

# Supervisors and Performance Management Systems

## Preliminary and Incomplete - Please do not circulate

Anders Frederiksen, Lisa Kahn, and Fabian Lange

Aarhus University, Yale School of Management, and McGill University\*

October 14, 2015

### Abstract

Supervisors occupy central roles in production and performance monitoring in a firm. We study how supervisor heterogeneity in performance evaluations affects career and firm outcomes using data on a 360 degree performance system of a Scandinavian service sector firm. We find a large amount of heterogeneity in performance ratings associated with supervisors. We write down a principal-agent model where supervisor heterogeneity can come in the form of real differences in the ability to elicit output from subordinates or from differences in a taste for leniency when rating subordinates and firms can be noisily informed about this heterogeneity. Within the context of this model, we can investigate the nature of supervisor heterogeneity and the degree to which firms are informed about this heterogeneity by relating supervisor heterogeneity in ratings to subordinate pay and pay for performance, objective performance measures, and supervisor pay. We find that worker pay and promotions are positively affected by a supervisor's propensity to rate highly, but that objective output is not related to the supervisors rating behavior. These findings, in conjunction with other supporting evidence, suggest that This suggests that supervisor heterogeneity is primarily driven by differences in leniency, about which the firm is relatively uninformed. Our research is the first to document the important variation in ratings across supervisors and to show that it is positively related to worker outcomes

---

\*Anders Frederiksen, Aarhus University, Department of Business Development and Technology, Birk Center Park 15, 7400 Herning, Denmark, Lisa Kahn Yale School of Management, PO Box 208200, New Haven, CT, 06520, Fabian Lange, McGill University, Department of Economics, Leacock Building, 855 Sherbrooke Street West,

# 1 Introduction

Modern cooperations increasingly install performance management systems that call upon supervisors to rate the performance of those subordinate to them. Supervisor ratings are used to determine who gets promoted, to allocate workers to tasks, and to compensate employees. Clearly though, supervisors are not passive instruments that provide unbiased measures of employee performance. Rather, the ratings of supervisors are biased and these biases that might or might not be correlated with differences across supervisors in how they manage their employees. Plausibly, supervisors differ in their managerial ability and better supervisors make their team members more productive. Part of these differences in managerial style might well be reflected in the rating behavior of supervisors. The willingness to rate employees highly might be part and parcel of a management style that results in higher productivity. Supervisors might also, irrespective of their management style, differ in how willing they are to report negatively on their team members. Some supervisors will be more inclined than others to rate their team members favorably rather than to accurately report on their performance. These supervisors will have subordinates who appear to perform much better in terms of subjective performance reviews, but may not perform better in terms of actual performance. Ratings in modern performance systems are thus likely affected in important ways by idiosyncratic factors governing the behavior of supervisors. And, these differences in rating behavior might or might not be correlated with differences in the ability of supervisors to manage teams to perform more highly.

Despite the important role supervisors play in performance management, there is little solid evidence on the nature of the heterogeneity in rating behavior across supervisors. The literature provides no answers on crucial questions about the nature of the heterogeneity. To begin with, how much heterogeneity across supervisors is there in ratings behavior? How important is this heterogeneity for determining employee's compensation and careers? Do supervisor ratings of individual performance differ because supervisors differ in their ability to raise worker productivity? Or do these ratings differ for reasons unrelated to worker performance?

The literature provides likewise little guidance on how ratings heterogeneity across supervisors affect the firms and how they use and design performance management systems? Does supervisor heterogeneity limit the usefulness of performance systems and counteract organizational goals? How do firms respond to supervisor heterogeneity when they design performance systems? Do firms succeed in removing biases introduced by idiosyncratic differences in reporting behavior? Answers to all of these questions are lacking.

In this paper we study the nature and importance of supervisor heterogeneity using data from a performance management system of a large Scandinavian service sector firm. These data include a 10-year panel (2004-2014) of employees including their salaries (including salary components), job level and function, and

surveys of job satisfaction. Important for our analysis is that we have access to measures of the financial and operational performance of branches within this firm that can serve as objective measures of performance at the team level. Crucial for our analysis is that the data also contains the records of the performance system including various performance ratings of individual employees and the identities of the supervisors that reported on the employees.

Below, we use a variety of empirical approaches to demonstrate that there is substantial heterogeneity in ratings across supervisors. And, we relate this heterogeneity to career outcomes of supervisors and subordinates as well as branch performance.

To interpret the data, we develop a simple behavioral model of behavior of supervisors, subordinates, and the firm. In particular, we follow a long tradition in personnel economics dating to at least Holmstrom (1979) and postulate that the central human resource challenge facing the firm is to incentivize workers to exert effort. We model this problem as a static problem - largely because of the challenges involved in expanding the analysis to a dynamic setting. The three actors in our model are the workers without supervisory function, the supervisors, and the firm. Workers choose to exert effort, but the chosen effort level is hidden from the firm and the supervisors. Supervisors observe worker output and report on this output to the firm. However, supervisors do not report truthfully on worker performance - rather they face a trade-off between reporting truthfully and reporting favorably about their team members. In our model, supervisors differ along two dimensions. First, they differ in how much weight they place on reporting truthfully as opposed to favorably. Second, they differ in their managerial ability, which we model as differences in the marginal costs of exerting effort on the part of their subordinates. Given this set-up, we consider the optimal linear compensation contracts of workers as well as salary contracts for supervisors. Our model is set up in a way that allows us to ask how the optimal contracts would depend on how informed firms are about the differences between supervisors.

We believe that our data broadly supports an interpretation according to which (a) the heterogeneity in supervisor rating behavior is largely associated with leniency bias as opposed to managerial effectiveness and (b) the firm is uninformed about the differences in reporting behavior across supervisors. We favor this interpretation due to a number of crucial findings. The first main piece of evidence supporting our interpretation of supervisor heterogeneity as reflecting leniency is that that branches managed by “high raters” (i.e., supervisors who give high ratings to their employees, on average) do not achieve higher scores on a number of objective Key Performance Indicators (KPI). If high raters also tended to manage more effectively, then we would expect the branches they manage to outperform those of other managers. Our data does not support this. Furthermore, the compensation and career outcomes of supervisors are only weakly associated with their proclivity to rate their subordinates highly. If managers differed primarily in their

managerial effectiveness and firms were informed about these differences, then we would expect supervisors compensation to correlate highly with their rating behavior. The data however shows that career outcomes of supervisors and their compensation are unrelated to their tendency to rate their subordinates highly.

What feature of the data support the interpretation that firms are uninformed about the heterogeneity in ratings behavior across supervisors? Wages of employees as well as their chances of being promoted are strongly associated with the ratings style of their supervisors. Those working for “high raters” tend to be paid higher wages and they tend to be promoted more rapidly. An informed firm would strive to eliminate such differences in outcomes associated with supervisors as they do not reflect differences in individual skill or effort. We also find that workers tend to be more satisfied with their jobs and tend to initiate fewer lateral moves if matched to a high rater, again suggesting economic rents associated with being matched to high raters. Informed firms would eliminate these rents. Thus, on balance, we interpret the heterogeneity in ratings as reflecting leniency biases about which the firm is imperfectly informed.<sup>1</sup>

In summary, our paper develops an analytic framework that allows investigating the nature of heterogeneity in ratings behavior across supervisors and the extend to which firms are informed about this heterogeneity. We show how to operationalize this framework using data from the Performance System of a Scandinavian service sector firm. Our results suggest that supervisors differ from each other in their tendency to report positively as opposed to accurately on their workers performance and that firms find it difficult to undo these biases when setting pay and determining promotions.

The remainder of the unfinished and incomplete paper proceeds as follows. We introduce the firm and the data at our disposal in the following Section (Section 2). Section 3 then describes the model and its implications for how career outcomes and performance are related to supervisor heterogeneity. Section 4 presents the empirical analysis. This section is the most incomplete as we are currently relying heavily on OLS regressions that we know face a number of empirical problems. Currently, we are engaged in developing and estimating specifications that address these problems. We expect those results to be ready within a few weeks time. Due to the preliminary state of the paper, Sections 5 and 6 are left empty. Section 5 will, in due course, contain a discussion and reduced form investigation of a number of hypothesis related to the dynamics of performance measurement. In particular we are planning to estimate the long run consequences of being assigned to a “high raters” and also whether there is evidence that the firm is learning about the heterogeneity of their supervisors. Section 6 will conclude.

---

<sup>1</sup>It is of course also possible that the firm ignores these biases for reasons outside of our model.

