Biases in Subjective Performance Evaluation

Daiji Kawaguchi

Hitotsubashi University, RIETI and IZA

Hideo Owan

The University of Tokyo and RIETI

Kazuteru Takahashi

The University of Tokyo

Very preliminary, do not quote.

December 15, 2014

Abstract

There are two different kinds of biases in subjective evaluation. First, a supervisor may give preferential rating to those in the same social category as himself/herself while giving discriminatory rating to those in other social categories. Second, supervisors may give more attenuated evaluation ratings to those in different social categories because they have less information for judging their subordinates' performance. Using a panel of personnel records from a large Japanese manufacturing company, we search for evidence of biases identified in the model that allow for both mean-shifting and mean-preserving biases. We account for unobservable ability using fixed effect models. Our findings are generally consistent with the existence of attenuation biases: (1) internally trained supervisors tend to refrain from giving low grades to mid-career hires; (2) married (unmarried) supervisors tend not to give high grades to unmarried (married) subordinates; (3) college-graduate supervisors give high grades less frequently to those with lower education (supervisors with master or Ph.D. degrees give low grades more frequently to those with only college degrees). We did not find any significant biases caused by gender differences partly due to very small sample of female supervisors. Finally, we did not find any strong evidence of "own-group effect"—tendency that supervisors give more favorable treatment to those in the same social category.

1. Introduction

In this paper, we examine the existence of bias in subjective performance evaluation using personnel records from large Japanese manufacturing company.

Prior research has found little evidence of bias in performance evaluation except sports (see Persons et al (2011), Price and Wolfers (2010) for exmaple). The only exception is Elvira and Town (2001) which, using personnel records from a large corporation, find that Caucasian (African-American) bosses tend to give lower grades to African-American (Caucasian) subordinates than those in their own race. Although their findings are consistent with Giuliano, Levine, and Leonard (2009, 2011) who find the similar own-race bias in hiring, quit, layoff, and promotion, Elvira and Town (2001) focus only on racial "match" and do not consider other characteristics that might be correlated with workers' racial backgrounds.

Finding the sources of bias is very important because the negative bias could demoralize the disadvantaged workers and lead to their quits. Takahashi, Owan, Tsuru, and Uehara (2014) in fact find strong evidence that negative bias led to a higher quit rate in a Japanese car dealership. Furthermore, if a systematic bias in evaluation causes one group of employees to be disadvantaged in pay and promotion, the usual tests of discrimination (where evaluation results are used to control for productivity) are biased toward finding nothing.

There are two different kinds of bias. First, a supervisor may give preferential rating to those in the same social category as himself/herself while giving discriminatory rating to those in other social catregories. Such own-group bias is reported for decisions in hiring, layoff, and promotion at a large US retail chain studied by Giuliano, Levine, and Leonard (2009, 2011). Such favoritism/discrimination is not necessarily caused by preferences but may reflect better communication between a supervisor and his/her subordinates in the same social category, which actually increases productivity. Or, it may be the case that supervisors in different social categories have different evaluation criteria that have adverse effects on workers in different groups. Second, supervisors may give more compressed evaluation ratings to those in different social categories because they have less information for judging their subordinates' performance. This view echos "the language theory of discrimination" by Lang (1984), who predicts that minority workers may receive attenuated evaluation from majority supervisors because the majority and minority do not share the same language, culture or social norm. The limited information on minority subordinate held by majority supervisor makes the majority supervisor less candid about evaluating minority subordinates. Accordingly, majority supervisors arguably give more attenuated grade to minority subordinates than to majority subordinates.[1]

The second type of bias is more subtle but nonetheless no less harmful than the first one. Prior works have shown that the optimal incentives scheme is lower-powered when there is more uncertainty in performance evaluation (Baker 2000, 2002), and the optimal incentive scheme exhibits pooling in a greater range at the top when the supervisor's assessment and the worker's assessment are less correlated (MacLeod 2003). Therefore, if the supervisor and the worker do not share the sources of information, the evaluation grades will be more compressed. Such situation is especially likely to be the case when he/she belongs to different social categories. This means that majority supervisors will be more likely to give attenuated grades to minority workers. If only high performers get promoted, fewer minority workers will get promoted due to compressed grades. Another problem arises when a minority worker gets promoted. Since she has a smaller network and can obtain less information about the performance of her subordinates, she will give similar grades to most of her subordinates, which might lower her subordinates' motivation. Over time, management will learn that she

[1] In this case, the mean evaluation received by minority subordinates does not differ from the evaluation received by majority subordinates, but the attenuated evaluation on minority subordinates has detrimental consequence on their career progression to upper level managerial positions because only a small numbers of exceptional performers are promoted to these positions.

cannot give candid evaluation and as a result often fail to identify talented workers or boost her subordinates' work morale. This further impedes the advancement of minority in the corporate ladder.

In order to investigate the above types of biases, we examine how the evaluation ratings are associated with supervisor-worker differences in five characteristics. We first develop a model of evaluation bias building on the work on favoritism by Prendergast and Topel (1996). The model is rich in its implications about how match or mismatch of characteristics between supervisor and worker affects the evaluation rating. Then, we use the model to interpret the results of linear regression model estimation to determine what characteristic differences between supervisors and their workers are affecting the evaluation rating and how such bias arise. We have chosen three dimensions of employee characteristics as the sources of potential bias including gender, marital status, and education.

2. A Model of Evaluation Bias

We consider the three-tier organization where the management employs a supervisor who supervises a worker. The supervisor privately observes the worker's performance given by

$$y_s = a + \varepsilon_s$$

where a is the worker's ability and ε is measurement error. The supervisor collects unorganized bits of information about the worker's contribution to the organization from his/her co-workers and customers and the precision of the aggregated information depends on the amount of communication that the supervisor has with the worker and other various sources of information. We assume that $\varepsilon_s \sim N(0, \sigma_s^2)$. The worker's ability is also drawn from a normal distribution $a \sim N(\bar{a}, \sigma_a^2)$. a is unknown to all parties but its distribution is public information.

Following Prendergast and Topel (1996), we assume that the supervisor's utility depends on his own pay, w_s, and on the pay of his subordinate, w_w:

$$v_s = w_s + \eta w_w$$

Here η is the intensity of the supervisor's preference for the worker. Since η is allowed to take both positive and negative values, bias may come from either favoritism or discrimination. Management

$$y_m = a + \varepsilon_m$$

where $\varepsilon_m \sim N(0, \sigma_m^2)$. There are two explanations for why management delegates the authority to evaluate the worker's performance to the supervisor. First, the supervisor may have greater advantage in evaluating the worker so $\sigma_s^2 < \sigma_m^2$. Second, it may take a lot of time for management to gather performance information for individual workers. For example, y_m may be the average of the assessments of the worker's performance by multiple supervisors over years including future ones. In this case, it is possible to have $\sigma_s^2 > \sigma_m^2$, but the need to motivate the worker in a timely manner may require the manager to delegate the right to the supervisor. We assume that a, ε_s and ε_m are all uncorrelated with each other. Management monitors the supervisor and penalizes biased assessment of the worker's performance. It does so by comparing the supervisor's report with its own assessment, given by:

$$w_s = w_0 - 0.5\lambda(\hat{y}_s - E(a|y_m))^2 \tag{1}$$

where $\hat{y}_s$ is the actual report of the supervisor's assessment and $E(a|y_m)$ is the best unbiased estimator of "a" conditional on y_m and can be shown to be

$$E(a|y_m) = \frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{a} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}y_m.[2] \tag{2}$$

