

The Value of Bosses

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

Stanford University,
Stanford University,
University of Utah

February 20, 2013

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, Mitch Hoffman, 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. A separate normalization 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 and if so, how large and how variable is it? Bosses generally earn more than the workers whom they supervise. Is the productivity that they generate worth the additional pay? Also, it is also clear from other studies of productivity that workers vary in their output even within the same job category and pay grade. Does boss productivity also vary and if so, how significant is the variation both in absolute terms and relative to the workers whom they supervise? Finally, what is the nature of the effect? Do bosses enhance productivity because they pass on valuable skills that are learned and retained or do bosses primarily serve to motivate workers contemporaneously and only fleetingly? 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

¹The 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;

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,

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).

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.

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.

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. Bosses in the lowest 10% of the quality distribution are over twice as likely to leave the firm as 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 i 's output at time t , q_{it} , depends on human capital, H_{it} , which reflects both innate ability and previously learned skills, and on effort, E_{it} . A natural (although not necessary) specification is multiplicative: harder work results in greater returns to human capital

$$(1) \quad q_{it} = H_{it} * E_{it}.$$

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_{it} , and some innate ability, denoted α_i . Then

$$(2) \quad H_{it} = H(X_{it}, \alpha_i, b_{it})$$

where b_{it} 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

$$(3) \quad b_{it} = b(d_{jt}/N_{jt}^\theta, d_{m,t-1} / N_{m,t-1}^\theta, \dots, d_{p0} / N_{p0}^\theta)$$

where d_{jt} is an index of the difference between the quality of boss j with whom worker i is paired at time t and the mean boss quality, N_{jt} is the size of that team, $d_{m,t-1}$ is the quality of boss m with whom the worker is paired at time $t-1$, $N_{m,t-1}$ is the size of that team, and so forth, and θ is a parameter that relates to the public or private nature of boss time. Note that the identity of boss m may be the same or may differ from that of boss j . Furthermore, this specification allows past bosses to affect the worker's output at time t because some of the knowledge and work habits acquired from those bosses may be retained.

If boss time is like individual tutoring, then $\theta=1$. Boss time is purely private so that time spent with one worker cannot be spent with another and has no spillover value to other workers. If boss time is like a lecture, then $\theta=0$. The boss's instruction or motivation improves all workers and there is no congestion. For $0 < \theta < 1$, there is some public good aspect to boss time and some private good aspect. A private good is one with total congestion.⁴

Analogously, effort is

$$(4) \quad E_{it} = Z(X_{it}, \alpha_i, b_{it})$$

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

$$q_{it} = H(X_{it}, \alpha_i, b(d_{jt}/N_{jt}^\theta, d_{m,t-1} / N_{m,t-1}^\theta, \dots, d_{p0} / N_{p0}^\theta)) \times \\ Z(X_{it}, \alpha_i, b(d_{jt}/N_{jt}^\theta, d_{m,t-1} / N_{m,t-1}^\theta, \dots, d_{p0} / N_{p0}^\theta))$$

or

$$(5) \quad q_{it} = f(X_{it}, \alpha_i, b(d_{jt}/N_{jt}^\theta, d_{m,t-1} / N_{m,t-1}^\theta, \dots, d_{p0} / N_{p0}^\theta)) \cdot$$

⁴ Lazear (2001) proposes a teaching model that has a public good structure, with congestion of a particular form. It relates more closely to classroom teaching, however, because the actions of one student have a direct and particular effect on another. The form used in (3) is less well-structured than that in Lazear (2001), but allows for a more general characterization of workplace interaction, where boss-worker instruction could be purely private. In Lazear (2001), there is no private instruction; all students are learning or no students are learning. Formally, there is no allowance for one student learning while the other is not unless classroom size was reduced to one student.

The linear form is a specific version of (5) that will be used in the empirical analysis. Then (5) becomes

$$(6) \quad q_{it} = \alpha_0 + \alpha_i + X_{it}\beta + (d_{0t}/N_{jt}^\theta + d_{0t-1}/N_{m\ t-1}^\theta + \dots + d_{0p}/N_{p\ 0}^\theta) + d_{jt}/N_{jt}^\theta + d_{m\ t-1}/N_{m\ t-1}^\theta + \dots + d_{p0}/N_{p0}^\theta$$

where α_0 is the ability level of the mean worker and d_{0t} is the ability level of the mean boss. Thus, the expectation of both the other α_i and of the other d_{ijt} is definitionally zero.

