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

The value of bosses

Working paper

Original citation:
Lazear, Edward P., Shaw, Kathryn L. and Stanton, Christopher (2014) *The value of bosses*. Working papers series, Social Sciences Research Network, Rochester, NY, USA.
DOI: 10.2139/ssrn.2131572

The Value of Bosses

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

Stanford University,
Stanford University,
University of Utah

June 13, 2014

We greatly appreciate the comments of seminar participants at the University of Chicago, Columbia, Yale, Stanford, Harvard, MIT, USC, Northwestern, 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 more than would adding one worker to a nine member team. Workers assigned to better bosses are less likely to leave the firm. A separate normalization implies that the average boss is about 1.75 times as productive as the average worker.

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? It is 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; if so, how significant is the variation both in absolute terms and relative to the workers whom they supervise? Even if bosses vary in their effects on worker output, do these variations persist or do they die out with time? Finally, are some bosses more likely to retain their workers than other bosses?

These questions merit examination. 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.[1] Despite the potentially important role that supervisors play, the economics literature has been largely silent on the effects that bosses actually have on affecting worker productivity.[2]

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.[3]

---

[1] 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.

[2] The literature has focused on CEOs or managers in detailed occupations. For work on CEOs' productivity, see Bennedsen, Perez-Gonzalez, Wolfenzon (2007), Bennedsen, Nielsen, Perez-Gonzalez, Wolfenzon(2007), 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).

[3] An exception is Garicano (2000) and Garicano and Rossi-Hansberg (2006). In these models, a supervisor is effective because of differences in knowledge, and the supervisor's task is to help production workers solve exceptional problems that arise. A supervisor's productivity and span of control is determined by the arrival rate of problems that can be solved directly by subordinates compared to the problems that can be solved by the worker

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.[4] 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.

By using data from a large services company with individual-level records of productivity, it is possible to examine the effect 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 company.

The primary findings are:

1. Bosses vary greatly in productivity. The difference between the best bosses and worst bosses is significant. Replacing a boss who is in the lower 10th percentile 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.
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.

---

under the supervisor's direction. The production environment analyzed here has some similarities in that bosses may teach workers how to deal with new problems, but the time spent solving problems for workers is limited.

[4] 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 component that differentiates the effect of particular bosses on workers is not highly persistent. About one-fourth of the boss-specific effect remains one year after the worker leaves a particular boss.
4. The worst bosses are more likely to separate from the firm. 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. Workers who are assigned to better bosses are more likely to remain with the firm, which is another aspect of boss productivity.
6. 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) $$q_{it} = H_{it} * E_{it} .$$

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) $$H_{it} = H(X_{it}, a_i, b_{it})$$

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

(3) $$b_{it} = b\left(d_{jt} \big/ N_{jt}^{\theta}, d_{mt-1} \big/ N_{mt-1}^{\theta}, \ldots, d_{p0} \big/ N_{p0}^{\theta} \right)$$

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_{mt-1} is the quality of boss m with whom the worker is paired at time t-1, N_{mt-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, as described below. 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 θ=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 θ=0. The boss's instruction or motivation improves all workers and there is no congestion. For 0< θ<1, there is some public good aspect to boss time and some private good aspect. A private good is one with total congestion.[5]

Analogously, effort is

(4) $$E_{it} = Z\left(X_{it}, a_i, b_{it}\right)$$

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

$$q_{it} = H\left(X_{it}, a_i, b\left(d_{jt} / N_{jt}^{\theta}, d_{mt-1} / N_{mt-1}^{\theta}, \ldots, d_{p0} / N_{p0}^{\theta} \right)\right) * Z\left(X_{it}, a_i, b\left(d_{jt} / N_{jt}^{\theta}, d_{mt-1} / N_{mt-1}^{\theta}, \ldots, d_{p0} / N_{p0}^{\theta} \right)\right)$$

or

(5) $$q_{it} = \left(f\left(X_{it}, a_i, b\left(d_{jt} / N_{jt}^{\theta}, d_{mt-1} / N_{mt-1}^{\theta}, \ldots, d_{p0} / N_{p0}^{\theta} \right)\right)\right).$$

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

(6) $$q_{it} = a_0 + a_i + X_{it}\beta + \left(d_{0t} / N_{jt}^{\theta} + d_{0t-1} / N_{mt-1}^{\theta} + \ldots + d_{00} / N_{p0}^{\theta} \right) + d_{jt} / N_{jt}^{\theta} + d_{mt-1} / N_{mt-1}^{\theta} + \ldots + d_{p0} / N_{p0}^{\theta}$$

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 α_i and of the d_{jt} is definitionally zero. Although this model writes α_0 and d_{0t} as separate parameters, identifying them separately is difficult in practice.

A contemporaneous-effects only version of (6) is

(7) $$q_{it} = a_0 + a_i + X_{it}\beta + d_{0t} / N_{jt}^{\theta} + d_{jt} / N_{jt}^{\theta}.$$

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

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

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

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.

Why would the boss effect persist over time? 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. A good boss may teach or motivate and these effects may persist over time.

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?

E3: Do past bosses influence workers' current output? Are boss effects persistent?

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.

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 maximize

(9) $$N_j q_j + (N - N_j) q_k$$

where q_j and q_k are given by (6).

Optimization requires that the first order condition holds

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

or that

(10) $$(d_0 + d_j) / N_j^{\theta} = (d_0 + d_k) / (N - N_j)^{\theta} .$$

Eq. (10) implies that as long as $\theta > 0$, team size increases in d_j.[6] The larger team is allocated to

[6]Using the implicit function theorem on (10),

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.[7]

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?

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. The Data, the Company, and the Assignment of Workers to Bosses

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.[8] The period covered is

---

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

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

[7] Rosen (1982) suggests that managers in top positions should be assigned more workers because the higher productivity of these managers will filter down in the organization when there is a recursive chain of control technology. For papers suggesting a similar scale-of-operations effect, see Mayer (1960), Lucas (1978), and Garicano (2000).

[8] In the data, 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.

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 job code classification where workers are involved in one task involving 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 the same activity.

To provide some context, consider an example of a technology-based service job where workers do 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 they 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 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 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.

The core measure of productivity is output-per-hour (OPH). The computer system 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.[9]

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 in productivity 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.

A final measure of productivity is quality. Quality is measured on a score of 0 to 5 and is collected approximately weekly for each worker using a randomly administered post-transaction survey. The quality data is sparse and was not used heavily by the management at the company, so for these reasons, it is de-emphasize in the analysis that follows.

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.

Table 1 provides summary statistics on the productivity measures and the extent of boss switching. Each boss interacts with, on average, about 50 different workers, and each worker has, on average, about 4 different bosses in the sample.

The identification of boss effects on workers’ productivity comes from the movement of workers between bosses so an important issue is the method of assignment of workers to bosses. Workers switch bosses about two to three times per year.[10] New workers who enter the company randomly fill in open spots due to worker attrition. The firm also has a policy to occasionally re-

[9] In Lazear, Shaw and Stanton (2013), the relation of productivity to demand conditions is explored in more detail.