We assume that the management and the supervisor perfectly knows σ_s^2 and σ_m^2 but η is a private information of the supervisor. By this pay scheme, the management penalizes the supervisor when the supervisor reports his subordinate's evaluation

---

[2] This expression is obtained by calculating $E(a|y_m) = \int_{-\infty}^{\infty} a f_{a|y_m}(a|y_m)\, da = \int_{-\infty}^{\infty} a \frac{f_a(a) f_\varepsilon(y_m - a)}{\int_{-\infty}^{\infty} f_a(\acute{a}) f_\varepsilon(y_m - \acute{a}) d\acute{a}} da$ where $f_{a|y_m}$ is the conditional probability density function of a given the value of y_m, f_a is the unconditional probability density function of a, and f_ε is the probability density function of ε. More details of this calculation is given in the appendix.

different from the management's. Therefore, the supervisor pay costs for discriminating or favoring his subordinate and the parameter λ determines the size of the cost.

Finally, we assume that the worker's pay depends linearly on the two pieces of information available to management: $w_w = \tau_0 + \tau_1 \hat{y}_s + \tau_2 y_m$. We treat this pay scheme as given as it is designed based on the factors (e.g. moral hazard) not considered in this model. The supervisor has an incentive to report $\hat{y}_s$ different from y_s because his subordinate's wage is partly determined by his report. The supervisor reports his subordinate's evaluation to the management considering both the cost and benefit of biasing the evaluation. It is also worth noting that the supervisor has an incentive to report attenuated evaluation when he/she does not have accurate information on the subordinate's performance to avoid his evaluation deviating from the management's evaluation.

Then, the supervisor's problem is to solve

$$\max_{\hat{y}_s} E[\mathrm{w}_s + \eta w_w \,|\mathrm{y}_s] = \mathrm{w}_0 - 0.5\lambda E[(\hat{y}_s - E(\mathrm{a}|\mathrm{y}_m))^2|\mathrm{y}_s]$$

$$+\eta(\tau_0 + \tau_1 E[\hat{y}_s|\mathrm{y}_s] + \tau_2 \mathrm{E}[y_m|\mathrm{y}_s])$$

The supervisor reports his subordinate's evaluation by adding bias to his best predictor of the subordinate's ability because there are two benefits for doing so. First, adding bias indulges the supervisor's taste for discrimination against (favoritism for) subordinates. Second, adding bias helps the supervisor's evaluation not standing out from the management's evaluation. The manager knows that he and the management do not share the same information on the same subordinate and attempt to conform his evaluation to the management's. Then,

$$\max_{\hat{y}_s} E[w_s + w_w\,|y_s] = w_0 - 0.5\lambda E\left[\left(\hat{y}_s - \frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{a} - \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}y_m\right)^2 \middle| y_s\right]$$

$$+\eta(\tau_0 + \tau_1\hat{y}_s + \tau_2 E[y_m|y_s])$$

$$= w_0 - 0.5\lambda\hat{y}^{s2} + \lambda\hat{y}^s\left(\frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{a} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}E[y_m|y_s]\right)$$

$$-0.5\lambda E\left[\left(\frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{a} - \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}y_m\right)^2 \middle| y_s\right] + \eta(\tau_0 + \tau_1\hat{y}^s + \tau_2 E(a|y_s))$$

By taking the first order condition,

$$\hat{y}_s = \frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{a} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}E[a|y_s] + \frac{\eta\tau_1}{\lambda}$$

$$= \bar{a} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}\left(\frac{\sigma_s^2}{\sigma_a^2+\sigma_s^2}\bar{a} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_s^2}y_s - \bar{a}\right) + \frac{\eta\tau_1}{\lambda}$$

$$= y_s - \left(1 - \frac{1}{1+\sigma_m^2/\sigma_a^2}\frac{1}{1+\sigma_s^2/\sigma_a^2}\right)(y_s - \bar{a}) + \frac{\eta\tau_1}{\lambda}$$

The last term is intuitive; the larger the benefit of giving biased evaluation, the larger the degree of the bias, while the larger the penalty of giving biased evaluation, the smaller the degree of bias. The second term expresses the bias in the form of attenuation; it takes negative value when the subordinate performs better than the average and takes positive value when the subordinate performs worse than the average in the supervisor's perception. The attenuation arises from the supervisor's desire to avoid penalty imposed on his/her biased assessment (Equation 1). This attenuation bias is smaller when the supervisor does not have precise information about the subordinate's performance because his/her unbiased predictor of a, $\mathrm{E}(a|\mathrm{y_s})$, is already sufficient close to $\bar{a}$ thus does not require much additional compression. When the management also has limited access to additional information about the worker's performance (i.e. large σ_m^2/σ_a^2), the attenuation bias gets larger because failing to set $\hat{y}_s$ close to $\mathrm{E}[y_m|\mathrm{y}_s]$ may result in huge penalty. Then,

$$\hat{y}_s = \bar{\mathrm{a}} + \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}\frac{\sigma_a^2}{\sigma_a^2 + \sigma_s^2}(\mathrm{y}_s - \bar{\mathrm{a}}) + \frac{\eta\tau_1}{\lambda}$$

Formally, we state this result as proposition.

Proposition 1

$$\begin{aligned}\hat{y}_s &= \mathrm{E}(a|\mathrm{y}_s) + \frac{\eta\tau_1}{\lambda} - \frac{\sigma_a^2\sigma_m^2}{(\sigma_a^2 + \sigma_m^2)(\sigma_a^2 + \sigma_s^2)}(\mathrm{y}_s - \bar{a}) \\ &= \bar{a} + \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}\frac{\sigma_a^2}{\sigma_a^2 + \sigma_s^2}(\mathrm{y}_s - \bar{a}) + \frac{\eta\tau_1}{\lambda} \\ &= \mathrm{y}_s - (1 - \frac{1}{1 + \sigma_m^2/\sigma_a^2}\frac{1}{1 + \sigma_s^2/\sigma_a^2})(\mathrm{y}_s - \bar{a}) + \frac{\eta\tau_1}{\lambda}\end{aligned}$$

Formal proof is in the appendix.

The final line of the equation renders a useful empirical prediction on the evaluation the supervisor gives. There is no attenuation bias, if both the manager and the supervisor observe the output of the worker perfectly, that is $\frac{\sigma_m^2}{\sigma_a^2} = \frac{\sigma_s^2}{\sigma_a^2} = 0$. On the contrary, the larger the unobservability of the worker either by the manager or the supervisor, the larger the attenuation bias. Minority supervisors with less information, that is large $\frac{\sigma_s^2}{\sigma_a^2}$, accordingly give less candid evaluation.