## 2 Firm and Data

### 2.1 Firm Overview

We rely on personnel data from a large Scandinavian service sector firm.<sup>2</sup> Our sample comprises all employees engaged in domestic activities between 2004 and 2014. In our data we have 22,688 unique employees with a total of 136,286 employee-year observations. Table 1 provides summary statistics for the full sample (all reported monetary figures are in 2010 US dollars). On average, workers earn 78,139 US dollars with a standard deviation of 51,487. In the data it is possible to distinguish between base pay and annual bonus. Roughly 30 percent of the workers receive a bonus and the bonus pool is close to 20 percent of the wage pool. In our sample, 51.9 percent are women, the average age is 43.5 years and tenure is 17.5 years. We also observe that 83.6 percent are working full time.

The firm is divided into an extensive branch network and a central corporate office (see Figure 1). The branches comprise 44 percent of workers. Across branches jobs are comparable and involve close client contact. Workers in the central corporate office have a variety of functions and there are more high level jobs (level 11). These differences are reflected in the compensation structures across branches and corporate functions in that both average compensation and the variance in compensation is higher in the corporate functions. In 2013, there were 269 branches and the median branch had 15 employees. The typical branch had a branch manager (level 9), a deputy branch manager (level 7), 5-7 senior workers in client-facing roles (levels 6), and 5-7 junior workers in client-facing roles (levels 4-5) and some times a trainee (level 1). Because of the qualitative differences between branch and corporate jobs in this firm and because of the availability of objective (financial and performance based) branch-level performance measures we will in part of our analysis pay particular attention to the branch network.

Just prior to the period covered by our data, the firm developed a performance management system. Each worker receives a rating that is meant to describe their aggregate performance. It ranges from 1 (unsatisfactory) to 5 (outstanding). In 2004, when our data begins, the system was still being rolled out, and 42 percent of the employees received performance ratings. In the following years, the system continued to spread so that by 2008 the system covered almost 82 percent of the employees and the coverage stayed at that level or slightly above throughout the remainder of the sample period.<sup>3</sup> In the branch network, ratings are typically given by the branch manager, but we also observe that deputy branch managers rate employees. In corporate functions, employees are typically rated by the worker with the highest job level within a given

---

<sup>2</sup>The firm is a market leader within the domestic market. It also has some international activities, but we focus on the domestic workforce here.

<sup>3</sup>There is no systematic variation in who gets rated when we look at full-time vs. part-time employees, corporate vs. branch employees or across job levels. Hence, we are not worried about any systematic reasons for missing ratings.

function. Overall the typical manager is rating 10 employees as the average span of control in the firm is 9.76 (s.d. 10.16).

The distribution of performance scores is shown in table 2. The lowest rating of 1 is rarely given and only 3 percent receive the second lowest rating of 2. The clear majority receive the ratings 3 and 4, with more than 50 percent of all employees receiving a 3. Only 5.7 percent of employees are rewarded the highest rating of 5. This range of ratings, as well as the effective range (of 3 to 5) is common among subjective performance systems, as shown in Frederiksen, Lange, and Kriechel (2015). So, because most ratings are either a 3 or a 4, we lose little information by using a “pass-fail” performance metric, which equals 1 if the rating is 4 or 5 and zero otherwise. The “pass-fail” performance metric makes it easy to interpret linear regression coefficients in that they represent marginal effects on the probability of receiving a “passing grade”. For these reasons, we will build our empirical investigation around this pass-fail metric. <sup>4</sup>

In addition to supervisor ratings we have obtained access to employee job satisfaction surveys for the years 2004 to 2010. These surveys contain information about the employees’ perceptions of supervisors’ performances. This information is elicited through 7 questions: 1) The professional skills of my immediate superior, 2) The leadership skills of my immediate superior, 3) My immediate superior is energetic and effective, 4) My immediate superior gives constructive feedback on my work, 5) My immediate superior delegates responsibility and authority so I can complete my work effectively, 6) My immediate superior helps me to develop personally and professionally, and 7) What my immediate superior says is consistent with what he/she does. These questions are answered on a 10-point scale and we use the average across the seven questions as the employees overall assessment of his/hers immediate superior. The minimum score is 1 (low) and the maximum score is 10 (high). On average employees rate their supervisors at 8.164 with a standard deviation of 1.373.

It is important to stress the uniqueness of the survey data. While employee satisfaction data at the individual level is made available in supplements to databases such as the the National Longitudinal Survey of Youth (NLSY), the German Socio-Economic Panel (GSEP), and the British Household Panel Survey (BHPS) such data is unavailable to managers in companies. The reason is that if the individual scores are made available to the company (and the immediate manager) it will violate the employees’ anonymity, which (most likely) will result in response bias (see the discussion in Frederiksen 2015). For this reason, employee satisfaction surveys are typically conducted by consulting companies outside the firm and the results are reported back to the firm as averages at the branch/business unit level. As researchers we have been able to obtain the survey data at the individual level and to merge it to the personnel records. Hence, we know how a given employee evaluates his/hers superior.

---

<sup>4</sup>Results when using the entire scale are available upon request and qualitatively and quantitatively very similar.

Our data also contains two measures of branch performance. The first measure of performance is based on Key Performance Indicators (KPIs). The first measures how the branches rank on a set of KPIs relative to their peer branches. The Key Performance Indicators (KPIs) include measures of financial performance of the branches, as well as other metrics (for example, customer satisfaction). The set of KPI changes from year to year as the firm’s focus evolves. Branches are placed into peer groups based on size and customer base, and these peer groups vary from year-to-year. The average peer group has 17 branches. These branch rankings, which we hereafter term “KPI rankings”, are available from 2007-2010. The second measure of branch level performance reflects the branches’ financial development between January in year  $t$  and  $t+1$ . We have successfully obtained this information for the year 2013 (i.e. the development in performance between jan. 2013 and jan. 2014). This measure is constructed such that a score of 100 implies no change in financial performance between the two years. A score of 110 implies a 10 percent improvement. Among the 160 branches for which we have financial performance information the average score is 102.6.

## 2.2 Variation in Performance Measures

There is substantial systematic variation in the incidence of passing grades across supervisors and workers. To illustrate this variation, we estimate the following regressions

$$p_{ist} = \alpha'_i + \phi'_s + \beta' X_{it} + \gamma' Y_{st} + \epsilon_{ist} \quad (1)$$

For an individual,  $i$ , being supervised by manager,  $s$ , at time  $t$ ,  $p$  is whether or not the worker “passed” his or her performance review. We allow  $p$  to be a function of a vector of worker effects ( $\alpha$ ), a vector of supervisor effects ( $\phi$ ), a vector of time-varying worker controls ( $X$ ), a vector of time-varying supervisor controls ( $Y$ ), and a time-varying error term.<sup>5</sup>

Estimating this regression requires the assignment of employees to supervisors to vary substantially over time. In our data, employees typically change supervisors repeatedly. Similarly, supervisors manage many different employees over time, with some employees joining or leaving their teams almost every year. Over the period 2004 to 2014 the average employee had 2.94 different supervisors (s.d. of 1.71). If we look at the employees who were with the firm throughout this period we find that they on average had 4.31 different supervisors (s.d of 1.59). The average supervisor manages 9.76 (s.d. of 10.16) employees in a given year. On average, supervisors manage a total of 21.48 different employees (s.d. of 37.28) while they are recorded

---

<sup>5</sup>The worker controls ( $X$ ) include cubics in age and tenure, and indicators for full-time status, gender, job level; supervisor controls ( $Y$ ) include a cubic in age of the supervisor and gender and job level of the supervisor dummies. We also control for business unit indicators (whether the worker is in a branch or the specific function in headquarters), and year fixed effects. The latter help control for differences in usage of performance ratings as they become more common in the firm.

as supervisors in our data. Those individuals who were supervisors throughout the entire sample period on average managed a total of 50.18 different employees.

Table 3 summarizes the variation in ratings explained by controls, worker fixed effects, and supervisor fixed effects. This table thus informs us about how much of the overall variation in ratings can be attributed to workers and supervisor identities respectively. A regression of just controls accounts for 15% of variation in the probability of passing. Worker fixed effects alone can account for 42% of the remaining variation in performance, while supervisor effects alone account for 7%. Combined, the worker and supervisor fixed effects account for 50% of the, and each are highly significant.

Thus both supervisors and workers contribute substantially to variation in performance ratings. However, interpreting the magnitudes of this variation is difficult because of notorious measurement error problems with large fixed effects regressions. To accommodate these issues we take a method of moments approach to correct for this sampling error and allow it to be correlated across workers and supervisors.

We estimate that supervisor's time-invariant propensity to pass his or her subordinates ( $\hat{\phi}_s$ ), holding constant time-invariant quality of workers and controls has a standard deviation of 0.28. That is, moving to a supervisor with a one standard deviation higher fixed effect increases a worker's probability of passing by an average of nearly 28 percentage points (ppt) of a base of roughly 45%. Remarkably this is true holding constant a worker's time-invariant propensity to pass, ( $\hat{\alpha}_i$ ). Being a one standard deviation more productive worker increases the probability of passing in a given period by 40 ppts. Finally, there is substantial period-to-period variation in performance. The residuals from the full regression have a standard deviation of 45 ppts.

