best candidates among the pool of potential bosses. However, the firm's forecast of future boss productivity is likely subject to error. As the firm learns about boss productivity, the worst bosses are likely to be replaced.

To test this prediction, boss attrition is analyzed. The approach is to select one calendar day in the data in each year. For each boss present on that calendar date, the dependent variable is an indicator that the boss is still present in the data at least one year in the future. The date chosen is January 10 for years 2007, 2008, and 2009. This dependent variable is then regressed on year fixed effects and an indicator that the boss's estimated fixed effect is in the bottom 10% of the distribution. The focus is on the lower quality bosses because the best bosses might leave on their own to move to better opportunities elsewhere.

The results are presented in Table 7. Bosses in the bottom 10% are much less likely to be in the data in the next year. The coefficient on being in the bottom 10% is -.236, compared with a constant of .645. The probability that a boss in the 10^{th} percentile of boss quality separates from the firm over a 1 year period is .59 compared to a baseline estimate of .36.

To ensure that this result is not due to noise (the concern being that the estimated boss fixed effects for short-lived bosses are most likely to be in either tail of the distribution), the second column of Table 6 includes an indicator for bosses in the top 10% where noise should also be an issue. While the coefficient is negative, it is small and not statistically different from zero.

These results suggest that the worst bosses do not survive. The exit rate of bad bosses is almost twice the exit rate of the average quality boss.

VIII. Peer Effects

There is a growing literature on peer effects.[23] If the best bosses are also likely to be matched with the best team members, peer effects may confound the estimates. To test for this, the basic specification with boss and worker fixed effects is run while adding a peer effect:

---

[23] Most current peer effects papers test whether workers learn from each other due to proximity, or adjust their effort in response to those who work around them (Falk and Ichnio, 2006) or who watch them (Mas and Moretti, 2009). Few papers test for the complementarity of skills within the teams that are formed among peers, because skills are

(10) $q_{ijt} = X_{it}\beta + \alpha_i + \delta_j + \xi\, p_{ijt} + \varepsilon_{ijt}$

where the peer effect, p_{ijt}, is specified in two ways.

One way to estimate peer effects is to use peers' fixed effects as measures of the peer output, estimated using a two-step non-linear least squares routine. The estimating equation for the joint model is

(10') $q_{ijt} = X_{it}\beta + \alpha_i + \delta_j + \xi_{\text{Peer}}(TeamSize - 1)^{-1}\sum_{k\in jt\backslash\{i\}}\alpha_k + \varepsilon_{ijt}$

where summation over $k \in jt\backslash\{i\}$ captures the fixed effects of worker i's team on day t with boss j while excluding worker i. This specification allows the estimated peer effect to depend only on the permanent effect of co-workers on the team, α_k, not on concurrent q_{ijt}. Estimation of the joint model is not feasible on the full set of data because of memory constraints. Storage of the matrix of peer-indicators, even in sparse form, requires an order of magnitude more memory than storage of the data with only worker and boss indicators. Because workers and bosses rarely move establishments, the joint procedure can be applied using subsets of establishments. The estimation algorithm is a two-step procedure. The outer-loop guesses a value of ξ_{Peer} and then computes the remaining parameters via a linear conjugate gradient procedure in an inner-loop conditioning on the value of ξ_{Peer}. Search is then over ξ_{Peer}.

The main result is that peer effects are not economically significant relative to boss and worker effects. The regressions in column 1 of Table 8 use a subset of the data corresponding to a typical region, because joint estimation of worker effects and unconstrained peer effects is only feasible on subsets of the data. The estimated peer effects are close to zero.

Another method to estimate peer effects uses a peer's first few months of output as a proxy for the peer's current output. These results are provided in column 2. Again, the coefficient is close to zero.

The conclusion is that peer effects are very small relative to boss effects.[24] Note that this production environment has relatively little teamwork because each worker primarily interacts

unobserved and most data has come from production functions (like store clerks) that are largely individual output, not team output. That is true of these data as well.