If $\bar{a}$ depends on education, the inequality also suggests that a minority supervisor's evaluation of a majority worker, or the evaluation of a minority worker in general should depend more on his/her education than in the case of a majority supervisor evaluating a majority worker.

How does this biased assessment affect the supervisor's own wage? By plugging in the solutions to Equation 1, we obtain

$$\mathrm{w}_s = \mathrm{w}_0 - 0.5\lambda E\left[\left(\hat{y}_s - E(\mathrm{a}|\mathrm{y}_m)\right)^2\right]$$

$$= \mathrm{w}_0 - 0.5\lambda E\left[\left(\bar{a} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}\frac{\sigma_a^2}{\sigma_a^2+\sigma_s^2}(\mathrm{y}_s - \bar{a}) + \frac{\eta\tau_1}{\lambda} - \bar{\mathrm{a}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}(\mathrm{y}_m - \bar{\mathrm{a}})\right)^2\right]$$

$$= \mathrm{w}_0 - 0.5\lambda\left[\left(\frac{\eta\tau_1}{\lambda}\right)^2 + \frac{\sigma_a^2+\sigma_s^2}{(1+\sigma_m^2/\sigma_a^2)^2(1+\sigma_s^2/\sigma_a^2)^2} + \frac{\sigma_a^2+\sigma_m^2}{(1+\sigma_m^2/\sigma_a^2)^2} - \frac{2\sigma_a^2}{(1+\sigma_m^2/\sigma_a^2)^2(1+\sigma_s^2/\sigma_a^2)}\right]$$

Some derivation of the equations reveal that: (1) the supervisor's wage is decreasing in his/her preference for favoritism/discrimination (η); (2) the supervisor's wage is increasing initially in the size of penalty (λ) but decreasing beyond the threshold (i.e. $\lambda > \tilde{\lambda}$); (3) the supervisor's wage is decreasing in the size of the noise in the supervisor’s information; and (4) the supervisor’s wage is increasing in the size of the noise in the management’s information.

The first result is immediate and easily understood. The second result indicates that the beneficial impact of decreasing bias in the face of penalty outweighs the direct impact of penalty on the supervisor’s wage when λ is sufficiently small but it is reversed when λ gets larger. The third result is intuitive because the supervisor benefits from having more precise information. Finally, informed management penalizes the bias in the supervisor’s assessment and lack of his/her information more precisely thus lowers the wage paid to the supervisor. The third result is especially important because it implies that minority supervisors may receive worse evaluation grade and receive less wage because they are likely to receive less information through her smaller network.

3. Hypotheses

We assume that the precision of information collected by management and the supervisor depends on the characteristics of the supervisor and the worker. Let us summarize the results on the effects of being in majority or informed group on $\hat{y}_S$. Here, majority does not mean real majority. We use the word to mean the group which has a larger network through which information is shared. So, we use majority (minority) and more informed (less informed) group interchangeably. In our context, majority or more informed groups are male employees, married employees, and college graduates. Let σ^2_{m,g_w} be σ^2_m when the worker's group is g_w where $g_w = \text{I}$ means the worker is in a more informed group and $g_w = \text{N}$ means the worker is in a less informed one. Similarly, let $\sigma^2_{s,g_s g_w}$ be σ^2_s when the supervisor's group is g_s $(g_s = \text{I or N})$ and the worker's group is g_w $(g_w = \text{I or N})$. Finally, let $\eta_{g_s g_w}$ be the bias parameter η when the supervisor's group is g_s and the worker's group is g_w.

We make the following assumptions. First, we assume that management obtains less information about minority workers because minority workers have less interactions with majority supervisors, who are the primary information source for management.

Assumption 1 $\sigma^2_{m,N} > \sigma^2_{m,I}$

Second, we assume that supervisors have less information about the performance of the workers in the different group because: (1) communication is less frequent; and (2) specific situations are harder to understand (so distinguishing ability and circumstances is harder). Minority supervisors will be further disadvantaged because they have a smaller network and the information obtained through peers is also limited. So, we assume the following.

Assumption 1 $\sigma^2_{s,NI} > \sigma^2_{s,IN} > \sigma^2_{s,II}$ **and** $\sigma^2_{s,NI} > \sigma^2_{s,NN} > \sigma^2_{s,II}$.

Under these hypotheses, how would the performance evaluation of a majority (minority) worker be different depending on whether his (her) supervisor belongs to the majority and minority group? Let $\mathrm{K}(g_s, g_w) = \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,g_w}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,g_s g_w}}$. Then, the supervisor's evaluation of the worker is then expressed as

$$E[\hat{y}_s | g_s, g_w] = (1-\mathrm{K}(g_s, g_w))\bar{\mathrm{a}} + \mathrm{K}(g_s, g_w)\mathrm{y}_s + \frac{\eta_{g_s g_w}\tau_1}{\lambda}$$

Note that $\mathrm{K}(N, I) - \mathrm{K}(I, I) = \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,I}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,NI}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,I}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,II}} < 0$.

Similarly, $\mathrm{K}(I, N) - \mathrm{K}(I, I) = \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,N}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,IN}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,I}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,II}} < 0$ And $\mathrm{K}(N, N) - \mathrm{K}(I, I) = \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,N}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,NN}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{m,I}}\frac{\sigma_a^2}{\sigma_a^2+\sigma^2_{s,II}} < 0$. They mean that the evaluation is attenuated when either supervisor or subordinate belongs to a minority group. If $\bar{\mathrm{a}}$ depends on education, the inequality also suggests that a minority supervisor's evaluation of a majority worker, or the evaluation of a minority worker in general should depend more on his/her education than in the case of a majority supervisor evaluating a majority worker. We do not have any a priori prediction about the bias on the average evaluation. $\eta_{II} > \eta_{NI}$ should hold if majority supervisors exhibit *endophilia* (preference for similar type) or minority supervisors exhibit *exophobia* (discrimination against different type).

Now, we formally state the hypotheses.

Hypotheses 1 Performance evaluation is attenuated when either the supervisor or the subordinate belongs to the minority group compared with the case in which both belongs to the majority group.

Hypothesis 1 implies the possibility that mean-preserving shift of the evaluation distribution can be caused by the supervisor-subordinate differences in characteristics. We also test the possibility of mean-shifting bias possibly caused by the own-group bias. As Giuliano, Levine, and Leonard (2009, 2011) show for the racial bias, people may exhibit own-group bias, namely they may offer favorable treatment for those who are similar to them. It is not necessarily caused by taste-based favoritism or discrimination (Becker 1957). It is possible that communication and coordination is easier between the two in the same group, and thus the subordinate who are in the same group as the supervisor may actually become more productive (Lang 1986). The above discussion leads to the second hypothesis.

Hypotheses 2 Performance evaluation grade is lower when the supervisor and the subordinate belong to different identify groups.

4. Data

In this research, we use personnel records from a large Japanese manufacturing company. We have the supervisor-worker matched information from 2006 to 2009. The supervisor information is not available for all workers partly because evaluation rating is optional for production workers and whether it is conducted or not is discretion of the management of each plant. They are also missing for some of the workers who are new (within one year), taking leaves, or transferred to subsidiaries. Typical evaluators for regular workers hold the G4 job grade (see Figure 1). Therefore, we restrict our analysis to those evaluated by the managers with the G4-G1 job grades. G4 is the lowest and G1 is the highest managerial rank. After dropping the observations that do not satisfy this requirement, the total number of observations is 23,965(ここは確認して必要ならば修正致します)[3].