The substantial variation in supervisor effects, ( $\hat{\phi}_s$ ), is the primary concern of this paper. In the next section, we develop a model that offers two possible sources of supervisor heterogeneity that can account for the systematic variation in performance, and empirical predictions to separate the two.

### 3 Model<sup>6</sup>

Supervisors play a crucial functional role in performance management systems. Besides managing and supervising teams, they provide subjective reports on the performance of their subordinates. Firms rely on these reports to set pay and to determine promotions. Naturally, supervisors may differ in both their rating behavior and in their ability to manage employees. Firms then face the problem of how to design performance management systems in the face of this heterogeneity. We analyze how heterogeneity in

---

<sup>6</sup>As we lay-out the model, we will focus on its implications and the intuitions embodied in it without presenting derivations in detail. Many results follow immediately from known results in the literature and need not be rederived here. A somewhat more formal treatment of the arguments is provided in the appendix.



managerial ability and supervisor rating behavior affects data generated by performance systems under different assumptions on how well informed firms are about the heterogeneity across supervisors.

Let the marginal product of an employee not in a supervisory role (a “worker”),  $i$ , at time,  $t$ , be  $q_{i,t}$ . As expressed in equation 2, we assume that this marginal product (“output”) depends on effort ( $e_{i,t}$ ) which is not directly observed by her supervisor or by the firm. Worker productivity also depends on the productive type  $\alpha_i$ , and random time-varying luck  $\varepsilon_{i,t}^q$ , distributed normally with mean 0, variance  $\sigma_q^2$  and independent of all other variables.

$$q_{i,t} = e_{i,t} + \alpha_i + \varepsilon_{it}^q \quad (2)$$

Supervisors observe  $q_{i,t}$ .<sup>7</sup>

We next specify utility functions for workers (v) and supervisors (u). Workers choose how much effort  $e$  to exert. Supervisors choose what to report to the firm about the performance of the worker. These choices are affected by two forms of supervisor heterogeneity: (a) heterogeneity in managerial ability, which impacts the worker’s cost of effort ( $\mu_s$ ), and (b) heterogeneity in rating behavior, whereby supervisors differ in their willingness to tradeoff a truthful rating with a more generous one ( $\beta_s$ ).

The timing of the model is as follows:

1. Workers and firms sign contracts that specify the type of supervisors individuals are matched with and the wage function. This wage function specifies how compensation depends on the supervisor characteristics and the rating that a worker receives.
2. Workers are matched to supervisors, make effort choices and production takes place.
3. Supervisors report ratings to the firm .
4. Workers are paid the wage corresponding to the reported rating that is specified in the wage contract.

Workers have Constant Absolute Risk Aversion (CARA) preferences over effort  $e$  and wages  $w$ :<sup>89</sup>

$$v(w, e) = -exp\left(-\psi\left(w - \frac{1}{2\mu_s}e^2\right)\right) \quad (3)$$

---

<sup>7</sup>During the remainder of this paper, we suppress individual and time subscripts unless required for understanding. However, we do subscript variables that vary across supervisors with  $s$ .

<sup>8</sup>During the remainder of this paper, we suppress individual and time subscripts unless required for understanding. However, we do subscript variables that vary across supervisors with  $s$ . These variables generally are constant over time.

<sup>9</sup>The functional form assumptions embodied in equation (3) keep the problem tractable. By assuming CARA, we abstract from income effects that might otherwise affect the trade-off between effort and risk. Quadratic effort costs result in linear first order conditions for effort and thus results in closed form solutions. Below, we make assumptions that ensure that wages are normally distributed conditional on worker choices and information. Combined with the exponential form in (3) this allows exploiting known results on expectations of log-normally distributed random variables (deGroot reference).

These preferences depend on the managerial ability of a worker’s supervisor in that the marginal cost of effort decline in when a worker is assigned to a good supervisors. We parametrize this idea using  $\mu_s$  so that better supervisors have higher  $\mu_s$ . Workers choose effort to maximize eq. (3) taking their supervisor type and the wage contract as given. We will specify the relation between wages and output below. For now, it is important to note that the optimal effort choice will depend on  $\mu_s$  and will thus vary systematically across supervisors. All else equal, workers for better supervisors will exert more effort.

We also allow for heterogeneity in reporting behavior. Some supervisors are more lenient or nicer than others, but all prefer to truthfully report on the performance of employees. This implies supervisors trade off the conflicting goals of being lenient and reporting accurately when they file reports,  $r$ , on their employee’s productivity. We embed this trade-off in supervisor preferences:

$$u(w_s, q, r) = w_s + \tilde{\beta}_s r - \frac{\tilde{\gamma}_s}{2} (r - q)^2 \quad (4)$$

Here the parameters  $(\tilde{\beta}_s, \tilde{\gamma}_s)$  allow for heterogeneity across supervisors in how they trade off leniency against accuracy. Supervisors choose a report,  $r$ , by maximizing their utility conditional on worker output,  $q$ . The result, shown in equation (5), is that supervisor reports  $r$  sum the observed output  $q$  and  $\beta_s = \frac{\tilde{\beta}_s}{\tilde{\gamma}_s}$ . The supervisor specific parameter  $\beta_s$  measures the strength of the motive to report favorably relative to the motive to report truthfully. We will call this parameter the “supervisor bias”.

$$r = q + \frac{\tilde{\beta}_s}{\tilde{\gamma}_s} = q + \beta_s. \quad (5)$$

Substituting (2) in (5) and denoting by  $e_s$  the equilibrium effort level that team members of the supervisor  $s$  exert, we get:

$$r = \alpha_i + (e_s + \beta_s) + \varepsilon_{it}^q = \alpha_i + \phi_s + \varepsilon_{it}^q \quad (6)$$

Variation in ratings attributable to the supervisor is summarized by  $\phi_s$ . As discussed above, this variation can arise either because supervisors differ in their managerial quality  $\mu_s$  or because they differ in their bias  $\beta_s$ .<sup>10</sup>

We now consider the contracts firms and workers would enter into in a static set-up under different assumptions on what is known to them. That is, we postulate that firms and workers have common, though

<sup>10</sup> Assuming  $r$  is directly observable in a long enough panel and also assuming that there is sufficient mobility of workers and supervisors, we could obtain estimates of  $\alpha_i$  and  $\phi_s$  by estimating a specification that relates ratings to worker and supervisor fixed effects. Data problems preclude this direct approach and force us to account for the ordinal nature of supervisor ratings, but ultimately we pursue a similar approach: we use variation in ratings within a worker independent of the supervisor to estimate  $\alpha_i$  and variation in ratings within a supervisor independent of the worker to estimate  $\phi_s$ . We will discuss the estimation approach in more detail below. As we proceed to analyze the model and its empirical predictions, we assume for now that  $\phi_s$  and  $\alpha_i$  can be measured.

possibly imperfect, information on  $\{\mu_s, \beta_s\}$  when they write contracts. These contracts specify all payoff relevant aspects of the employment relationship. Thus, they consist of assignments  $\{\mu_s, \beta_s\}$  as well as a mapping of observed ratings to wages.

We make a number of assumptions to keep the analysis tractable. First, as is common in the literature, we restrict attention to wage contracts for workers that are linear in the ratings that supervisors report. The parameters of these wage contracts are allowed to vary with each individual and supervisor assignment. Thus, we consider contracts of the form  $w_{i,s,t} = a_{i,s} + b_{i,s}r_{it}$ .<sup>11</sup> In addition, we make the necessary assumptions that ensure that the exponent in equation (3) is normally distributed conditional on the available information *both* at the contracting stage and when individuals decide on effort. This allows us to use well-known results on the expectation of log normal random variables (deGroot (1970)) to represent worker preferences using the certainty equivalent. That is, we can express the participation constraints as

$$E\left[w - \frac{1}{2\mu_s}e^{*2}|I_C\right] - \frac{1}{2}\psi var\left(w - \frac{1}{2\mu_s}e^{*2}|I_C\right) \geq \underline{u}(\alpha) \quad (7)$$

where  $I_C$  represents the information available during the contracting stage and  $e^*$  is the optimal effort level chosen by the worker.<sup>12</sup> Workers observe  $\mu_s$  when choosing this effort and face a linear wage. The solution of maximizing equation (3) subject to the linear contract delivers the optimal effort choice  $e^*$ :

$$e^* = b_{s,i}\mu_s \quad (8)$$

We now solve for the optimal terms  $(a_{i,s}, b_{i,s})$  of the wage contract. The solution to the contracting problem depends on what firms and workers know about supervisors. We begin by assuming that firms and workers are perfectly informed about the supervisors types.

