European Economic Review 47 (2003) 877–890
EUROPEAN ECONOMIC REVIEW
www.elsevier.com/locate/econbase

Do new brooms sweep clean? When and why dismissing a manager increases the subordinates' performance

Felix Höffler[a], Dirk Sliwka[b,*]

[a] *Deutsche Telekom, Corporate Strategy, Bonn, Germany*
[b] *Betriebswirtschaftliche Abteilung II, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany*

Abstract

If a manager stays in office for a long time he will have learned much about his subordinates. Thus competition among them will be weak as the manager has made up his mind who is suited best for which position. With a new manager the "race" for favorable tasks is restarted leading subordinates to exert higher effort. But for the firm-owner the trade-off arises that with a new manager effort is larger but the quality of task allocation is worse since information is lost. The optimal dismissal policy will be nonmonotonic in the expected heterogeneity of the subordinates' abilities.

JEL classification: C73; D82; L22

Keywords: Career concerns; Tournaments; Managerial turnover; Asymmetric information

1. Introduction

In this paper, we want to investigate the consequences of a replacement of managers. As a starting point we focus on one important effect: Managerial turnover increases the incentives of subordinates.

A manager who has been in office for some time knows the abilities of her subordinates well and therefore will typically be quite sure who is suited best for which position or task. But positions within a certain department are more or less attractive.

* Corresponding author.
E-mail addresses: felix.hoeffler@telekom.de (F. Höffler), dirk.sliwka@uni-bonn.de (D. Sliwka).

0014-2921/03/$ - see front matter
PII: S0014-2921(02)00272-6

As has been stressed for instance by tournament theory[1] or in the literature on promotions,[2] the competition for attractive positions is an important incentive device. Such competition will be weak when the manager already made up his mind on the optimal assignment of positions and tasks. If a new manager comes in, however, he knows less on the subordinates' abilities and hence, the race for attractive positions is restarted to some extent.

This "new brooms" effect can be illustrated for instance by the following statement of Michael Owen, European Footballer of the Year 2001, after the dismissal of Kevin Keegan as England's Team Manager:

> It'll be another challenge now for all the lads to impress the new manager, whoever he may be. Even if he knows what you can do, you'll still have to convince him you should be in his plans.[3]

However, note that the mechanism indicates a direct drawback of managerial turnover: As the new manager has less information on the subordinates' abilities, initially his task assignment decisions will be worse in expected terms than the old manager's. Both effects, the increase of effort and the reduction of the quality of the task assignment, are two sides of the same coin, as they are caused by the loss of information implied by managerial turnover. The optimal dismissal decision trades off these two effects.

From a more general perspective our results may indicate some consequences of job rotation, limited maximum times in office or mandatory age limits. All those practices lead to regular managerial turnover and, hence, confront subordinates with new superiors from time to time. As our theory indicates, this has positive effects on effort incentives but may well have negative consequences as information is lost on the subordinates qualifications.[4]

To analyze the consequences of managerial turnover formally, we set up a model with three hierarchical levels and three periods: A firm owner who decides whether to keep or to dismiss the old manager, a manager who selects a subordinate for an important position or task and finally two subordinates competing for this favorable task in every period. We will show that the dismissal of a manager on the one hand increases the subordinates' equilibrium efforts and on the other reduces the quality of the task allocation. In a third step, we analyze the owner's optimal dismissal policy in the second period. It turns out that the manager will optimally be dismissed only if the expected abilities of the subordinates are neither too homogenous nor too heterogenous.

[1] Compare for instance Lazear and Rosen (1981), Nalebuff and Stiglitz (1983) or Green and Stokey (1983).

[2] See Waldman (1990), Prendergast (1993) or Fairburn and Malcomson (1994).

[3] Interview at www.liverpoolfc.tv, October 2000.

[4] Previous explanations have been e.g. that job rotation improves incentives to reveal new ideas (Carmichael and MacLeod, 1993), prevents collusion (Bolton, 1992) or the ratchet effect (Ickes and Samuelson, 1987).