A contemporaneous-effects only version of (6) is

$$(7) \quad q_{it} = \alpha_0 + \alpha_i + X_{it}\beta + d_{0t}/N_{jt}^\theta + d_{jt}/N_{jt}^\theta$$

The contemporaneous boss effect on any single worker is then $(d_{jt} + d_{0t})/N_{jt}^\theta$ and the effect of boss j on all workers that she supervises is

$N_{jt} ((d_{jt} + d_{0t})/N_{jt}^\theta)$ or

$$(8) \quad \text{Boss effect on productivity} = (d_{jt} + d_{0t})N_{jt}^{1-\theta}$$

The boss effects can vary over time for three reasons. First, the worker's boss today may differ from the one that he had in the past. Second, the influence of a boss may diminish (or even possibly increase) as time passes. Third, the second day with a boss does not necessarily have the same value as the first day. It may be that most of what is to be learned gets learned quickly, in which case the marginal effect of boss time on worker productivity diminishes with time spent with that boss. An alternative is that it takes time to learn to communicate with a boss, which would mean that the second day with her is more valuable than the first. Note that this time effect is different from that of boss effects diminishing with time that has passed since the boss encounter. Because (6) allows the identity of the boss at time $t-q$ to differ from that at time $t-q-1$, the structure allows for diminishing or increasing returns to spending time with given boss as well as allowing bosses who were encountered longer ago to have different effects from those encountered more recently.

Bosses, in the context in which we study, are most important in their ability to teach and

to 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 and that the teaching role of bosses works primarily through skill level, 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. The Allocation of Bosses to Workers

Allocating bosses to workers may have significant effects on productivity. There are two aspects of allocation that may be important. The first involves team size. The second is pairing a given worker with the right boss for him or her.

Consider team size first and the simplest problem of allocating N identical workers among two bosses, j and k . Boss j has a team of N_j workers, which leaves boss k with $N-N_j$ workers. As before, since the workers are identical, the i subscript is dropped. Then the goal is to choose N_j so as to maximize the output of the two teams taken together. Thus, choose N_j to

$$(9) \text{ Max } N_j q_j + (N-N_j) q_k$$

where q_j and q_k are given by (6).

Optimization requires that the first order condition hold or that

$$(1+\theta) (d_0+d_j)/N_j^\theta - (1+\theta)(d_0+d_k)/(N-N_j)^\theta = 0$$

(10) or

$$(d_0+d_j)/N_j^\theta = (d_0+d_k)/(N-N_j)^\theta$$

Eq. (10) implies that as long as $\theta > 0$, team size increases in d_j .⁵ The larger team is allocated to the boss with the greater effect on productivity. This makes sense. If there were no constraints on boss time, all workers would be allocated to the best boss. But spreading a boss too thin hurts worker productivity so the lower quality boss gets some workers as well as long as $\theta > 0$. Were $\theta = 0$ so that boss time was completely public, it would make sense to choose the corner solution of assigning all worker to the highest quality boss.

A second question is whether good bosses should 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. Below, the assumption of no interaction effects between boss quality and worker quality is tested and found to be very close to true.

Additional empirical questions associated with worker allocation are:

E4: Are team sizes adjusted in a way consistent with optimality that gives the higher quality bosses larger teams? If so, how close is that relation to the structure implied by equation (10)?

E5: 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)?

II. Data

The data are from an extremely large service company that has daily records on worker

⁵Using the implicit function theorem on (10),

$$\begin{aligned} \frac{\partial N_j}{\partial d_j} \Big|_{(10)} &= \frac{-\partial / \partial d_j}{\partial / \partial N_j} \\ &= \frac{-d_j / N_j^\theta}{S.O.C.} \end{aligned}$$

which is positive since the second-order condition is negative for an interior maximum.

output, linked to the boss to which the worker was assigned on each day.⁶ 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 idiosyncrasies. 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 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

⁶ 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.

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. As in the grading example, there is no obvious interaction among team members. Workers switch bosses about four times per year.⁷ It is these switches to different bosses that permits estimation of the effect of bosses on workers’ productivity.

The measure of productivity is output-per-hour (OPH). The core data measures the time it takes for each transaction, and from this the number of transactions per hour is calculated. Slack time, when the worker is not facing a transaction, is not measured because OPH is calculated as (60 / average minutes per transaction). Each worker handles about 10.3 transactions per hour.