### 3.1 The Informed Firm and the Performance Management System

We now consider workers and firms that are perfectly informed about the supervisors and the workers types  $\mu_s, \beta_s$  and  $\alpha_i$ . Firms offer workers both an assignment to a supervisor with characteristics  $(\mu_s, \beta_s)$  and a wage contract that maps observed signals  $r$  onto wages. The terms of the wage contract are allowed to vary with  $I_C = \{\mu_s, \beta_s, \alpha_i\}$ .<sup>13</sup> Thus, wage contracts are:

<sup>11</sup>In a closely related setting with normal signals and with preferences of the type provided here Holmstrom and Milgrom [1987] find that the optimal contract does take the linear form. We suspect but have not proven that our setting could be specialized further to map into Holmstrom and Milgrom [1987] and that linear contracts are therefore at least conceivably optimal. For now, I think that exercise is besides the point.

<sup>12</sup>The outside opportunity  $\underline{u}(\alpha)$  depends on the productive type of the worker, since firms compete for workers and are symmetrically informed about the type of workers.

<sup>13</sup>In Section @.@, we consider firms that are imperfectly informed about supervisor heterogeneities  $(\mu_s, \beta_s)$ .

$$w = a(\mu_s, \beta_s, \alpha) + b(\mu_s, \beta_s, \alpha) r$$

Substituting the optimal effort  $e^*$  from eq. (8) into the certainty equivalent (7) and simplifying, we obtain the participation constraint:

$$a + b(\alpha + \beta_s) + \frac{1}{2}b^2\mu_s - \frac{\psi}{2}b^2\sigma_q^2 \geq \underline{u}(\alpha) \quad (9)$$

It is straightforward to show that the piece-rate  $b$  maximizes the sum of the expected profit and the certainty equivalent subject to workers choosing effort optimally (eq.8).<sup>14</sup> Thus, the optimal piece rate solves

$$b_s^* = \underset{\{b\}}{\operatorname{argmax}} \left\{ \alpha + b\mu_s - \frac{b^2}{2} (\mu_s + \psi\sigma_q^2) \right\} \quad (10)$$

This results in the standard solution familiar from the literature:

$$b_s^* = \frac{\mu_s}{\mu_s + \psi\sigma_q^2} \quad (11)$$

Because firms compete for workers, they will make zero profits for any worker-supervisor pair. This zero profit condition amounts to

$$\alpha + b\mu_s - a - b(\alpha + \beta_s + b\mu_s) - w_s(\mu_s, \beta_s) = 0 \quad (12)$$

where  $w_s(\mu_s, \beta_s)$  is the wage paid to a supervisor with characteristics  $(\mu_s, \beta_s)$ .

Consider now how the compensation of employees and supervisor varies with  $(\alpha, \beta_s, \mu_s)$ .

We begin with  $\beta_s$ . The optimal piece-rate (11) does not depend on the generosity of the supervisor  $\beta_s$ . The reason is that the firm extracts the entire surplus from workers using base compensation - workers with more generous supervisor will simply see their base compensation reduced. Similarly, the effort choice  $e^*$  does not vary with  $\beta_s$ .

Rearranging the certainty equivalent in eq (7) to isolate expected compensation we have  $E[w|I_C] = \underline{u}(\alpha) + \frac{1}{2\mu_s}e^{*2} + \frac{1}{2}\psi\operatorname{var}(w|I_C)$ . The loading  $b_s$  and thus the variable part of compensation does not depend on  $\beta_s$ . Furthermore,  $e^*$  does not vary with  $\beta_s$ .<sup>15</sup> Thus, the two terms on the right hand side are independent of  $\beta_s$ . This implies that the expected compensation of employees will not vary with  $\beta_s$  since the firm extracts

<sup>14</sup>For this, set up the profit maximization of the firm subject to the Participation constraint. The first order condition with respect to the intercept can be used to show that the Lagrange multiplier on the participation constraint equals 1, from which the statement in the text follows.

<sup>15</sup>See eq. (8).

the entire surplus using the intercept of the wage contract.

Continuing with  $\alpha_i$ , we note that competition ensures that the expected compensation of workers increases one-for-one with  $\alpha_i$ . It is obvious that competition will not permit for wage differences across supervisors with  $\alpha_i$ .

Consider now  $\mu_s$ . From equation (11), we have that the optimal loading increases in  $\mu_s$ . To determine the effect on average compensation, consider the certainty equivalent restated here substituting the expected wage of an employee:

$$E[w|\alpha, \mu_s, \beta_s] - \frac{1}{2\mu_s}e^2 - \frac{\psi}{2}b^2\sigma_q^2$$

Since the entire surplus is extracted from workers and all workers earn the outside option, we obtain

$$\frac{d\left(E[w|\alpha, \mu_s, \beta_s] - \frac{1}{2\mu_s}e^2 - \frac{\psi}{2}b^2\sigma_q^2\right)}{d\mu_s} = 0$$

Workers maximize the certainty equivalent by choice of  $e$ . We can thus apply the envelope condition and ignore any variation in effort in response to variation in  $\mu_s$ . However, as  $\mu_s$  varies, so will the optimally chosen piece-rate  $b$  (see eq. 11).<sup>16</sup> Thus, we obtain

$$\begin{aligned} \frac{d(E[w|\alpha, \mu_s, \beta_s])}{d\mu_s} &= \frac{\partial\left(\frac{1}{2\mu_s}e^2\right)}{\partial\mu_s} + \frac{\partial\left(\frac{\psi}{2}b^2\sigma_q^2\right)}{\partial b} \frac{\partial b}{\partial\mu_s} = -\frac{1}{2\mu_s^2}e^2 + \psi\sigma_q^2 b \frac{\partial b}{\partial\mu_s} \\ &= -\frac{1}{2}b^2 + b\left(\frac{\psi\sigma_q^2}{\mu_s + \psi\sigma_q^2}\right)^2 = -\frac{1}{2}b^2 + b(1-b)^2 \\ \Rightarrow \text{sign}\left(\frac{d(E[w|\alpha, \mu_s, \beta_s])}{d\mu_s}\right) &= \text{sign}\left(-\frac{1}{2}b^2 + b(1-b)^2\right) \end{aligned}$$

This expression cannot generally be signed. When incentives are low-powered ( $b < \frac{1}{2}$ ), total pay increases in  $\mu_s$ , while the opposite is true when incentives are high-powered ( $b > \frac{1}{2}$ ). Workers with better managers face lower costs of providing any given effort level. This effect tends to lower compensation of workers with better managers. At the same time, when managers are better the optimal piece rate increases and the risk borne by workers increases. Wages thus need to increase to compensate workers for the increase in risk this implies. When incentives are high ( $b > \frac{1}{2}$ ), much effort is provided. Since the effort cost function is convex, this implies that the marginal costs of providing additional effort is also high. Thus, better managers reduce the efforts costs born by workers significantly when incentives are high. Therefore wages for workers with better managers decline if incentives are high. When incentives are low, effort provision is low, and convex effort costs imply the savings from a small reduction in effort costs will also be low. Thus, pay increases with

<sup>16</sup>The piece rate is not chosen to maximize the certainty equivalent, so no envelope condition applies here.

$\mu_s$  when incentives are low ( $b < \frac{1}{2}$ ) because workers need to be compensated for the extra risk they bear. By contrast, better managers reduce effort costs by more when incentives are high ( $b > \frac{1}{2}$ ) and workers exert a lot of effort.

Regarding the compensation of the supervisor, note that the surplus generated by any supervisor-worker match increases in  $\mu_s$ . As firms compete for supervisors, any differences in the surplus across  $\mu_s$  are paid to the supervisor. Thus the compensation of the supervisor increases in her managerial ability:  $\frac{\partial w_s(\mu_s)}{\partial \mu_s} > 0$ .

We have so far considered the problem of how wages depend on supervisor and worker heterogeneity without considering the problem of assigning workers to supervisors. Since worker type  $\alpha$  enters additively in the production function and does not affect the risk-effort trade-off as summarized by worker preferences (3), we have no predictions for how  $\alpha$  and  $(\mu_s, \beta_s)$  are assigned to each other. Both positive and negative assortative matching are thus entirely consistent with this set-up.

To summarize, we have that

1. The optimal piece-rate  $b(\mu_s, \beta_s, \alpha)$  is determined by equation (11). It is independent of  $(\alpha, \beta_s)$  and increases in  $\mu_s$ .
2. The average compensation received by employees increases one-for-one in  $\alpha$  and is independent of supervisor generosity  $\beta_s$ . It is not possible to sign the relation between average compensation of employees and  $\mu_s$ .
3. Expected output  $E[q|\mu_s, \beta_s, \alpha]$  increases in  $\mu_s$  and  $\alpha$  and is independent of  $\beta_s$ .
4. Earnings of supervisors  $w_s(\mu_s, \beta_s)$  are independent of  $\beta_s$  and increase in  $\mu_s$ .

