More Dispersion, Higher Bonuses? The Role of Differentiation in Subjective Performance Evaluations*

Patrick Kampkötter

Dirk Sliwka

March 18, 2017

Abstract

We investigate the claim that supervisors do not differentiate enough between high and low performing employees when evaluating performance. In a first step, this claim is illustrated in a formal model showing that rating compression reduces performance and subsequent bonus payments. The effect depends on the precision of performance information and may be reversed when cooperation is important. We then investigate panel data spanning different banks and find that stronger differentiation indeed increases subsequent bonus payments. The effect tends to be larger for larger spans of control and at higher hierarchical levels, but is reversed at the lowest levels.

*We thank Bernd Irlenbusch, Andrew Kinder, Michael Kramarsch, as well as seminar and conference participants at Amsterdam, Bonn, Cologne, Frankfurt am Main, Freiburg, Mannheim, New York, Oslo, and Osnabrück for valuable comments and suggestions. We also thank Towers Watson, especially Martin Emmerich, Florian Frank and Katharina Hessenauer, for their support. We thank the German Research Foundation (DFG) for financial support through priority program SPP 1764 (SL 46/2-1) and the research unit "Design and Behavior – Economic Engineering of Firms and Markets" (FOR 1371).

1 Introduction

Most bonus contracts for employees, in practice, are not based on objective measures of performance but rather on a subjective performance assessment by a supervisor. But it has often been stressed (compare e.g., Murphy and Cleveland (1995), Prendergast and Topel (1996), Moers (2005)) that supervisors tend to give performance ratings that are too compressed relative to the true performance of their employees. In that case, bonus payments will presumably not adequately reward high performance or sanction low performance. A straightforward conjecture is that this should lead to lower levels of performance incentives.

In this paper we investigate this conjecture and, in particular, guided by a formal theoretical model, we empirically study the impact of differentiation in bonus payments within work units on subsequent individual bonuses. The research question we address is *to what extent* and *under which organizational circumstances* differentiation in bonus payments indeed affects future bonuses of employees. This question is of substantial practical relevance as many firms still struggle with the question of whether to enforce more differentiation. For instance, as Jack Welch, who has put a large emphasis on establishing a culture of differentiation as CEO of General Electric, put it: "*Differentiation comes down to sorting out the A, B, and C players. [..]" (Welch (2003),* pp. 195). He also admits "*Differentiation isn't easy"* (p. 153) and *"[..] we spent over a decade building a performance culture with candid feedback at every level*" (p. 199). And indeed there is an ongoing discussion on whether enforcing differentiation is beneficial or harmful. Recently, the debate again received attention when, in 2013, Yahoo set up a forced ranking and Microsoft abandoned it.[1] Indeed, it is sometimes claimed in the management literature that differentiated ratings may destroy employee motivation

[1] See, for instance, "Yahoo is ranking employees. When Microsoft did that, it was a disaster." in *The Washington Post*, November 12, 2013 or "Forced Ranking Is Just As Bad For Yahoo As It Was For Microsoft" in the *Chicago Tribune*, November 13, 2013.

(compare, for instance, the discussion in Pfeffer and Sutton (2006), pp. 125). From a behavioral economics perspective, fairness and equity considerations (see e.g., Akerlof and Yellen (1990), Fehr et al. (1997), Fehr and Schmidt (1999), Bolton and Ockenfels (2000)) may play a role, and supervisors may be reluctant to differentiate for these reasons. In a recent worldwide survey among employees from a broad set of firms, only 41% of the respondents stated that supervisors differentiate enough between low and high performers.[2]

We first illustrate the connection between past degrees of differentiation and future bonus payments by analyzing a formal model of subjective performance evaluations. The model addresses how evaluators use available performance information and how this affects rating compression and incentives. We consider a supervisor who evaluates the performance of a group of agents. Supervisors have a preference for the accuracy of ratings and a preference for equity among the agents, and they face a trade-off between both. The agents' joint performance determines the size of a bonus pool, which is then allocated to the agents based on the supervisor's performance assessments. Supervisors observe signals on the agents' individual performance, and these signals may vary in their precision. We show that (i) when agents work independently, a stronger preference for equity reduces rating differentiation and, in turn, performance incentives; (ii) as a consequence, the size of the bonus pool and thus average bonuses are reduced; (iii) this detrimental effect of equity concerns is stronger when the span of control is larger and when the available performance measures are more precise; but (iv) rating compression can be beneficial when there are interdependencies and scope for cooperation among employees.

We then empirically study the relationship between differentiation and subsequent bonus payments by analyzing a large panel data set spanning

[2]See Towers Watson *Global Workforce Study 2010*. Frederiksen et al. (2012) analyze a number of firm-level data sets on subjective assessments and consistently find that performance ratings are concentrated on a subset of the applied rating scale.

many different firms in one industry, the majority of larger banks in Germany, in which we can track individual bonus payments over time for a large subset of the employees and have detailed information on specific functions and hierarchical levels. As the banks use bonus systems in which the size of the bonus increases in the financial performance of the unit, we use subsequent bonus payments as measures of performance. The key idea of our approach is the following: We investigate to what extent a higher variation in bonus payments in a certain unit and a given year leads to higher bomus payments in this unit in the subsequent year. Of course, unobserved individual heterogeneity will be an important issue as differentiation is also driven by the specific amount of heterogeneity in abilities in the different units. We therefore construct a balanced panel data set and make use of the *within-department variation* in the degree of differentiation to identify its effects on subsequent bonus payments.

We find that an increase in differentiation, on average, is associated with significantly positive increases in future bonus payments. The effect is also economically significant: When ranking units by their degree of differentiation, we estimate that future bonuses increase by 31% when moving from the lowest to the highest quintile of differentiation. In a next step we analyze to what extent these gains depend on the type of job, in particular the span of control, hierarchical level, and functional area. We find that the larger the span of control, the stronger is the association between differentiation and future bonuses. Moreover, the effects are strongest at the highest and intermediate levels. Surprisingly, we observe a strongly reduced or even reversed effect of differentiation at the lowest hierarchical levels. Additionally, differentiation has the strongest effects in retail banking and the capital market-based functions, which is in line with a complementary survey indicating that underlying performance information is more objective in these areas as compared to, for instance, the support functions.

The question of how subjective assessments make use of and complement objective performance information has received some attention both in the economics and accounting literatures. Ittner et al. (2003) study detailed data on a bonus plan and investigate how supervisors use objective performance signals when they have discretion on how to weight this information. Gibbs and Hendricks (2004) empirically explore the determinants of the use of subjective bonuses to complement objective performance information, concluding that subjective bonuses are used to compensate for perceived weaknesses in the objective performance measures. The use of bonus pools and subjective assessments to allocate these pools has been studied theoretically by, among others, Baiman and Rajan (1995), Rajan and Reichelstein (2006), or Rajan and Reichelstein (2009). Bol and Smith (2011) experimentally study how supervisors' subjective assessments are affected by an objective performance measure of a separate performance dimension. They find evidence to support the importance of fairness concerns as supervisors adapt their subjective assessments to apparently compensate for bad luck in the realization of an objective performance measure. Manthei and Sliwka (2015) investigate a natural experiment where a bank introduced objective performance measures in a subset of its branches and find that this led to a significant increase in firm performance.

Several studies empirically evaluate the relationship between differentiation in *fixed wages* of non-executive employees and firm performance[3] or individual productivity[4]. The effect of inequality in the compensation of board members on firm performance has been studied by Leonard (1990), Main et al. (1993), Eriksson (1999), Kale et al. (2009), and Bebchuk et al. (2011), with rather mixed results. Only recently have researchers started to study the connection between differentiation in bonuses among employees and performance (Bol (2011), Engellandt and Riphahn (2011)). In a recent

[3]See e.g., Winter-Ebmer and Zweimüller (1999), Heyman (2005), Jirjahn and Kraft (2007), Grund and Westergaard-Nielsen (2008), Martins (2008).

[4]See e.g., Becker and Huselid (1992), Pfeffer and Langton (1993), Drago and Garvey (1998), Bloom (1999), Depken II (2000).

experimental study, Berger et al. (2013) analyze the impact of a forced distribution system on individual performance and show that productivity is significantly higher if supervisors are forced to differentiate between employees who work independently.

The paper proceeds as follows: Section 2 gives a brief overview of the structure of bonus plans used in banks. In section 3, the illustrative formal model is introduced. Section 4 provides an overview of the data set, and the empirical strategy is described in detail. In section 5, we then investigate the performance effects of bonus dispersion for the whole data set, as well as for separate subsamples. Section 6 reports some extensions and robustness checks, applying alternative identification strategies, and we study the connection between bonus payments and financial performance. Finally, section 7 concludes.

2 The Structure of Bonus Plans

We first describe the structure of typical bonus plans in banks and then develop a simple formal model that captures key aspects of these plans. In order to do so, we conducted a survey in collaboration with the consultancy Towers Watson in 2013 among banks in Germany, Austria and Switzerland.[5] The bonus plans are typically characterized by one of the following three related types. When *bonus pools* are used, the bank assigns a fixed amount of money that depends on the unit's financial performance to each supervisor, who then has to allocate this pool to her subordinate employees according to subjective assessments. In so-called *additive bonus systems*, the individual bonus is usually the sum of three components: One part depends on individual performance, another on the performance of a unit or team and

[5] The survey collects information on 36 bonus plans from 25 different banks. We asked the respondents not only to characterize their current plan but also to state important changes made since 2004. However, the key underlying structures remained very stable during this time frame.

the third part on the profits of the whole bank. In *multiplicative bonus systems*, the supervisor subjectively assesses the performance of an employee according to a given scale, and the performance evaluation is subsequently multiplied by a certain factor, which depends on the profitability of the specific unit and the bank as a whole. All three systems have in common that bonus payments, ceteris paribus, increase when the financial performance of the relevant unit increases. Bonus plans can also combine bonus pools with additive or multiplicative systems, for example, when one part of the bonus comes from a pool distributed in a discretionary manner and another part is a function of a certain performance metric. In banking, bonus pools are the most commonly used element for allocating individual bonuses: 64% of the surveyed plans include bonus pools, 42% follow an additive logic and 47% a multiplicative logic. Moreover, only two of the plans (6%) are neither additive, nor multiplicative, nor use a bonus pool but are purely discretionary. Hence, financial performance of the unit affects individual bonus payments in nearly all of the considered plans. Almost all plans include individual performance assessments, which are mainly based on a mix of qualitative and quantitative indicators. Second, in all of these plans, individual performance evaluations include qualitative or discretionary assessments. But in 86% of these plans, supervisors also use objective performance indicators when making their assessments.[6]

3 An Illustrative Model

We consider a formal model of subjective performance evaluations which captures key aspects of typical bonus plans, as described above. Consider the

[6] Note that the structure of the bonus plans is also rather similar to those in other countries. A survey by Mercer in North America, Europe, and Emerging Markets (Mercer's Global Financial Services Executive Compensation Snapshot Survey 2013) concluded that "*the top-down pool approach is predominant in the banking industry*" with 62% of the responding banks. The other respondents all use either an additive or a multiplicative approach.

situation of n risk neutral agents $i = 1, 2, ..., n$ with initially unknown ability $a_i \sim N(m_i, \sigma_a^2)$ who work in a certain organizational unit. Each agent exerts effort e_i at cost $c(e_i)$ and generates a performance outcome $y_i = e_i + a_i$. The individual performance contributions of all agents are assessed subjectively by a supervisor who observes a vector of performance signals $s = (s_1, ..., s_n)$ where

$$s_i = y_i + \eta_i \text{ for } i = 1, 2, ..., n$$

with $\eta_i \sim N\left(0, \sigma_\eta^2\right)$. There is a further verifiable performance signal on the revenue of the unit, which is a function of the sum of the individual signals

$$\Pi = \pi \cdot \sum_{i=1}^{n} s_i$$

such that π measures the marginal returns to the agents' efforts. The supervisor observes the signals s and then reports a vector of performance ratings $r = (r_1, .., r_n)$. Each agent receives a wage that is linear in his performance assessment r_i such that

$$w_i = \alpha + \beta \cdot r_i.$$

The variable wage components $\beta \cdot r_i$ are paid from a bonus pool $B = \kappa \cdot \Pi$ which is a linear function of the verifiable measure of the financial success of the unit. Hence, the supervisor has to take a budget constraint

$$\beta \cdot \sum_{i=1}^{n} r_i \overset{!}{=} B = \kappa \cdot \Pi$$

into account when assigning performance ratings.[7]

Supervisors differ in their personal characteristics.[8] As in Prendergast

[7]A typical argument given for the use of bonus pools is that such an arrangement allows principals to commit to payments based on verifiable information, while at the same time incorporating unverifiable subjective information (see, for instance, Baiman and Rajan (1995), Rajan and Reichelstein (2006)).