A second measure of performance is uptime. 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 empirical analysis focuses on output per hour.

III. The Basic Results

The empirical specification is to estimate the stochastic version of equation (6) above written as

$$(11) \quad q_{it} = \alpha_0 + \alpha_i + X_{it}\beta + d_{0t}/N_{jt}^0 + d_{0t-1}/N_{m\ t-1}^0 + \dots + d_{0p}/N_{p\ 0}^0 + d_{jt}/N_{jt}^0 + d_{m\ t-1}/N_{m\ t-1}^0 + \dots + d_{p0}/N_{p0}^0 + v_{it}$$

⁷ 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.

Each boss j 's quality is assumed in (11) to be invariant over time, although it is possible conceptually, but not econometrically (the boss tenure data is censored), to allow boss effects to vary with boss tenure just as worker productivity varies with worker tenure.

The error term, v_{it} , may simply be classical error or it may be composed of two components: classical error, ε_{it} , and a term, φ_{ijt} , which captures interaction effects between the worker and the boss with whom the worker is paired at time t . It is conceivable that worker i is better suited to boss j than to boss k and φ_{ij} allows that generality. In that case, (11) is written as

$$(12) \quad q_{it} = \alpha_0 + \alpha_i + X_{it}\beta + d_{0t}/N_{jt}^\theta + d_{0t-1}/N_{m\ t-1}^\theta + \dots + d_{0p}/N_{p\ 0}^\theta + d_{jt}/N_{jt}^\theta + d_{m\ t-1}/N_{m\ t-1}^\theta + \dots + d_{p0}/N_{p0}^\theta + \varphi_{ijt} + \varphi_{ijt-1} + \dots + \varphi_{ij0} + \varepsilon_{it}$$

A contemporaneous-effects only version of (12) that will be used in some of the estimation is

$$(13) \quad q_{it} = \alpha_0 + \alpha_i + X_{it}\beta + d_{0t}/N_{jt}^\theta + d_{jt}/N_{jt}^\theta + \varphi_{ijt} + \varepsilon_{it}$$

Estimation begins with equation (13). Also estimated is a contemporaneous version, which assumes that bosses have only public good effects, implying $\theta=0$, and also assumes that match effects are zero, or $\varphi_{ijt}=0$.

$$(14) \quad q_{it} = \alpha_0 + \alpha_i + X_{it}\beta + d_{0t} + d_{jt} + \varphi_{ijt} + \varepsilon_{it}$$

A. A Preview 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 VII 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 will be used and described in more detail in the subsequent analysis of section VI. One method is to use mixed effects estimation, which allows for random interaction effects, denoted φ_{ij} above in (12), between workers and bosses. All approaches yield the same qualitative conclusions. Bosses are both important to worker productivity and vary in their effectiveness.

B. Regression Results

The first step in estimating the impact of bosses on productivity is to estimate the productivity regression (13) for the effect of contemporaneous bosses on output. Equation (13) restricts the effect of past bosses on current worker output to be zero, but permits bosses effects to have a public and private component and also allows the effect of boss j on one worker to differ from that on another worker.

The first set of estimates shown in Table 2 employs a mixed model specification. The mixed effects specification treats α_i and d_j as random effects but allows arbitrary correlation between the random effects design matrix $[A \ D \ F]$ and X , where A , D , and F are matrices of worker, boss and match indicators and X is the matrix of other right hand side variables. This is in contrast to the usual random effects estimator that requires orthogonality between the random effects design matrix and X . The identifying assumptions are: $E(\alpha|X) = E(d|X) =$

$$E(\varphi|X) = E(\varepsilon|X, A, D, F) = 0, \quad \text{and} \quad Cov \begin{pmatrix} \alpha \\ d \\ \varphi \\ \varepsilon \end{pmatrix} \Big| X = \begin{bmatrix} \sigma_\alpha^2 I_{\#W} & 0 & 0 & 0 \\ 0 & \sigma_d^2 I_{\#B} & 0 & 0 \\ 0 & 0 & \sigma_\varphi^2 I_{\#M} & 0 \\ 0 & 0 & 0 & R \end{bmatrix} \quad \text{where } I_{\#W}, I_{\#B}, \text{ and}$$

$I_{\#M}$ are identity matrices with sizes corresponding to the numbers of workers and bosses and the number of distinct matches in the data, respectively. R is the covariance matrix of the errors, $\sigma_\varepsilon^2 I_N$. 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 . The covariance parameters σ_d^2 , σ_α^2 , and σ_φ^2 are estimated via restricted maximum likelihood, using the procedure detailed in Abowd, Kramarz, and Woodcock (2006).