Workers are ranked as C, B, A3, A2, A1, or S where C and S are the lowest and the highest, respectively. Tables 2 and 3 show the distribution of the evaluation grades. As you see in Table 3, the evaluation scales for managers were changed in 2008: A1, A2

---

[3] As we discuss in a subsequent part of this section, we made three subsamples. The number of observations of whole non-managerial workers is 16686.And professional white-collar workers and managers account for 4644 observations and 7279, respectively.

and A3 were consolidated into A , and the standards for S and B were also adjusted accordingly.

As dependent variables, we created two evaluation grade dummies with *A1 and over* indicating A1 and better grades (S for managers in 2008-2009 due to the above change in the scale), and *A3 and below* indicating A3 and lower ones (B for managers in 2008-2009 for the same reason). Given that A2 (or A for managers in 2008-2009) accounts for 55-80% of the total observations, using both indicator variables allows us to evaluate the impact of the supervisor-worker "match" in characteristics on the evaluation share of high performers and low performers, separately.

Explanatory variables indicating the supervisor-worker social category "match" are used to identify possible sources of evaluation bias. We focus on the "match" along four employee characteristics that potentially define the identity groups of supervisors and subordinates. These four category variables are[4];

[4] We earlier used new graduate hire vs. mid-career hire matching variable. *insider vs. outsider* is a category variable of two-by-two combination of supervisor's status and worker's status at the time of each entry—new graduate (supervisor) & new graduate (subordinate), new graduate & mid-career hire, mid-career hire & new graduate, and mid-career hire & mid-career hire. However, we found no significant effects on evaluation in our analysis, so we omitted the variable in our final analysis.

gender match is a category variable of two-by-two combination of supervisor's gender and worker's gender: male (supervisor) & male (subordinate), male & female, female & male, and female & female.

marital status match is a category variable of two-by-two combination of supervisor's marital status: married (supervisor) & married (worker), married & unmarried, unmarried & married, and unmarried & unmarried.

education match is a category variable indicating whether the supervisor has higher, equal, or lower education than the worker: supervisor's education > subordinate's education, supervisor's education = subordinate's education, and supervisor's education < subordinate's education.

school match is a dummy variable indicating whether the subordinate graduated the same school as the supervisor.

We use these categorical "match" variables in our estimation to examine possible evaluation bias.

Since criteria for evaluation are likely to differ between white-collar and blue-collar workers, and managerial and non-managerial positions, we use three subsamples to conduct regression analyses.[5] They are all non-managerial workers, white-collar

---

[5] Recall that the evaluation scale was changed only for managers in 2008.

workers defined as those holding college degrees (including those from colleges of technology but not those from two-year colleges), and managers. [6] In this company, there is no clear job classification distinguishing between (management track) professionals and non-professionals, the latter of which may include production workers and administrative assistants. As you see in Figure 1, all employees at the entry level, including both college graduates and high school graduates, start at the J1 grade. Management track white-collar (college graduate) workers quickly move up to the SA level, while non-professional (non-college graduate) workers move up the ladder for non-professional workers, J-labeled grades, very slowly. We do not know exactly who in J1 and J2 are on the managerial track but we make inference based on the worker's education.

Managers are defined as those who hold the management grades, G6-G1. Descriptive statistics by samples are in the Table 4.

5. Empirical Strategy

[6] Note that almost all graduates from two-year liberal-art colleges were hired as administrative assistants, they are not included in professionals.

In this paper, we examine the existence of evaluation bias due to the "mismatch" of characteristics between the supervisor and the worker, by testing Hypotheses 1-2. We estimate both linear probability (OLS) models and the Fixed Effect models. The OLS model and the FE model are specified as follows for worker *i*, supervisor *j*, and year t, respectively:

$$y_{it} = \boldsymbol{X_{ijt}}\boldsymbol{\beta} + \boldsymbol{Z_{ijt}}\boldsymbol{\gamma} + c_i + d_t + \epsilon_{ij}$$

where y_{it} takes either *A1 & over* or *A3 & below* given to worker *i* and year *t*, $\boldsymbol{X_{ijt}}$ is a vector of control variables including worker's job tenure, supervisor's job tenure, the interaction between the worker's gender and marital status dummies, job grade dummies and education dummies. In the fixed effects specification $\boldsymbol{X_{ijt}}$ does not include time-invariant variables such as gender and education are dropped from the latter. $\boldsymbol{Z_{ijt}}$ is a vector of social category match between the supervisor and the worker including *gender match* (the interaction between the worker's and the supervisor's gender), *marital status match* (the interaction between the worker's and the supervisor's marital status), *education match* (the indicator of whose education level is higher between the supervisor and the worker), and *school match* (the indicator of the case when the worker and the supervisor went to the same college). The vector $\boldsymbol{Z_{ijt}}$ does not include time-invariant terms in the fixed effects specification. For example, β_{mf} $male_j$ & $female_i$, + β_{ff} $female_j$ & $female_i$ = (β_{mf} -β_{ff}) $male_j$ & $female_i$, +β_{ff} $female_i$. Therefore, once the worker fixed effects are included, β_{mf} and β_{ff} cannot be identified and onlyβ_{mf} -β_{ff} is estimated as a coefficient of $male_j$ & $female_i$ (or the coefficient of - $female_j$ & $female_i$ depending on which variable is chosen as the reference group). c_i is the worker effect, d_t is the year effect, and ϵ_{ij} and ϵ'_{ij} are error terms uncorrelated with the rest of the terms. .

As in the Persons et al. (2011), readers might see it necessary to control for the supervisor's fixed effect[7]. However, this company requires all evaluators to keep the average at A2 and our null hypothesis that supervisor effects are all zero cannot be rejected.[8] Another issue that arises in the model with both supervisor and worker fixed effects is that the interpretation of the results becomes very difficult because, for example, the following four cases cannot be separately identified: (1) favoritism of women toward women; (1) favoritism of men toward men; (3) discrimination of women toward men, and (4) discrimination of men toward women.

Thus, we have decided not to control for the supervisor's fixed effect.

---

[7] Persons controlled not only pitcher's (worker's) fixed effect but also umpire's (evaluator's) fixed effect to capture the evaluation bias.

[8] Numerical representation of S-C grades were not revealed to the researchers.

6. Empirical Results

First, we performed the base regressions specified in the previous section without school matching dummy variable. Results are in the Table 5 (for non-managerial worker), Table 6 (for white collar worker), and Table 7(for manager). Hypothesis 1 generally implies that the evaluation grade is more likely to take medium evaluation (A2), and thus good (A1&Over) and bad (A3& Below) are less likely, when either supervisor or worker belongs to a minority group, or relatively less informed group. In our context, this is the case when either the supervisor or the worker is female, unmarried, or less educated. As we show in Tables 5-7, some results are consistent with the hypothesis.