2. The model

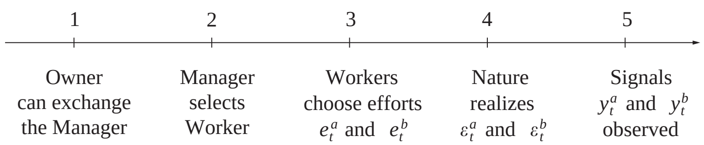
There are three kinds of players in the model. The firm owner, a firm manager, and two subordinate workers, a and b. At the beginning of each period the owner has the opportunity to replace the manager by a new one. The workers cannot be replaced and stay in the firm for all three periods. The only decision the manager has to make is to select one of the two workers every period for an important position or a more attractive task. The importance of the position is reflected in two ways. First, how well the selected worker performs in this leading position will determine the firm's profits. Second, workers like being promoted: The worker holding the important position receives a utility gain Δu. We assume that the only incentive to exert effort stems from the desire to get promoted to the attractive position. No payments contingent on the performance are feasible.

The workers' performance y_t^i $(i{=}a,b)$ in period t is given by the sum of the worker's ability η^i, his effort in that period e_t^i, and random noise ε_t^i:

$$y_t^i = \eta^i + e_t^i + \varepsilon_t^i. \tag{1}$$

As in the career concerns literature, we assume that initially all players are symmetrically informed about the workers' types, which are independently drawn from a normal distribution with mean m_0 and precision h_0, i.e. $\eta^i \sim \mathrm{N}(m_0, 1/h_0)$. The ε_t^i are uncorrelated and also follow a normal distribution $\varepsilon_t^i \sim \mathrm{N}(0, 1/h_\varepsilon)$. All players observe both workers' aggregate signals y_t^i and use them to update their beliefs about the workers' abilities. For simplicity we assume that the signal of the worker not selected for the important position is as informative as the one from the competitor who holds the position in period t. Exerting effort is costly for the workers, according to a strictly convex cost function $c(e)$.

Each of the three periods has exactly the same form:



We are interested in the effect of replacing an informed manager by a less informed one. Therefore, we make the extreme assumption that a new manager has no knowledge about the firm's past performance signals. Additionally, we assume that the owner is not able to communicate her information to the new manager. [5]

All players are risk neutral and there is no discounting. The owner's payoff consists of the sum of the profits of all three periods, where the profit Y_t in period t equals the signal y_t^i of the worker i who is selected in period t. [6] We assume that the manager's

[5] These assumptions can be relaxed as we will discuss in the conclusion.

[6] This assumption can easily be generalized such that the profit is a linear combination of both signals with a stronger weight on the important task.

interests are perfectly aligned with the owner's as long as he is in office, e.g. because he receives a fixed share of the owner's profits.

3. The incentive problem

We now look for a perfect Bayesian equilibrium. As workers invest effort only in order to increase their chances to be picked for the leading position in the following periods, clearly zero effort is optimal in period $t = 3$ in any equilibrium. Given that workers do not invest effort, the manager will select the worker he believes to be of higher ability.

We begin by analyzing the workers' optimal behavior in period two for the case where the manager has not been dismissed. As the manager has no information at the beginning of $t=1$, he is indifferent between the two workers. At the beginning of period $t = 2$ manager, owner, and workers have already observed the period one signals y_1^a and y_1^b and updated their beliefs. Call m_1^i worker i's expected ability after period one. This belief is shared by all players, since they have observed the same information and know the efforts chosen in equilibrium. We assume that a pure strategy equilibrium in the first period effort choice game exists. Denote by $\hat{e}_t^i$ the equilibrium effort of player i in period t, which will depend upon first period performance signals. Applying a result for the updating of normally distributed random variables (see e.g. DeGroot, 1970, p. 167) we get the following expected values for the workers' types after period 1:

$$m_1^i := \mathrm{E}\left[\eta^i | y_1^i - \hat{e}_1^i\right] = \mathrm{E}\left[\eta^i | \eta^i + \varepsilon_1^i\right] = \underbrace{\frac{h_0}{h_0 + h_\varepsilon}}_{=: \alpha_1} m_0 + \underbrace{\frac{h_\varepsilon}{h_0 + h_\varepsilon}}_{=: 1-\alpha_1} \left(y_1^i - \hat{e}_1^i\right). \tag{2}$$