This provides a set of predictions that can be tested in our empirical framework. In Section 4, we will exploit the firm level data to test these implications. To do so, we will first use the panel on worker performance to estimate how performance ratings workers receive depend on worker and supervisor fixed effects. Under a set of (strong) assumptions, the estimated worker and supervisor fixed effects in performance equations will converge to  $\alpha$  and  $(e(\mu_s) + \beta_s)$  respectively. We can then use the estimated worker and supervisor fixed effects together with the compensation variables and the observed objective measures of team output to test the predictions of the model summarized by points 1-4 above.

### 3.2 The Partially Informed Firm and the Performance Management System

So far we assumed that  $(\mu_s, \beta_s)$  are known to the firm. Next, we analyze contracts when firms and workers are only partially informed about supervisor types. We continue to assume that the only information asymmetry

in the model is about the hidden effort  $e$ . Thus, we assume that supervisors and employees share the same information about  $(\mu_s, \beta_s)$  during the contracting stage. We proceed in much the same fashion as when analyzing the problem faced by the informed firm.

To begin, assume that  $(\mu_s, \beta_s)$  are independent normally distributed random variables with variances  $\sigma_\beta^2$  and  $\sigma_\mu^2$ . Firms and employees hold beliefs  $(\beta_s^E, \mu_s^E)$  about the supervisor characteristics such that

$$\begin{aligned}\beta_s &= \beta_s^E + \varepsilon_\beta \\ \mu_s &= \mu_s^E + \varepsilon_\mu\end{aligned}$$

Let the expectation errors  $(\varepsilon_\beta, \varepsilon_\mu)$  also follow a normal distribution and be independent of each other. We parametrize the share of total variation in  $\beta$  and  $\mu$  unknown to firms as  $\theta_\beta$  and  $\theta_\mu$  so that

$$\begin{aligned}\sigma_\beta^2 &= \text{var}(\beta_s^E) + \text{var}(\varepsilon_\beta) = (1 - \theta_\beta) \sigma_\beta^2 + \theta_\beta \sigma_\beta^2 \\ \sigma_\mu^2 &= \text{var}(\mu_s^E) + \text{var}(\varepsilon_\mu) = (1 - \theta_\mu) \sigma_\mu^2 + \theta_\mu \sigma_\mu^2\end{aligned}$$

During the contracting stage, the marginal cost of effort is not known to anybody. However, employees observe the marginal cost of effort after having been assigned to a supervisor and before they decide on their optimal effort level. The optimal level of effort conditional on the piece rate  $b$  is obtained in the same manner as before (see eq. 8):  $e^* = b\mu_s$ .

During the contracting stage, the parties share information on  $(\mu_s^E, \beta_s^E)$ . A work contract consists of an assignment of a worker  $\alpha_i$  to a supervisor with  $(\mu_s^E, \beta_s^E)$  and a wage contract that depends on  $(\mu_s^E, \beta_s^E, \alpha)$ :  $w(r; \mu_s^E, \beta_s^E, \alpha) = a(\mu_s^E, \beta_s^E, \alpha) + b(\mu_s^E, \beta_s^E, \alpha)r$ .

As before, we can use the employee's certainty equivalent to write the participation constraint:

$$a + b(\alpha_i + \beta_s^E) + b^2 \frac{\mu_s^E}{2} - \frac{\psi}{2} \left( b^2 (\theta_\beta \sigma_\beta^2 + \sigma_q^2) + \frac{b^4}{4} \theta_\mu \sigma_\mu^2 \right) \geq \underline{u}(\alpha) \quad (13)$$

The firm problem is still to maximize profits from any given worker-supervisor pair:<sup>17</sup>

$$\Pi(\mu_s^E, \beta_s^E, \alpha) = \underset{\{a, b\}}{\text{Max}} \{ \alpha + b\mu_s^E - a - b(\alpha + \beta_s^E + b\mu_s^E) - w_s(\beta_s^E, \mu_s^E) \} \quad (14)$$

s.t. the participation constraint (13).

And, as before, competition in the labor market for workers and supervisors will ensure that expected

---

<sup>17</sup>We have already imposed the optimal effort choice  $e = b\mu_s^E$ .

profits conditional on  $(\alpha, \beta_s^E, \mu_s^E)$  will equal zero.

### Wage contracts between partially informed firms and employees

The optimal loading is implicitly determined by the FOC of eq. 14:

$$\mu_s^E = b \left( \mu_s^E + \psi \left( \theta_\beta \sigma_\beta^2 + \sigma_q^2 + b^2 \frac{\theta_\mu \sigma_\mu^2}{2} \right) \right) \quad (15)$$

The RHS of this expression increases monotonically in  $b$  and there is thus a unique loading that solves the firms problem.

It is instructive to compare (15) with the optimal loading of the informed firm:  $b = \frac{\mu_s}{\mu_s + \psi \sigma_q^2}$  stated in eq. (11). Besides replacing  $\mu_s^E$  with  $\mu_s$ , there are two differences. First, the signal becomes less informative as the share of the variation in  $\beta_s$  that is unknown to the firm increases. Thus, the optimal loading declines in  $\theta_\beta \sigma_\beta^2$ . Second,  $\theta_\mu \sigma_\mu^2$  measures differences in managerial ability that are unobserved by both workers and the employer during the contracting stage. However, once the worker has been assigned to a supervisor she observes the marginal cost of effort  $\mu_s$  associated with this supervisors. At that point, she will exploit this additional information and will “game” the performance system in the sense of supplying disproportionately more effort in low marginal cost stages than in high marginal cost states of the world. Therefore, the usefulness of setting incentives using performance signals declines in  $\theta_\mu \sigma_\mu^2$  and so does the optimal loading.

As before, firms extract any surplus from workers during the contracting stage. Again, expected compensation will be independent of  $\beta_s^E$  since  $\beta_s^E$  only enters the workers certainty equivalent through the expected wage. And, as before, we have that competition for employees implies that productive differences across employees are paid to workers so that we have

$$a = a_0(\mu_s^E) + \alpha - b\beta_s^E \quad (16)$$

It is again not possible to sign the relationship between average employee compensation and  $\mu_s^E$ . Furthermore, as before, we have that expected output net of the wage for the employee is independent of  $\beta_s^E$  and increases in  $\mu_s^E$ . Thus, earnings of the supervisor are independent of  $\beta_s^E$  and increases in  $\mu_s^E$ . Thus, we have the following results that are analogous to those stated at the end of Section 2.1:

1. The optimal piece rate  $b(\mu_s^E, \beta_s^E, \alpha)$  is independent of  $(\beta_s^E, \alpha)$  and increases in  $\mu_s^E$ .
2. Expected compensation increases one-for-one in  $\alpha$  and is independent of  $\beta_s^E$ . It is not possible to sign the relationship between expected compensation of the employee and  $\mu_s^E$ .
3. Expected output  $E[q|\mu_s^E, \beta_s^E, \alpha]$  increases in  $\mu_s^E$  and  $\alpha$  and is independent of  $\beta_s^E$ .



4. Earnings of supervisors  $w_s (\mu_s^E, \beta_s^E)$  are independent of  $\beta_s^E$  and increase in  $\mu_s^E$ .

These results mirror those in the previous section. We also have an additional result on the relation between the piece rate and the unobserved variation in supervisor heterogeneity.

5. The optimal piece rate declines in  $\theta_\beta \sigma_\beta^2$  and  $\theta_\mu \sigma_\mu^2$ .

Besides these results, we can ask how employee and supervisor salaries as well as output depend on those components not observed by the firm. This question is empirically of interest because we have access to a panel of ratings and pay. We thus have an information advantage relative to the firm when it is setting pay. Furthermore, it is conceivable that firms do not use the available data optimally. Firms might therefore act as if they are uninformed about  $(\beta_s, \mu_s)$  even though they might have inferred  $(\beta_s, \mu_s)$  from the available data.

Thus, consider what predictions are obtain for how wages of an employee vary with  $(\beta_s, \mu_s, \beta_s^E, \mu_s^E)$ :

$$\begin{aligned} w(\beta_s, \mu_s, \beta_s^E, \mu_s^E, \alpha) &= a_0(\mu_s^E) + \alpha - b\beta_s^E + b(\beta_s + b\mu_s) \\ &= a_0(\mu_s^E) + \alpha_i + b\varepsilon_\beta + b^2\mu_s = a_0(\mu_s^E) + \alpha + b\theta_\beta\beta_s + b^2\mu_s + b\varepsilon_\beta \end{aligned}$$

where we substitute the linear projection of  $\varepsilon_\beta = \frac{\text{cov}(\varepsilon_\beta, \beta_s)}{\text{var}(\beta_s)}\beta_s + \epsilon_\beta = \frac{\text{cov}(\varepsilon_\beta, \beta_s^E + \varepsilon_\beta)}{\text{var}(\beta_s)}\beta_s + \epsilon_\beta = \theta_\beta\beta_s + \epsilon_\beta$ .