First, in the FE model estimations for non-managerial workers and white collar workers, female supervisors tend to give more attenuated evaluation for male subordinates than male counterparts. This may be because female supervisor tends to acquire less information about her subordinate's ability. Lack of significance for other gender match variables or the OLS results does not necessarily negate Hypothesis 1 because: (1) the OLS results are biased due to unobservable worker characteristics; and (2) the coefficient of $male_i$*$female_j$ is basically the difference in the effects between

$male_i*female_j$ and $female_i*female_j$ but the relative size of the two terms cannot be clarified according to the derivation for Hypothesis 1. One puzzle is that, in the subsample of managers, female evaluators tend to give significantly more A3 and lower grades to male subordinates than male counterparts but this effect disappear once worker effects are controlled for. This may indicate that less competent male managers are more likely to be assigned under female managers or in the female-dominant workplaces.

Second, although the results are not significant, unmarried supervisors tend to give more attenuated grades for married subordinates than married supervisors in all sample and subsamples both in the OLS and FE models, which is consistent with Hypothesis 1. One interpretation is that unmarried supervisors don't know well about the family requirements and circumstantial factors surrounding married subordinates, thus find it more difficult to distinguish among ability, efforts, and luck. We also find that married supervisors tend to give lower grades to unmarried workers than married workers. But, this is likely to be caused by unobservable worker ability because the inclusion of workers fixed effects eliminated this effect.

Third, for non-managerial workers, an evaluator whose education is higher than his/her subordinate is apt to give more compressed evaluation than a supervisor whose education level is the same as the subordinate. Although including the worker effects makes the result less significant, the signs of the coefficients remain the same. This result may sound inconsistent with Hypothesis 1 because we tend to believe that more educated people should be more informed or should have a greater network. However, note that a majority of less educated workers are high-school graduates who work as production workers. It is very likely that high-school graduate plant managers who were once production workers can better evaluate the productivity of his subordinates and have a greater network through which information is acquired than college-graduate managers who have no experience as production workers. Therefore, the result for non-managerial workers about the education match variables seems to be consistent with Hypothesis 1. Such explanation does not hold any more once production workers are dropped from the sample. The FE results for white collar workers suggest that an evaluator with more education tends to evaluate subordinates with less education significantly lower. This is more consist with the own-group effect expressed in Hypothesis 1.

Other important findings in Table 5-7 include:

i) Evaluation diverges with job tenure but at a decreasing rate as more information is revealed to the supervisor

ii) As supervisor’s job tenure becomes longer, evaluation becomes more candid for low performers, but this effect cannot be observed for manager’s evaluation.

Evaluation ratings for employees in the headquarters are more likely to fall in A2—medium rating. This may imply that employees in the headquarters engage more in coordination tasks which are presumably more difficult to evaluate. Note that result (i) is consistent with Hypothesis 2 because the more the supervisor has information, the more diverse the evaluation becomes. As for results (ii) for non-managerial worker and white collar worker suggest that newly appointed front-line managers (not those who manage managers) tend to give lenient rating to their subordinates. This may mean that inexperienced managers are easy to impress, or that they fear to alienate their subordinates by giving harsh ratings because they rely more on information and support from their subordinates.

One implication from Hypothesis 1 is that if female supervisors tend to give more attenuated evaluation for male workers because female supervisors have less information about male subordinates, female supervisors may be putting more weight on prior expectation and less weight on current assessment. If prior expectation of worker performance heavily depends on the worker’s education, the evaluation grade of male workers should be more correlated with the worker’s education, which includes both the level of academic degree and the quality of school the worker graduated from, under female supervisors. We next run the FE model regressions with school quality index interacted with the supervisor's gender, and school matching dummy variable.[9] Results are presented in Table 8.

In the table, we find that the cross term between the school quality index and the supervisor's gender is significantly positive whereas the interaction between $female_j$ *and* $male_i$ is significantly negative in the equation for *A1&over*. This is consistent with our expectation. Namely, female supervisors use more explicit information—school quality index—as proxy information about the worker's ability. One concern is that it may be the case that female managers are graduates from some selected universities while male managers are more dispersed, and the former's more strict assessment toward male workers is a simple reflection of favoritism toward those from the same university (alma mater effect) or discrimination against those from different universities. In order to rule out such effect, we include school matching dummy variable, which turns out to have weakly significant negative coefficient for *A3&below*. This may imply that a supervisor from the same school as his/her subordinate may give a preferential treatment for the latter.

[9] School quality index is demeaned.

7. Conclusion and Further Issues

In this research, we have searched for evidence of biases in subjective evaluation using personnel records from a large Japanese manufacturing company. We have found some evaluation biases that might have been caused by asymmetry in information acquisition capability due to differences in background or differences in the size of network that the supervisor belongs to. For example, we find that female supervisors give more attenuated evaluation grades to male workers. This finding is disturbing because evaluating and motivating the performance of workers is one of the most important roles of middle managers. If female managers cannot give candid assessment of worker performance and fail to motivate their subordinates, their promotion opportunities to higher level positions may be limited.

Interestingly, we did not find any evidence of “own-group effect” or any taste-based discrimination in the company. Social category divide by gender, marital status, or education may not cause substantial favoritism or discrimination as observed among racial groups in previous studies using personnel records from U.S. firms.

These results will contribute to the economic analyses of discrimination, and gender pay.

But there are some remaining issues. First, although we control for job grades of workers in our analyses, we may still need to account for differences in jobs because job assignment is not random. If more women are sorted into “specialists” jobs where outstanding performance or poor performance is less conspicuous, managers in such jobs are less likely to give highest or lowest grades than those in jobs where performance is more easily measured such as sales.

Second, since our dataset is restricted to four years from 2006 to 2009, within-worker variation of supervisor characteristics is rather limited. Therefore, it is rather difficult to find the relationship between evaluation and the supervisor-worker match in characteristics precisely. If we could obtain more data for longer periods in the future, we should be able to identify more precisely how much of the variation in evaluation ratings is caused by biases. Furthermore, we might be able to explore for evaluating consequences of biases such as whether biased evaluation tends to end with transfers or quits of workers or whether supervisors who tend to make biased evaluation are punished or not.

Appendix.