[10] 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.

assign experienced workers and bosses. This policy was in place to deal with two issues related to attrition. First, teams needed to be re-balanced, where the goal of re-balancing is to balance the number of experienced relative to inexperienced on the team. Second, many employees required accommodations due to changing schedules outside of work, and re-assignment allowed employees to adjust their shift rather than leave the firm. In interviews with the company, we were told that when assignments to bosses were rotated that management played no centralized role in assigning workers to bosses. Workers with superior recent production tended to be given precedent in choosing their shifts, but the process was described as haphazard and subject to change over time. The high amount turnover at the firm, the need to keep team sizes balanced, and the need to switch shifts are the primary force behind worker mobility across teams.

One concern is that the potential for non-random sorting affects the ability to estimate boss effects. The theory above suggests that the number of workers assigned to a boss should depend on the quality of the boss. This does not imply that there should be any correlation between boss quality and worker quality, which is a separate issue and depends on comparative advantage. It is the latter issue, not team size, that is the primary concern. There are several formal statistical tests for non-random sorting in later sections that suggest that random assignment prevails.

III. The Basic Results

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

$$(11)\quad q_{it} = a_0 + a_i + X_{it}\beta + {d_{0t}}/{N^{\theta}_{jt}} + {d_{0t-1}}/{N^{\theta}_{mt-1}} + \ldots + {d_{0p}}/{N^{\theta}_{p0}} + {d_{jt}}/{N^{\theta}_{jt}} + {d_{mt-1}}/{N^{\theta}_{mt-1}} + \ldots + {d_{p0}}/{N^{\theta}_{p0}} + \nu_{it}\,.$$

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 not observed), to allow boss effects to vary with boss tenure just as worker productivity varies with worker tenure. To ensure notation is clear, the t subscript on d_{jt}/N^{θ}_{jt} is the effect of boss quality on current productivity, q_{it}, when a worker is matched with boss j at time t. Indexing by t captures the history of past boss assignments, allowing the effect of past bosses to persist in a general way.

The error term, ν_{it}, may simply be classical error or it may be composed of two components: classical error, $\varepsilon_{it,}$ and a term, φ_{ijt}, which captures interaction or match 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)

$$q_{it} = a_0 + a_i + X_{it}\beta + d_{0t}/N_{jt}^{\theta} + d_{0t-1}/N_{mt-1}^{\theta} + \ldots + d_{0p}/N_{p0}^{\theta} + d_{jt}/N_{jt}^{\theta} + d_{mt-1}/N_{mt-1}^{\theta} + \ldots + d_{p0}/N_{p0}^{\theta} + \varphi_{ijt} + \varphi_{ijt-1} + \ldots + \varphi_{ij0} + \varepsilon_{it} .$$

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

$$\text{(13)} \quad q_{it} = a_0 + a_i + X_{it}\beta + d_{0t}/N_{jt}^{\theta} + d_{jt}/N_{jt}^{\theta} + \varphi_{ijt} + \varepsilon_{it} .$$

Estimation begins with equation (13). A version of (13) that constrains θ to be equal to 0, i.e., assumes that boss effects are completely public, is also estimated as

$$\text{(14)} \quad q_{it} = a_0 + a_i + X_{it}\beta + d_{0t} + d_{jt} + \varphi_{ijt} + \varepsilon_{it} .$$

A. A Preview of Estimation Issues

Before discussing the estimates, it is important to flag a few potential problems that may arise in estimation. First, as already mentioned, there may be non-random assignment of workers to bosses. As an empirical matter, assignment is close to random. Even it were not, if worker quality is measured well and the functional form of the estimating equation is properly specified[11], then non-random assignment is does not present a problem. A standard control for worker quality is the inclusion of worker effects. If good workers are more frequently assigned to good bosses, there is no bias in the estimates of the boss effect as long as the model controls adequately for worker quality. That is, as long as sorting is on the time-invariant worker or boss

[11] Required is that the error term after accounting for worker quality is not correlated with boss assignment. In many contexts with linked employer-firm data, one may be concerned that the workers who switch firms differ from the workers who do not switch. In the company studied here, workers switch bosses frequently and these switches occur at pre-specified times, alleviating some concern that only a selected set of workers switch bosses.

effect and not on the recent history of residuals, then assortative matching does not bias the estimates. In addition to adding worker effects to control for worker quality, there are more sophisticated and more comprehensive ways both to test for the extent of the non-random assignment problem and to treat it. A variety of methods will be used and described in more detail in the subsequent analysis of Section VII. All approaches yield the same qualitative conclusions that non-random assignment is unlikely to be a serious problem in interpreting the estimates.

B. Estimation Methods

The first step in estimating the impact of bosses on productivity is to estimate the productivity regression (13) to recover the effects 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, θ; the equation also allows the effect of boss j on one worker to differ from that on another worker through the match effects, φ_{ijt}.

The first set of estimates shown in Table 2 employs a mixed model specification.[12] The mixed effects specification treats α_i and d_j as random effects but allows arbitrary correlation between the random effects design matrix Z = [A B M] and X, where A, B, and M 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 more widely known random effects estimator that requires orthogonality between the random effects design matrix Z and X. The identifying assumptions are: $E(a \mid X) = E(d \mid X) = E(\varphi \mid X) = E(\varepsilon \mid X, A, B, M) = 0$,

and $$Cov\begin{pmatrix} a \\ d \\ \varphi \\ \varepsilon \end{pmatrix} | X = \begin{bmatrix} \sigma_a^2 I_{\#W} & 0 & 0 & 0 \\ 0 & \sigma_d^2 I_{\#B} & 0 & 0 \\ 0 & 0 & \sigma_\varphi^2 I_{\#M} & 0 \\ 0 & 0 & 0 & R \end{bmatrix}$$

where $I_{\#W}$, $I_{\#B}$, 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, $\sigma_\varepsilon^2 I_N$. The covariance parameters σ_d^2, σ_a^2, and σ_φ^2 are estimated via restricted maximum likelihood, using the procedure detailed in Abowd, Kramarz, and Woodcock (2006).

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

In some later applications, recovering individual boss, worker, and match effects rather than just the distributional parameters is necessary. To do so, we use Henderson's mixed model equations to recover best linear unbiased predictors (BLUPS) of the individual effects (see Abowd, Kramarz, and Woodcock (2006) for a discussion on the history of this method). Letting $G = \begin{bmatrix} \sigma_a^2 I_{\#W} & 0 & 0 \\ 0 & \sigma_d^2 I_{\#B} & 0 \\ 0 & 0 & \sigma_\varphi^2 I_{\#M} \end{bmatrix}$, then the BLUPS are the solutions for the random effects from

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}\begin{bmatrix} A & B & M \end{bmatrix} \\ \begin{bmatrix} A' \\ B' \\ M' \end{bmatrix} R^{-1}X & \begin{bmatrix} A' \\ B' \\ M' \end{bmatrix} R^{-1}\begin{bmatrix} A & B & M \end{bmatrix} + G^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ a \\ d \\ \varphi \end{bmatrix} = \begin{bmatrix} X'R^{-1}oph \\ \begin{bmatrix} A' \\ B' \\ M' \end{bmatrix} R^{-1}oph \end{bmatrix}.$$

C. Results

The key message in Table 2 is that bosses matter and differ substantially. Column 1 reports the results from the estimating equation (13). The standard deviation of the boss effect is large, equaling 4.74 units of output, 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 affect their productivity or their wages rates. Here, the variance of the boss effects dwarfs that of worker effects.