Similarly, after period $t = 2$, the expectations will be

$$m_2^i := \mathrm{E}\left[\eta^i | y_1^i - \hat{e}_1^i, y_2^i - \hat{e}_2^i\right] = \underbrace{\frac{h_0 + h_\varepsilon}{h_0 + 2h_\varepsilon}}_{=: \alpha_2} m_1^i + \underbrace{\frac{h_\varepsilon}{h_0 + 2h_\varepsilon}}_{=: 1-\alpha_2} \left(y_2^i - \hat{e}_2^i\right). \tag{3}$$

The manager selects worker i in period $t = 3$ if and only if $m_2^i > m_2^j$. We are now looking for an equilibrium in the game between the two workers in $t = 2$, given the beliefs and the manager's strategy for $t = 3$.

A worker knows that he will gain Δu if the manager believes him to be of higher ability. His expected payoff therefore depends on the probability that this will be the case given the information I_1 known after period 1 and his colleague's effort level $\hat{e}_2^j$. Therefore a worker i will solve

$$\max_{e_2^i} P\left\{m_2^i > m_2^j | I_1, \hat{e}_2^j\right\} \Delta u - c(e_2^i).$$

Let $\phi(x)$ be the density of the standardized normal distribution. We obtain the following result:

Lemma 1. *If a pure strategy equilibrium in the effort choice game in period 2 exists, there is a unique perfect Bayesian equilibrium in the case where the old manager is kept. It has the following properties*: (i) *In period 3 the workers exert no effort and the old manager chooses worker i if* $m_2^i > m_2^j$.

(ii) *Both workers choose identical effort levels* e_2^{O} *in period* $t = 2$ *which are defined by*

$$\frac{1}{\sigma_\delta}\phi\left(\frac{1}{(1-\alpha_2)\sigma_\delta}\Delta m_1\right)\Delta u = c'\left(e_2^{\mathrm{O}}\right), \tag{4}$$

where $\Delta m_1 = |m_1^a - m_1^b|$ *and* $\sigma_\delta^2 = (2h_0 + 4h_\varepsilon)/((h_0 + h_\varepsilon)h_\varepsilon)$.

Proof. See the appendix.

The symmetric effort choice result is well known from the tournament literature. Note that the equilibrium effort level depends on the difference in the agents expectations about the types Δm_1. It is easy to see from Eq. (4) that equilibrium effort levels decrease in Δm_1. The workers are in a situation of relative performance comparison. Eq. (4) states that the closer the race, the stronger the incentives to exert high effort.[7] The existence of such an equilibrium is ensured if the cost function is sufficiently convex.[8]

We now turn to the case where the manager is dismissed at the beginning of period two. The only information a new manager has is that his predecessor has been dismissed. As in equilibrium the owner will dismiss the manager only for certain first period outcomes, the new manager learns something about the workers' types from the fact that he is employed. However, his expectations in period 2 will be the same for both workers. Hence, the new manager can only randomly select one of both workers in period 2. In period 3 the most natural decision will be to select the worker with the higher output in period 2 and indeed there is an equilibrium with this property:[9]

Lemma 2. *Suppose that a pure strategy equilibrium in the effort choice game in period 2 exists. There is a perfect Bayesian Equilibrium when a new manager has been hired with the following properties:* (i) *In period 3 the workers exert no effort and the new manager chooses worker i if* $y_2^i > y_2^j$. (ii) *Both workers choose identical effort levels* e_2^N *in period* $t = 2$ *which are defined by*

$$\frac{1}{\sigma_\delta}\phi\left(-\frac{\Delta m_1}{\sigma_\delta}\right)\Delta u = c'\left(e_2^N\right). \tag{5}$$

[7] Compare for instance Meyer (1992).

[8] More precisely, existence can be ensured if $(1/\sqrt{2\pi})(\Delta u/\sigma_\delta^2)e^{-1/2} < \inf_e c''(e)$. More details can be obtained from the authors. Compare Lazear and Rosen (1981, p. 845), or Bhattacharya and Guasch (1988, p. 871).