Lemma $E(\mathrm{a}|\mathrm{y}_m) = \frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{\mathrm{a}} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}\mathrm{y}_m,\ E(\mathrm{a}|\mathrm{y}_s) = \frac{\sigma_s^2}{\sigma_a^2+\sigma_s^2}\bar{\mathrm{a}} + \frac{\sigma_a^2}{\sigma_a^2+\sigma_s^2}\mathrm{y}_s.$

Proof: We will show the proof for y_m because the two equations are identical. This expression is obtained by calculating

$$\mathrm{E}(a|\mathrm{y_m}) = \int_{-\infty}^{\infty} a f_{a|y_m}(a|y_m)\,da = \int_{-\infty}^{\infty} a \frac{f_a(a) f_\varepsilon(y_m - a)}{\int_{-\infty}^{\infty} f_a(\acute{a}) f_\varepsilon(y_m - \acute{a}) d\acute{a}}\,da \qquad \text{(A1)}$$

where $f_{a|y_m}$ is the conditional probability density function of a given the value of y_m, f_a is the unconditional probability density function of a, and f_ε is the probability density function of ε. Now,

$$f_a(a) f_\varepsilon(y_m - a) = \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[-\frac{(a-\bar{a})^2}{2\sigma_a^2}\right] \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left[-\frac{(y_m - a)^2}{2\sigma_m^2}\right]$$

$$= \frac{1}{\sqrt{2\pi(\sigma_a^2+\sigma_m^2)}} \exp\left[-\frac{(y_m-\bar{a})^2}{2(\sigma_a^2+\sigma_m^2)}\right] \frac{1}{\sqrt{2\pi\frac{\sigma_a^2\sigma_m^2}{\sigma_a^2+\sigma_m^2}}} \exp\left[-\frac{1}{2\frac{\sigma_a^2\sigma_m^2}{\sigma_a^2+\sigma_m^2}}\left(a - \frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{\mathrm{a}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}\mathrm{y}_m\right)^2\right].$$

Then, by substituting this into (A1), we obtain

$$\mathrm{E}(a|\mathrm{y_m}) = \int_{-\infty}^{\infty} a \frac{1}{\sqrt{2\pi\frac{\sigma_a^2\sigma_m^2}{\sigma_a^2+\sigma_m^2}}} \exp\left[-\frac{1}{2\frac{\sigma_a^2\sigma_m^2}{\sigma_a^2+\sigma_m^2}}\left(a - \frac{\sigma_m^2}{\sigma_a^2+\sigma_m^2}\bar{\mathrm{a}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma_m^2}\mathrm{y}_m\right)^2\right] da.$$ This concludes the proof.

Proof of Proposition 1

First, from the above lemma,

$$E(\mathrm{a}|\mathrm{y}_s) - E(\mathrm{a}|\mathrm{y}_m) = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_s^2}(\mathrm{y}_s - \bar{\mathrm{a}}) - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}(\mathrm{y}_m - \bar{\mathrm{a}}).$$

Then,

$$\mathrm{E}[E(\mathrm{a}|\mathrm{y}_s) - E(\mathrm{a}|\mathrm{y}_m)|\mathrm{y}_s] = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_s^2}(\mathrm{y}_s - \bar{\mathrm{a}}) - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}(\mathrm{E}(\mathrm{y}_m|\mathrm{y}_s) - \bar{\mathrm{a}})$$

$$= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_s^2}(\mathrm{y}_s - \bar{\mathrm{a}}) - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}\left(\frac{\sigma_s^2}{\sigma_a^2 + \sigma_s^2}\bar{\mathrm{a}} + \frac{\sigma_a^2}{\sigma_a^2 + \sigma_s^2}\mathrm{y}_s - \bar{\mathrm{a}}\right)$$

$$= \frac{\sigma_a^2\sigma_m^2}{(\sigma_a^2 + \sigma_m^2)(\sigma_a^2 + \sigma_s^2)}(\mathrm{y}_s - \bar{\mathrm{a}})$$

where the first equality is obtained from $\mathrm{E}(\mathrm{y}_m|\mathrm{y}_s) = \mathrm{E}(a|\mathrm{y}_s)$ and the above lemma. We further calculate

$$\mathrm{E}[\left(E(\mathrm{a}|\mathrm{y}_s) - E(\mathrm{a}|\mathrm{y}_m)\right)^2|\mathrm{y}_s] = \mathrm{E}[\left(E(\mathrm{a}|\mathrm{y}_s) - \frac{\sigma_m^2}{\sigma_a^2 + \sigma_m^2}\bar{\mathrm{a}} - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}(\mathrm{a} + \varepsilon_m)\right)^2 |\mathrm{y}_s]$$

$$= \mathrm{E}[\left(\frac{\sigma_m^2}{\sigma_a^2 + \sigma_m^2}(E(\mathrm{a}|\mathrm{y}_s) - \bar{\mathrm{a}}) - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}(\mathrm{a} - E(\mathrm{a}|\mathrm{y}_s) + \varepsilon_m)\right)^2 |\mathrm{y}_s]$$

Since $\mathrm{a} - E(\mathrm{a}|\mathrm{y}_s)$ and ε_m are independent and their conditional means are zero,

$$\mathrm{E}[\left(E(\mathrm{a}|\mathrm{y}_s) - E(\mathrm{a}|\mathrm{y}_m)\right)^2|\mathrm{y}_s]$$

$$= \frac{\sigma_a^4\sigma_m^4}{(\sigma_a^2 + \sigma_m^2)^2(\sigma_a^2 + \sigma_s^2)^2}(\mathrm{y}_s - \bar{\mathrm{a}})^2 + \left(\frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2}\right)^2 (V(\mathrm{a}|\mathrm{y}_s) + \sigma_m^2) =$$

$$= \frac{\sigma_a^4\sigma_m^4}{(\sigma_a^2 + \sigma_m^2)^2(\sigma_a^2 + \sigma_s^2)^2}(\mathrm{y}_s - \bar{\mathrm{a}})^2 + \frac{\sigma_a^4(\sigma_a^2\sigma_s^2 + \sigma_a^2\sigma_m^2 + \sigma_s^2\sigma_m^2)}{(\sigma_a^2 + \sigma_m^2)^2(\sigma_a^2 + \sigma_s^2)}$$

Taking a derivative to derive the first-order condition is straight-forward. This concludes the proof.

References

Baker, George P.(2000) “The Use of Performance Measures in Incentive Contracting”, *The American Economic Review*, Vol.90, No.2, pp.415-420

Baker, George P.(2002) “Distortion and Risk in Optimal Incentive Contracts”, *Journal of Human Resources*, Vol.37, No.4, pp.728-751

Becker, Gary S.(1957) “The Economics of Discrimination” The University of Chicago Press.

Elvira, Marta, and Robert Town (2001) “The Effects of Race and Worker Productivity on Performance Evaluations”, *Industrial Relations*, Vol.40, Issue 4, pp.571-590

Giuliano, Laura, David I. Levine, and Jonathan Leonard (2011) "Racial Bias in the Manager-Employee Relationship: An analysis of Quits, Dismissals, and Promotions at a Large Retail Firm”, *Journal of Human Resources,* Vol.46, No.1, pp.26-52.

Lang, Kevin (1986) “A Language Theory of Discrimination”, *Quarterly Journal of Economics*, Vol. 101, No. 2, pp. 363-382.