The estimate of θ equal to .3 in column (1) implies that bosses are engaged in both public and private mentoring.[14] Were boss time completely private, θ would equal 1 and, from (7), the total effect of the boss would simply be $d_{jt.}$ At the other extreme, if θ=0, then the effect of boss time would be $d_{jt}\times N_{jt.}$ With an estimated θ of 0.30 and an average team size of 9.04, the boss effect equals 4.67 $d_{jt.,}$ which means that boss time is about half way between being purely private and purely public. One interpretation is that about half of what a boss teaches is done in a common setting, with the rest taking the form of private tutoring.[15]

---

[13] The standard deviation of boss effects is calculated as the standard deviation of the boss random coefficients, a parameter that is estimated directly, times (average team size)$^{1-\theta}$. 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 $d_{jt.}$ It is that standard deviation that is reported in Table 2.

[14] The estimate of θ is recovered using an outer loop to search over θ; in an inner loop the REML estimates condition on each guess of θ. This is repeated until a value of θ is found that maximizes the REML likelihood.

[15] The variance of worker and match effects is not sensitive to the value of θ. To assess the sensitivity of the

Columns 2 through 4 present other variations of the model. All specifications contain tenure controls given by a fifth order polynomial in tenure.[16] [17] Not surprisingly, and consistent with prior work in other industries,[18] worker output is increasing and concave in tenure. Regressions that include only contemporaneous boss effects include day of the week dummies and month dummies, which capture technological change and demand conditions, to the extent that they are relevant.[19]

Column 2 estimates equation (12) introducing lagged boss effects. The lags are based on the identity of the bosses six and twelve months ago.[20] Because the data do not include complete production histories for experienced workers at the time our sample begins, we use only observations on experienced workers with greater than six months of tenure. In column 2, the value of θ is constrained to be .3 from column 1.[21] When the lagged boss effects are included, the standard deviation of contemporaneous boss effects is 4.08. Although the standard deviation drops somewhat from that estimated in column 1, the estimate is in the same range and still much greater than that for worker effects.[22] It is possible to infer the degree to which boss effects

---

estimates, θ was also constrained to 0.25 and 0.35. The standard deviation of the boss effects is, respectively, 5.10 and 2.90. The model fit according to AIC and BIC is worst when θ is constrained to 0.35.

[16] 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.

[17] For fixed effects models, because the month dummies and the tenure profile are nearly collinear within person, we estimate the tenure profile as $g(ten) = 1\{ten < 365\} * f(ten) + L1\{ten \geq 365\} + 1\{ten \geq 365\} * f(365)$, 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.

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

[19] 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.

[20] The data have gaps, so if the worker was not observed in the data exactly 6 or 12 months previously, we use the identity of the worker's last boss prior to this date. We have also estimated the model after aggregating to the monthly level, using the modal boss in a month as an indicator for the lagged boss. Results using daily data or data aggregated to the monthly level are qualitatively similar.

[21] The value of θ is constrained rather than estimated in models with lags because of computational complexity. As the number of random effects dimensions increases, the computational difficulty in fitting the model increases, making it difficult to estimate θ.

[22] In this specification, θ is constrained to be .3 to be consistent with that obtained in column (1).

persist over time by comparing contemporaneous effects to lagged effects. More is said on this below in section F.

Column 3 of Table 2 constrains θ to be zero and deletes lagged boss effects for the estimation of equation (14). The variation of bosses on productivity remains high, with the standard deviation of boss effects now equal to 4.104.

For comparability with other techniques, column 4 presents boss effects estimated using the more widely-known fixed effects method. In this case, a productivity regression is run that includes boss and worker fixed effects. This specification imposes the assumption that the boss effect is completely public, that is, that $\theta=0$. Now, the standard deviation of the boss effects is 3.44, which is still about 2.5 times as large as the standard deviation of the worker fixed effects themselves.[23] It is reassuring that the basic conclusion is not sensitive to the version of the model that is estimated or to the estimation method used. That said, there are numerous advantages to the mixed effects method over the fixed effects method. First, the mixed effects specification allows estimation of 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 permits estimation of the variance of boss, worker, and match effects directly. Still, 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.

The last column of Table 2 presents results where quality is the dependent variable. The most important finding is that the boss effects on the output-per-hour dimension are positively correlated with the boss effects on the quality dimension. A good boss appears to be a good boss in general and does not reduce quality to get faster transaction times. In addition, the magnitude of boss effects on quality compared to worker effects remains large and is large relative to the standard deviation of the quality effect (equal also to .78).

D. The Impact of Bosses

[23] This is calculated by taking the standard deviation of the estimated boss fixed effects, weighted by the number of worker-day observations. 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. Weighting the fixed effects by the observations in the sample reduces the influence of sampling error on estimates of the distribution of the true fixed effects. This is done in the last column of Table 2.

The fact that there is variation in the effects of bosses on output means that bosses must affect output in the first place. 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.

Mixed estimation assumes a normal distribution of boss effects, which implies that the boss who is at the 90^{th} percentile of the boss quality distribution increases productivity by 6.07 units/hour more than the boss at the 50^{th} percentile. Comparably, replacing a boss who is in the lowest 10^{th} percentile of boss quality with one who is at the 90^{th} 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 on 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.[24] 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 from worker to boss are superior to the best workers. Bosses obtain and retain their jobs only by being

[24] 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. The major concern is that team size variation is modeled to affect output only through the boss effect, but any other effects would also be captured in d_0.

more productive as bosses than they would be as workers. Otherwise, comparative advantage would dictate that they operate as workers rather than bosses.[25] 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).[26] 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 implies that the average boss is about 1.75 times as productive as the average worker. 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 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

[25] For a model that analyzes the choice of boss versus production worker and other hierarchies see Garicano (2000).
[26] See Lazear (2004) for a theoretical exposition of promotion decisions under uncertainty and the effects of error.

the graph-theoretic sense.[27] 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 4.7 times during this interval. Therefore, when the boss is the unit of analysis, her 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.

F. Persistence of Boss Productivity Effects

Some of the knowledge or motivation that bosses pass to their subordinates may be fleeting, but some may persist for extended periods of time. Skills that are taught might stay with the worker longer than contemporaneous motivation that comes from encouragement or threats that vanish once the boss is no longer present. It is tempting to refer to the part of the boss effect that persists as teaching and that which dies out quickly as motivation. Unfortunately, no evidence on the cause or nature of the boss productivity is available. Consequently, we choose to refer simply to persistence, which is the amount of the boss effect that is retained for six or twelve months. The results in column 2 of Table 2 shed some light on this issue.