Table 2 reports the results.⁸ All models contain tenure controls given by a fifth order polynomial in tenure.⁹ ¹⁰ Not surprisingly, and consistent with prior work in other industries,¹¹ worker output is increasing and concave in tenure. The regression also includes day of the week dummies and month dummies, which capture technological change and demand conditions.¹²

The key message in Table 2 is that bosses matter: the standard deviation of the boss effect is large, equaling 4.74, whereas the standard deviation of worker effects is only 1.33.¹³ There is a large literature in labor economics that emphasizes how differences in workers' underlying ability affects their productivity or their wages rates. Here, the variance of the boss effects dwarfs that of worker effects.

The second message is that bosses provide both public and private mentoring. The estimates in column 1 of Table 2 include an estimate of $\theta=0.30$, which speaks to the private or public nature of boss time on worker productivity. Were boss time completely private, $\theta=1$ and from (7), the total effect of the boss would simply be \underline{d}_{jt} . At the other extreme, if $\theta=0$, then the effect of boss time would be $\underline{d}_{jt}N_{jt}$. With an estimated θ of 0.30 and an average team size of 9.04, the boss effect equals $4.67 \underline{d}_{jt}$, which means that boss time is about half way between being purely private and purely public. One interpretation is that half of what a boss teaches is done in a common setting, with the rest taking the form of private tutoring.

⁸ Throughout the remainder of the paper, details about the estimation procedures are contained in the notes accompanying each table.

⁹ 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.

¹⁰ For fixed effects models, because the month dummies and the tenure profile are nearly collinear within person, we estimate the tenure profile as $g(\text{ten}) = 1\{\text{ten} < 365\} * f(\text{ten}) + \mathcal{L}1\{\text{ten} \geq 365\} + 1\{\text{ten} \geq 365\} * f(365)$,

where f is a fifth order polynomial over the first year. 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.

¹¹ See Lazear (2000), Shaw and Lazear (2008) for examples of estimated productivity-tenure profiles.

¹² Although the measure of output is average transaction time (from which output-per-hour is inferred), it is possible that workers might speed up when there is a long queue of customers waiting for service. 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.

¹³ The expectation of the mixed effects is zero over the entire sample, so pairing a boss with the typical worker yields a boss effect that has an expected value equal to \underline{d}_{jt} . It is that standard deviation that is reported in Table 2.

The more comprehensive productivity equation contains lagged boss effects on current productivity, estimating equation (12). It is important to control for these lagged effects in estimating the effect of current bosses, but a subsequent regression will estimate the persistence and depreciation of past bosses on current productivity (in section IV). In column (2) of Table 2, it is shown that with these lagged boss effects, the standard deviation of current boss effects is 5.04 when the θ is constrained to be .3 based on the column 1 results.

Column 3 of Table 2 assumes that $\theta=0$. The variation of bosses on productivity is only slightly diminished, with the standard deviation of boss effects now equal to 4.104. Another way to estimate boss effects is the standard method, which employs standard fixed effects estimation. Typical productivity or wage regressions introduce worker fixed effects; boss fixed effects can be added. In fixed effects estimation, the standard deviation of the boss effects, shown in column (4), is 3.44, which is about 2 ½ times larger than the standard deviation of the worker fixed effects themselves. This is calculated by taking the standard deviation of the estimated boss fixed effects, weighted by the number of worker-day observations.¹⁴ Although these estimates are somewhat diminished from those using mixed effects estimation, qualitatively the result is the same. Boss effects are significant and varied, even more so than worker effects.

There are numerous advantages to the mixed effects methods used above. First, the mixed effects specification allows us to estimate worker-boss match effects. Second, fixed effects suffer from a problem with sampling error, so determining the true variance of the boss effects is difficult. The mixed effects specification allows us to estimate the variance of boss, worker, and match effects directly.

No matter the estimation method, 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.

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

¹⁴ A boss effect estimated with a small numbers of workers for that boss will have more sampling error than a boss fixed effect estimated off a large number of workers. Because of the inconsistency of the individual fixed effects estimates in short panels, sampling variation is non-negligible. The simplest way of correcting for this excess variance in the boss and worker fixed effects is to weight the fixed effects by the observations in the sample. This is done in the last column of Table 2.