[24] There is also possible sorting of workers into teams of correlated peers, because good workers will work together if given the choice of their preferred shift and there are similar preferred shifts for all workers. If this sorting is temporal, based on recent performance (as it is), introducing worker fixed effects for peer effects will reduce the bias. If the sorting is based on permanent performance, there will be an upward bias in the estimated peer effects. Given that the peer effects are zero or negative, this is not a concern.

with a customer, not with other workers.[25] Although the workers can see each other and may learn from each other or compete with each other, the workers do not appear to be complements in production.

IX. Two Measures of Output

Table 9 provides results that treat both output-per-hour and uptime as dependent variables, again using the structure of equation (7). First note that the range of boss effects is larger, both in absolute and percentage terms, on output-per-hour than on uptime. The weighted standard deviation of boss fixed effects is 3.31 for output-per-hour and 0.58 for uptime. This is in part a function of the fact that output per hour varies much more than uptime. The unconditional standard deviation of output per hour is 30.7% of the mean whereas the standard deviation of uptime is 2.8% of the mean. There is a limitation to how important bosses can be in monitoring uptime. For service jobs of this type, where monitoring is easily performed by the information technology, the incremental effect of bosses on workers through monitoring uptime is low.

Good bosses appear to be good along both dimensions. The correlation between the boss effect on output-per-hour and the boss effect on uptime can be computed. The simple correlation of the two fixed effects (bosses on uptime and bosses on output-per-hour) is .16, which is significant at standard levels. Bosses who are better at increasing output-per-hour are better at increasing uptime in their workers as well.[26]

X. Conclusion

Supervision and management are a fundamental in personnel economics and in the theory of the firm. Although we take as given that managers matter, neither the mechanisms through which they affect productivity nor the actual size of the effects have been documented previously. By using a data set that reports daily output on workers and records the supervisors to which they are assigned on each day, it is possible to examine the effects of bosses on worker productivity.

Boss effects are large and significant. The value of the average boss is about 1.75 times

[25] The same is true in Mas and Moretti (2009), who also find significant, but small peer effects.
[26] Worker fixed effects are already held constant so this is not a result of good workers sorting to good bosses.

that of a worker. Furthermore, bosses vary substantially in their quality. A very good boss increases the output of the supervised team over that supervised by a very bad boss by about as much as adding one member to the team. Additionally, in this production context, peer effects are trivial. The only "peer" who matters in this work environment is the boss. The primary means by which bosses matter is through teaching; motivating is less important. Finally, good bosses increase the output of the better workers by slightly more than that of poorer workers, which implies that productivity can be enhanced through strategic assignment of workers to bosses.

References

1. Abowd, John, Robert Creecy, and Francis Kramarz. 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Working Paper.
2. Abowd, John, Francis Kramarz, and Simon Woodcock. 2006. "Econometric Analysis of Linked Employer-Employee Data." Working Paper.
3. Bennedsen, M., Nielsen, K.M., Perez-Gonzalez, F., Wolfenzon, D. 2007a. "Inside the family firm: the role of families in succession decisions and performance." *Quarterly Journal of Economics 122*, 647-691.
4. Bennedsen, M., Perez-Gonzalez, F., Wolfenzon, D. 2007b. "Do CEOs matter?" Working paper. Copenhagen Business School.
5. Bertrand, M., Schoar, A. 2003. "Managing with style: the effect of managers on firm policies." *Quarterly Journal of Economics 228*, 2279-2308.
6. Borghans, L., ter Weel, B., Weinberg, B.A. 2008. "Interpersonal styles and labor market outcomes." *Journal of Human Resources 43*, 815-858.
7. Branch, Gregory, Eric Hanushek, and Steven Rivkin. 2012. "Estimating the Effect of Leaders on Public Sector Productivity," working paper, Stanford University.
8. Dawson, P., Dobson, S., Gerrard, B. 2000. "Estimating coaching efficiency in professional team sports: evidence from English Association football." *Scottish Journal of Political Economy 47*, 399-421.
9. Engelberg, Joseph, Raymond Fisman, Jay Hartzell, and Christopher Parsons. 2012. "Human Capital and the Supply of Religion," Working Paper, Columbia University, March.
10. Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics*, 24(1), 39-57.
11. Frick, B., Simmons, R. 2008. "The impact of managerial quality on organizational performance: evidence from German soccer." *Managerial and Decision Economics 29*, 593-600.
12. Garicano, Luis and Thomas Hubbard. 2007. "Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyers' Human Capital." *Journal of Law and Economics*, 50(1), 1-43.
13. Goodall, Amanda H. 2009a. "Highly cited leaders and the performance of research universities." *Research Policy 38*, 1079-1092.
14. Goodall, Amanda H. 2009b. "Socrates in the boardroom: why research universities should be led by top scholars." Princeton University Press, Princeton and Oxford.
15. Hermalin, Ben E. 1998. "Toward an economic theory of leadership: leading by example." *American Economic Review 88*, 1188-1206.