[8]See e.g., Murphy and Cleveland (1995) for an overview on the psychology of per-

and Topel (1996) or Prendergast (2002), supervisors have a preference for accurate evaluations.[9] However, based on recent research on subjective evaluations, we also allow for the possibility that supervisors care about equity among their subordinates.[10] A supervisor's overall expected utility is

$$-\nu_A \cdot \sum_{i=1}^{n} \left(E\left[(r_i - y_i)^2 \middle| s_i \right] \right) - \nu_E \cdot \left(\sum_{i=1}^{n} \left(\beta \left(r_i - \frac{B}{n} \right) \right)^2 \right),$$

where ν_A measures the importance of accuracy[11] and ν_E that of equity in the supervisor's preferences. Or, in other words, ν_E can be interpreted as the extent to which supervisors are prone to the 'centrality bias' in performance assessments.

In a slight reinterpretation of the model, ν_A and ν_E measure two concepts of "fair evaluations." A supervisor with a higher v_A follows a contribution-based fairness norm, according to which an agent should receive a bonus that reflects his or her contribution to the team success. A supervisor with a higher ν_E follows a pure distributional fairness concept, according to which the agents should equally benefit from the joint output.

---

formance appraisals. Kane et al. (1995), for example, show that there are substantial differences between the ratings given by different supervisors to the same employees.

[9] One interpretation for the preference for accuracy is that the firm can verify the report with a certain probability and then impose a fine $(r_i - y_i)^2$.

[10] A recent literature in behavioral economics has stressed the importance of equity concerns (see, for instance, Fehr and Schmidt (1999), Bolton and Ockenfels (2000)). Ockenfels et al. (2015) study reactions to subjective performance assessments in a multinational company and find strong evidence for harmful effects of differentiation on job satisfaction when this leads to violations of a reference point for "fair" bonus payments. Berger et al. (2013) find in a lab experiment that more equity-oriented supervisors assign less differentiated ratings.

[11] Golman and Bhatia (2012) analyze a model where there is asymmetry in the supervisor's preference for accuracy, such that supervisors dislike unfavorable errors in performance assessments more than favorable errors.

3.1 Reporting Behavior

When the signals s and the overall bonus budget are realized, the supervisor chooses the reports r by maximizing this utility, taking into account the budget constraint. When $\nu_E = 0$ and without budget constraint, the supervisor would just report her conditional expectation $E\left[y_i | s_i\right]$. For normally distributed random variables, the conditional expectation is identical to the least squares estimator. But here the supervisor has to trade-off accuracy and equity motives and also take into account that the sum of the bonuses must be equal to the size of the bonus pool.

Note that

$$E\left[(r_i - y_i)^2 \middle| s_i\right] = V\left[r_i - y_i | s_i\right] + \left(E\left[r_i - y_i | s_i\right]\right)^2$$

such that (see the appendix for details) we can compute

$$E\left[(r_i - y_i)^2 \middle| s_i\right] = \tfrac{\sigma_a^2 \sigma_\eta^2}{\sigma_a^2 + \sigma_\eta^2} + \left(m_i + \hat{e}_i - r_i + \tfrac{\sigma_a^2}{\sigma_a^2 + \sigma_\eta^2}\left(s_i - m_i - \hat{e}_i\right)\right)^2$$

where $\hat{e}_i$ are the agents' equilibrium efforts. The supervisor's optimization problem is thus

$$\begin{aligned} \max_{r_1, r_2, .., r_n} \quad & -\nu_A \cdot \sum_{i=1}^{n} \left(\tfrac{\sigma_a^2 \sigma_\eta^2}{\sigma_a^2 + \sigma_\eta^2} + \left(m_i + \hat{e}_i - r_i + \tfrac{\sigma_a^2}{\sigma_a^2 + \sigma_\eta^2}\left(s_i - m_i - \hat{e}_i\right)\right)^2 \right) \quad (1) \\ & -\nu_E \cdot \left(\sum_{i=1}^{n} \beta^2 \left(r_i - \frac{B}{n} \right)^2 \right) \\ \text{s.t.} \quad & \beta \cdot \sum_{i=1}^{n} r_i = \kappa \cdot \pi \cdot \sum_{i=1}^{n} s_i \end{aligned}$$

Solving this problem gives us the performance assessments reported by the supervisor as a function of the observed performance signals:

Proposition 1 *The supervisor's performance assessment of an agent i is* *given by*

$$r_i(s) = \frac{\kappa\pi}{\beta n}\left(\sum_{j=1}^{n} s_j\right) + \frac{\sigma_a^2\left(s_i - \left(\frac{\sum_{j=1}^{n} s_j}{n}\right)\right) + \sigma_\eta^2\left((m_i + \hat{e}_i) - \frac{\sum_{j=1}^{n}(m_j+\hat{e}_j)}{n}\right)}{\left(1 + \left(\frac{v_E}{v_A}\right)\beta^2\right)\left(\sigma_a^2 + \sigma_\eta^2\right)}. \tag{2}$$

The bonus payment to an agent i increases with the absolute performance of the team $\sum_{j=1}^{n} s_j$ and with this agent's relative performance $s_i - \frac{1}{n}\left(\sum_{j=1}^{n} s_j\right)$. For a given signal vector s, the within-team standard deviation of ratings is decreasing with the supervisor's preference for equity ν_E and increasing with her preference for accuracy ν_A.

Proof: See appendix.

Higher team performance thus leads to better evaluations and, in turn, higher bonuses because it increases the size of the bonus pool paid out to the agents. Relative performance matters because the budget is limited: Higher relative performance of a colleague increases the "pressure" on the supervisor to reward this accurately and thus decreases the bonus that can be paid to an agent. Moreover, the result reflects the equity/accuracy trade-off described above: Stronger degrees of inequity aversion (relative to the preference for accuracy) lead to rating compression, which benefits low performers but harms high performers.[12] We denote $\nu_D = \frac{\nu_A}{\nu_E}$ as the supervisor's *preference for differentiation*, which is large when her preference for accuracy is strong but her preference for equity is weaker.

Expression (2) also illustrates the impact of precision in the underlying performance information: The less precise the performance measurement, i.e., the higher σ_η^2, the more the supervisors stick to their prior expectations about the agent's relative performance $(m_i + \hat{e}_i) - \frac{1}{n}\sum_{j=1}^{n}(m_j + \hat{e}_j)$, which

[12]Grund and Przemeck (2012) analyze a model where supervisors do not directly care for equity but internalize the agents' well-being to some extent. When the agents themselves are inequity averse supervisors then also have an interest in rating compression.

reflects the a priori expected difference in abilities (and equilibrium efforts), and the less emphasis they put on observed performance.[13] Hence, under symmetric priors about ability and effort, there is a second reason for rating compression in the model: When performance signals are imprecise, supervisors anticipate their own errors in observing the agent's performance and report less differentiated ratings. In the extreme case, i.e., when $\sigma_\eta^2 \to \infty$, the performance assessments of all agents are identical when the prior is symmetric.[14]

3.2 Effort Incentives

As the agents anticipate this reporting behavior, it affects work incentives and, in turn, the size of the bonus pool that can be paid out to the agents. To study this, consider an agent i's expected utility

$$E\left[\alpha + \beta \cdot r_i\left(s\right)\right] - c\left(e_i\right).$$

Substituting (2), we can directly derive the following result from the first order condition of the agent's incentive problem.

Proposition 2 *The agents' efforts are*

$$e_i = c'^{-1}\left(\frac{\kappa\pi}{n} + \beta\frac{\left(1-\frac{1}{n}\right)\sigma_a^2}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)\left(\sigma_a^2+\sigma_\eta^2\right)}\right) \tag{3}$$

[13] In a recent paper, Woods (2012) empirically investigates performance evaluations in a firm that introduced objective performance measures and studies how supervisors made subjective adjustments to the objective performance information. He finds that subjective adjustments that bring evaluations closer to the previous subjective assessments are indeed more likely the more supervisors perceive deficiencies in the objective measures.

[14] Note that when precision were endogenously chosen by a supervisor, the precision effect could reinforce the preference effect. A supervisor with a stronger preference for differentiation should invest more resources to avoid measurement error and the increased precision then allows for an even stronger differentiation. In section 3.4 we will analyze the optimal degree of precision from the firm's perspective.

for $i = 1, 2, ...n$. Agents exert higher efforts and the expected size of the bonus pool is larger when the supervisor has a stronger preference for differentiation ν_D. Moreover, incentives are higher powered when the available performance signals are more precise (higher $1/\sigma_\eta^2$).

Hence, rating compression, which is here either due to preferences for equity or due to imprecise performance information, directly affects incentives and, in turn, the expected size of the bonus pool. When there is more rating compression, agents anticipate that high performance will be rewarded and low performance sanctioned to a weaker extent. Hence, marginal returns to effort are smaller, and thus the amount of money paid out in the bonus pools will also be smaller. This reasoning thus implies a testable hypothesis: An increase in the degree of differentiation should increase subsequent average bonus payments.[15]

In the real world employees may not be able to directly observe a supervisor's preferences for differentiation, but they can infer information about this from the assignment of bonuses. To illustrate this point, note that when priors are symmetric we can use equation (8) in the proof of Proposition 2 to reformulate the equilibrium effort choice as a function of the standard deviation of reports and obtain

$$e_i = c'^{-1}\left(\frac{\kappa\pi}{n} + \beta\left(1 - \frac{1}{n}\right)\frac{SD\,(r)}{SD\,(s)}\right) \tag{4}$$

where $SD\,(r)$ and $SD\,(s)$ denote the within-team standard deviations of reports and underlying signals, respectively. For a given signal structure, a higher standard deviation in ratings should thus be associated with higher efforts.

[15]The bonus payment to a lower-performing agent may decrease when differentiation increases, but it may also increase depending on the strength of the incentive effect. But, average bonuses must go up as the pool size increases.

In the following we explore specific context factors that affect the strength of the association between a supervisor's preferences for differentiation and the incentives of the agents to exert effort.

3.3 Span of Control

An important context factor is the supervisor's span of control. First of all, there is a "scale of operations" effect as, for instance, analyzed by Smeets et al. (2015). If, in our context, a supervisor has a higher willingness to differentiate and, in turn, generates higher incentives for each of her subordinates, the total benefits of a stronger preference for differentiation should be larger for larger spans of control.