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 estimation results (column 1, Table 2), the boss who is at the 90th percentile of boss quality distribution increases productivity by 6.07 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.

There are a number of ways to estimate the absolute productivity level of the boss effect and none is without problems.¹⁵

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 and 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

¹⁵ For example, implicit in (13) is an estimate that comes from d_0 , but this places very heavy weight on variations in team size to identify the effect of the boss on workers.

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).¹⁶ 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 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. The Boss Effects are Identified

The intuition behind identifying the boss effects comes from the fact that workers switch bosses frequently. The change in worker productivity associated with the switch to a new boss provides the relevant information for identifying the boss effect. 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. 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.¹⁷ 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.

¹⁶ See Lazear (2004) for a theoretical exposition of promotion decisions under uncertainty and the effects of error.

¹⁷ 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).

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 4.7 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 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), but in a more constrained fashion. Assume teaching occurs in a public fashion, thus setting $\theta=0$. Specify the persistent (teaching) portion of past boss effects as λ . Assume the past boss’s effect is independent of the length of the spell with that boss, so that

$$(15) \quad q_{it} = X_{it}\beta + \alpha_i + \delta_j + \sum_k^j 1\{t > \tau_{ik}\} \gamma^{t-\tau_{ik}} \lambda \delta_k D_{ik,t-\tau} + \varepsilon_{ijt},$$

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 $t-\tau$. 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 $(t-\tau_{ik})/30$. In sum, equation (15) 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 $\lambda=0.67$. 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. Worker - Boss Match Effects

The treatment effect of boss quality on worker productivity may vary with the quality of the worker. Heterogeneity in the treatment effect was permitted in equation (12) and Table 2, through the φ_{ij} effect. At a conceptual level, 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 and the best among them, 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 $\partial E/\partial b$ and $\partial H/\partial b$ 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.

With estimates of φ_{ij} in hand from (12) estimated in Table 2, 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, d_j , is above or below the median. Workers are also classified as "star" or "laggard" according to whether their estimated worker effect, α_i , is above or below the median.

The designations of good/bad bosses and star/laggard workers are formed from the

distributions of the random boss and worker effects holding constant the match effect. Because the match effects are unbiased, so too are the designations. There are four cells of (good-boss, good-worker), (good-boss, bad worker), (bad-boss, good-worker), and (bad-boss, bad-worker). To obtain estimates, 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 top panel of Table 4 provides cell means for regression (12) with $\theta=.30$ (based on column 1 of Table 2) and the bottom panel provides cell means for $\theta=0$ (based on column 3 of Table 2). The results do not differ between the two panels, so concentrate on the more flexible model in the top panel.

The issue here is one of comparative advantage: how best to allocate the bosses. The results in Table 4 provide a clear answer. There are two choices. Either good bosses are paired with stars, which implies that bad bosses are paired with laggards, or bad bosses are paired with stars, which implies that good bosses are paired with laggards. Combining good bosses with stars and bad with laggards yields an average match effect of $.100-.063 = .037$. Combining bad bosses with stars and good with laggards yields an average match effect of $-.083+.051 = -.032$. The value of bosses is maximized by assigning the better bosses 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 net average gain from proper assignment over incorrect assignment is $.037 - (-.032) = .069$ on a mean output-per-worker hour of 10.26. This is less than 1% of output.

VI. Determination of Team Size Across Bosses

Equation (10) implies that better bosses should have larger teams and yields a specific functional form for the relation of team size to boss effect. It is impractical for the firm to adjust team size on a minute-by-minute basis, but it is reasonable to expect that over long periods of time and within a particular site, team size could be altered to assign more workers to the better bosses. If (10) holds, then team size is an endogenous variable, which means that the estimates might not be consistent. That issue is ignored and it is assumed that the d_j boss effects are

estimated appropriately in (14), as reported in Table 2. Using the d_j coefficients from that, it is possible to compute the correlation between the d_j (which reflects the entire four year time period d_j) and the within-site-within-time period team size. The correlation is essentially zero, at -.019. Notwithstanding the endogeneity issue, there does not seem to be much of a relationship between the number of workers assigned to a boss and that boss's productivity.