16. Ichniowski, Casey, and Kathryn Shaw. 2003. "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices." *The Journal of Economic Perspectives*, 17(1), 155-180.
17. Jackson, C. Kirabo. 2012. "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence From Teachers." NBER Working Paper No. 15990.
18. Jenter, Dirk and K. Lewellen. 2010. "Performance Induced CEO Turnover," working paper, Stanford University.
19. Jones, B.F., Olken, B.A. 2005. "Do Leaders matter" National leadership and growth since World War II." *Quarterly Journal of Economics 120*, 835-864.
20. Kahn, L.M. 1993. "Managerial quality, team success and individual player performance in Major League Baseball." *Industrial & Labor Relations Review 46*, 531-547.
21. Kandel, Eugene, and Edward Lazear. 1992. "Peer Pressure and Partnerships." *Journal of Political Economy*, 100(4), 801-817.
22. Kaplan, S.N., Klebanov, M.M., Sorensen, M. 2008. "Which CEO characteristics and abilities matter?" In: *National Bureau of Economic Research Working Paper No. 14195*, National Bureau of Economic Research, Cambridge, MA.
23. Kuhn, P., Weinberger, C. 2005. "Leadership skills and wages." *Journal of Labor Economics 23*, 395-436.
24. Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5), 1346-1361.
25. Lazear, Edward P. 2004. "The Peter Principle: A Theory of Decline," *Journal of Political Economy* 112 (1), part 2: S141-S163.
26. Lazear, Edward P. 2012. "Leadership: A Personnel Economics Approach." *Labour Economics*, 19(1): 92-101.
27. Lazear, Edward P. and Kathryn L. Shaw. "Tenure and Output," *Labour Economics*, 15 (2008): 710-724.
28. Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1), 112-145.
29. Matsa, David and Amalia Miller. 2012. "A Female Style in Corporate Leadership? Evidence from Quotas." Working paper, Northwestern University.
30. Perez-Gonzalez, F. 2006. "Inherited control and firm performance." *American Economic Review 96*, 1559-1588.
31. Perez-Gonzalez and Daniel Wolfenzon, 2012. "Death of CEOs Impact on Firm Performance." Working paper, Stanford University.
32. Porter, P.K., Scully, G.W. 1982. "Measuring managerial efficiency: the case of baseball." *Southern Economic Journal 48*, 642-650.
33. Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics*, 13(2), 311-323.
34. Rotemberg, J.J., Saloner, G. 2000. "Visionaries, managers, and strategic direction." *RAND Journal of Economics 31*, 693-716.

35. Schoar, Antoinette, and Luo Zuo. 2008. "Shaped by Booms and Busts: How the Economy Impacts CEO Careers and Management Styles." Working paper, MIT.
36. Shaw, Kathryn, and Edward Lazear. 2008. "Tenure and Output." *Labour Economics,* 15, 710-724.
37. Simon, Herbert. 1957. "The Compensation of Executives." *Sociometry*, 20(1), 32-35.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Output Per Hour	5,729,508	10.26	3.16	0.1	40.0
Uptime	4,870,610	0.96	0.03	0.5	1.0
Output Per Hour* Uptime	4,870,610	10.01	3.00	0.4	40.0
Tenure	5,729,508	648.91	609.83	1.0	4,235.0
Number of Workers	23,878				
Number of Unique Bosses Per Worker	23,878	3.99	2.78	1.0	19.0
Daily Team Size	633,818	9.04	4.54	1.0	29.0
Number of Bosses	1,940				
Number of Unique Workers Per Boss	1,940	49.15	35.41	1.0	250.0
Mean Number of Other Bosses for Each Worker	1,940	4.69	1.51	0.0	11.3