As we show in the following, the span of control also directly matters for the incentives of *individual* agents. It is again instructive to conduct a simple thought experiment: If we could increase a supervisor's intrinsic willingness to differentiate (i.e., reduce ν_E or increase ν_A) and therefore increase differentiation, where would the gain in incentives of an individual agent be the largest? Consider the impact of a change in ν_D on an agent's marginal returns to effort

$$\begin{aligned}\frac{\partial}{\partial v_D}\left(\frac{\kappa\pi}{n}+\beta\frac{\left(1-\frac{1}{n}\right)\sigma_a^2}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)\left(\sigma_a^2+\sigma_\eta^2\right)}\right)\\ =\frac{\left(1-\frac{1}{n}\right)\sigma_a^2\beta^3}{\left(\sigma_a^2+\sigma_\eta^2\right)\left(v_D+\beta^2\right)^2}>0.\end{aligned} \tag{5}$$

Taking the cross derivative with respect to n yields

$$\frac{1}{n^2}\frac{\sigma_a^2\beta^3}{\left(\sigma_a^2+\sigma_\eta^2\right)\left(v_D+\beta^2\right)^2}>0.$$

and thus we can conclude:

Proposition 3 *The effect of an increase in the preference for differentiation ν_D on individual efforts and average bonuses is the larger the larger the span of control n.*

Thus a stronger willingness to differentiate has a stronger incentive effect when there is a larger span of control. To understand this effect note the following: if the span of control is small, the budget balancing constraint implies a stronger "cost of differentiation" than in a larger team, i.e. in smaller teams supervisors will differentiate less and therefore induce weaker incentives. Consider, for instance, a team of two agents. Suppose that one of them has shown an exceptionally high performance. Rewarding this is only possible by reducing the bonus of the colleague which is costly in terms of equity concerns as this other agent then receives a bonus which would have to be substantially smaller than the equitable one. In a larger team the supervisor can instead reallocate smaller amounts from more other agents which facilitates differentiation. This is anticipated by the agents who in turn exert higher efforts.

3.4 Precision

We now investigate the role of precision in performance measurement $(1/\sigma_\eta^2)$.[16] Taking the cross derivative of the agent's marginal returns to effort (5) with respect to σ_η^2 we obtain

$$-\frac{\left(1-\frac{1}{n}\right)\sigma_a^2\beta^3}{\left(\sigma_a^2+\sigma_\eta^2\right)^2\left(v_D+\beta^2\right)^2}<0$$

which directly yields:

[16] In the accounting literature the role of precision for objective performance measurement has been studied by, for instance, Banker and Datar (1989) or Lambert (2001).

Proposition 4 *The effect of an increase in the preference for differentiation ν_D on efforts and the size of the bonus pool is the larger the more precise the underlying performance measures are (i.e., the smaller σ_η^2).*

If performance measures are not very accurate, a lack of willingness to differentiate does not cause great harm, as even a supervisor with a low v_E and a high ν_A will not differentiate much according to observed performance. Note that even without any differentiation, each agent's interest in having a larger bonus pool still represents an incentive to exert effort. However, effort incentives are in that case lower-powered due to free rider effects. If, on the other hand, performance measurement is rather accurate, supervisors are in principle able to make precise assessments of individual performance. A supervisor who does not fully make use of this information "unnecessarily" inhibits differentiation and thus undermines incentives to a stronger extent.

In our model precision so far has been exogenously given. However, organizations are often able to invest in the information structure (for example, IT systems to track performance indicators, time resources for monitoring, etc.). It is instructive to study a simple extension where we allow that the firm ex-ante endogenously determines the precision of the available performance information. Suppose that the firm chooses the level of precision $\psi = 1/\sigma_\eta^2$ which induces monitoring costs $z(\psi)$ where $\frac{\partial z}{\partial \psi} > 0$ and $\frac{\partial^2 z}{\partial \psi^2} > 0$ and $lim_{\psi \to \infty} z(\psi) \to \infty$. For simplicity assume that the agents' costs of effort are $c(e) = \frac{1}{2}e^2$. The firm's profits are then

$$(1-\kappa)\pi \cdot \sum_{i=1}^{n} \left(\frac{\kappa\pi}{n} + \beta \frac{\left(1-\frac{1}{n}\right)\sigma_a^2}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)\left(\sigma_a^2+\frac{1}{\psi}\right)} + m_i \right) - z(\psi).$$

Using this expression we can show:

Proposition 5 *When the firm can endogenously determine the precision of performance measurement $\psi = 1/\sigma_\eta^2$, the optimal precision ψ is increasing in the profitability π of the agents' tasks.*

Proof: See appendix.

This, for instance, suggests that precision is larger on higher hierarchical levels since here the marginal returns to effort exerted by an agent are larger, and, therefore, inducing stronger incentives has higher benefits for the firm. In turn, by Proposition 4 the effect of an increase in the supervisor's willingness to differentiate should also lead to a stronger increase in efforts on these levels.

3.5 Interdependence

Finally, we study a slight extension of the model where there is an interdependency between the agents' work. Suppose that each agent can choose not only a productive effort e_i but also an unobservable helping effort h_i at costs $c_h(h_i)$. This effort may not only reflect direct help but also, among other things, the sharing of information, i.e., any activity that is costly for an individual agent but increases the performance of his coworkers. Suppose that the performance outcome of an individual agent is now $y_i = e_i + a_i + \zeta \cdot \sum_{j \neq \mathrm{i}} h_j$ such that helping increases all coworkers' performance. The analysis of the supervisor's rating behavior proceeds as before such that now

$$r_i^h(s) = \frac{\kappa\pi}{\beta n}\left(\sum_{j=1}^{n} s_j\right) + \frac{\sigma_a^2\left(s_i - \left(\frac{\sum_{j=1}^{n} s_j}{n}\right)\right) + \sigma_\eta^2\left(\left(m_i + \hat{e}_i + \zeta \cdot \sum_{j\neq\mathrm{i}} \hat{h}_j\right) - \frac{\sum_{j=1}^{n}\left(m_j + \hat{e}_j + \zeta \cdot \sum_{k\neq\mathrm{j}} \hat{h}_k\right)}{n}\right)}{\left(1 + \left(\frac{v_E}{v_A}\right)\beta^2\right)\left(\sigma_a^2 + \sigma_\eta^2\right)}$$

An agent i now maximizes

$$E\left[\alpha + \beta \cdot r_i^h(s)\right] - c(e_i) - c_h(h_i).$$

Due to the additive separability of the cost and production functions, the choice of e_i remains unchanged. To derive the optimal helping effort, note that $\frac{\partial E[s_i]}{\partial h_i} = 0$ and $\frac{\partial E[s_j]}{\partial h_i} = \zeta$ for any $j \neq i$ such that

$$\frac{\partial E\left[r_i^h(s)\right]}{\partial h_i} = \frac{\kappa\pi\,(n-1)\,\zeta}{\beta n} - \frac{\sigma_a^2 \frac{n-1}{n}\zeta}{\left(1+\left(\frac{v_E}{v_A}\right)\beta^2\right)\left(\sigma_a^2+\sigma_\eta^2\right)}$$

and we can conclude:

Proposition 6 *The agent's helping efforts are*

$$h_i = c_h'^{-1}\left(\frac{n-1}{n}\zeta\left(\kappa\pi - \frac{\beta\sigma_a^2}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)\left(\sigma_a^2+\sigma_\eta^2\right)}\right)\right)$$

for $i = 1, 2, \ldots n$. A higher preference for differentiation v_D reduces helping efforts.

When there are interdependencies between the agents' work differentiation is thus accompanied by a countervailing effect: It may reduce incentives for mutual help and cooperation. Note that when ζ is sufficiently large, this negative effect of reduced cooperation will outweigh positive incentive effects of differentiation. This is in line with the argument by Lazear (1989), who shows that relative performance pay through promotion tournaments reduces cooperative behavior.

4 Empirical Approach and the Data

We now investigate the connection between past differentiation and future bonus payments. The discussion will be guided by the insights from the illustrative model: 1) Differentiation in bonus payments should be positively associated with the size of subsequent bonus payments; 2) this association is stronger when there is a larger span of control; 3) this association is stronger in areas where individual performance can be assessed more accurately; and 4) this association is weaker when there is a stronger interdependence between the agents. While we can directly test the first and second statements, we will use the latter two statements to guide the discussion of heterogeneity of the effects in the different levels and functions.

In the empirical analysis we investigate a panel data set on compensation in the German banking and financial services sector for the years 2005-2007. As bonus payments in 2007 are based on performance in 2006, this time frame covers a rather stable period before the start of the financial crisis. The data set, which is owned by the management consultancy Towers Watson, the market leader in wage benchmarking in the German banking and financial services industry, is used for professional compensation benchmarking and thus spans a large part of the industry.[17]

The data set covers more than 50 German banks and financial services companies[18] and contains detailed individual information on base salary, bonus payments, age, firm tenure, hierarchical level (6 levels), functional area, and specific function.[19] For a subset of 18 of these banks we can track individual employees over time. The functional areas represent a broad classification of the main sectors in the banking and financial services industry: retail banking, asset management, corporate banking and private banking, investment banking, treasury and capital markets, the service functions, as well as the cross-divisional functions. We make also use of a much more detailed classification of industry-specific jobs, as these functional areas are subdivided into about 60 specific functions.[20] A useful feature of the data set

[17]Towers Watson (formerly Towers Perrin) data sets have also been used by Abowd and Kaplan (1999), Murphy (1999), and Murphy (2001).

[18]Sparkassen (publicly owned savings banks), Volks- and Raiffeisenbanken (cooperative banks) and the German central bank are not part of the sample. Due to confidentiality reasons, company names had to be anonymized.

[19]Note that top executive positions are not included in the data set. Hence, stock options only play a minor role in incentive compensation of the observed employees.

[20]Most of the employees in the data set work in retail banking and in the service and corporate functions, followed by corporate banking. A list of exemplary functions is given

is the systematic comparability of employee positions across and within different firms. As the consultancy offers compensation benchmarking services, it applies a standardized job evaluation method to determine the specific function and hierarchical level of a job.

The empirical strategy is as follows: In our core specifications we analyze a balanced panel data set comprising about 12,000 individuals in 18 banks over the three-year period to investigate the effects of differentiation within a department on individual bonus payments in the subsequent year. In a first step, we generate cells capturing the organizational units of a company. A cell is characterized by a unique combination of year, company, functional area, detailed function, career ladder[21], and hierarchical level. We observe 1,455 unique cell-year combinations and an average (median) size of 31 (7) employees per cell. Note that we restrict our analysis to cells with a minimum of three observations. Then we compute different measures of bonus dispersion within each unit and for each year: the coefficient of variation, i.e., the ratio of the standard deviation to the mean, the P90/P10 ratio, i.e., the ratio of the 90th to the 10th percentile, and the standard deviation of logs. There are differences in the degree of variation between the broader functional areas[22] and between the more detailed specific functions. The average coefficient of variation in human resources is 0.43, in marketing 0.37, in corporate finance 0.76, in equity trading 0.58, in treasury trading 0.67, and in IT generalist functions 0.27. A descriptive overview on the variables used in the analyses is shown in table A.2.

The dependent variable is the logarithm of the bonus payment an individual receives in the subsequent year.[23] We only include observations with

in table A.1 in the appendix.

[21] The data set differentiates between four career ladders (management, professionals, sales, and support).

[22] Due to a small number of observations, the functional areas investment banking, asset management, and treasury and capital markets are pooled.