This is both good and bad. It is good because to the extent that the d_j are close to what they would be taking endogeneity into account, ignoring endogeneity is not much of an issue. It is bad because it raises the question as to why the firm is not adjusting team size appropriately. One possibility is that the firm is unaware of the d_j , which is difficult to infer without sophisticated statistical analyses.

VII. Non-random Assignment of Workers to Bosses

There may be non-random assignment of experienced workers to bosses. This section presents evidence to assess whether non-random assignment is a concern for the estimates.

A. A Specification Test

The mixed effects estimator 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 based on the fixed effects estimator. 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, 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 not based on the idiosyncratic match quality component. If the assignment of workers to bosses is not random, then the estimated match effects from Table 2 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 0.0014 with a standard error of the mean of 0.0011 and the mean of the bosses' average match effect across bosses is 0.0014 with a standard deviation of 0.0018.¹⁸ 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.

The most likely source of non-random sorting is through assignment based on match-specific productivity, which is captured already.¹⁹ Still, it is interesting to examine non-random assignment and to consider any possible sensitivity of the estimates of the boss effects to non-random assignment that might result because of a specification different from the one assumed in (11)-(14). A series of further tests suggests that non-random assignment, in this context, is unlikely to be a significant problem for the estimates of boss effects.

A. Using Randomly Assigned Workers to Validate the Estimates

The first test examines whether the estimated boss effects predict well out of sample. 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 is driven by the stochastic nature of quits, reflecting the fact that new workers randomly fill open slots.²⁰

The experiment uses new workers who were allocated to their first boss based on the timing of quits and vacancies to conduct an out of sample validation exercise. We first estimate boss quality using data from older workers with three or more prior bosses and then assess whether the boss quality measures recovered from the partitioned sample of experienced workers

¹⁸ 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 .

¹⁹ Recall that unbiased estimates of the mixed effects model do not hinge on random assignment of workers to bosses on match-specific productivity because that model includes boss-worker interactions (the ϕ_{ijt} terms).

²⁰ 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.

predict the productivity of new workers. Interviews with company management indicated that workers on their first boss assignment are assigned haphazardly in a process based on random vacancies.²¹ If non-random sorting is confounding estimation of individual boss quality, the estimated boss effects should have little predicted power on a partitioned sample of nearly randomly assigned workers. The maintained hypothesis is that new workers are assigned randomly to fill slots that have been vacated by other workers mid-match spell.

In step one, boss effects were estimated on a subset of matched worker-boss data for older workers who have had at least two previous bosses. The estimated individual boss quality measures were recovered through Henderson's mixed model equations. These BLUPS (best linear unbiased predictors) were saved. In step two, daily output per hour for new workers' on their first boss assignment was regressed on estimates of boss quality. In the public model with $\theta=0$, boss quality is measure using the bosses' BLUP, $\widehat{d_j^{\theta=0}}$, estimated from the set of experienced workers. The estimating equation is $oph_{i,Boss1,t} = \alpha + X_{it}\beta + \delta_1\widehat{d_j^{\theta=0}} + \varepsilon_{it}$, where X_{it} contains year x month dummies, day of week dummies, and a fifth order tenure polynomial.

In the model with $\theta=0.3$, the measure of boss quality is $\frac{\widehat{d_j^{\theta=0.3}}}{team\ size^{0.3}}$ yielding an estimating equation $oph_{i,Boss1,t} = \alpha + X_{it}\beta + \delta_1\frac{\widehat{d_j^{\theta=0.3}}}{team\ size^{0.3}} + \frac{\delta_2}{team\ size^{0.3}} + \varepsilon_{it}$.

Assessing whether the boss quality measures predict well out of sample (i.e., beyond the older group on which they were estimated) implies a null hypothesis that $\delta_1 = 1$.

In practice, testing this hypothesis raises at least two difficulties. First, while the boss BLUPS contain very little sampling error, measurement error is not eliminated entirely. This will bias estimates of δ_1 toward zero, resulting in over-rejection of the null that $\delta_1 = 1$. Second, the standard errors from estimating the above models will be smaller than the true standard error because the boss quality measures are generated regressors, also resulting in over-rejection of the null.

The results in Table 5, Columns 1 and 2, suggest that the estimated boss quality measures predict well out of sample, but they are statistically different from 1 at the 5% level using a t-test with standard errors clustered by boss. However, given the difficulties with inference, the boss

²¹ We are unable to test if observable characteristics of new workers are balanced across bosses because the data contain only worker identifiers, their start dates, and production histories.

quality measures 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. The parameter estimates are 0.8 and 0.72 for the model with $\theta=0$ and $\theta=0.3$, respectively. On average, a good boss is always a good boss, validating the measures.