Persons, Chiristopher A., Johan Sulaeman, Michael C. Yates, and Daniel S. Hamermesh (2011) “Strike Three: Discrimination, Incentives, and Evaluation”, *The American Economic Review*, Vol.101, No.4, pp.1410-1435

Prendergast, Canice and Robert H. Topel (1996) “Favoritism in Organization” *Journal of Political Economy*, Vol.104, No.5, pp.958-978

Price,Joseph, and Justin Wolfers (2010) “Racial Discrimination among NBA Referees”, *Quarterly Journal of Economics*, Vol.125, Issue 4, pp.1859-1887

Figure 1 Promotion Path Chart

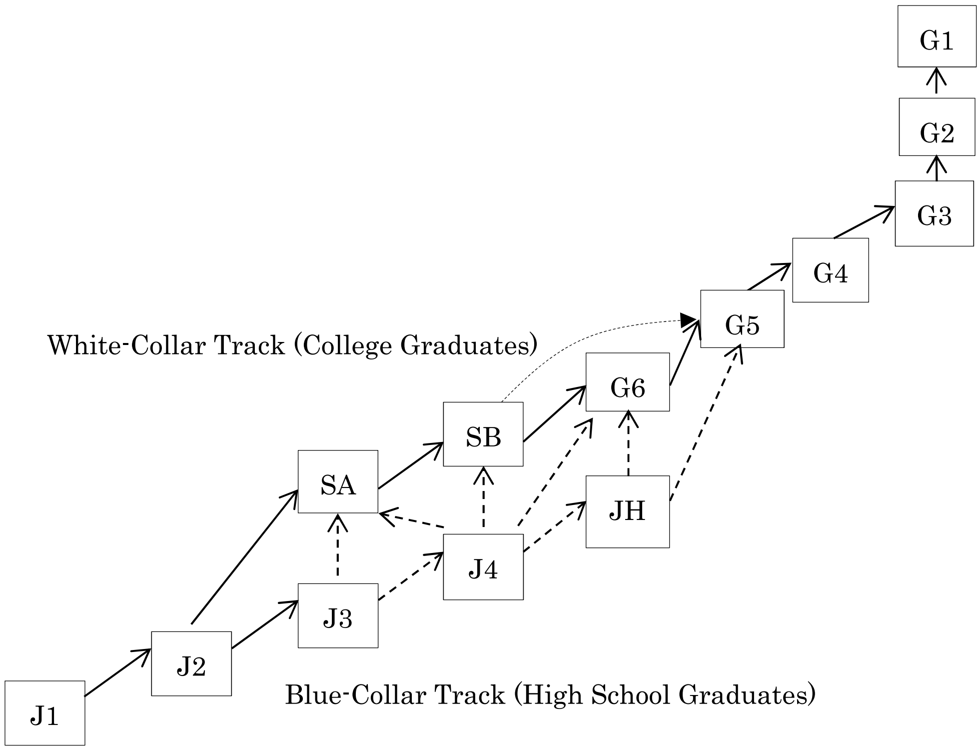

Table 2: Distribution of Evaluation Grades (non-managerial workers)

	2006	2007	2008	2009	2010	2011	2012	2013	Total
SS	0	0	0	0	0	1	0	0	1
	0.0%	0.0%	0.0%	0.0%	0.0%	0.03%	0.0%	0.0%	0.0%
S	53	59	69	84	499	515	417	301	1,997
	1.9%	2.1%	2.1%	2.3%	13.1%	14.0%	15.8%	14.3%	8.0%
A1	554	496	501	594	0	0	0	0	2,145
	19.7%	17.9%	15.2%	16.0%	0.0%	0.0%	0.0%	0.0%	8.6%
A2	1,755	1,827	2,328	2,498	3,072	2,897	1,998	1,617	17,992
	62.3%	65.8%	70.6%	67.3%	80.5%	78.9%	75.6%	77.0%	72.5%
A3	382	310	314	420	0	0	0	0	1,426
	13.6%	11.2%	9.5%	11.3%	0.0%	0.0%	0.0%	0.0%	5.7%
B	56	73	72	91	227	238	195	156	1,108
	2.0%	2.6%	2.2%	2.5%	5.9%	6.5%	7.4%	7.4%	4.5%
C	16	11	12	27	18	20	34	25	163
	0.6%	0.4%	0.4%	0.7%	0.5%	0.5%	1.3%	1.2%	0.7%
Total	2,816	2,776	3,296	3,714	3,816	3,671	2,644	2,099	24,832
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 3: Distribution of Evaluation Grades (managers)

	2006	2007	2008	2009	2010	2011	2012	2013	Total
S	35	30	179	178	230	207	171	148	1,178
	2.8%	2.3%	11.7%	11.2%	14.2%	12.7%	11.7%	12.4%	10.2%
A1	293	257	0	0	0	0	0	0	550
	23.4%	19.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4.8%
A2	699	802	1,266	1,254	1,265	1,297	1,113	889	8,585
	55.9%	62.1%	82.6%	78.7%	78.1%	79.9%	76.3%	74.5%	74.3%
A3	190	168	0	0	0	0	0	0	358
	15.2%	13.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.1%
B	28	29	81	149	113	107	150	138	795
	2.2%	2.2%	5.3%	9.4%	7.0%	6.6%	10.3%	11.6%	6.9%
C	5	5	6	12	12	13	25	18	96
	0.4%	0.4%	0.4%	0.8%	0.7%	0.8%	1.7%	1.5%	0.8%
Total	1,250	1,291	1,532	1,593	1,620	1,624	1,459	1,193	11,562
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 4: Descriptive Statistics

Dummy or Categorical Variavles	Non-Managerial Workers		White Collar(Management Track)		Managers	
	Observations	Percentage	Observations	Percentage	Observations	Percentage
Evaluation						
A1&Over						
Yes	3218	19.3	903	19.4	1291	17.7
No	13468	80.7	3741	80.6	5988	82.3
A2&Over						
Yes	14269	85.5	4119	88.7	6451	88.6
No	2417	14.5	525	11.3	828	11.4
Total	16686	100	4644	100	7279	100
Gender Matching						
Male×Male	14411	86.4	4079	87.8	7000	96.2
Male×Female	2221	13.3	531	11.4	208	2.9
Female×Male	18	0.1	17	0.4	37	0.5
Female×Female	36	0.2	17	0.4	34	0.5
Total	16686	100	4644	100	7279	100
Experience Matching						
New Graduate×New Graduate	10086	60.5	2782	59.9	6545	89.9
New Graduate×Midcareer	6205	37.2	1780	38.3	484	6.7
Midcareer×New Graduate	244	1.5	61	1.3	197	2.7
Midcareer×Midcareer	151	0.9	21	0.5	53	0.7
Total	16686	100	4644	100	7279	100
Mariage Matching						
Married×Married	9928	59.5	2181	47.0	6270	86.1
Married×Not Married	5616	33.7	2172	46.8	737	10.1
Not Married×Married	795	4.8	158	3.4	222	3.1
Not Married×Not Married	347	2.1	133	2.9	50	0.7
Total	16686	100	4644	100	7279	100
Education Matching						
Supervisor=Worker	3919	23.8	1543	33.2	3197	43.9
Supervisor>Worker	10343	62.7	1319	28.4	2405	33.0
Supervisor<Worker	2235	13.6	1782	38.4	1676	23.0
Total	16497	100	4644	100	7278	100
School Matching						
Equal	261	1.6	103	2.2	415	5.7
Not Equal	16283	98.4	4525	97.8	6829	94.3
Total	16544	100	4628	100	7244	100
Gender						
Male	14429	86.5	4096	88.2	7037	96.7
Female	2257	13.5	548	11.8	242	3.3
Total	16686	100	4644	100	7279	100
Experience						
New Graduate	10330	61.9	2843	61.2	6742	92.6
Midcareer	6356	38.1	1801	38.8	537	7.4
Total	16686	100	4644	100	7279	100
Marital Status						
Married	10723	64.3	2339	50.4	6492	89.2
Not Married	5963	35.7	2305	49.6	787	10.8
Total	16686	100	4644	100	7279	100
Workplace						
Head Office	697	4.2	429	9.2	1903	26.1
Others	15989	95.8	4215	90.8	5376	73.9
Total	16686	100	4644	100	7279	100
Education						
Junior High School	934	5.6	—	—	35	0.5
High School	9560	57.3	—	—	1364	18.7
Technological College	657	3.9	657	14.2	874	12.0
College: Undergraduate	2853	17.1	2853	61.4	2219	30.5
College: MA	1022	6.1	1022	22.0	2531	34.8
College: Ph.D.	108	0.7	108	2.3	229	3.2
Others	1552	9.3	4	0.1	27	0.4
Total	16686	100	4644	100	7279	100

Note: For categorical variables, observation and percertage are repoted.
Note: For continuous variables, mean,standard deviation, min and max are reported.