[23] Absolute bonus payments exhibit a substantially more skewed distribution than log bonus payments. Moreover, the log specifications allow us to estimate relative performance

non-missing and positive actual bonus payments in order to only capture positions that are eligible for a bonus payment.[24] The formal model illustrates why we expect a positive effect of the bonus dispersion on subsequent bonuses: Stronger differentiation comes along with higher-powered incentives, which in turn should lead to a better financial performance of the unit and thus larger bonus pools.[25] The underlying idea is that employees learn about supervisors' future evaluation behavior from past evaluation behavior and adapt their efforts accordingly.

Of course, there are other factors beyond the mere incentive effect that may drive a correlation between dispersion and bonus payments that are closely related to the type of the considered cell.[26] Hence, we use changes in the degree of dispersion over time *within* a cell for a given employee to estimate its effect on subsequent bonus payments. We thus run regressions with employee fixed effects, where the log of the individual bonus payment of a person i in a year t is the dependent variable. The key independent variable is the measure of dispersion $s_{c_i t-1}$ (coefficient of variation, P90/P10 ratio, standard deviation of logs) of bonus payments in the relevant cell c_i in year $t-1$. Hence, our key specification is

$$\ln b_{it} = \beta \cdot s_{c_i t-1} + X_{it}' \gamma + a_i + \lambda_t + \varepsilon_{it}$$

effects of our normalized measures of dispersion.

[24]Bonus eligibility levels in the financial services industry are very high, with an average eligibility rate of 90% in our data set and rates of more than 97% at the higher levels considered in our analyses. We exclude newly hired employees from the data set as employees in their first year are often not eligible for bonus payments.

[25]As laid out above, even the plans that do not entail bonus pool arrangements nearly all have the property that bonus payments are based both on a subjective assessment of individual performance and the financial success of the unit or the bank and, hence, variations in bonus payments reflect variations in profit contributions. We will show in section 6 that within-bank variations of financial success over time strongly predict variations in the actual bonus payments.

[26]For instance, as illustrated in the model, firms have an incentive to invest more in the precision of performance measures or to employ supervisors with a stronger preference for differentiation in positions where the employees' efforts are more profitable.

where X_{it} is a vector of additional (time-variant) covariates including the log of base salary and age squared, $a_i$are individual and λ_t are year fixed effects.

The underlying identifying assumption of this fixed effects approach is that potential omitted variables are time-invariant, or that

$$E\left[b_{it}\left(s\right)|a_i, X_{it}, t, s_{c_i t-1}\right] = E\left[b_{it}\left(s\right)|a_i, X_{it}, t\right], \tag{6}$$

i.e., the actually observed lagged dispersion $s_{c_i t-1}$ contains no information important for estimating the (counterfactual) bonus $b_{it}\left(s\right)$ of the same person under a different degree of dispersion, beyond the information contained in the time-invariant characteristics and observable time-variant covariates. Hence, in order to reduce the possibility that changes in dispersion are driven by organizational changes, promotions, functional rotation, entry, exit, or changes in the team composition (which may then also affect future bonus payments), we restrict our samples to employees who remained at the same hierarchical level, functional area, specific function, and career ladder throughout all the years at the same employer. With this approach, we exclude all employees who were promoted and those who joined or left the organization. Note that this also excludes the possibility that changes in team composition drive our results. However, we also run robustness checks where we include employees who are promoted in the considered time frame. We also apply alternative identifying assumptions and analyze lagged dependent variable models, as well as instrumental variables regressions, in subsection 6.1 (thus relaxing assumption (6)).

5 Differentiation and Bonus Payments

We start by estimating the average effect of differentiation on future bonus payments across the whole sample and then study potential heterogeneity across hierarchical levels and functions.

5.1 Aggregate Effects

Table 1 reports estimation results of the baseline regressions, with individual fixed effects and the logarithm of bonus payments as the dependent variable. The key independent variable is the respective measure of dispersion for the relevant cell in the previous year. To account for potential within-cell correlation in the error terms, robust standard errors clustered on the cell-level are reported. All models include the time-varying logarithm of base salary, age squared, and year dummies as further control variables.

The results in table 1 show that there is a highly significant, positive relationship between differentiation and future bonus payments, i.e., an increase in the degree of differentiation in a department's bonus payments in one year is associated with significantly higher individual bonus payments in the subsequent year for all three indicators.[27] A one standard deviation increase in the coefficient of variation (P90/P10 ratio) is associated with an increase in bonus payments of about 10% (7%). To give some further indication about the economic significance of this effect, we ranked all cells by the degree of differentiation and then created dummy variables for each quintile in the distribution of the measures of dispersion. The coefficient for the $5th$ quintile now gives an estimate of the percentage change in bonuses when a supervisor who is among the 20% of weakest differentiators moves to the degree of differentiation applied by the 20% strongest differentiators. Note that these effects are quite sizeable. For the coefficient of variation, the model in table 1 predicts a 31% increase[28] in bonuses when moving from rather undifferentiated incentives to highly differentiated bonus payments. The coefficients for the P90/P10 ratio and the standard deviation of logs are even slightly higher, with a predicted increase in subsequent bonuses of 33% and 36%,

[27] The results are robust when we exclude retail banking, as shown in table A.3 in the appendix. Recall that most of the employees in our data set work in retail banking. Furthermore, the structure of the units we consider in retail banking is different as in this area we observe fewer cells comprising a larger number of employees per cell.

[28] Note that $e^{0.2680} = 1.31$. See e.g., Halvorsen and Palmquist (1980) for details.

respectively. The effects are roughly monotonic in all specifications, i.e., the effects increase when moving from the lowest quintile to the highest one.[29]

Using the panel data set which excludes any employees who moved between areas and levels over the observed period of time, should eliminate effects of changes in team composition, but excludes promoted (and, thus on average, more able) employees. To check whether the results are robust, we replicated the baseline regressions with the panel data set including movers. As can be seen in table A.4 in the appendix, the results remain broadly unchanged.[30]

We have thus shown that differentiation, on average, indeed has a substantial positive effect on subsequent bonus payments. However, this effect may be heterogeneous across different units or job types. In a next step we therefore explore to what extent the overall effect depends on specific organizational characteristics.

5.2 Span of Control

Our formal model suggests that the individual performance effects are larger the larger the span of control, i.e., the number of employees evaluated by a supervisor. For a given information structure a larger direct span of control makes it easier to assign differentiated ratings without violating the budget

[29]Wald tests for the regression with the coefficient of variation show that the coefficients for quintile 3 and 4 ($p<0.1$) are statistically significantly different from each other. For the P90/P10 ratio, coefficients for quintile 2 and 3 ($p<0.1$), for quintile 2 and 4 ($p<0.05$) and for quintile 2 and 5 ($p<0.01$) are significantly different from each other. For the standard deviation of logs, coefficients for quintile 2 and 5 ($p<0.05$), for quintile 3 and 4 ($p<0.01$) and for quintile 3 and 5 ($p<0.01$) are significantly different from each other.

[30]We still prefer the analysis based on the data set that excludes movers. A potential concern in this panel data set including movers is the following: Suppose that an employee of high ability joins a team and immediately gets a higher bonus than the others, and thus dispersion increases. In the next year the team may benefit from the presence of this high performer, and thus team performance and the size of the bonus pool may increase. Or, an agent of very low ability joins a team and gets a low bonus (thus increasing dispersion) and then quits the firm. If then team performance increases, so does again the size of the bonus pool. Excluding movers thus seems to be the more conservative approach.

Table 1: Fixed effects regressions with measures of dispersion

Dependent variable:	Logarithm of bonus payments					
	Coef. of variation		P90/P10 ratio		Std. dev. of logs	
$\text{Differentiation}_{t-1}$	0.5026**		0.0238***		0.6575***	
	(0.2168)		(0.0092)		(0.2008)	
2^{nd} Quintile_{t-1}		0.2044***		-0.0542		-0.0540
		(0.0691)		(0.0413)		(0.1441)
3^{rd} Quintile_{t-1}		0.1547**		0.1935		0.0555
		(0.0659)		(0.1265)		(0.0634)
4^{th} Quintile_{t-1}		0.2265***		0.2519**		0.1665**
		(0.0750)		(0.1198)		(0.0742)
5^{th} Quintile_{t-1}		0.2688***		0.2876***		0.3070***
		(0.0958)		(0.1100)		(0.0901)
Ln Base salary_t	-0.4489	-0.3372	-0.3346	-0.3671	-0.3815	-0.6174**
	(0.3062)	(0.3333)	(0.2883)	(0.2906)	(0.2879)	(0.2522)
Age squared_t	0.0001***	0.0002***	0.0001***	0.0001***	0.0001**	0.0002***
	(0.00005)	(0.00005)	(0.00004)	(0.0001)	(0.00005)	(0.00005)
Observations	25587	25587	25587	25587	25587	25587
R^2within	0.09	0.06	0.08	0.08	0.09	0.08
1 std. dev. increase	10%		7%		12%	

Additional year dummies included. Robust standard errors clustered on cell-level in parentheses.

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

balancing constraint. Hence, a change in the willingness to differentiate should have a stronger incentive effect and thus be accompanied by a stronger increase in subsequent bonus payments when the span of control is larger.[31]

We constructed a measure for the span of control in a cell in the following way: We know whether employees in our data set have managerial authority and are allocated to the career ladder "management." Exploiting this information, we define our measure for the span of control as the total number of employees at a given hierarchical level l divided by the number of managerial staff at hierarchical level $l + 1$ (i.e., one level higher) in the same bank, functional area, and function. Hence, this measure provides information about the average number of subordinates a supervisor has at each level. Similar to the generation of our cells, we restrict the minimum number of observations when computing the span of control to 3 and drop cells where this measure exceeds 40 employees per supervisor.

We include an interaction term between our measure for span of control and the coefficient of variation in the baseline regression. The results are shown in table 2. In line with our hypothesis, the interaction term between our measure of span of control and the level of differentiation is indeed positive and (weakly) statistically significant in two of the three specifications.

5.3 The Role of the Hierarchical Level

In a next step we investigate the role of the hierarchical level. We acknowledge that our analysis here is more exploratory in nature as differences in hierarchical levels may affect different context parameters. According to our formal model, differentiation should have a stronger positive effect on performance and subsequent bonus payments if performance measurement is more precise. And, as we have shown in Proposition 5, a firm should invest more

[31] In the formal model, this can also be directly seen by inspecting equation (4) showing that (under a symmetric prior) effort is increasing in $\beta\left(1-\frac{1}{n}\right)\frac{SD(r)}{SD(s)}$ such that $\frac{\partial^2 e}{\partial SD(r)\partial n} > 0$.

Table 2: Span of control and differentiation

Dependent variable:	Logarithm of bonus payments		
	Coef. of variation	P90/P10 ratio	Std. dev. of logs
Differentiation$_{t-1}$	0.5342***	0.0242***	0.4353***
	(0.0636)	(0.0058)	(0.1622)
Diff. * Span of control	0.0071*	0.0007	0.0362**
	(0.0039)	(0.0007)	(0.0171)
Observations	10225	10225	10225
R^2within	0.14	0.12	0.09

Additional controls include ln base salary, age squared, and year dummies. Robust standard errors clustered on cell-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

in collecting accurate ratings when the marginal returns to agents' efforts are larger, which should be the case at higher levels. Moreover, firms may have an incentive to use "stricter" evaluators (i.e. supervisors who have a stronger preference for accuracy as opposed to equity) on higher levels.

A first implication is that differentiation should be stronger at higher levels. As table A.5 in the appendix shows, this tends to be the case. In each of the considered years, all three dispersion measures exhibit larger values on the highest three levels as compared to the lowest three.