Notes:

The data contain daily worker productivity records from June 2006 to May 2010. Output per hour is the daily average of the number of transactions per hour. Uptime is the daily percent of time that the worker is available to handle transactions. These measures are recorded by computer software. There is some missing data on uptime. The missing uptime data is concentrated toward the beginning of the sample period. The mean of output per hour when restricting the sample to the 4,870,610 worker-days with non-missing uptime is 10.38 with standard deviation 3.08.

Table 2: Regressions of Output-per-hour on combinations of fixed effects

	OLS	Worker Effects	Boss Effects	Worker and Boss Effects
PANEL A: FIXED EFFECTS				
R-squared	0.061	0.237	0.092	0.243
Standard Deviation of Worker Fixed Effects				
(1) Weighted by worker-days (frequency)		1.34		1.32
(2) Adjusted for Sampling Error (1 observation per worker)		1.24		1.22
(3) Unweighted (1 observation per worker)		1.87		1.85
F statistic		55.6***		47.7***
Standard Deviation of Boss Fixed Effects **(Multiplied by Average Team Size of 9.04)**				
(4) Weighted by worker-days (frequency)			5.24	3.44
(5) Adjusted for Sampling Error (1 observation per boss)			5.6	3.53
(6) Unweighted (1 observation per boss)			8.86	7.5
F on Joint Fixed Effects				53.3***
PANEL B: MIXED EFFECTS	OLS	Worker Effects	Boss Effects	Worker and Boss Effects
Standard Deviation of Worker Mixed Effects		1.52		1.45
Standard Deviation of Boss Fixed Effects **(Multiplied by Average Team Size of 9.04)**			6.69	4.61
SAMPLE SIZES				
Number of Observations	5,729,508	5,729,508	5,729,508	5,729,508
Number of Workers	23,878	23,878	23,878	23,878
Number of Bosses	1,940	1,940	1,940	1,940
Percent of sample in largest connected group				99.99

Notes:

All specifications contain a fifth order polynomial function of tenure (with a 365 day cutoff and cutoff indicator), monthly time dummies, and day of week dummies. Within each connected group of workers and bosses, worker effects are mean zero and one boss effect is restricted to be zero. In Panel A, the estimated fixed effects are weighted by the number of worker-days (rows 1 and 4), are adjusted for sampling error (rows 2 and 5), or are unweighted (rows 3 and 6). The adjustment for sampling error assumes the true boss and worker fixed effects are normally distributed. The estimated boss and worker fixed effects are then normally distributed with variance equal to the sum of sampling error and the true variance. Maximum likelihood estimates of the true variance are reported. The sampling variance is computed as the residual variance for each worker or boss, which slightly understates the sampling variance because there is no accounting for off-diagonal elements in the normal equations matrix of regressors. The fixed effects R-squared values are reported; the R-squared values from mixed effects are similar in each specification.

Table 3: Teaching and Motivation

Teaching (λ)	0.67
Monthly Rate of Decay (λ)	0.75
Amount of Boss Effect Remaining After Six Months (λ^6)*λ	0.18
Standard Deviation of Worker Fixed Effects	
Weighted by worker-days	1.30
Standard Deviation of Boss Fixed Effects	
(Multiplied by Average Team Size of 9.04)	
Weighted by worker-days	3.35
Number of Observations	5,729,508
Number of Workers	23,878
Number of Bosses	1,940

Notes:

The specification contains a fifth order polynomial function of tenure (with a 365 day cutoff and cutoff indicator), monthly time dummies, day of week dummies, boss fixed effects, and worker fixed effects. Estimation is conducted via nonlinear least squares, where search over the reported parameters involves an “outer” loop, while an inner loop conditions on the outer loop values to solve for the other parameters. Estimation on the full data set is infeasible because storing the matrix of past boss histories is not possible for the full sample. Instead, the reported results are from a set of regressions in which the data is divided into regional subsamples and then aggregated by taking weighted averages across these subsamples.