Table 4: Descriptive Statistics(cont.)

Continuous Variables	Mean	Std. Dev.	Min	Max	# of Obs.
Job Tenure	2.0386	1.5390	0	5	16686
Job Tenure Squared	6.5243	7.3417	0	25	16686
Supervisor's Job Tenure	1.3901	1.2698	0	5	16686
Job Tenure:White Collar	1.6602	1.4985	0	5	4644
Job Tenure Squared:White Collar	5.0013	6.6919	0	25	4644
Supervisor's Job Tenure:White Collar	1.3096	1.2897	0	5	4644
Job Tenure:Managers	1.4241	1.3905	0	5	7279
Job Tenure Squared:Managers	3.9613	5.8523	0	25	7279
Supervisor's Job Tenure:Managers	1.1067	1.1771	0	5	7279

Note: For categorical variables, observation and percertage are repoted.

Note: For continuous variables, mean,standard deviation, min and max are reported.

Table 5. Regression results for non-managerial workers

	A1&Over		A3&Below	
	OLS	FE	OLS	FE
INSIDER VS. OUTSIDER				
Midcreer×New Graduate	-0.0179	-0.0099	-0.0278	0.0090
	(0.0262)	(0.0291)	(0.0187)	(0.0230)
New Graduate×Midcreer	-0.0210 **	-0.0106	-0.0063	-0.0653 **
	(0.0083)	(0.0377)	(0.0073)	(0.0298)
Midcreer×Midcreer	-0.0128		0.0589	
	(0.0348)		(0.0382)	
MARITAL STATUS				
Married×Not Married	-0.0665 ***	-0.0433 ***	0.0361 ***	-0.0043
	(0.0082)	(0.0119)	(0.0078)	(0.0094)
Not Married×Married	-0.0303 **	-0.0053	-0.0051	-0.0183
	(0.0145)	(0.0154)	(0.0127)	(0.0122)
Not Married×Not Married	-0.0294		0.0292	
	(0.0219)		(0.0189)	
EDUCATION				
Evaluator>Evaluated	-0.0085	-0.0334 ***	-0.0249 ***	-0.0021
	(0.0083)	(0.0089)	(0.0076)	(0.0070)
Evaluator<Evaluated	-0.0014	0.0093	0.0039	-0.003
	(0.0116)	(0.0122)	(0.0094)	(0.0097)
# of Obs.	24249	24249	24249	24252
(Adjusted/Within) R Squared	0.0419	0.0150	0.0334	0.0164

Table 6. Regression results for white-collar workers

	A1&Over		A3&Below	
	OLS	FE	OLS	FE
INSIDER VS. OUTSIDER				
Midcreer×Internally-trained	-0.0498	-0.0072	-0.0595 **	-0.0151
	(0.0362)	(0.0547)	(0.0255)	(0.0427)
Internally-trained×Midcreer	-0.0266	-0.0757	-0.0005	-0.0338
	(0.0167)	(0.0740)	(0.0155)	(0.0579)
Midcreer×Midcreer	-0.0718		0.0089	
	(0.0616)		(0.0567)	
MARITAL STATUS				
Married×Not Married	-0.0856 ***	-0.0527 ***	0.0474 ***	0.0042
	(0.0147)	(0.0184)	(0.0122)	(0.0144)
Not Married×Married	-0.0614 **	-0.0665 **	-0.0354 **	-0.0262
	(0.0270)	(0.0306)	(0.0169)	(0.0239)
Not Married×Not Married	-0.0724 **		0.0778 **	
	(0.0317)		(0.0324)	
EDUCATION				
Evaluator>Evaluated	-0.0427 ***	-0.0236	0.0169	0.0322 ***
	(0.0147)	(0.0161)	(0.0120)	(0.0126)
Evaluator<Evaluated	-0.0120	0.0168	0.0159	0.0098
	(0.0141)	(0.0151)	(0.0110)	(0.0118)
# of Obs.	7292	7292	7292	7292
(Adjusted/Within) R Squared	0.0562	0.0207	0.0227	0.0163

Table 7. Regression results for managers

	A1&Over		A3&Below	
	OLS	FE	OLS	FE
INSIDER VS. OUTSIDER				
Midcreer×New Graduate	-0.0028	0.0051	-0.0082	0.0313
	(0.0199)	(0.0272)	(0.0193)	(0.0223)
New Graduate×Midcreer	0.0359 **	-0.0894	-0.0300 *	0.0102
	(0.0179)	(0.0703)	(0.0164)	(0.0575)
Midcreer×Midcreer	0.0041		-0.0482 *	
	(0.0430)		(0.0274)	
MARITAL STATUS				
Married×Not Married	-0.0493 ***	0.0283	0.0738 ***	-0.0189
	(0.0136)	(0.0307)	(0.0179)	(0.0251)
Not Married×Married	-0.0574 ***	-0.0176	-0.0284 *	-0.0279
	(0.0154)	(0.0219)	(0.0161)	(0.0179)
Not Married×Not Married	-0.0750 **		-0.0120	
	(0.0354)		(0.0430)	
EDUCATION				
Evaluator>Evaluated	-0.0035	0.0106	0.0096	0.0044
	(0.0105)	(0.0133)	(0.0098)	(0.0108)
Evaluator<Evaluated	0.0011	0.0054	0.0009	0.0000
	(0.0113)	(0.0126)	(0.0100)	(0.0103)
# of Obs.	11354	11354	11354	11354
(Adjusted) R Squared	0.0331	0.0492	0.0421	0.0290

Table 8. Gender match effect for non-managerial workers (includes R&D subsidiary companies)

	A1&Over		A3&Below	
	OLS	FE	OLS	FE
Male*Female	-0.0594	0.0239	0.0316	0.0262
	(0.0294) **	(0.0524)	(0.0312)	(0.0403)
Female*Male	0.0133	-0.0354	0.0580	-0.0354
	(0.0412)	(0.0692)	(0.0635)	(0.0532)
Female*Female	-0.1176		0.1188	
	(0.0378) ***		(0.0701) *	
# of Obs.	2393	2624	2393	2624
R Squared	0.0607	0.0139	0.0352	0.0023