In a next step, we investigate the role of the hierarchical level by including interaction terms between the measures of differentiation and each of the six hierarchical levels in the baseline regression model. The reference category is level 1, the lowest level in the data set. Results are reported in table 3. The effect of differentiation on subsequent bonus payments indeed becomes stronger at higher hierarchical levels. From level 4 upwards we find a highly significant positive relationship between differentiation and future bonuses, whereas the coefficients for the bottom levels are mostly not significantly different from zero.[32]

[32]Wald tests for the coefficient of variation show that the coefficients for level 2 and 3 (p<0.1), for level 2 and 4 (p<0.001), for level 2 and 5 (p<0.001), for level 2 and 6

Table 3: Interactions between measures of dispersion and hierarchical levels

Dependent variable:	Logarithm of bonus payments		
	Coef. of variation	P90/P10 ratio	Std. dev. of logs
Differentiation $_{t-1}$	-0.3793	-0.0222**	-0.1083
	(0.3234)	(0.0108)	(0.3185)
Differentiation$_{t-1}$ $\times$ Level 2[a]	-0.5199	-0.0349	0.7703
	(0.4123)	(0.0565)	(1.0120)
Differentiation$_{t-1}$ $\times$ Level 3	0.2569	0.0296*	0.1734
	(0.4757)	(0.0166)	(0.3654)
Differentiation$_{t-1}$ $\times$ Level 4	1.1724***	0.0427***	0.9313**
	(0.3952)	(0.0152)	(0.4498)
Differentiation$_{t-1}$ $\times$ Level 5	1.2041***	0.0804***	1.0419**
	(0.3665)	(0.0181)	(0.5290)
Differentiation$_{t-1}$ $\times$ Level 6	2.2380***	0.1213***	1.6416***
	(0.7129)	(0.0242)	(0.5757)
Observations	25587	25587	25587
R^2 within	0.15	0.10	0.11

Age squared and year dummies included. [a] Reference category: Level 1 (lowest level).

Robust standard errors clustered on cell-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

But interestingly the regressions reported in table 3 also show that the positive effects of differentiation are reversed at the lowest levels: Stronger degrees of differentiation are accompanied by a lower subsequent performance on level 1, and the effect is even significant for the P90/P10 ratio.[33] As shown in the formal model, differentiation may cause within-team competition and therefore can have detrimental effects when cooperation is very important.[34] Berger et al. (2013), for instance, find in a lab experiment that enforced differentiation in performance appraisals is beneficial when workers work independently but is detrimental when they have an opportunity to harm each other. However, it is not clear that interdependencies are indeed stronger on lower levels. Another interpretation is that, because firms have a stronger incentive to invest in resources to track performance at higher levels, they also should have a stronger incentive to assess cooperative behavior. In turn, it may be easier to avoid detrimental effects of differentiation at higher levels.[35]

An alternative explanation for the observation that bonus differentiation has weaker benefits on lower levels posits that, at these levels, career concerns may dominate incentives generated through a bonus plan. If firms either intentionally use promotions to set incentives, as in Lazear and Rosen (1981), or infer information on agents' abilities from past performance (Holmström (1982)), which in turn affects promotion decisions (Waldman (2013)), the pursuit of a promotion can indeed be a powerful incentive device. In line

($p<0.001$), for level 3 and 4 ($p<0.05$), for level 3 and 5 ($p<0.05$), and for level 3 and 6 ($p<0.01$) are statistically significantly different from each other. Referring to the P90/P10 ratio, the coefficients for level 3 and 5 ($p<0.01$), coefficients for level 3 and 6 ($p<0.001$), coefficients for level 4 and 5 ($p<0.05$), and coefficients for level 4 and 6 ($p<0.001$) are significantly different from each other. For the standard deviation of logs, the coefficients for level 3 and 4 ($p<0.05$), level 3 and 5 ($p<0.1$) and the coefficients for level 3 and 6 ($p<0.01$) are significantly different from each other.

[33] As table A.6 in the appendix shows, this is particularly driven by a drop in performance for very high degrees of differentiation at the lowest hierarchical levels.

[34] In the model, expected bonus payments will increase in ν_E when ζ is sufficiently large.

[35] One potential instrument used to assess cooperative behavior is so-called 360º feedback, which is an appraisal format where managers are appraised not only by their supervisors but also by their peers at the same hierarchical level and their subordinates.

with the arguments by Gibbons and Murphy (1992), career concerns may then substitute direct incentives generated through bonus plans (and thus differentiation in the assignment of bonuses).

In order to explore this mechanism, we analyze the nexus between promotion decisions, bonus payments and the incentive effects of differentiation in more detail. Indeed, the promotion probability per cell, measured as the 3-year average of all individual promotion frequencies within a cell, monotonously decreases with the hierarchical level, which indicates that promotion opportunities or career concerns are more important at lower levels.[36]

In a next step, we replicate the regressions from table 3 on a larger data set that includes movers, i.e., employees who switched cells for instance because of promotions. In these regressions we include an interaction term between the measures of dispersion and the promotion probability in the respective cell. The results are shown in table 4. The interaction terms between the promotion probability and the coefficient of variation (column (1)) and the standard deviation of logs (column (3)) are positive and significant. Hence, *bonus dispersion* and promotions seem to be complements rather than substitutes.[37] Our preferred interpretation is the following: It seems likely that supervisors who are more willing to differentiate between high and low performers also are more willing to make performance-contingent promotion decisions. Hence, an increase in differentiation may generate an additional incentive effect, reinforcing the direct effect of bonus dispersion, as it signals a stronger performance-based culture.[38]

[36]The average promotion probabilities are as follows: level 1: 37%, level 2: 11%, level 3: 10%, level 4: 8%, and level 5: 3%.

[37]The direct coefficient of the promotion probability is significantly negative in column (1), which may suggest that absolute bonus payments and promotions are substitutes. Note, however, that - as these are fixed effects models - the coefficient is identified only from the movers between cells. This pattern can thus, for instance, be a consequence of the common observation that employees that just have been promoted typically earn lower bonuses (relative to their fixed wages) in the first year in their new role.

[38]In line with this reasoning, we also find that higher bonus payments predict future promotions (essentially replicating the observation in Frederiksen et al. (2012) or Smeets et al. (2015) that subjective evaluations predict promotions).

Table 4: Interaction with promotion probability

Dependent variable:	Logarithm of bonus payments		
	Panel including movers		
	Coef. of variation	P90/P10 ratio	Std. dev. of logs
Differentiation $_{t-1}$	-0.8853***	-0.0333***	-0.4959***
	(0.1645)	(0.0128)	(0.1433)
Diff. * Promotion probability	0.5397***	0.0179	0.4458**
	(0.1412)	(0.0129)	(0.1991)
Promotion probability	-0.2641**	0.1372	-0.0155
	(0.1069)	(0.1004)	(0.1149)
Differentiation$_{t-1}$ $\times$ Level 2[a]	0.0208	0.0029	0.1149
	(0.1626)	(0.0110)	(0.1534)
Differentiation$_{t-1}$ $\times$ Level 3[a]	0.6594***	0.0255**	0.4905***
	(0.1688)	(0.0111)	(0.1501)
Differentiation$_{t-1}$ $\times$ Level 4	1.5386***	0.0472***	1.1862***
	(0.1703)	(0.0119)	(0.1597)
Differentiation$_{t-1}$ $\times$ Level 5	1.4766***	0.0564***	1.3180***
	(0.1666)	(0.0129)	(0.1714)
Differentiation$_{t-1}$ $\times$ Level 6	2.1775***	0.0626***	2.0358***
	(0.2220)	(0.0149)	(0.2348)
Observations	34091	34098	34091
R^2 within	0.22	0.16	0.19

Ln base salary, age squared and year dummies included. [a] Reference category: Level 2.

Robust standard errors clustered on cell-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

5.4 Differences across Functional Areas

We now estimate the performance effects of differentiation for subsamples comprising different functional areas. Besides *retail banking (RB)*, we consider the following broader areas: *private and corporate banking (PB/CB)* covers banking services for corporations and wealthy private clients and *corporate production* and *corporate services (CP/CS)* includes back office jobs, such as facility manager and secretary positions as well as cross-divisional functions such as human resources, finance, or accounting. Furthermore, we look at the subsample including the capital market-based functions *investment banking, asset management*, and *treasury and capital markets (IB/AM/TCM),* which include jobs in, for instance, money markets, trading, corporate finance, and fund management. It is important to note that the areas differ in the composition of the hierarchical levels.[39] Therefore, we first report regressions including all levels in panel A of table 5 but then in panel B focus only on jobs at intermediate levels 3 and 4, which exist in all these areas, in order to exclude capturing level effects. The last row in each panel again shows the estimated effect of a one standard deviation increase in the coefficient of variation.

Table 5 shows that there are sizeable differences between the functions. However, these are indeed to some extent driven by the level composition, as becomes clear when investigating the more comparable subsamples, which are reported in panel B. In these middle management positions, differentiation has positive effects in all four areas. Both the absolute coefficient and the standardized effects are largest in *retail banking,* and *investment banking, asset management and treasury and capital markets.* Furthermore, differentiation has weaker effects in *private and corporate banking* as well as in *corporate production and services*, where the standardized coefficient is the

[39] There are nearly no employees on level 1 and 2 in IB/AM/TCM and PB/CB, but 39% of the employees in RB and 35% in CP/CS are at these lower levels. However, 50% of the employees in RB, 60% in IB/AM/TCM, 55% in CP/CS, and 35% in PB/CB are on levels 3 and 4.

smallest.

Table 5: Subgroups of functional areas across levels (coefficient of variation)

Panel A: All levels				
Dependent variable:	Ln bonus payments			
	CP/CS	IB/AM/TCM	PB/CB	RB
CV Bonus$_{t-1}$	0.2315	0.5118*	0.6553***	1.8398***
	(0.2922)	(0.2691)	(0.1753)	(0.6565)
Observations	9172	1002	1169	14244
R^2 within	0.08	0.21	0.28	0.17
1 std. dev. increase	6%	14%	21%	19%
Panel B: Levels 3 and 4				
CV Bonus$_{t-1}$	0.4628**	0.7929**	0.4414***	2.2756***
	(0.2272)	(0.3484)	(0.0652)	(0.3169)
Observations	4943	624	397	7191
R^2 within	0.19	0.22	0.17	0.40
1 std. dev. increase	11%	22%	17%	18%

Additional controls: Ln base salary, age squared, and year dummies included. Robust standard errors clustered on cell-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The formal model again suggests a potential explanation when the functional areas differ in the precision of the available performance measures. While we caution that we cannot test this conjecture directly, we collected complementary information by running a short online survey among employees in German banks. The survey invitations were sent out to subscribers of the electronic newsletter of the Association of German Banks in 2013.[40] Our key survey item focused on the perceptions about the potential precision of

[40] Participation was incentivized by a lottery in which participants could win an iPad. In total, 121 bankers participated in this survey, with 46% belonging to cross-divisional and service functions, 21% to investment banking, asset management and treasury and capital markets, 21% to private and corporate banking, and 11% to retail banking positions. Average firm tenure was 9.1 years (sd=7.3), and 49% of the participants were in supervisory positions.

performance measurement in the different functional areas. To be specific, the survey item stated "*To what extent do you think it is, in principle, possible in your functional area to evaluate individual performance objectively?*" The respective scale ranged from 1 (not at all objectively) to 5 (very objectively). Figure 1 shows the means for the perceived objectivity in the functional areas we investigated above.