Table 4: Heterogeneous Boss Effects

	Output-per-hour No Match Effects	Output-per-hour With Match Effects
R-squared	0.233	0.26
Standard Deviation of Worker Mixed Effects	1.50	1.35
Standard Deviation of Boss Mixed Effects **(Multiplied by Average Team Size of 9.04)**	3.75	2.97
Standard Deviation of Match Effects		0.68
Means of Match Effects by Groups		
Good Bosses and Good Workers		0.112
Good Bosses and Bad Workers		-0.063
Bad Bosses and Good Workers		0.032
Bad Bosses and Bad Workers		-0.060
Number of Observations	5,729,508	5,729,508
Number of Workers	23,878	23,878
Number of Bosses	1,940	1,940

Notes:

All specifications contain a fifth order polynomial function of tenure (with a 365 day cutoff and cutoff indicator), monthly time dummies, and day of week dummies. Estimation on the full data set is infeasible because storing the matrix of match effects is not possible for the full sample. Instead, the reported results are from a set of regressions in which the data is divided into regional subsamples and then aggregated by taking weighted averages across these subsamples. The first column of results purges across-region heterogeneity from the mixed effects estimates, and is comparable to the mixed effects estimates in Table 2. In the second column, bosses are classified as “good” if the estimated “best linear unbiased predictor” (BLUP) of their mixed effect is above the median; bosses are classified as “bad” if their mixed effect is below the median. A similar definition applies to workers. The means of the match effects are then calculated for each boss-worker cell formed from the mixed estimates of boss and worker effects.

Table 5: Regressions of Output-per-hour with Lagged Dependent Variables

Number of Lags	(1)	(2)
R-squared	0.2502	0.2537
Coefficient on the first lag	0.104	0.096
Coefficient on the second lag		0.074
Standard Deviation of Worker Fixed Efforts		
Weighted by worker-days	1.18	1.09
Standard Deviation of Boss Effects (Multiplied by Average Team Size of 9.04)		
Weighted by worker-days	3.04	2.87
NPV of a Standard Deviation of Boss Effects for an Average Team	3.39	3.17
Durbin-Watson Statistic	2.01	2.01
Number of Observations	5,705,630	5,682,019

Notes:

The specifications correspond to Table 2, Column 4, but add either 1 lag (column 1) or 2 lags (column 2) of output per hour as right-hand-side variables. These models rely on large T asymptotics. All specifications contain a fifth order polynomial function of tenure (with a 365 day cutoff and cutoff indicator), monthly time dummies, and day of week dummies. Worker fixed effects are mean zero, and one boss fixed effect is restricted to be zero. To calculate the standard deviations of boss and worker effects, the estimated fixed effects are weighted by the sample frequency of worker-days. The NPV of the boss effect is the standard deviation of the boss effect divided by $(1 - \eta_{i,t-1})$ where $\eta_{i,t-1}$ is the coefficient on the lagged dependent variable.

Table 6: Out of Sample Validation

Dependent Variable	Mean OPH 30 days prior To the first boss switch (1)	Residuals from Column 1 (2)
2nd Quartile	0.07	0.11
	(0.06)	(0.06)
3rd Quartile	0.18	-0.00
	(0.06)	(0.06)
4th Quartile	0.22	-0.01
	(0.06)	(0.06)
Constant		-0.03
		(0.04)
F Statistic that all parameters are zero		1.59
P-Value		0.18
Type of Boss Quartiles	Concurrent	Future
R-squared	0.19	0
Number of Observations	5,950	5,950

Notes:

The sample contains workers on their first assignment prior to the first boss switch. To be included, workers must have had at least 15 days of productivity data before and after the boss switch. Boss quartiles are calculated using a partitioned set of workers as follows. First, using a sample including only workers after their second boss switch, boss fixed effects for this sample are computed by regressing oph on worker fixed effects, boss fixed effects, a fifth order polynomial function of tenure (with a 365 day cutoff and cutoff indicator), monthly time dummies, and day of week dummies. Second, the estimated boss fixed effects are merged onto the sample of workers on their first boss. We then compute quartiles of the boss distribution for the initial and subsequent boss. Bosses who do not work with older workers (after 2 switches) are not included in the sample. The regression in Column 1 has controls for tenure and monthly time dummies. Day of the week dummies are not included in the second stage regressions because average output is taken over several days.