To see this, recall that $d_{0t-x} + d_{jt-x}$ is defined as the effect of the boss with whom the worker was paired at time t-x on output at time t. Define $d^*_{0t-x} + d^*_{jt-x}$ as the effect of the boss with whom the worker was paired at time t-x on output at time t-x. Then the persistence parameters, λ_{0x} and λ_x, are the part of the effect of the boss on output at time t-x that remains at time t, or

$$d_{0t-x} + d_{jt-x} = \lambda_{0x} d^*_{0t-x} + \lambda_x d^*_{jt-x}$$

[27] 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 two persistence parameters have different interpretations. λ_{0x} is the amount of the average boss effect from time t-x that persists until time t whereas λ_x is the part of the idiosyncratic boss effect from time t-x that persists until time t. It follows that

$$\sigma_{dj\,t-x} = \lambda_x \sigma_{d*j\,t-x}$$

If the d_j distribution is stationary, then $\sigma_{d*j\ t\text{-}x} = \sigma_{d*j\ t}$, which implies that

$$\sigma_{d*j\,t-x} = \sigma_{dj\,t}$$

so an estimate of λ_x is given by

(15) $$\lambda_x = \sigma_{dj\,t-x} \big/ \sigma_{dj\,t}$$

which can be obtained from column 2 of Table 2.

Because there are six and twelve months lags included in the specification, λ_6 and λ_{12} can be estimated from the ratio of the standard deviations in the relevant $d_{j\ t\ -\ x}$. The estimates of λ_6 and λ_{12} suggests that about 35 percent of the current boss effect persists for six months and 26 percent persists for one year.

This result does not imply that what bosses do does not persist. For example, it might be the case that most of what bosses do is convey skills that workers learn early and keep with them for a long time. If all bosses are equally good at providing persistent skills, then this effect is contained in the d_0 component, rather than in the d_j component. Although there is persistence, it does not differ in the cross-section of bosses and therefore is not reflected in the ratio of the standard deviations. It is also possible that bosses differ in their ability to motivate workers. Once the worker leaves the boss, contemporaneous motivation is lost. That would be picked up in the d_j factor. These contemporaneous motivation effects might die out rapidly, which would be consistent with two-thirds being gone after six months.

Even if most of the effect, including that which does not vary across bosses, was contemporaneous, this would not imply that bosses are unimportant. If the effect of having a good boss is fleeting, keeping the worker paired with a good boss is all the more important. Were the skills that were learned kept forever, then it would only be necessary to make sure that a worker encountered a good boss at one point, preferably early, in his career. If the boss effects die out quickly, then having a larger stock of good bosses is required to achieve the same level of output.

IV. Another Measure of Boss Productivity: Worker Retention

The adage, "workers don't leave bad firms, they leave bad bosses" can be tested. Since turnover is costly, a boss who causes workers to leave when the firm would profit by having them retained is implicitly a low productivity boss. Let us define "bad boss" by the d_j that is estimated by examining the effect of the boss on worker output-per-hour. If workers dislike bosses with low d_j, then one might expect that there would be a correlation between worker departure hazard rates and d_j. This is not unreasonable. Workers may prefer to work with bosses who make them productive, perhaps because good bosses enhance worker promotion probabilities the most and thus raise the worker's future wage.

Table 3 reports the results of a Cox proportional hazard model where worker departure is the dependent variable. Because the data do not contain details about attrition versus promotions out of the sample, the analysis is conducted on two separate samples. The first four columns of Table 3 provide estimates of the model on a sub-sample of workers with fixed effects α_i below the mean. These workers are unlikely to be promoted, alleviating concerns about classification error in measuring departures. The remaining columns include the full sample of workers.

Coefficients and standard errors are presented in Table 3, but the exponentiated coefficients are most easy to interpret. The first row of Table 3 (labeled "Boss Effect x Team Size ^ -θ") shows that the effect of d_j on turnover is negative. Good bosses as measured by their best linear unbiased predictor of the boss effect do indeed reduce worker turnover. This is true in every specification and the effect is significant. The effect is not only statistically significant, it is sizeable. A boss who is one standard deviation above the mean quality in d_j (the standard deviation of d_j equals 1.015 at the individual worker level) experiences a 12 percent reduction in the turnover hazard among her workers.

It is also true (see columns 2 and 4), that workers who are in the lowest decile of productivity as measured by their individual effects (α_i) are more likely to leave the firm. There is no significant interaction effect. It is not the case that good bosses are better at losing the worst workers in the firm (row 3).

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 (13) and Table 2, through the match effect, φ_{ijt}. 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 the best 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 $\partial E/\partial b$ 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_{ijt} in hand from Table 2, it is possible to calculate whether good bosses should be matched to good workers or to bad workers.[28] Note that the assumption of a diagonal covariance matrix of the boss, worker, and match effects does not imply that the effect of a boss is uniform on each worker. The possibility that better bosses should be matched with better workers can be assessed through the match effects. There is nothing in the estimation that restricts $\text{cov}(\alpha+\varphi, d+\varphi)$ to be zero.

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).

[28] We thank John Abowd for suggesting this approach and making clear its statistical properties.

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 (13) with θ=.30 (based on column 1 of Table 2) and the bottom panel provides cell means for regression (14) for θ=0 (based on column 3 of Table 2). The results do not differ between the two panels, so concentrate on the first 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=-0.032. 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. To test this, assume the d_j boss effects are estimated appropriately in (13), as reported in column 1 of Table 2. Using the d_j BLUPS from that model, it is possible to compute the correlation between the estimated d_j (which reflects the entire four year time period) and the within-site-within-time period team size. The correlation is essentially zero, at -.019. [29] The firm does not appear to be allocating workers appropriately to bosses, perhaps because it is difficult to determine the $d_{j,}$. The effect of

[29] In contrast to these results, Smeets, Waldman, and Warzynski (2013) use data from a high tech manufacturing firm and find that the span of control of a manager is positively correlated with his performance rating.

a particular boss on workers is challenging to separate from worker effects and it could take considerable time to learn this, simply through observation, of worker output.

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. The most likely source of non-random sorting is through assignment based on match-specific productivity, which is captured already.[30] Still, it is useful 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 (12)-(14). 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. 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 pair after fixed effect estimation, where the mean is over all periods during which the boss and worker are together. The orthogonal match fixed effects estimator imposes that the mean of the match effects for each worker and each boss (although not for each boss-worker pair) is zero by construction. In contrast, the mixed effects estimator allows the observed, actual match effects to deviate from zero for each boss and each worker. The mixed effects estimator instead imposes that the *potential* match effects are zero. This means that if a boss and worker were paired at random, the expected match effect for bosses and for workers 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.[31]

The implication is that the mean of the match effects for each worker and each boss

[30] 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).

[31] If the data were balanced, meaning that every boss and every worker were paired, then the match effects recovered from the orthogonal fixed effects estimator and the mixed effects estimator would both be mean zero.

(taking unique workers and then unique bosses as the units of analysis, respectively) 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 are likely to deviate from zero within workers and within bosses because the worker to boss assignment process will reflect match specific productivity gains. That is, when averaging over actual matches, the mean of the actual matches may differ from the population mean of the matches if there is systematic sorting on idiosyncratic match-specific productivity. It is important to note that this is not a test for assortative matching on the common boss and worker effect. This is a test for matching on idiosyncratic match quality that does not covary with respect to the common boss effect and the common worker effect. Sorting on the common worker and boss effect—both of which are general and invariant to the identity of the match—is not a problem for our estimates.

Using the boss as the unit of analysis (which means taking the mean of the actual match effects BLUPs across workers for a given boss), the average boss match effect is 0.0014 with a standard error of the mean of 0.0018 (for estimates from column 1 of Table 2). When using the individual worker as the unit of analysis (which means taking the mean of the match effects across bosses that a given worker has had), the mean of the workers' average match effect across workers is 0.0014 with a standard error of the mean of 0.0011.[32] 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.