And, we have that a workers output is given by

$$q = b\mu_s + \alpha + \varepsilon^q$$

These two equations show how expected output and wages vary with  $(\beta_s, \mu_s, \beta_s^E, \mu_s^E)$  in the partially informed firm:

1. Expected compensation increases in  $\beta_s$ , where the coefficient on  $\beta_s$  is given by the product of the optimal piece-rate multiplied by the proportion of the variation of supervisor heterogeneity that is unknown to the firm.
2. Output does not vary with  $\beta_s$ , but does vary with  $\mu_s$ .

### 3.3 A 2-by-2 Matrix to Distinguish Types of Heterogeneity and How Informed the Firm is

Above we analyzed a structure that allows for different assumptions of how supervisors differ from each other and how informed the firm is about the types of supervisors employed. Supervisors could differ in

their ability to manage their employees as well as in their bias. And, firms could differ in how informed they are about the differences between supervisors. Depending on the assumptions made, we obtain different predictions that we can test in the firm data available to us.

At this point, we find it useful to consider extreme assumptions on the source of heterogeneity and the information available to firms in order to build intuition about how the fundamentals of the model map into the data on ratings, compensation, and output. In particular, we will consider the situation where firms are perfectly informed ( $\theta_\beta = \theta_\mu = 0$ ) or completely ignorant ( $\theta_\beta = \theta_\mu = 1$ ). And, we will distinguish the case when supervisors differ primarily in how lenient they are ( $\sigma_\beta^2 > 0, \sigma_\mu^2 = 0$ ) from the case when supervisors differ primarily in their ability to elicit effort from their team members ( $\sigma_\beta^2 = 0, \sigma_\mu^2 > 0$ ). Combining, we obtain 4 different sets of assumptions on how supervisors differ from each other and how informed the firm is.

Recall, empirically we will strive to measure the heterogeneity  $\phi_s$  in ratings associated with supervisors using the panel of performance ratings and the supervisor identifiers included in the data. We will then related worker and supervisor compensation as well as a measure of expected productivity of workers in a given team to  $\phi_s$ . The table below summarizes what implications these four different sets of assumptions have for worker and supervisor compensation and expected productivity of a worker.

Information \ Heterogeneity		Leniency ( $\sigma_\beta^2 > 0, \sigma_\mu^2 = 0$ )	Effectiveness ( $\sigma_\beta^2 = 0, \sigma_\mu^2 > 0$ )
Fully Informed Firms ( $\theta_\mu = \theta_\beta = 0$ )	Wages: $\frac{\partial \mathbf{E}[\mathbf{w} \phi_s]}{\partial \phi}$	0	$\neq 0^*$
	Piece rate: $\frac{\partial \mathbf{b}}{\partial \phi}$	0	$> 0$
	Productivity: $\frac{\partial \mathbf{E}[\mathbf{q} \phi_s]}{\partial \phi}$	0	$> 0$
	Supervisor Wages: $\frac{\partial \mathbf{w}}{\partial \phi}$	0	$> 0$
Uninformed Firms ( $\theta_\mu = \theta_\beta = 1$ )	Wages: $\frac{\partial \mathbf{E}[\mathbf{w} \phi_s]}{\partial \phi}$	$> 0$	$> 0$
	Piece rate: $\frac{\partial \mathbf{b}}{\partial \phi}$	0	0
	Productivity: $\frac{\partial \mathbf{E}[\mathbf{q} \phi_s]}{\partial \phi}$	0	$> 0$
	Supervisor Wages: $\frac{\partial \mathbf{w}}{\partial \phi}$	0	0

*\*The model does not make a clear prediction about the relationship between employee wages and  $\phi_s$ .*

The above table reveals that the four different set of assumptions can indeed be distinguished.

It is intuitive that informed firms will undo any differences between supervisors in how lenient they are. Thus, wages of workers and supervisor, productivity and piece rates will not vary with  $\phi_s$  if it reflects only differences in leniency. By contrast, the informed firm will be very responsive to differences in the managerial effectiveness of supervisors. Thus, supervisor wages, piece rates, productivity and potentially average em-

ployee compensation will vary with effectiveness of the supervisor when firms are well informed. Assuming that firms are perfectly informed, we can thus determine whether supervisors differ primarily in leniency or in managerial effectiveness by testing whether supervisor and employee compensation, productivity, and piece rates co-move with  $\phi_s$ .

By contrast, if firms are uninformed, then the piece rates and the wages of supervisors will not vary across supervisors, regardless of why supervisors differ from each other (leniency or effectiveness). However, if firms are uninformed, we will find that employee wages will vary with  $\phi_s$ , regardless whether it reflects leniency or managerial effectiveness. However, if the firm is uninformed, then expected productivity will only vary with  $\phi_s$  if it indeed represents differences in managerial effectiveness  $\mu_s$ .

Inspection of the above table reveals that observing how employee compensation varies with  $\phi_s$  is particularly important to distinguish informed from uninformed firms if the main source of heterogeneity across supervisors is how lenient they are toward their team members. In uninformed firms, such variation increases average compensation of workers since the firm can not undo this variation. The informed firm by contrast will simply undo this source of variation. Similarly, observing how productivity varies with  $\phi_s$  is necessary to distinguish between heterogeneity in leniency  $\beta_s$  and effectiveness  $\mu_s$  if firms are uninformed.

Overall, we have developed a structure in this Section that allows for two fundamentally distinct interpretations of supervisor heterogeneity. We can distinguish between these sources of heterogeneity and can also empirically test how well informed the firm is about the supervisor heterogeneity within this structure.

## 4 Testing the Model

The previous sections analyzed the implications of heterogeneity across managers in: ability ( $\mu_s$ ) and leniency ( $\beta_s$ ). Able managers are those that lower the effort costs of their supervisors, while lenient managers are those with a greater innate desire to give high ratings as opposed to reporting the truth. Both types of heterogeneity will be positively correlated with performance ratings. The analysis then showed how one can distinguish between the source of heterogeneity ( $\mu_s$  or  $\beta_s$ ) as well as the amount of information held by the firm by exploring how total pay, pay-for-performance, actual productivity, and supervisor pay vary with the empirical heterogeneity in ratings behavior observed across supervisors.

Central to our empirical analysis is identifying the heterogeneity in ratings behavior observed across supervisors. From equation 6 we have defined this heterogeneity as  $\phi_s = e_s(\mu_s) + \beta_s$ . Table 4 summarized what the model implies about how  $\phi_s$  relates to (1) total pay of subordinates, (2) the pay for performance component of subordinate pay, (3) actual productivity, and (4) total pay of supervisors. We take two

approaches for testing these predictions. First, we directly estimate  $\phi_s$  by regressing performance (“passing”) on worker fixed effects, supervisor fixed effects, and controls, specified above in equation 1. Recall that this regression produces  $\hat{\alpha}_i$ ,  $\hat{\phi}_s$ , and  $\hat{\epsilon}_{ist}$ , which represent time-invariant worker propensity to pass his or her performance review, time-invariant supervisor propensity to pass his or her subordinates, and an idiosyncratic shock impacting the probability of passing this period. We then project these estimates onto the four outcomes of interest in a second stage regression.

However, this method suffers from a number of econometric problems. Of primary concern is that  $\hat{\alpha}_i$ ,  $\hat{\phi}_s$ , and  $\hat{\epsilon}_{ist}$  are generated regressors. Since the panel is relatively short, they are estimated with sampling error. Furthermore, the sampling error in these estimates will be correlated. In particular, in a regression of performance on worker and supervisor fixed effects, the sampling error in the fixed effects of workers and supervisors in the same team are likely to be negatively correlated. Below we therefore also propose a method of moments approach that is not subject to this problem of small sample error that plagues the regression based empirical approach.

We describe both of these methods and how they are applied to the specific outcome variable of interest in each of the following subsections.