Table 7: The Probability that Bosses are Retained for 1 Year

Bottom 10% of Boss Effects	-0.236	-0.239
	(.082)***	(.082)***
Top 10% of Boss Effects		-0.044
		(.050)
Constant	0.645	0.645
	(.021)***	(.022)***
R-squared	0.018	0.018
Number of Observations	1,444	1,444

Notes:

The data are repeated cross sections of bosses, present in the data on January 10 of each year, and matched with their estimated boss fixed effect from Table 2. The dependent variable is an indicator that the boss is present in the data 1 year in the future. Year fixed effects are included. The distribution of boss effects uses each unique boss as the unit of analysis. Standard errors are in parentheses and triple asterisks indicate significance at the 1% level.

Table 8: The Effect of Peer Quality on Output-per-hour

Estimation method:	Joint	Peer Proxies
R-squared	0.2356	0.243
Coefficient on Peers' Mean Ability	0.001	-0.022
Standard Deviation of Peer Effects	0.022	0.0009
Standard Deviation of Boss Effects (Weighted by worker-days)	0.31	0.38
Standard Deviation of Worker Effects (Weighted by worker-days)		1.32
Number of Workers	1,679	23,878
Number of Bosses	155	1,940
Number of Observations	391,730	5,729,508

Notes:

All specifications contain a fifth order polynomial function of tenure (with a 365 day cutoff and cutoff indicator), monthly time dummies, day of week dummies, and boss and worker fixed effects. In column 1, the joint estimation procedure uses non-linear least squares, taking the mean of the team members' individual fixed effects as a measure of peer quality. The joint estimation procedure is computationally demanding; an "outer" loop is used to search over the peer effect coefficient, while an inner loop conditions on the outer loop value and solves for the parameters using a conjugant gradient procedure. The joint procedure is not possible on the full data because of memory issues in Matlab; storage of the matrix of peer fixed effects requires an order of magnitude more memory than using a single-dimensional index of peer quality. In column 2, the peer proxies use mean output on the first three months on the job as the value of peer quality. If a worker's first three months are not observed, then the mean value of all observed workers' first three months is used. To calculate the standard deviation of peer effects, it is assumed that one peer's output increases by a standard deviation change in output per hour, or 3.16 units. This is then multiplied by the Coefficient on Peer's Mean Ability and divided by (9.04-1), the mean number of other team members.

Table 9: Comparison of Output-per-hour and Uptime

	Output-per hour	Uptime
Mean of Dependent Variable	10.26	0.96
R-squared	0.243	0.0994
Standard Deviation of Worker Fixed Effects		
(1) Weighted by worker-days (frequency)	1.32	0.1
(2) Adjusted for Sampling Error (1 observation per worker)	1.22	0.12
(3) Unweighted (1 observation per worker)	1.85	0.21
F statistic	47.7***	18.6***
Standard Deviation of Boss Fixed Effects (Multiplied by Average Team Size (9.04))		
(4) Weighted by worker-days (frequency)	3.44	0.58
(5) Adjusted for Sampling Error (1 observation per boss)	3.53	1.61
(6) Unweighted (1 observation per boss)	7.50	2.58
F statistic	20.5***	16.9***
F on Joint Fixed Effects	53.3***	20.7****
Number of Observations	5,729,508	4,870,610
Number of Bosses	1,940	1,726
Percent of sample in largest connected group	99.99	99.99
Correlation of Worker Oph and Uptime Fixed Effects	0.15	
Correlation of Boss Oph and Uptime Fixed Effects	0.16	

Notes:

For notes on estimation, see Table 2. Some data on uptime is missing toward the beginning of the sample period, accounting for differences in the number of observations.