Under the implications of standard matching models,[33] workers would also be expected to spend more time with those bosses with whom they are better matched. Figure 1 provides estimates of the probability that a given worker switches bosses as a function of lagged residual productivity of the worker. The logic is as follows: if the firm or individual bosses and workers were sorting to maximize the productivity of boss-worker matches, then one would expect highly productive matches to endure. For the exercise in Figure 1, a highly productive match can be

[32] 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.

[33] See Jovanovic (1979).

operationalized as the within-worker residual from a restricted version of equation (6) that is estimated without boss effects or match effects.[34] The residual captures the influence of a boss in a very flexible way that relies on few assumptions. Results from a local polynomial regression, contained in the figure, show that the probability that a worker switches bosses in the next month does not vary much as a function of lagged residual productivity in the prior month. The kernel density plot provides a measure of the variability of residual productivity. Workers with high lagged productivity—possibly due to a match with a good boss—are about as likely to switch bosses in the next month as workers with low lagged productivity. This suggests that taking assignment as random conditional on a worker effect is a reasonable assumption for estimating boss effects.

B. Using Randomly Assigned Workers to Validate the Estimates

Other tests are available. The first 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.[35] 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.[36]

The assessment uses new workers who were allocated randomly to their first boss to conduct an out-of-sample validation exercise. We estimate boss quality using data from older workers and then assess whether the boss quality measures recovered from these experienced workers predict the productivity of new workers. If non-random sorting is confounding estimation of individual boss quality, the estimated boss effects should have little predictive power on the sample of randomly assigned new workers. High predictive power, here, means that a regression of productivity for new-workers on the boss effects for older workers should have a coefficient of 1.

---

[34] The residuals are taken from estimates of the equation $oph_{it} = a_0 + a_i + X_{it}\beta + \varepsilon_{it}$ where X_{it} contains a fifth order polynomial in tenure, year-by-month fixed effects, and establishment fixed effects.

[35] There remains the possibility that even the first assignment is not completely random if departures are more common for workers who are assigned to low quality bosses. But even if this were true, new workers would be more likely to be assigned to low quality bosses, but the quality of the workers assigned to those bosses would still be random.

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

To test this, in step one, boss effects are estimated on a subset of data for experienced workers who have had at least two previous bosses. The BLUPS (best linear unbiased predictors) from step one are saved and used in step two, where daily output per hour of new workers is regressed on estimates of boss quality obtained from the experienced worker sample.

The test is conducted for the models with θ=0 and with θ=.3 in the public model with θ=0 from column 3 of Table 2, boss quality is measured using the bosses' BLUP $\hat{d}_j^{\theta=0}$, estimated from the set of experienced workers. The estimating equation for the test is then

$$(16)\; oph_{i,Boss1,t} = a + X_{it}\beta + \delta_i \hat{d}_j^{\theta=0} + \varepsilon_{it},$$

where X_{it} contains year x month dummies, day of week dummies, and a fifth order tenure polynomial. In the model with θ=0.3 from column 1 of Table 2, boss quality is measured as $\frac{\hat{d}_j^{\theta=0.3}}{team\ size^{0.3}}$ yielding an estimating equation for the test

$$(16')\; oph_{i,Boss1,t} = a + X_{it}\beta + \delta_1 \frac{\hat{d}_j^{\theta=0.3}}{team\ size^{0.3}} + \frac{\delta_2}{team\ size^{0.3}} + \varepsilon_{it}.$$

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

Testing this hypothesis raises two difficulties. First, each individual boss BLUP may contain measurement error. This will bias estimates of δ_1 toward zero, resulting in over-rejection of the null that $\delta_1 = 1$. To correct for this, modified null hypothesis is created that accounts for measurement error; the correction is performed for the estimates where θ=0. This correction is not performed in the case where θ=0.3 because the correction for measurement error depends on team size as well as the conditional variance of the boss random coefficients. The variance of measurement error is assumed to be the mean of the conditional sampling variance of the individual boss BLUPs. The result is 0.048. The true variance of the BLUPS is just given by the estimated variance of the boss random effects for the experienced sample: 0.14. If the true coefficient is 1, the estimated coefficient given the measurement error should be 1/(1+0.048/0.14) = 0.745.

Second, the standard errors from estimating the above models will be smaller than the true standard errors because there is no accounting for the fact that the boss quality measures are generated regressors; this also results 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. The parameter estimates are 0.8 and 0.72 for the models with θ=0 and θ=0.3, respectively, are statistically different from 1.[37] However, the estimate is not significantly different from the measurement-error-corrected coefficient. After accounting for the expected influence of measurement error, the data do not reject a coefficient of 1 in the regression of inexperienced worker productivity on the boss effects estimated from the experienced sample. It is reassuring that the boss quality measures estimated from the experienced sample of workers do a reasonable job of predicting the productivity of the new hires who are assigned randomly. On average, a good boss for experienced workers is a good boss for new, randomly assigned workers.

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.

C. 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 determine whether the quality of future bosses predicts the mean residual calculated for each worker after estimating equation (16) or (16'). The test is implemented by regressing the mean worker residuals on dummy variables for the quartiles of the distribution of the subsequent boss. That is, the mean residual for each worker from (16) or (16') is regressed on dummies for quartiles of the quality distribution of a worker's second boss.

Under the null hypothesis of no sorting on unobservables, the dummy variables for quartiles of the distribution of future bosses should not predict the mean residuals from a worker's first spell.[38] A Wald test cannot reject that the quartile indicators for the second boss

[37] The observation counts in Columns 1 and 2 reflect the fact that only workers on their first boss spell are in the sample.
[38] Using quartile indicators for boss quality is a stronger test than regressing the residuals on linear boss quality

assignment are zero in columns 3 and 4, providing additional reassurance that the boss quality measures are not contaminated by sorting on unobservables.[39]

A second test assesses whether the mean residual by worker from the first boss spell, estimated from (16) or (16'), predicts second period boss quality. These results are contained in the last two columns of Table 5. The dependent variable is the boss BLUP (column 5) or boss random coefficient (column 6) on each worker's second boss spell. Some statistical evidence for predictability is detected, as the parameter estimates associated with the residuals from (16) and (16') are statistically different from zero. However, the parameter estimates are very small, suggesting that non-random the extent of non-random sorting is minimal compared to the variability in worker and boss ability.

While these tests cannot speak to the allocation of bosses 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 problem for estimation.

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.[40]

To test this prediction, boss attrition is analyzed. The approach is to estimate Cox proportional hazard models of the probability of boss exit. The model includes indicators that the boss's estimated fixed effect is below the 10^{th} percentile of the distribution or above the 90^{th} percentile. The prediction is that bad bosses leave and good bosses stay.[41] Results are presented using two versions of the estimated boss effects – those with the public/private boss effect

because additional information about changes over the distribution can be captured.

[39] Although this test may under-reject because of problems with measurement error regarding the quality of future bosses, using quartile indicators alleviates some of this concern. Suppose that the measurement error is independent of the true boss effect. Then if the concern is that the best bosses are assigned workers with the best residuals, it is very unlikely that measurement error is responsible for the failure to reject the null because the parameter estimates are non-monotonic. This suggests that attenuation bias is not driving these results.