[9] Note that here we do not show that this equilibrium is the unique pure strategy equilibrium, as we are not able to exclude the existence of equilibria where due to "strange" beliefs the agent with the lower performance may be promoted, leading to different efforts supporting those beliefs.

Proof. See the appendix.

Although the manager has no information about the period 1 outcome, it still influences the optimal effort choice in period 2 as not only the manager's but also the workers' own expectations about the type difference matter for their effort choice. If the workers think that their abilities differ strongly they will exert lower efforts even if this difference is not perceived by the new manager.

4. Consequences of a dismissal

Dismissing the manager does make sense in our model only at the beginning of period 2 as in the terminal period effort will be zero in any case and job allocation can only be worse with a new manager. First we investigate the impact of a dismissal at that stage on the efforts exerted by the workers, then on the expected ability of the selected worker, and finally adding the two opposing effects we derive the owner's optimal decision.

4.1. The effort effect

By comparing (4) with (5) we obtain the following result:

Proposition 1. *If the expected ability difference Δm_1 is strictly positive second period efforts will be higher with a new manager.*

With the old manager the information disclosed in period 1 lowers the period 2 effort choice in two ways: The workers have learned something about their respective abilities but so has the manager. Both effects will cause the period 2 efforts to be smaller. With the new manager only the first effect is present. Therefore effort will be higher with a new manager than with the old precisely because the new one has less information. [10]

Now, we analyze the size of this *effort effect* $\Delta e = e_2^{\mathrm{N}} - e_2^{\mathrm{O}}$, i.e. the expected effort gain in period 2 from dismissing the manager.

Lemma 3. (i) *If there is no expected ability difference* ($\Delta m_1 = 0$), *the effort effect* Δe *is zero and* $\partial \Delta e / \partial \Delta m_1 = 0$ *for* $\Delta m_1 = 0$. *For* $\Delta m_1 \to \infty$ *the effort effect tends to* 0. (ii) *With quadratic effort costs* $c(e) = (k/2)e^2$ *the effort effect* Δe *linearly increases in the size of the intrinsic utility gain* Δu *the workers receive from holding the important position and decreases in* k. *There exists a single peak of* Δe *at some strictly positive value* $\Delta \hat{m}_1$.

Proof. See the appendix.

[10] Related "less information is better" effects are observed for example by Crémer (1995) who shows that learning less on an agent's type makes tougher incentive schemes credible or Meyer and Vickers (1997) who show that better information from relative performance evaluation may weaken incentives.

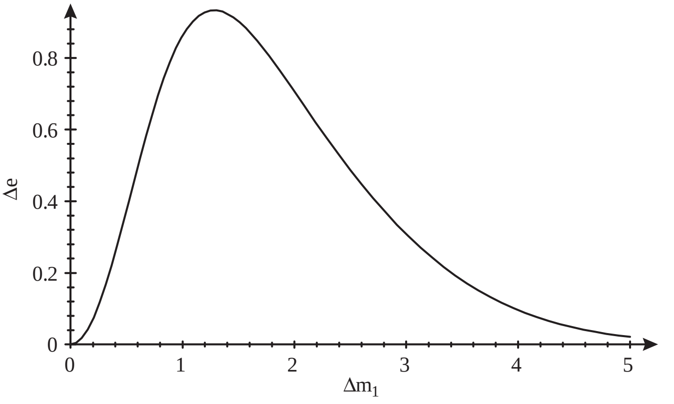


Fig. 1. The effort effect.

In Fig. 1 the effort effect is plotted for an example.[11] The effort gain from dismissing the manager is not monotonically increasing in the type difference Δm_1. For small differences the effort gain from exchanging the manager is small because competition works relatively well with the old manager. For intermediate values of Δm_1, the effort gain from exchanging the manager is big, as competition is weak with the old manager but competition is restarted to some extent when a new manager is hired. This is no longer true for very large values of Δm_1, since the workers know that they are very different and even a new manager cannot induce strong competition. Furthermore, the effort effect increases in Δu, as gaining Δu is the only reason to exert effort at all. For higher values of Δu the effort effect is shifted upwards.

4.2. The ability effect

But the owner also has to take into account that a new manager makes more mistakes when choosing the worker since he knows less about their abilities. She compares the expected ability of the worker selected by a new manager to the expected ability of the selected worker when the old manager is kept.