Of course, it is possible that the initial assignment is not random, invalidating the maintained hypothesis and the test. As a result, two additional tests of non-randomness are presented that do not rely on this maintained hypothesis.

B. Testing for Non-random Boss Transitions

The remaining columns of Table 5 test for non-random sorting on unobservables. Consistent estimation of the individual boss quality measures requires orthogonality between the design matrix of boss assignments and the concurrent and lagged residuals in the productivity equation. 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. Two tests are implemented. The first is a test to see whether the quality of future bosses predicts the mean residual from Columns 1 or 2 of Table 5. Because these residuals are calculated after random assignment to a boss, sorting does not contaminate the estimates of these residuals. The test is implemented by regressing the residuals on indicators for the quartiles of the distribution of the subsequent boss. Under the null hypothesis of no sorting on unobservables, future bosses should not predict these residuals and the quartiles of future boss effects should be unrelated to the lagged residual. Using quartile indicators for boss quality is a stronger test than regressing the residuals on linear boss quality because additional information about changes over the distribution can be captured. A Wald test cannot reject that the quartile indicators are zero in columns 3 and 4, providing some additional reassurance that the boss quality measures are not contaminated by sorting on unobservables. While this test cannot speak to the allocation of bosses and workers that occur 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.

Just as before, this test is based on the maintained hypothesis that workers' initial assignment is random. One test that does not maintain this assumption assesses whether the residuals on the first boss spell predict second period boss quality. These results are contained in the last two columns of Table 5. Some statistical evidence for predictability is detected.

However, the parameter estimates are essentially zero, suggesting that non-random sorting is unlikely to be problematic economically.

VIII. 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 estimate Cox proportional hazard models of the probability of boss exit. The exit rate regresses an indicator that the boss's estimated fixed effect is below the 10th percentile of the distribution or above the 90th percentile. The prediction is that bad bosses leave and good bosses stay. Results are presented using two estimated boss effects – those with the public/private boss effect estimated to be $\theta=.30$ and those with the public boss effect in the regression in which $\theta=0$.

The results are presented in Table 6. Coefficients and standard errors are presented, but the exponentiated coefficients are most easy to interpret. Bosses in the bottom 10% are more than twice as likely to exit than bosses outside of the bottom 10%. This is true for all specifications.²² To ensure that this result is not due to noise (the concern being that the estimated boss quality measures for short-lived bosses are most likely to be in either tail of the distribution), the specifications include indicators for bosses in the top 10%. These coefficients are small and are not statistically different from zero.

IX. 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:

²² Regressions are also run with boss tenure, where boss tenure is inferred from the time that we observe bosses, and thus is left-censored. The hazard rate models are unchanged.

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 Ichimura, 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 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.

$$(16) \quad q_{ijt} = X_{it}\beta + \alpha_i + \delta_j + \xi p_{ijt} + \varepsilon_{ijt}$$

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

$$(16') \quad q_{ijt} = X_{it}\beta + \alpha_i + \delta_j + \xi_{Peer}(TeamSize - 1)^{-1} \sum_{k \in jt \setminus \{i\}} \alpha_k + \varepsilon_{ijt}$$

where summation over $k \in jt \setminus \{i\}$ 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 7 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 with a customer, not with other workers.²⁵ Although the workers can see each other and may

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.

²⁵ The same is true in Mas and Moretti (2009), who also find significant, but small peer effects.

learn from each other or compete with each other, the workers do not appear to be complements in production.

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. Most important, 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. Using one normalization, the value of the average boss is about 1.75 times that of a worker. 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. Good bosses have a comparative advantage in working with the better workers, but the differences are not economically important.

References

1. Abowd, John, Robert Creedy, 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 Boss Effects

	<u>Mixed Effects</u>			<u>Fixed Effects</u>
	(1) ⁺	(2) ⁺⁺	(3) ⁺⁺⁺	(4) ⁺⁺⁺⁺
Standard Deviation of Boss Effects*	4.740	5.040	4.104	3.440
θ	0.300	0.300		
Standard Deviation of Worker Effects	1.330	1.320	1.350	1.320
Standard Deviation of Match Effects	0.758	0.673	0.752	
R-squared				0.243
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

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.

Notes:

⁺ Column 1 estimates equation (13) using mixed effects estimation.