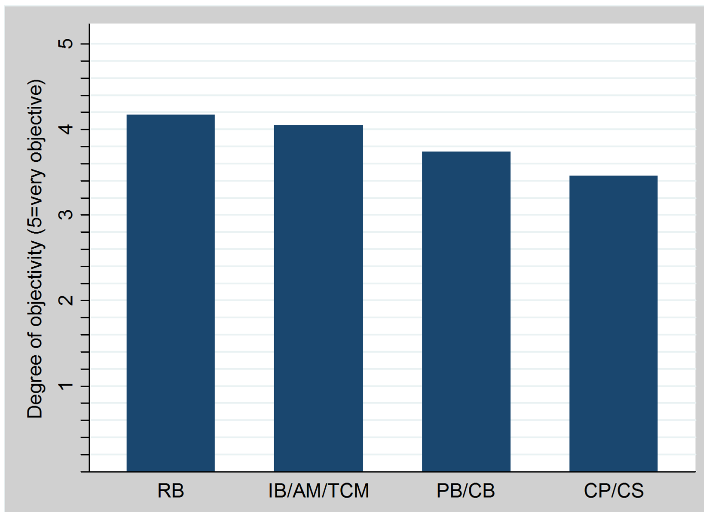

Figure 1: Survey answers on objectivity of performance appraisals

The ranking of the perceived objectivity tends to reflect the ordering of the regression coefficients: Perceived objectivity is the highest in retail banking and the capital market-based functions, whereas we observe the lowest degree of perceived objectivity in the support and cross-divisional functions (CP/CS). Perceived objectivity in CP/CS is significantly smaller than in RB ($p = 0.024$, Mann-Whitney U/Wilcoxon rank-sum test) and in IB/AM/TCM ($p = 0.020$).[41]

[41] Between RB and IB/AM/TCM ($p = 0.564$), RB and PB/CB ($p = 0.115$), as well as IB/AM/TCM and PB/CB ($p = 0.187$), we observe no statistical differences.

6 Further Robustness Checks and Extensions

6.1 Alternative Identification Strategies

As laid out in section 4, the underlying identifying assumption of our approach is the time invariance of potential omitted variables that may affect the size of bonus payments and are correlated with the lagged dispersion. In other words, the lagged dispersion is not correlated with unobservable time-variant characteristics that affect the size of bonus payments. We estimate the causal effects of differentiation if this assumption holds. We already have laid out above that we use a balanced panel and restrict the sample to employees who stay at the same hierarchical level, functional area, specific function, and career ladder throughout all the years in order to ensure that variations in bonus dispersion are not driven by changes in the team composition, which can affect the size of bonus payments. Hence, we exclude that team composition effects drive our results. While we are not aware of other potential confounding mechanisms, in principle we cannot exclude that such time-variant omitted variables exist. We therefore consider two alternative identification strategies.

First, we re-estimate the models reported in table 1 with a lagged dependent variable (LDV) approach, i.e., we estimate how the bonus of an employee is affected by the lagged dispersion controlling for her own bonus in the last period and all other observable characteristics.[42] In other words, we ask the question: When we compare two employees who received the same bonus in the last period and worked under similar organizational characteristics but experienced different degrees of dispersion, how do their current bonuses differ? As argued by Angrist and Pischke (2008), the fixed effects and lagged dependent variable approaches have a useful bracketing property: When the lagged dependent variable is the better-suited specification

[42] The identifying assumption is then $E\left[b_{it}\left(s\right)|b_{it-1}, X_{it}, t, s_{c_i t-1}\right] = E\left[b_{it}\left(s\right)|b_{it-1}, X_{it}, t\right]$.

(i.e., there is time-varying unobserved heterogeneity that can be captured by the past bonus), a fixed effects model will tend to overestimate the true effects. On the other hand, when the fixed effects model is more appropriate (i.e., omitted variables are constant over time), the lagged dependent variable model will tend to underestimate the true effects. The coefficients of the LDV model should thus give a lower boundary to the true effects. Table 6, which reports the LDV regressions, shows that indeed the coefficients tend to be smaller (with the exception of the P90/P10 ratio) but are still sizeable.

Table 6: Lagged dependent variable regression results

Dependent variable:	Logarithm of bonus payments					
	Lagged dependent variable model					
	Coef. of variation		P90/P10 ratio		Std. dev. of logs	
$\text{Differentiation}_{t-1}$	0.3995***		0.0249***		0.5829***	
	(0.1412)		(0.0087)		(0.1196)	
2^{nd} Quintile_{t-1}		0.0965***		-0.0866**		0.0013
		(0.0371)		(0.0423)		(0.0356)
3^{rd} Quintile_{t-1}		0.1093***		0.1490***		0.1542***
		(0.0416)		(0.0530)		(0.0375)
4^{th} Quintile_{t-1}		0.1481***		0.1631***		0.1960***
		(0.0356)		(0.0482)		(0.0433)
5^{th} Quintile_{t-1}		0.1158***		0.2478***		0.1990***
		(0.0442)		(0.0416)		(0.0460)
Ln Bonus_{t-1}	0.4375***	0.4160***	0.4300***	0.4065***	0.4370***	0.4149***
	(0.0444)	(0.0494)	(0.0478)	(0.0493)	(0.0461)	(0.0484)
Observations	25587	25587	25587	25587	25587	25587
Adj. R^2	0.82	0.82	0.82	0.82	0.82	0.82
1 std. dev. increase	8%		7%		11%	

Additional controls: Ln base salary, age squared, and year dummies. Robust standard errors clustered on cell-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

As a second alternative identification strategy, we apply an instrumental variable 2SLS approach, where we allow for time-variant omitted variables but try to identify factors that affect the dispersion in a unit but are exogenous to this unit's performance. Indeed, we can construct a potentially useful instrument from our data by measuring the average degree of differentiation of all *other* cells within a functional area of the respective company and hierarchical level for each year (excluding the cell we are looking at). The key idea of the approach is to capture changes in a firm's general evaluation policies and guidelines or the "culture of differentiation," which affect the degree of differentiation in all departments. Consider, for instance, the announcement of stricter guidelines for performance evaluations, the introduction of a forced distribution scheme[43], or the announcement that managers should tie bonuses to individual performance to a stronger extent. Such a change would influence supervisors' differentiation behavior across different units. If this is indeed the case, the change in the degree of differentiation in the other areas should be informative about the change in the degree of differentiation in a particular area. In other words, we would expect to observe a substantial correlation between changes in the degrees of differentiation in a cell with changes in the average level of differentiation of the other cells in the same firm and level. This can be checked on the basis of the first-stage estimation results, where we regress the coefficient of variation of a particular cell on the average coefficient of variation in the other cells.

The instrumental variables approach also imposes an identifying assumption which we cannot directly test, namely that the level of differentiation in *other* functional areas does not have a direct impact on the bonus payments in a particular area beyond the influence through the dispersion in the area itself.

Note that the first-stage dependent variable varies only on the level of a cell. Hence, we created a "collapsed" data set, where a cell is the unit of observation. The results of the 2SLS instrumental variables procedure applied to this reduced data set with cell fixed effects are reported in table 7. As can

[43]When a firm uses a forced distribution scheme, supervisors have to stick to a predetermined distribution when assigning performance evaluations (for instance, 10% of employees should belong to the top category, 20% to the next, ...).

be seen from the results of the first-stage regression (panel B), for all three measures of dispersion, changes in the dispersion in the other cells are highly correlated with changes in the dispersion in a particular cell. Moreover, the first-stage F-statistic is larger than 10 for all three measures of dispersion. Hence, the instrument is sufficiently strong, and there must be factors that jointly affect the degree of variation in all areas of a company. Additionally, the effect of all three measures of dispersion on the bonus payments in the subsequent year is again significant and substantial, as the second-stage regression results (panel A) reveal. These results are in a similar order of magnitude as the fixed effects and LDP estimates. For the coefficient of variation, the estimated coefficient of *0.32* is smaller than the FE estimate of *0.5* and the LDV estimate of *0.4*. For the P90/P10 ratio, the coefficient is with *0.036* larger than the FE estimate of *0.02* and the LDV estimate of *0.025*. For the standard deviation of logs, the three estimates (FE, LDP, IV) are very close to each other, with values of *0.66*, *0.58* and *0.69*.

6.2 Bonus Payments and Financial Performance

Finally, we take a closer look at bonus payments as the key dependent variable. As laid out above, it is impossible to access a comparable set of financial performance measures on the level of individual units across the different banks. However, the financial success of banks as a whole is publicly observable from balance sheets and profit and loss statements. We collected the financial performance measures return on equity (ROE), return on assets (ROA), and net income for a subset of the considered banks from the Bankscope database provided by Bureau van Dijk.

We used this data to validate whether the bonus paid out to the employees is in fact a good proxy for the performance contribution as our arguments above build on the assumption that higher bonuses are paid out in areas where there is also higher performance. For these banks, we computed the pay-for-performance sensitivity, i.e., the relationship between financial

Table 7: IV 2SLS fixed effects regression with cell-level data set

Panel A: IV 2SLS fixed effects estimates (Cell-level panel)			
Dependent variable:	Log. of average bonus payments		
	Coef. of variation	P90/P10 ratio	Std. dev. of logs
$\text{Differentiation}_{t-1}$	0.3214**	0.0360**	0.6913***
	(0.1580)	(0.0149)	(0.2450)
Ln avg. base salary_t	-0.1148	0.0283	-0.0782
	(0.2327)	(0.2512)	(0.2355)
Panel B: First-stage estimates			
Dependent variable:	Coef. of variation_{t-1}	P90/P10 ratio	Std. dev. of logs
Diff. other cells_{t-1} (Instr.)	0.6385***	0.4204***	0.5031***
	(0.0579)	(0.0660)	(0.0541)
Ln avg. base salary_t	0.0390	-2.763	-0.0762
	(0.1334)	(2.5990)	(0.1036)
Observations	688	688	688
Adj. R^2	0.27	0.12	0.21
F-Test (H_0 : IV coeff.=0), weak ID	121.50	40.60	86.51

Additional control variables include average cell age and year dummies. Weighted estimates (weight: cell size). Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Pay-for-performance sensitivity

	Bank-level FE regression		
Dependent variable:	Log. of average bonus payments		
	ROE	ROA	Net Income
Ln performance$_{t-1}$	0.1590**	0.2480***	0.1960***
	(0.0736)	(0.0922)	(0.0736)
Observations	109	108	109
R^2within	0.24	0.27	0.27

Additional year dummies included. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

performance indicators of a bank and individual payments, by estimating a bank-level fixed effects model with the log average bonus payment per bank as the dependent variable and the log of the different financial performance measures as well as year dummies as independent variables. As table 8 shows, these elasticities indicate a pronounced and statistically significant relationship between firm performance and subsequent average bonus levels. Hence, bonus payments are directly affected by changes in corporate financial indicators.

7 Conclusion

We study the performance effects of between-employee differentiation in bonus payments. In a formal model of subjective performance evaluations and bonus pools, we show that a higher willingness to differentiate should have a positive effect on subsequent bonus payments, if interdependencies between employees are not too strong. Moreover, a reluctance to differentiate is more detrimental in areas where more precise performance information is available and where the span of control is larger. In our empirical analysis, we indeed find a highly significant and economically meaningful average effect of differentiation on individual bonuses. However, there is heterogeneity when we look at different subsamples. The positive effect of differentiation is indeed the strongest when there is a larger span of control. It is also stronger at higher hierarchical levels, whereas differentiation may be even harmful at the lowest levels. We also find differences between functional areas and career ladders, which we contrasted with insights from a survey among bankers on the perceived objectivity of performance measurement in the different functions. The results provide some evidence that differentiation has a stronger effect in areas where performance assessments are perceived to be more objective.