[40] For theoretical models that would imply the layoff or voluntary separation of bad bosses, see Gibbons and Katz (1991), Gibbons and Waldman (1999a, b), Jovanovic (1979), and Lazear and Rosen (1981).

[41] In this firm, bosses are not demoted.

estimated to be θ=.30 and those with the public boss effect in the regression in which θ=0.

The results are in Table 6. The exponentiated coefficients imply that bosses in the bottom 10% are more than twice as likely to exit the firm as bosses outside of the bottom 10%. This is true for both specifications in which θ=.30 and θ=0.[42] 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 above the 90th percentile. These coefficients are small and are not statistically different from zero.

IX. Peer Effects

There is a growing literature on peer effects.[43] 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:

$$(17) \qquad q_{ijt} = X_{it}\beta + a_i + \delta_j + \xi\ p_{ijt} + \varepsilon_{ijt}$$

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.[44] The estimating equation for the joint model is

$$(17') \qquad q_{ijt} = X_{it}\beta + a_i + \delta_j + \xi_{Peer}\left(Team\ Size - 1\right)^{-1}\sum_{k \in jt \setminus \{i\}} a_k + \varepsilon_{ijt}$$

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.[45] Because

[42] 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.

[43] 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 Ichino, 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.

[44] It is not possible to use a mixed effects model, as estimates of the individual worker fixed effects are necessary to ξ_{Peer}.

[45] 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.

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}, the effect of increasing peers' mean innate ability on productivity, 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.[46] Note that this production environment has relatively little teamwork because each worker primarily interacts with a customer, not with other workers.[47] 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.

X. Conclusion

Supervision and management are 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 that 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

[46] 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.
[47] The same is true in Mas and Moretti (2009), who also find significant, but small peer effects.

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. There is some documented persistence of boss effects, but much of the idiosyncratic aspect of what bosses do is fleeting. Additionally, the poorer quality bosses are less likely to remain with the firm and workers who are paired with the better bosses are more likely to remain with the firm. Finally, comparative advantage does not seem to be a major factor in this firm. Sorting of the good bosses to good workers only slightly increases the firm's total output.

Of course, the results here pertain to only one firm. The magnitudes of the boss effects are found to be large in this firm, but it is not possible to say whether boss effects of this size would be found in other firms. In the course of conducting this research, we interviewed managers and workers at many firms. All emphasized the significant effects that bosses have in coaching and motivating workers. In this paper, we provide calculations that suggest that managerial productivity is roughly 75 percent greater than that of workers, which is in line with the pay premium that managers receive over workers. The prevalence of higher managerial pay in the market is consistent with the findings from this one firm.[48]

[48] See evidence in Cardiff, Lafontaine, and Shaw (2013) that managers (who are often first line supervisors) in the retail and manufacturing sectors earn from seven to twenty percent higher wages than workers.

References

Abowd, John, Robert Creecy, and Francis Kramarz. 2002. Computing person and firm effects using linked longitudinal employer-employee data. LEHD Program Technical Paper no. TP-2002-06, U.S. Census Bureau, Washington, D.C.

Abowd, John, Francis Kramarz, and Simon Woodcock. 2006. Econometric analysis of linked employer-employee data. Unpublished manuscript, Department of Economics, Cornell University.

Bennedsen, Morten, Kasper M. Nielsen, Francisco Perez-Gonzalez, and Daniel Wolfenzon. 2007. Inside the family firm: The role of families in succession decisions and performance. *Quarterly Journal of Economics* 122, no. 2: 647-691.

Bennedsen, Morten, Francisco Perez-Gonzalez, and Daniel Wolfenzon. 2007. Do CEOs matter? Working Paper No. 13-2007, Copenhagen Business School, Frederiksberg, DK.

Bertrand, Morten, and Antoinette Schoar. 2003. Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics* 118, no. 4: 1169-1208.

Borghans, Lex, Bas ter Weel, and Bruce A. Weinberg. 2008. Interpersonal styles and labor market outcomes. *Journal of Human Resources* 43, no. 4: 815-858.

Branch, Gregory, Eric Hanushek, and Steven Rivkin. 2012. Estimating the effect of leaders on public sector productivity: The case of school principals. Working Paper no. 17803, National Bureau of Economic Research, Cambridge, MA.

Cardiff, Brianna, Francine Lafontaine, and Kathryn Shaw. 2013. The spread of modern retail: Implications for wages. Unpublished manuscript, Graduate School of Business, Stanford University.

Dawson, Peter, Stephen Dobson, and Bill Gerrard. 2000. Estimating coaching efficiency in professional team sports: Evidence from English association football. *Scottish Journal of Political Economy* 47, no. 4: 399-421.

Engelberg, Joseph, Raymond Fisman, Jay Hartzell, and Christopher Parsons. 2012. Human capital and the supply of religion. Unpublished manuscript, Department of Economics, Columbia University.

Falk, Armin, and Andrea Ichino. 2006. Clean evidence on peer effects. *Journal of Labor Economics* 24, no. 1: 39-57.

Frick, Bernd, and Robert Simmons. 2008. The impact of managerial quality on organizational performance: Evidence from German soccer. *Managerial and Decision Economics* 29, no. 7: 593-600.

Garicano, Luis. 2000. Hierarchies and the organization of knowledge in production. *Journal of Political Economy* 108, no. 5: 874-904.

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, no. 1: 1-43.

Garicano, Luis, and Esteban Rossi-Hansberg. 2006. Organization and inequality in a knowledge economy. *Quarterly Journal of Economics* 121, no. 4: 1383-1435.

Gibbons, Robert and Lawrence Katz. 1991. Layoffs and lemons. *Journal of Labor Economics* 9, no. 4: 351-380.

Gibbons, Robert and Michael Waldman. 1999a. Careers in organizations: Theory and evidence. In *Handbook of Labor Economics*, ed. Orley Ashenfelter and David Card, vol. 3B, 2373-2437. Amsterdam: North-Holland.

------. 1999b. A theory of wage and promotion dynamics inside firms. *Quarterly Journal of Economics* 114, no. 4: 1321-58.

Goodall, Amanda H. 2009a. Highly cited leaders and the performance of research universities. *Research Policy* 38, no. 7: 1079-1092.

------. 2009b. *Socrates in the Boardroom: Why Research Universities Should Be Led by Top Scholars*. Princeton, New Jersey: Princeton University Press.

Hermalin, Ben E. 1998. Toward an economic theory of leadership: Leading by example. *American Economic Review* 88, no. 5: 1188-1206.

Ichniowski, Casey, and Kathryn L. Shaw. 2003. Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *The Journal of Economic Perspectives* 17, no. 1: 155-180.

Jackson, C. Kirabo. 2012. Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics* 95, no. 4: 1096-1116.

Jenter, Dirk, and Katharina Lewellen. 2010. Performance induced CEO turnover. Unpublished manuscript, Graduate School of Business, Stanford University.

Jones, Benjamin F., and Benjamin A. Olken. 2005. Do leaders matter: National leadership and growth since World War II. *Quarterly Journal of Economics* 120, no. 3: 835-864.

Jovanovic, Boyan. 1979. Job matching and the theory of turnover. *Journal of Political Economy* 87: 972-90.