#### 4.1 Supervisor effects and total earnings

To test the first comparative static, we estimate the following regression:

$$\log(\text{earnings})_{ist} = \beta_0 + \beta_1 \hat{\alpha}_i + \beta_2 \hat{\phi}_s + \beta_3 \hat{\epsilon}_{ist} + \beta' X_{it} + \gamma' Y_{st} + \nu_{ist} \quad (17)$$

We regress the log of earnings (which includes both base pay and bonus) on the three components of performance, estimated above, as well as the same control variables. The coefficient,  $\beta_2$ , informs us on how having a supervisor with a higher likelihood of “passing” subordinates impacts a worker’s pay, holding constant the worker’s own average performance and the worker’s idiosyncratic performance this period. Note that  $\hat{\alpha}_i$  and  $\hat{\epsilon}_{ist}$  are important to control for since total pay is increasing in both. They also help to control for factors outside the model, such as nonrandom sorting of workers to supervisors, and correlated shocks.<sup>18</sup>

Recall, from above that if managers differ primarily in how lenient they are and firms are fully informed then we will find no systematic relationship between  $\hat{\phi}_s$  and total pay. If instead firms are imperfectly informed about supervisor leniency, then compensation will on average increase in the leniency of the su-

---

<sup>18</sup>Standard errors are obtained allowing  $\nu$  to cluster by supervisor to take into account that time-varying observations may be correlated within supervisor.

pervisors. If managers differ primarily in their managerial ability ( $\mu_s$ ), rather than in their leniency, and firms are informed, then employee compensation will be affected by supervisor heterogeneity, but we can not predict whether average compensation increases or decreases in  $\phi_s$ . If in contrast, firms are completely uninformed about variation in  $\mu_s$ , then higher  $\mu_s$  will result in higher average output and compensation.

Table 5 presents regression results of equation 17 for a number of subsamples. Column 1, is the full sample, and we see that all components of performance significantly impact pay. The coefficient on  $\hat{\phi}_s$  can be interpreted as the impact on earnings for moving from a supervisor who never passes subordinates to one who always passes, a 5% pay increase. A more natural interpretation would be the impact on pay for moving to a supervisor whose probability of passing is one standard deviation higher (0.28), or a 1.4% pay increase. By comparison, impacts of worker effects and residual performance are all the same order of magnitude. For example, a worker with a one standard deviation higher propensity to pass earns on average almost 3% more, while a one standard deviation higher idiosyncratic shock produces a 1% earnings gain.

For robustness, table 5 includes additional regressions based on a number of subsamples. One problem with the results reported in column 1 is that they potentially suffer from correlated sampling error in the two regression stages. To evaluate the importance of this issue we split the sample, using the first four years to identify the fixed effects and pass residuals and the second part to establish the relationship between the performance components and earnings. The results are reported in column 2. The results are very similar to those reported in column one. Below we will further address the issue of correlated sampling error using a method of moments strategy. Reassuringly this approach produces qualitatively similar results.

In table 5, column 3 restricts attention to employees working in the branches, where work activities are more homogenous. Column 4 further restricts the sample to branches with KPI rankings (only years 2010-2014), and column 5 excludes the years 2010 and 2013 where bonus information is unavailable. The results are similar across all specifications. Hence, supervisor effects are positively correlated with earnings. This finding is inconsistent with firms being fully informed about managerial heterogeneity that is driven primarily by leniency, but consistent with the other 3 quadrants of table 4.

## 4.2 Supervisor effects and pay for performance

The second comparative static relates  $\hat{\phi}_s$  to the piece rate,  $b$ . Recall from table 4 that if the firm is informed and manager heterogeneity stems primarily from managerial ability ( $\mu_s$ ), then pay for performance should be increasing in the supervisor effect. That is, when supervisors reduce the marginal cost of effort, subordinates should have stronger pay for performance incentives. However, if firms are uninformed about managerial ability or if managerial heterogeneity stems primarily from leniency, there will be no relationship.

To estimate how the strength of the piece rate varies, we estimate regressions of the following form:

$$outcome_{ist} = \beta_0 + \beta_1 \hat{\alpha}_i + \beta_2 \hat{\phi}_s + \beta_3 \hat{\epsilon}_{ist} + \beta_4 [\hat{\phi}_s * p_{ist}] + \nu_{ist}$$

Outcome variables are total pay, the probability of receiving a bonus and the size of the bonus conditional on receiving one. The coefficient on the interaction effect,  $\beta_4$ , identifies whether pay fluctuates more for supervisors who give higher ratings.

Table 6 summarizes these regression results. Column 1 repeats column 1 from table 5, showing the effect of performance components on earnings. In column 2, we add the interaction between supervisor effects and whether or not the worker passed.<sup>19</sup> The coefficient of interest,  $\beta_4$  is negative and significant. Hence, with high confidence we can rule out a positive effect. In columns 3 and 4 the relationships with  $\log(\text{bonus})$  is explored. The  $\log(\text{bonus})$  results show that bonuses increase in the three performance components, but bonuses are unaffected by the interaction term between supervisor FE and the pass variable. The results presented in columns 5 and 6 reveal that the probability of receiving a bonus is also positively related to the three performance components, and, as for  $\log(\text{earnings})$ , the interaction term is negative and significant. Based on these results we conclude that pay for performance is not increasing in supervisor fixed effects; a result which is inconsistent with firms being fully informed and manager heterogeneity stemming from ability differences

### 4.3 Supervisor effects and productivity

The third comparative static relates  $\hat{\phi}_s$  to objective productivity. We have access to a correlate of objective productivity for a subset of our data in the form of the KPI ranking. These rankings are at the branch level from 2007-2010. They represent the ranking of the branch in a given period, relative to a set of peers. We estimate regressions of the following form at the branch-time level.

$$\begin{aligned} KPIoutcome_{bt} &= \beta_0 + \beta_1 \overline{p_{ist}} + \beta' \bar{X}_{it} + \gamma' \bar{Y}_{st} + \nu_{bt} \\ KPIoutcome_{bt} &= \beta_0 + \beta_1 \bar{\alpha} + \beta_2 \bar{\phi}_s + \beta_3 \bar{\epsilon}_{ist} + \beta' \bar{X}_{it} + \gamma' \bar{Y}_{st} + \nu_{bt} \end{aligned} \tag{18}$$

$KPIoutcome$  is a summary measure of the relative ranking of branch,  $b$ , in year,  $t$ . As mentioned above, we experiment with a number of measures to ensure our results are not sensitive to functional form. The first regression relates the KPI ranking to the average pass rate among all workers in the branch-year ( $\overline{p_{ist}}$ ), controlling for average worker- and supervisor- characteristics in the branch-year ( $\bar{X}_{it}$  and  $\bar{Y}_{st}$ , respectively).

<sup>19</sup>We obtain similar results throughout when we instead include an interaction of the supervisor fixed effect and the pass residual ( $\epsilon$ ).

This regressions helps build intuition for both the subjective and objective performance measures, by estimating their correlation. Regressions are weighted by the number of workers in the branch  $b$ , in year  $t$  with non-missing pay and performance data.

Panel A of table 7 summarizes these results. For all measures, subjective performance and KPI ranking are positively correlated. When an entire branch moves from failing to passing, the inverse rank score increases by 0.124, or by roughly two ranking spots (column 1); the probability of being the top branch in the peer group increases by 9 ppts (column 2); the probability of being in the top 5 branches increases by 22 ppts (column 3); the probability of being in the top half of the branches in the peer group increases by 19 ppts. All these correlations are highly significant. This is thus reassuring that subjective performance and objective performance are designed to pick up the same thing.<sup>20</sup>

The second regression equation is a direct test of our model. The explanatory variables are branch-time averages of the typical variables used.  $\bar{\alpha}$  is the average propensity to pass among all workers in the branch in that year;  $\bar{\phi}_s$  is the average propensity of supervisors to pass employees among supervisors employed at the branch in that year;  $\bar{\epsilon}_{i,t}$  is the average idiosyncratic component of performance among all workers in the branch in that year. As noted above, there is typically only one supervisor giving ratings at the branch level and in those cases,  $\bar{\phi}_s$  is the supervisor fixed effect for that supervisor. In cases where there is more than one rater,  $\bar{\phi}_s$  is the average fixed effect across raters, weighted by the number of subordinates they rated this period.

The model predicts that if leniency ( $\beta_s$ ) is the primary driver of supervisor heterogeneity then objective performance will be unrelated to the supervisor effect ( $\beta_2 = 0$ ) because then supervisors do not influence actual productivity. If instead manager ability ( $\mu_s$ ) is the primary driver of supervisor heterogeneity then we should see a positive relationship between objective performance and supervisor effects ( $\beta_2 > 0$ ). This is because worker effort is increasing in  $\mu_s$  and is true for both informed and uninformed firms.

Panel B of table 7 shows the relationship between KPI ranking and the components of performance. We find that the primary driver of the relationship between subjective and objective performance measures is via the idiosyncratic shock to performance ( $\epsilon$ ) and not through the worker or supervisor effects. For each KPI variable this is the only coefficient that is statistically significant and it is also an order of magnitude larger than the other coefficients.

Standard errors are fairly large so that we cannot rule out even a fairly sizable positive relationship between supervisor effects and objective performance. However, the evidence so far on this dimension supports leniency, rather than ability, as being the primary driver of the supervisor effects.

---

<sup>20</sup>We have also explored the relationship between branch performance and performance of the highest-ranking person in the branch and obtained similar results.