Our results also shed some light on the quite controversial debate among practitioners about methods to increase differentiation in performance appraisals, such as forced distribution systems. As recent surveys show, many firms are still adapting the degree of differentiation among high and low performers on the same job, and most firms are aiming to increase it. A study by the consultancy Mercer, for instance, finds "*companies widening performance differentials for short-term incentive payouts [..]. The highest-performing management level employees are expected to receive average short-term incentive payouts of 36 percent compared to just 8 percent for the lowest performers.*" A similar survey by Towers Perrin concluded: "*In 2010, a full 48% of companies indicated they will continue with the same differentiation strategies they used in 2009 for their 2010 salary review process, while an additional 40% will differentiate more than in prior years.*"[44]

Our results indicate that, for positions in the middle or at the top of the corporate hierarchy, firms should indeed strive to achieve differentiated performance ratings, for example, through the introduction of distribution guidelines. This should positively affect incentives and performance. On the other hand, at lower levels firms should be careful when considering

[44]See *Mercer 2008/2009 US Compensation Planning Survey* and *Towers Perrin 2009 Survey on Compensation Strategies*.

enforcement of differentiation.

References

Abowd, J. M. and D. S. Kaplan (1999). Executive compensation: Six questions that need answering. *Journal of Economic Perspectives 13*(4), 145–168.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics 105*(2), 255–283.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Baiman, S. and M. V. Rajan (1995). The informational advantages of discretionary bonus schemes. *The Accounting Review*, 557–579.

Banker, R. D. and S. M. Datar (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research 27*, 21–39.

Bebchuk, L. A., K. M. Cremers, and U. C. Peyer (2011). The ceo pay slice. *Journal of Financial Economics 102*(1), 199–221.

Becker, B. E. and M. A. Huselid (1992). The incentive effects of tournament compensation systems. *Administrative Science Quarterly 37*(2), 336–350.

Berger, J., C. Harbring, and D. Sliwka (2013). Performance appraisals and the impact of forced distribution: An experimental investigation. *Management Science 59*(1), 54–68.

Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal 42*(1), 25–40.

Bol, J. C. (2011). The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review 86*, 1549–1575.

Bol, J. C. and S. D. Smith (2011). Spillover effects in subjective performance evaluation: Bias and the asymmetric influence of controllability. *The Accounting Review 86*(4), 1213–1230.

Bolton, G. E. and A. Ockenfels (2000). ERC - a theory of equity, reciprocity and competition. *American Economic Review 90*, 166–193.

Depken II, C. A. (2000). Wage disparity and team productivity: evidence from major league baseball. *Economics Letters 67*(1), 87–92.

Drago, R. and G. T. Garvey (1998). Incentives for helping on the job: Theory and evidence. *Journal of Labor Economics 16*(1), 1–25.

Engellandt, A. and R. T. Riphahn (2011). Evidence on incentive effects of subjective performance evaluations. *Industrial and Labor Relations Review 64(2)*, 241–257.

Eriksson, T. (1999). Executive compensation and tournament theory: Empirical tests on danish data. *Journal of Labor Economics 17*(2), 262–280.

Fehr, E., S. Gächter, and G. Kirchsteiger (1997). Reciprocity as a contract enforcement device - experimental evidence. *Econometrica 65*(4), 833–860.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics 114*, 817–868.

Frederiksen, A., F. Lange, and B. Kriechel (2012). Subjective performance evaluations and employee careers. *IZA Discussion Paper No.6373*.

Gibbons, R. S. and K. J. Murphy (1992). Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *Journal of Political Economy 100*, 468–505.

Gibbs, M. and W. Hendricks (2004). Do formal salary systems really matter? *Industrial and Labor Relations Review 58*, 71–93.

Golman, R. and S. Bhatia (2012). Performance evaluation inflation and compression. *Accounting, Organizations and Society 37*(8), 534–543.

Grund, C. and J. Przemeck (2012). Subjective performance appraisal and inequality aversion. *Applied Economics 44*(17), 2149–2155.

Grund, C. and N. Westergaard-Nielsen (2008). The dispersion of employees' wage increases and firm performance. *Industrial and Labor Relations Review 61*(4), 485–501.

Halvorsen, R. and R. Palmquist (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review 70*(3), 474–475.

Heyman, F. (2005). Pay inequality and firm performance: evidence from matched employer-employee data. *Applied Economics 37*(11), 1313–1327.

Holmström, B. (1982). Managerial Incentive Problems - A Dynamic Perspective. In *Essays in Economics and Management in Honor of Lars Wahlbeck*. Helsinki.

Ittner, C. D., D. F. Larcker, and M. W. Meyer (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review 78*(3), 725–758.

Jirjahn, U. and K. Kraft (2007). Intra-firm wage dispersion and firm performance - is there a uniform relationship? *Kyklos 60*(2), 231–253.

Kale, J. R., E. Reis, and A. Venkateswaran (2009). Rank-order tournaments and incentive alignment: The effect on firm performance. *Journal of Finance 64*(3), 1479–1512.

Kane, J. S., H. J. Bernardin, P. Villanova, and J. Peyrefitte (1995). Stability of rater leniency: Three studies. *Academy of Management Journal 38*(4), 1036–1051.

Lambert, R. A. (2001). Contracting theory and accounting. *Journal of Accounting and Economics 32*(1), 3–87.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economy 97*(3), 561–580.

Lazear, E. P. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy 89*(5), 841–864.

Leonard, J. S. (1990). Executive pay and firm performance. *Industrial and Labor Relations Review 43*(3), 13–S–29–S.

Main, B. G. M., C. A. O'Reilly III, and J. Wade (1993). Top executive pay: Tournament or teamwork? *Journal of Labor Economics 11*(4), 606–628.

Manthei, K. and D. Sliwka (2015). Multitasking and the benefits of objective performance measurement - evidence from a field experiment. *University of Cologne Working Paper*.

Martins, P. S. (2008). Dispersion in wage premiums and firm performance. *Economics Letters 101*(1), 63–65.

Moers, F. (2005). Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations & Society 30*(1), 67–80.

Murphy, K. J. (1999). Executive compensation. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Amsterdam. North-Holland.

Murphy, K. J. (2001). Performance standards in incentive contracts. *Journal of Accounting and Economics 30*(3), 245–278.

Murphy, K. R. and J. N. Cleveland (1995). *Understanding Performance Appraisal*. Thousand Oaks: Sage.

Ockenfels, A., D. Sliwka, and P. Werner (2015). Bonus payments and reference point violations. *Management Science 61*(7), 1496–1513.

Pfeffer, J. and N. Langton (1993). The effect of wage dispersion on satisfaction, productivity, and working collaboratively: Evidence from college and university faculty. *Administrative Science Quarterly 38*(3), 382–407.

Pfeffer, J. and R. Sutton (2006). *Hard facts, dangerous half-truths, and total nonsense: Profiting from evidence-based management.* Harvard Business Press.

Prendergast, C. and R. Topel (1996). Favoritism in organizations. *Journal of Political Economy 104*(5), 958–978.

Prendergast, C. J. (2002). Uncertainty and incentives. *Journal of Labor Economics 20*(2), 115–137.

Rajan, M. V. and S. Reichelstein (2006). Subjective performance indicators and discretionary bonus pools. *Journal of Accounting Research 44*(3), 585–618.

Rajan, M. V. and S. Reichelstein (2009). Objective versus subjective indicators of managerial performance. *The Accounting Review 84*(1), 209–237.

Smeets, V., M. Waldman, and F. Warzynski (2015). Performance, career dynamics, and span of control. *Working Paper*.

Waldman, M. (2013). Classic promotion tournaments versus market-based tournaments. *International Journal of Industrial Organization 31*(3), 198–210.

Welch, J. (2003). *Jack- Straight from the Gut.* Warner Books.

Winter-Ebmer, R. and J. Zweimüller (1999). Intra-firm wage dispersion and firm performance. *Kyklos 52*(4), 555–572.

Woods, A. (2012). Subjective adjustments to objective performance measures: The influence of prior performance. *Accounting, Organizations and Society 37*(6), 403–425.

8 Appendix

Computation of $E\left[\left.(r_{ti}-y_i)^2\right| s_i\right]$:
Using that $E\left[Y\,|X\right]=E\left[Y\right]+\frac{Cov[X,Y]}{Var[X]}\left(X-E\left[X\right]\right)$ and $V\left[Y\,|X\right]=V\left[Y\right]-\frac{(Cov[X,Y])^2}{Var[X]}$ we obtain

$$\begin{aligned} E\left[\left.y_i-r_i\right| s_i\right] &= E\left[y_i-r_i\right]+\tfrac{Cov[y_i-r_i,s_i]}{Var[s_i]}\left(s_i-E\left[s_i\right]\right) \\ &= m_i+\hat{e}_i-r_i+\tfrac{\sigma_a^2}{\sigma_a^2+\sigma_\eta^2}\left(s_i-m_i-\hat{e}_i\right) \end{aligned}$$

$$\begin{aligned} V\left[\left.y_i-r_i\right| s_i\right] &= V\left[y_i-r_i\right]-\tfrac{(Cov[\hat{e}_i+a_i-r_i,\hat{e}_i+a_i+\eta_i])^2}{Var[\hat{e}_i+a_i+\eta_i]} \\ &= \sigma_a^2-\tfrac{\sigma_a^4}{\sigma_a^2+\sigma_\eta^2}=\tfrac{\sigma_a^2\sigma_\eta^2}{\sigma_a^2+\sigma_\eta^2} \end{aligned}$$

and thus

$$E\left[\left.(r_{ti}-y_i)^2\right| s_i\right]=\tfrac{\sigma_a^2\sigma_\eta^2}{\sigma_a^2+\sigma_\eta^2}+\left(m_i+\hat{e}_i-r_i+\tfrac{\sigma_a^2}{\sigma_a^2+\sigma_\eta^2}\left(s_i-m_i-\hat{e}_i\right)\right)^2.$$

Proof of Proposition 1:
We solve program (1). The Lagrangean is

$$\begin{aligned} &-\nu_A\sum_{i=1}^{n}\left(\frac{\sigma_a^2\sigma_\eta^2}{\sigma_a^2+\sigma_\eta^2}+\left(m_i+\hat{e}_i-r_i+\frac{\sigma_a^2}{\sigma_a^2+\sigma_\eta^2}\left(s_i-m_i-\hat{e}_i\right)\right)^2\right) \\ &-\nu_E\cdot\left(\sum_{i=1}^{n}\left(\beta r_i-\frac{B}{n}\right)^2\right)+\lambda\left(\beta\sum_{i=1}^{n}r_i-B\right) \end{aligned}$$

Hence, we must have that

$$\nu_A\left(2\left(m_i+\hat{e}_i-r_i+\tfrac{\sigma_a^2}{\sigma_a^2+\sigma_\eta^2}\left(s_i-m_i-\hat{e}_i\right)\right)\right)-\nu_E\cdot 2\beta\left(\beta r_i-\frac{B}{n}\right)+\lambda\beta=0\quad\forall i=1,..n$$

which is equivalent to

$$r_i = \frac{2\nu_A\left(m_i + \hat{e}_i + \frac{\sigma_a^2}{\sigma_a^2+\sigma_\eta^2}\left(s_i - m_i - \hat{e}_i\right)\right) + 2\nu_E\beta\frac{B}{n}}{2\left(\nu_A + \nu_E\beta^2\right)} + \lambda\frac{\beta}{2\left(\nu_A + \nu_E\beta^2\right)} \quad \forall i = 1, ..n \tag{7}$$