Kahn, Lawrence M. 1993. Managerial quality, team success, and individual player performance in major league baseball. *Industrial & Labor Relations Review* 46, no. 3: 531-547.

Kandel, Eugene, and Edward P. Lazear. 1992. Peer pressure and partnerships. *Journal of Political Economy* 100, no. 4: 801-817.

Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen. 2012. Which CEO characteristics and abilities matter? *Journal of Finance* 67, no. 3: 973-1007.

Kuhn, Peter, and Catherine Weinberger. 2005. Leadership skills and wages. *Journal of Labor Economics* 23, no. 3: 395-436.

Lazear, Edward P. 2000. Performance pay and productivity. *American Economic Review* 90, no. 5: 1346-1361.

Lazear, Edward P. 2004. The Peter principle: A theory of decline. *Journal of Political Economy* 112, S1: S141-S163.

Lazear, Edward P. 2012. Leadership: A personnel economics approach. *Labour Economics* 19, no. 1: 92-101.

Lazear, Edward P. and Sherwin Rosen. 1981. Rank-order tournaments and optimum labor contracts. *Journal of Political Economy* 89, no. 5: 841-64.

Lazear, Edward P., and Kathryn L. Shaw. 2008. Tenure and output. *Labour Economics* 15, no. 4: 710-724.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton. 2013. Making do with less: Working harder during recessions. Working Paper no. 19328, National Bureau of Economic Research, Cambridge, MA.

Lucas, Robert E., Jr. 1978. On the size distribution of firms. *Bell Journal of Economics* 9, no. 2: 508-523.

Mas, Alexandre, and Enrico Moretti. 2009. Peers at work. *American Economic Review* 99, no. 1: 112-145.

Matsa, David, and Amalia Miller. 2012. A female style in corporate leadership? Evidence from quotas. *American Economic Journal: Applied Economics* 5, no. 3: 136-69.

Mayer, Thomas. 1960. The distribution of ability and earnings. *Review of Economics and Statistics* 42: 189-95.

Perez-Gonzalez, Francisco. 2006. Inherited control and firm performance. *American Economic Review* 96, no. 5: 1559-1588.

Perez-Gonzalez, Francisco, and Daniel Wolfenzon. 2012. Death of CEOs: Impact on firm performance. Unpublished manuscript, Graduate School of Business, Stanford University.

Porter, Phillip K., and Gerald W. Scully. 1982. Measuring managerial efficiency: The case of baseball. *Southern Economic Journal* 48, no. 3: 642-650.

Rosen, Sherwin. 1982. Authority, control, and the distribution of earnings. *Bell Journal of Economics* 13, no. 2: 311-323.

Rotemberg, Julio J., and Garth Saloner. 2000. Visionaries, managers, and strategic direction. *RAND Journal of Economics* 31, no. 4: 693-716.

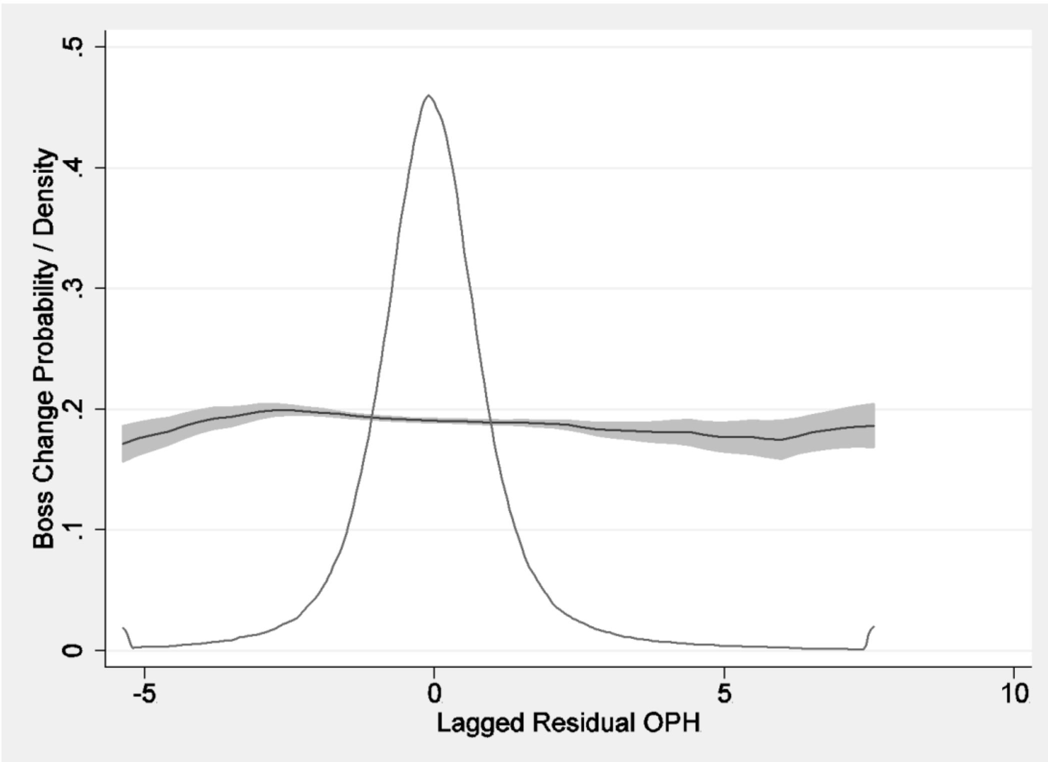
Schoar, Antoinette, and Luo Zuo. 2011. Shaped by booms and busts: How the economy impacts CEO careers and management styles. Working Paper no. 17590, National Bureau of Economic Research, Cambridge, MA.

Shaw, Kathryn L., and Edward P. Lazear. 2008. Tenure and output. *Labour Economics* 15, no. 4: 710-724.

Simon, Herbert. 1957. The compensation of executives. *Sociometry* 20, no. 1: 32-35.

Smeets, Valerie, Michael Waldman, and Frederic Warzynski. 2013. Performance, career dynamics, and span of control. Working paper no. ICS 2013-001, Institute for Compensation Studies, ILR School, Cornell University.

Figure 1. Local Polynomial Smoothed Estimates of Boss Change Probabilities as a Function of Lagged Residual Productivity



Note: The figure displays the results of a local polynomial regression of the probability of a boss change in month t on the mean productivity residual for worker i in month t-1. This residual is taken from a regression of output-per-hour on worker fixed effects, establishment fixed effects, year-by-month fixed effects, and a fifth order polynomial in tenure. Confidence intervals are given in the shaded region. The density plot shows the distribution of the lagged productivity residuals.

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
Quality	221,954	2.16	0.78	0.5	5.0
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

Note: 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. Quality is collected approximately weekly from a randomly administered survey on the service quality that a worker provides.