#### 4.4 Supervisor effects and supervisor pay

The fourth comparative static relates  $\hat{\phi}_s$  to supervisor pay. We estimate the following regression:

$$pay_{st} = \beta_0 + \beta_1 \bar{\epsilon}_{it} + \beta_2 \hat{\phi}_s + \beta_3 \bar{\epsilon}_{ist} + \beta' \bar{X}_{it} + \gamma' Y_{st} + \nu_{bt}$$

We regress pay measures for a supervisor,  $s$ , in year,  $t$ , on that supervisor's propensity to pass subordinates ( $\hat{\phi}_s$ ). We control for average quality of workers in the pool that year, the average idiosyncratic performance of subordinates this period, average time-varying characteristics of subordinates and time-varying supervisor characteristics. Standard errors are clustered at the supervisor level, the level of variation underlying the key explanatory variable. Observations are weighted by the number of subordinates to a given supervisor in time  $t$ .

If supervisor heterogeneity is driven primarily by ability, and firms are informed about this ability, then supervisors should capture it in their pay ( $\beta_2 > 0$ ). If firms are uninformed, or if supervisor heterogeneity is primarily driven by leniency, then high-rating supervisors will not be paid differently compared to low-rating supervisors ( $\beta_2 = 0$ ).

Results are reported in table 8. We find that supervisors earnings are positively influenced by average team quality and the average idiosyncratic performance of the team, but is not affected by the supervisor effect. This is the case both for  $\log(\text{earnings})$  and for  $\log(\text{bonuses})$ . Hence, our results are in line with leniency and an uninformed firm as the explanation for supervisor heterogeneity.

#### 4.5 Discussion

To summarize we find that for supervisors who give higher ratings, holding constant the ability of their workers:

1. their workers receive more in total earnings
2. their workers do not have higher pay for performance components
3. their branch-level objective performance is not higher
4. they themselves receive marginally higher pay.

Returning to table 4, our model would rationalize this set of results as primarily supporting the bottom-left quadrant: manager heterogeneity is driven by leniency which the firm is uninformed about. When workers receive higher ratings, the firm does not know whether true performance or leniency is driving these ratings so they reward them with higher pay and do not adjust the strength of incentives. But since differences



performance ratings are driven largely by leniency and not actual ability, these differences will not show up in objective measures of productivity. We do find small positive impacts on pay of the supervisors themselves. These effects are only marginal and could again reflect the fact that firms are uninformed. <sup>21</sup>

## 5 Reduced Form Exploration: Dynamics

[NOT DONE YET]

## 6 Conclusion

[NOT DONE YET]

## 7 Appendix: Estimating Heterogeneity in Ratings and Wage Function Using Method of Moments

@NOT DONE YET.

## References

B. Holmstrom and P. Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, 55:303–328, 1987.

---

<sup>21</sup>Though not shown, we also find that managers who rate their subordinates more highly receive higher ratings by their own supervisors.

Table 1: Summary Statistics

	N	Mean	St Dev
<i>Outcomes:</i>			
Earnings	136286	78139	51487
Wages	136286	73964	31497
Bonuses	40113	14186	55379
Bonus received	136286	0.294	0.456
Wage growth in pct.	112992	0.018	0.089
<i>Controls:</i>			
Full-time	136286	0.836	0.370
Tenure	136286	17,463	13,454
Age	136286	43,504	11,263
In Branch	136286	0.445	0.497
Woman	136286	0.519	0.500
Supervisor Female	136286	0.278	0.448
Supervisor age	136286	44,173	10,358
Supervisor tenure	136286	19,639	11,788

Table 2: Performance Distribution

Fail			Pass	
1	2	3	4	5
0.12%	3.07%	51.21%	39.89%	5.71%
54.40%			45.60%	

Table 3: Variation in Probability of Passing

Specification	R <sup>2</sup>	Std. Error	F-stat	Degrees of Freedom
Worker Effects ( $\alpha$ )	0.42	0.0012	3.24	18.050
Supervisor Effects ( $\phi$ )	0.07	0.0014	3,573	2.857
Worker & Supervisor Effects	0.48		worker = 3.07 supervisor = 2.69	worker = 18,050 supervisor = 2,821

Note: Probability of passing is first residualized on time-varying worker characteristics (cubics in age and tenure, indicators for full-time status, gender, and job level), time-varying supervisor characteristics (cubic in age of the supervisor, gender and job level) indicators for business unit (whether the worker is in a branch or the specific function in headquarters), and year fixed effects. We then regress the residuals on worker and/or supervisor effects to obtain R-squares and F-stats.

Table 5: Impact of Performance Rating Components on Log(Earnings)

	mean = 11.2				
	(1)	(2)	(3)	(4)	(5)
Supervisor FE ( $\phi$ )	0.050*** (0.007)	0.036*** (0.011)	0.037*** (0.007)	0.048*** (0.012)	0.048*** (0.008)
Worker FE ( $\alpha$ )	0.079*** (0.003)	0.060*** (0.004)	0.076*** (0.003)	0.067*** (0.005)	0.085*** (0.003)
Pass Residual ( $\epsilon$ )	0.027*** (0.001)	0.035*** (0.003)	0.021*** (0.002)	0.020*** (0.003)	0.031*** (0.001)
Observations	97.299	31.144	43.103	12.606	76,338
Partial R-squared	0.036	0.024	0.055	0.042	0.040
All	X				
Split sample		X			
Branch subsample			X		
Branches with KPI's				X	
Years with bonus <sup>^</sup>					X

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

<sup>^</sup>bonus data not available in 2010 and 2013

Notes: Regressors are obtained from a first stage regression of the probability of passing a performance review on supervisor fixed effects, worker fixed effects, and controls (see table 3). We then regress log earnings on the fixed effects and the residuals obtained from this regression along with the full set of controls. Standard errors clustered by supervisor in parentheses. In column 4, fixed effects are obtained from the first four years of data. We use these fixed effects to generate an alternative pass residual then use all three variables in an earnings regression on the next four years of data.

Table 6: Supervisor Effects and Pay for Performance

	Log(Earnings) mean = 11.2		Log(Bonus) mean = 8.35		Pr(Bonus>0) mean = 0.35	
	(1)	(2)	(3)	(4)	(5)	(6)
Supervisor FE ( $\varphi$ )	0.050*** (0.007)	0.071*** (0.008)	0.310*** (0.0554)	0.378*** (0.0742)	0.176*** (0.0183)	0.215*** (0.0199)
Worker FE ( $\alpha$ )	0.079*** (0.003)	0.079*** (0.003)	0.458*** (0.0209)	0.455*** (0.0208)	0.258*** (0.00680)	0.259*** (0.00680)
Pass Residual ( $\epsilon$ )	0.027*** (0.001)	0.027*** (0.013)	0.237*** (0.0133)	0.233*** (0.0135)	0.120*** (0.0053)	0.120*** (0.0053)
$\varphi$ *Pass		-0.045*** (0.012)		-0.0993 (0.0782)		-0.0760*** (0.0242)
Observations	97,299	97,299	27,058	27,058	76,338	76,338
Partial R-squared	0.036	0.037	0.053	0.053	0.049	0.049

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Notes: See table 5.

Table 7: Supervisor Effects and Branch-Level Productivity (KPI's)

Dependent Variable: (mean)	(1) Inverse Rank Score (0.53)	(2) Pr(Top) (0.06)	(3) Pr(Top 5) (0.30)	(4) Pr(Top half) (0.46)	(5) Financial performance (102.9)
Panel A: Average Performance					
Average Pass Rate	0.124*** (0.046)	0.086** (0.039)	0.218*** (0.075)	0.193** (0.081)	1,766 (2.047)
Panel B: Average Components of Performance					
Branch-Level Average:					
Supervisor FE ( $\phi$ )	0.028 (0.081)	0.025 (0.068)	-0.001 (0.130)	0.102 (0.141)	2,281 (3.306)
Worker FE ( $\alpha$ )	0.017 (0.079)	0.083 (0.067)	0.0252 (0.127)	-0.003 (0.138)	0.449 (3.497)
Pass Residual ( $\varepsilon$ )	0.242*** (0.075)	0.118* (0.064)	0.429*** (0.121)	0.392*** (0.131)	2,194 (3.227)
Observations	766	766	766	766	156

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Observations are at the branch-year level, weighted by number of workers with non-missing pay and performance variables. Inverse rank score is -1 times the branche's KPI ranking in that year divided by the number of branches it is ranked against.

Table 8: Supervisor Effects and Supervisor Pay

Dependent Variable: (mean)	Log(earnings)	Log(bonus)
	(11.7)	(9.2)
	(1)	(2)
Supervisor FE ( $\varphi$ )	0.0112 (0.0191)	0.0588 (0.0704)
Team-Average Worker FE	0.103*** (0.0201)	0.307*** (0.0663)
Team-Ave Pass Resid	0.0248*** (0.00885)	0.0782 (0.0485)
Observations	9,444	6,816
Partial R <sup>2</sup>	0.018	0.007

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$