Substituting this expression into the budget constraint $\beta\sum_{j=1}^{n} r_j = B$ we obtain

$$\beta\sum_{j=1}^{n}\left(\frac{2\nu_A\left(m_j + \hat{e}_j + \frac{\sigma_a^2}{\sigma_a^2+\sigma_\eta^2}\left(s_j - m_j - \hat{e}_j\right)\right) + 2\nu_E\beta\frac{B}{n}}{2\left(\nu_A + \nu_E\beta^2\right)} + \lambda\frac{\beta}{2\left(\nu_A + \nu_E\beta^2\right)}\right) - B = 0$$

and thus

$$\lambda = 2\frac{B\left(\nu_A + \nu_E\beta^2\right)}{n\beta^2} - \frac{2\nu_A}{\beta n}\sum_{j=1}^{n}\left(m_j + \hat{e}_j + \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\eta^2}\left(s_j - m_j - \hat{e}_j\right)\right) - 2\frac{\nu_E B}{n}.$$

By substituting λ and $B = \kappa\cdot\pi\cdot\sum_{i=1}^{n} s_i$. into (7) and simplifying we obtain

$$r_i(s) = \frac{\kappa\pi}{\beta}\left(\frac{1}{n}\sum_{j=1}^{n} s_j\right) + \left(\frac{\sigma_a^2\left(s_i - \frac{1}{n}\sum_{j=1}^{n} s_j\right) + \sigma_\eta^2\left(\left(m_i + \hat{e}_i\right) - \frac{1}{n}\sum_{j=1}^{n}\left(m_j + \hat{e}_j\right)\right)}{\left(1 + \frac{\nu_E}{\nu_A}\beta^2\right)\left(\sigma_a^2 + \sigma_\eta^2\right)}\right).$$

Using that the mean of the ratings must be equal to

$$\frac{1}{n}\sum_{i=1}^{n} r_i = \frac{\kappa\cdot\pi}{\beta n}\cdot\sum_{i=1}^{n} s_i$$

we obtain the within-team variance of ratings

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\sigma_a^2\left(s_i - \left(\frac{\sum_{j=1}^{n} s_j}{n}\right)\right) + \sigma_\eta^2\left(\left(m_i + \hat{e}_i\right) - \frac{\sum_{j=1}^{n}\left(m_j + \hat{e}_j\right)}{n}\right)}{\left(1 + \left(\frac{v_E}{v_A}\right)\beta^2\right)\left(\sigma_a^2 + \sigma_\eta^2\right)}\right)^2, \tag{8}$$

such that the standard deviation is

$$\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\sigma_a^2\left(s_i-\left(\frac{\sum_{j=1}^{n}s_j}{n}\right)\right)+\sigma_\eta^2\left((m_i+\hat{e}_i)-\frac{\sum_{j=1}^{n}(m_j+\hat{e}_j)}{n}\right)\right)^2}}{\left(1+\left(\frac{v_E}{v_A}\right)\beta^2\right)\left(\sigma_a^2+\sigma_\eta^2\right)} \tag{9}$$

■

Proof of Proposition 5:

The first order condition of the optimization problem is

$$(1-\kappa)\,\pi\frac{\beta n\left(1-\frac{1}{n}\right)\sigma_a^2}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)}\left(\psi\sigma_a^2+1\right)^{-2}-z'(\psi)=0.$$

The second derivative of the objective function

$$-2\,(1-\kappa)\,\pi\frac{\beta n\left(1-\frac{1}{n}\right)\sigma_a^4}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)}\left(\psi\sigma_a^2+1\right)^{-3}-z''(\psi)<0$$

such that the function is strictly concave. By the implicit function theorem we obtain

$$\frac{\partial\psi}{\partial\pi}=\frac{(1-\kappa)\,\frac{\beta n\left(1-\frac{1}{n}\right)\sigma_a^2}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)}\left(\psi\sigma_a^2+1\right)^{-2}}{2\,(1-\kappa)\,\pi\frac{\beta n\left(1-\frac{1}{n}\right)\sigma_a^4}{\left(1+\left(\frac{1}{v_D}\right)\beta^2\right)}\left(\psi\sigma_a^2+1\right)^{-3}+z''(\psi)}>0$$

■

Table A.1: Examples of functional areas and functions

Functional areas and functions	
Retail Banking	*Corporate and Private Banking*
Retail Banking Product Development	Corporate Banking Product Development
Retail Sales	Corporate / Institutional Relationship
Telebanking Sales	Client Relationship Management
Financial Advice	Portfolio Management
Investment Banking/Asset Management	*Corporate Production*
Treasury and Capital Markets	Human Resources
Asset Allocation	Legal / Economics
Credit Syndication	Risk Management
Money Markets	Sales & Marketing
Hedge Funds	Finance / Accounting
Asset Management Product Development	Project Management
Money Transfers	
Fund Management	*Corporate Services*
Structured Finance	IT Administration / Support
Corporate Finance	IT Architecture
Commodity Trading	Customer Service
Fixed Income	Asset Management Support
Equity	Foreign Operations

Table A.2: Descriptive statistics (sample used in baseline regression)

Variables	Obs.	Mean	Std. Dev.	Min	Max
CV Bonus	25587	0.32	0.20	0	2.03
P90/P10 ratio	25587	2.59	2.83	1	43.67
Standard dev. of logs	25587	0.33	0.17	0	1.90
Ln Bonus	25587	8.77	0.89	4.32	12.80
Ln Base Salary	25587	10.80	0.29	9.85	12.11
Age squared	25587	1833.40	685.03	361	4225
Tenure squared	25587	385.01	378.24	1	2116
Hierarchical level:					
Level 6 (highest)	25587	0.008	0.088	0	1
Level 5	25587	0.131	0.337	0	1
Level 4	25587	0.249	0.432	0	1
Level 3	25587	0.265	0.442	0	1
Level 2	25587	0.304	0.460	0	1
Level 1 (lowest)	25587	0.044	0.204	0	1
Functional area:					
Corp. Production/ Corp. Services	25587	0.358	0.480	0	1
Inv. Bank. / Asset Man./Treas. & Cap. Mkts.	25587	0.039	0.194	0	1
Priv. Banking / Corp. Bank.	25587	0.046	0.209	0	1
Retail Banking	25587	0.557	0.497	0	1
Career Ladder:					
Management	25587	0.088	0.284	0	1
Professional	25587	0.223	0.416	0	1
Sales	25587	0.494	0.500	0	1
Support	25587	0.195	0.396	0	1

Table A.3: Baseline regressions excluding retail banking area

Dependent variable:	Logarithm of bonus payments					
	Coef. of variation		P90/P10 ratio		Std. dev. of logs	
Differentiation$_{t-1}$	0.3837**		0.0191**		0.5851***	
	(0.1891)		(0.0075)		(0.1661)	
2^{nd} Quintile$_{t-1}$		0.1841**		0.0916		-0.0911
		(0.0924)		(0.1134)		(0.1374)
3^{rd} Quintile$_{t-1}$		0.2338***		0.3968***		0.1700*
		(0.0838)		(0.1168)		(0.0904)
4^{th} Quintile$_{t-1}$		0.2190***		0.2997**		0.2581***
		(0.0788)		(0.1219)		(0.0938)
5^{th} Quintile$_{t-1}$		0.2736***		0.4214***		0.3684***
		(0.0809)		(0.1007)		(0.1007)
Ln Base salary$_t$	-0.4957	-0.3641	-0.4106	-0.5732*	-0.3629	-0.6797**
	(0.4074)	(0.3953)	(0.3891)	(0.3051)	(0.3948)	(0.2717)
Age squared$_t$	0.0012	0.0017	0.0014	0.0007	0.0007	0.0008
	(0.0013)	(0.0015)	(0.0014)	(0.0011)	(0.0012)	(0.0011)
Observations	11343	11343	11343	11343	11343	11343
R^2 within	0.12	0.10	0.11	0.15	0.13	0.14
1 std. dev. increase	10%		8%		14%	

Additional year dummies included. Robust standard errors clustered on cell-level in parentheses.

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

Table A.4: Baseline regressions with panel data set including movers

Dependent variable:	Logarithm of bonus payments					
	Panel including movers					
	Coef. of variation		P90/P10 ratio		Std. dev. of logs	
$\text{Differentiation}_{t-1}$	0.4882***		0.0160***		0.6057***	
	(0.0445)		(0.0023)		(0.0516)	
2^{nd} Quintile_{t-1}		0.2138***		-0.1256***		-0.0912***
		(0.0155)		(0.0084)		(0.0075)
3^{rd} Quintile_{t-1}		0.0891***		0.0448***		0.0498***
		(0.0156)		(0.0143)		(0.0149)
4^{th} Quintile_{t-1}		0.1451***		0.1251***		0.1328***
		(0.0172)		(0.0150)		(0.0150)
5^{th} Quintile_{t-1}		0.2232***		0.2004***		0.2741***
		(0.0238)		(0.0196)		(0.0194)
Ln Base salary_t	0.3481***	0.4099***	0.4000***	0.5784***	0.3678***	0.3496***
	(0.0696)	(0.0696)	(0.0685)	(0.0751)	(0.0683)	(0.0713)
Age squared_t	0.0001	-0.00001	0.00006	-0.00002	0.00001	-0.00006
	(0.0001)	(0.00005)	(0.00005)	(0.00005)	(0.0001)	(0.00005)
Observations	34438	34438	34443	34443	34438	34438
R^2within	0.18	0.15	0.16	0.16	0.17	0.16

Additional controls include hierarchical level, career ladder, functional area, function, company, and year dummies.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Measures of differentiation across hierarchical levels

Level	Median of measures of differentiation								
	Coef. of variation			P90/P10 ratio			Std. dev. of logs		
	2005	2006	2007	2005	2006	2007	2005	2006	2007
6 (highest)	0.33	0.29	0.27	2.24	2.12	2.40	0.30	0.38	0.31
5	0.36	0.34	0.34	2.29	2.35	2.32	0.34	0.33	0.34
4	0.35	0.30	0.33	2.33	2.17	2.27	0.33	0.30	0.33
3	0.21	0.22	0.23	1.76	1.80	1.85	0.23	0.24	0.26
2	0.23	0.25	0.26	1.77	1.82	1.91	0.24	0.26	0.28
1 (lowest)	0.22	0.25	0.20	1.77	1.93	1.47	0.23	0.27	0.28

Table A.6: Hierarchical levels (coefficient of variation)

Dependent variable:	Logarithm of bonus payments					
	Level 5+6		Level 3+4		Level 1+2	
CV Bonus$_{t-1}$	0.9236***		0.6588***		-1.0701***	
	(0.1327)		(0.2006)		(0.4003)	
2^{nd} Quintile$_{t-1}$		0.1903		0.0559		0.1351
		(0.1338)		(0.1048)		(0.1513)
3^{rd} Quintile$_{t-1}$		0.1921		0.0636		0.1541
		(0.1628)		(0.1035)		(0.1357)
4^{th} Quintile$_{t-1}$		0.2429		0.2227**		0.1757
		(0.1748)		(0.0924)		(0.2864)
5^{th} Quintile$_{t-1}$		0.3094		0.3241***		-0.0292
		(0.2233)		(0.1028)		(0.2845)
Observations	3540	3540	13155	13155	8892	8892
R^2 within	0.25	0.12	0.17	0.12	0.10	0.03

Age squared, base salary and year dummies included. Robust standard errors clustered on cell-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1