Table 2. Variance Components Estimates of Boss, Worker, and Match Effects

	Output-per-hour				Quality Score
	(1)	(2)	(3)	(4)	(5)
St. Dev. of Worker Effects	1.33	1.33	1.35	1.32	0.09
St. Dev. of Match Effects	0.758	0.59	0.752		0.06
St. Dev. of Boss Effects x Avg. Team Size^(1-θ)	4.74	4.08	4.104	3.44	0.78
θ	0.3	0.3	0	0	0
St. Dev. of 6 Month Lagged Boss Effects x Average Team Size^(1-θ)		1.41			
St. Dev. of 12 Month Lagged Boss Effects x Average Team Size^(1-θ)		1.06			
Lambda 6: St. Dev. of 6 month lagged boss effect / St. Dev. of Current Boss Effect		0.35			
Lambda 12: St. Dev. of 12 month lagged boss effect / St. Dev. of Current Boss Effect		0.26			
Correlation of Boss Blups from (5) and (3)					0.143
Mixed or Fixed Effects	Mixed	Mixed	Mixed	Fixed	Mixed
Theta Estimated	Yes	Constrained	Constrained	Constrained	Constrained
Number of Observations	5,729,508	4,747,015	5,729,508	5,729,508	221,954
Number of Workers	23,878	21,886	23,878	23,878	20,155
Number of Bosses	1,940	1,857	1,940	1,940	1,769

Notes: For all mixed effects models, likelihood ratio tests reject a constrained model without boss effects. All results weight the boss effects by the average team size of 9.04. In specifications with non-zero θ, the weighting multiplies the standard deviation of boss random coefficients by 9.04^(1-θ). All models contain a fifth order polynomial function of tenure and year x month dummies. To be included in the sample for column 2, a worker either had to have less than 6 months of tenure or the boss 6 months previously must have been observed. For workers with greater than 1 year of tenure, the boss 12 months previously must have been observed. Notes for particular specifications follow. Column (1): Mixed effects estimates of equation (13) in the text. Θ is estimated in a two-step procedure with search in an outer loop over θ. The inner loop uses the lme4 package in R to estimate the model. For all mixed models, lme4 is used. Column (2): Mixed effects estimates of equation (12) in the text. Θ is constrained to the estimate from column 1. Match effects lagged 6 months and 12 months were also included in the specification. The standard deviations of these match effects are 0.62 and 0.61, respectively. The correlation between current and 6 month and 12 month lagged boss BLUPS is 0.10 and 0.16, respectively. Column (3): Mixed effects estimates of equation (14) in the text. Θ is constrained to be 0. Column(4): Fixed effects estimates of equation (14) in the text. Θ is constrained to be 0. The standard deviation of worker and boss fixed effects is calculated based on weighting by observations in the sample. A conjugate gradient algorithm is used to recover the parameters. Column (5): Mixed effects estimates of equation (14) in the text, replacing the dependent variable with a post-transaction survey of service quality on a 0-5 scale.

Table 3. Estimates of Worker Attrition from Cox Proportional Hazard Models

Sample	Workers with α below the mean				Full Sample			
	θ=0.30	θ=0.30	θ=0	θ=0	θ=0.30	θ=0.30	θ=0	θ=0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Boss Effect BLUP x	-0.111	-0.099	-0.180	-0.164	-0.123	-0.117	-0.197	-0.189
Team Size ^ - θ	0.895	0.906	0.084	0.085	0.884	0.890	0.821	0.828
	(0.020)	(0.022)	(0.041)	(0.045)	(0.014)	(0.014)	(0.029)	(0.030)
Dummy for worker in bottom		0.187		0.192		0.134		0.139
10% of distribution of α		1.205		1.212		1.143		1.149
		(0.027)		(0.027)		(0.026)		(0.026)
Interaction of boss effect		-0.043		-0.027		-0.032		-0.011
and bottom 10% worker		0.958		0.974		0.969		0.990
		(0.033)		(0.083)		(0.030)		(0.078)
Number of Observations	3,012,703	3,012,703	3,012,703	3,012,703	5,573,233	5,573,233	5,573,233	5,573,233

Notes: Estimates of Cox proportional hazard models of worker attrition. Each cell contains the coefficient, the exponentiated coefficient below and the standard error in parentheses. All specifications contain establishment and year x month fixed effects. The boss effects in columns 1, 2, 5 and 6 are the BLUPS from column 1 of Table 2. The boss effects in the remaining columns are from column 3 of Table 2. Sample sizes differ from prior tables because the last month of data is excluded. This restriction is necessary to characterize exits separately from the end of the sample. Results are qualitatively similar after dropping left censored workers.

Table 4. Analysis of Match Effects by Boss and Worker Quality Cells

Panel A: Match Effects from Column 1, Table 2 (θ=0.3)

		Boss: Good	Boss: Bad
Worker	Star	0.100	-0.083
	Laggard	0.05	-0.063

Panel B Match Effects from Column 3, Table 2 (θ=0)

		Boss: Good	Boss: Bad
Worker	Star	0.104	-0.077
	Laggard	0.04	-0.066

Notes: Good bosses and star workers are those above the median of the distribution of boss and worker effects, respectively. The cells contain the means of the estimated match effects.

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

	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?	
Dependent Variable:	Oph	Oph	Mean residual by worker from Column (1)	Mean residual by worker from Column (2)	Boss effect from experienced sample (θ= 0)	Boss effect from experienced sample (θ = 0.3)
Sample:	New workers' on their 1st boss assignment		New workers' on 1st boss assignment		New workers' on 2nd boss assignment	
	(1)	(2)	(3)	(4)	(5)	(6)
Boss BLUP estimated from experienced sample (θ = 0)	0.8008 (0.0814)					
Boss random coeff. estimated from experienced sample x Team Size ^ -0.3 (θ = 0.3)		0.7163 (0.0669)				
Dummy for boss in the bottom 25% of Blups (column 3) or random coefficients (column 4)			-0.1039 (0.0620)	-0.0560 (0.0698)		
Dummy for boss in 25%-50%			-0.0382 (0.0582)	-0.0895 (0.0667)		
Dummy for boss in 50%-75%			-0.1115 (0.0618)	-0.0404 (0.0618)		
Mean residual by worker from Column (1)					0.0036 (0.0016)	
Mean residual by worker from Column (2)						0.0109 (0.0038)
Wald statistic that quartile differences are zero			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 are clustered by boss. Wald tests are also 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 removes 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, boss effects for this sample are computed 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). The individual boss BLUPs (or random coefficients) are then recovered via the method in the text. Second, the boss quality measures from the experienced sample are merged with 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 fixed effects.

Table 6. Estimates of Boss Attrition from Cox Proportional Hazard Model

	θ=0.30	θ=0
	(1)	(2)
Dummy for Boss Effect Below 10th Percentile	1.000	0.591
	2.720	1.805
	(0.090)	(0.090)
Dummy for Boss Effect Above 90th Percentile	0.040	0.083
	1.040	1.086
	(0.100)	(0.100)
N	620,130	620,130

Notes: Each cell contains the coefficient, the exponentiated coefficient below and the standard error in parentheses. The boss effects in column 1 are from estimation of equation 13 in column 1 of Table 2. The boss effects in column 2 are from the estimation of equation 14 in column 5 of Table 2. Sample sizes differ from Table 1 because the last month of data is excluded to be able to characterize boss exits. Results are qualitatively similar after dropping left censored bosses and in specifications with and without left censored bosses.

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	0.0004	0.009
Standard Deviation of Boss Effects (Weighted by worker-days)	2.85	3.44
Standard Deviation of Worker Effects (Weighted by worker-days)	1.65	1.32
Number of Workers	1,679	23,878
Number of Bosses	155	1,940
Number of Observations	391,730	5,729,508

Note: 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.