⁺⁺ Column 2 estimates equation (12) using mixed effects estimation. θ is constrained to be .3.

⁺⁺⁺ Column 3 estimates equation (14), assuming $\theta = 0$ using mixed effect estimation.

⁺⁺⁺⁺ Column 4 estimates equation (14), assuming $\theta = 0$ using fixed effects estimation and weighting the estimated fixed effects by worker-days.

* The Standard Deviation of the Boss Effect is weighted by average team size of 9.04.

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: Boss-Worker Matching

Panel A: Match Effects from Column 1, Table 1 ($\theta = .300$)		Boss	
		Good	Bad
Worker	Star	0.100	-0.083
	Laggard	0.051	-0.063
Panel B: Match Effects from Column 2, Table 1 ($\theta = 0$)		Boss	
		Good	Bad
Worker	Star	0.104	-0.077
	Laggard	0.044	-0.066

Table 5: Out of Sample Validation of Boss Effects and Tests for Non-Random Assignment

Dependent Variable:	Out of Sample Validation		Sorting Test 1: Do quantiles of future boss quality predict residuals?		Sorting Test 2: Do lagged residuals predict future boss quality?	
	Oph	Oph	Mean residuals from Column (1)	Mean residuals from Column (2)	Boss BLUP estimated from experienced sample (Theta = 0)	Boss random coefficient estimated from experienced sample (Theta = 0.3)
Sample:	New workers' on their first boss assignment		New workers' on their first boss assignment		New workers' on their second boss assignment	
	(1)	(2)	(3)	(4)	(5)	(6)
Boss BLUP estimated from experienced sample (Theta = 0)	0.8008 (0.0814)					
Boss random coeff. estimated from experienced sample (Theta = 0.3) x Team Size ^ -0.3		0.7163 (0.0669)				
Bosses in the bottom 25% of either Blups (column 3) or random coefficients (column 4)			-0.1039 (0.0620)	-0.0560 (0.0698)		
Bosses in 25%-50%			-0.0382 (0.0582)	-0.0895 (0.0667)		
Bosses in 50%-75%			-0.1115 (0.0618)	-0.0404 (0.0618)		
Mean residual from Column (1)					0.0036 (0.0016)	
Mean residuals from Column (2)						0.0109 (0.0038)
Wald statistic that parameters on boss quartiles are zero (Chi-squared 3)			4.9012	2.0311		
P-Value			0.1792	0.5660		
R-squared	0.1058	0.1076	0.0366	0.0299	0.2404	0.4010
Number of Observations	782,778	782,778	10,935	10,935	10,935	10,935
Notes:						
Standard errors in parentheses, clustered by boss. Wald tests are calculated using a variance-covariance matrix clustered by boss. The sample in columns 1-4 contains workers on their first assignment prior to the first boss switch. The sample in columns 5 and 6 track the workers from columns 1-4 after their first boss change. To be included, workers must have had at least 20 days of tenure on the first boss spell; this choice was to remove data from a training period. Boss BLUPS or random coefficients are calculated using a partitioned set of workers as follows: First, using a sample including only workers after their second boss switch, we compute boss effects for this sample by regressing oph on a tenure polynomial, month, and day of week fixed effects, along with boss, worker, and match random effects (or random coefficients in the case of Theta = 0.3). We then recover the individual boss BLUPs (or random coefficients). Second, we merge the boss quality measures onto the sample of workers on their first boss. The models in columns 1 and 2 have controls for tenure and monthly time dummies. The models in columns 3-6 add establishment						
Notes for internal use: The Wald statistic in column 3 is 5.45 (p-value 0.14) if we instead use residuals from regressing oph on quartiles of initial boss quality as the dependent variable.						

Table 6: Cox Proportional Hazard Model

	$\theta = .30^+$	$\theta = 0^{++}$
Boss Effect Below 10th Percentile	1.000	0.591
	2.720	1.800
	-0.090	-0.090
Boss Effect Above 90th Percentile	0.040	0.083
	1.040	1.086
	-0.100	-0.100
N.	620,130.000	620,130.000

Notes:

Each cell contains the coefficient, the exp (coefficient), and the standard error in parentheses.

⁺ The Boss Effects are from the estimation of equation (13) in column 1 of Table 2.

⁺⁺ The Boss Effects are from the estimation of equation (14) in column 3 of Table 2.

Table 7: 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.