In period 1 the owner has observed the signals y_1^a and y_1^b. Therefore, the owner knows that in period 2 the old manager picks a worker with expected ability:

$$A_2^{\text{old}} = \max(\mathrm{E}[\eta^a|I_1], \mathrm{E}[\eta^b|I_1]) = \max\left(m_1^a, m_1^b\right). \tag{6}$$

Given the owner's information at the beginning of $t=2$, her expectation of this player's ability equals:[12]

$$A_3^{\text{old}} = \mathrm{P}\left\{m_2^a > m_2^b|I_1\right\}\mathrm{E}\left[\eta^a|m_2^a > m_2^b; I_1\right] + \mathrm{P}\left\{m_2^a < m_2^b|I_1\right\}\mathrm{E}\left[\eta^b|m_2^a < m_2^b; I_1\right]. \tag{7}$$

[11] The figure shows Δe for values $h_\varepsilon = h_0 = 1$, $\Delta u = 6$, and $k = 1$. For these values the existence of the equilibrium is ensured.

[12] The calculation of A_3^{old} and in the following A_3^{new} is given in the appendix.

The new manager does not know which worker performed better in period 1 and, hence, picks a worker with expected ability of

$$A_2^{\text{new}} = \frac{1}{2}\mathrm{E}\left[\eta^a|I_1\right] + \frac{1}{2}\mathrm{E}\left[\eta^b|I_1\right] = \frac{m_1^a + m_1^b}{2}. \tag{8}$$

In period 3 he will have observed the second period signals and will use this information. The owner knows that the new manager will select a worker with the expected ability of A_3^{new}:

$$\begin{aligned} A_3^{\text{new}} = {} & \mathrm{P}\left\{\eta^a + \varepsilon_2^a > \eta^b + \varepsilon_2^b | I_1\right\}\mathrm{E}\left[\eta^a|\eta^a > \eta^b + \varepsilon_2^b - \varepsilon_2^a; I_1\right] \\ & + \mathrm{P}\left\{\eta^a + \varepsilon_2^a < \eta^b + \varepsilon_2^b | I_1\right\}\mathrm{E}\left[\eta^b|\eta^b > \eta^a + \varepsilon_2^a - \varepsilon_2^b; I_1\right]. \end{aligned} \tag{9}$$

Hence, the total expected ability loss when the old manager is replaced is given by $\Delta a = A_2^{\text{old}} + A_3^{\text{old}} - A_2^{\text{new}} - A_3^{\text{new}}$. We obtain the following result:

Proposition 2. *If the ability difference Δm is strictly positive the expected ability of the selected worker is lower with a new manager than with the old one, i.e. $\Delta a > 0$. Furthermore, $\Delta a = 0$ if $\Delta m_1 = 0$ and Δa is strictly increasing in the expected type difference Δm_1 for all values of Δm_1. It tends to infinity for $\Delta m_1 \to \infty$.*

Proof. See the appendix.

The old manager is better informed about the workers' types, thus Δa is positive. To see that the expected ability loss strictly increases in the type difference Δm_1, note that the larger the difference in first period signals the higher the costs of installing a new manager as the owner can be less sure that the better worker is chosen in the last two periods.

4.3. The optimal dismissal decision

The size of *both* effects depends only on the expected type difference Δm_1. Given her information at the beginning of $t = 2$, the owner will dismiss the old manager if and only if $\Delta e - \Delta a \geqslant 0$. This will happen only if the effort effect is large enough, which depends on how strongly the workers are interested in the important position, as Δe strictly increases in the size of the intrinsic utility Δu. For simplicity we now impose a quadratic cost function, but the result qualitatively holds for more general convex cost functions.[13]

The optimal dismissal policy is given in the following proposition:

Proposition 3. *For very large or very small values of the expected ability difference Δm_1 it is never optimal to dismiss the manager*. (i) *For values of the players' utility*

[13] For all cost functions with $c'''(e) \leqslant 0$ and $c''(0) > 0$ the effort effect will always dominate the ability effect for some intermediate values of Δm_1 if only Δu is sufficiently large. The proof can be obtained from the authors.

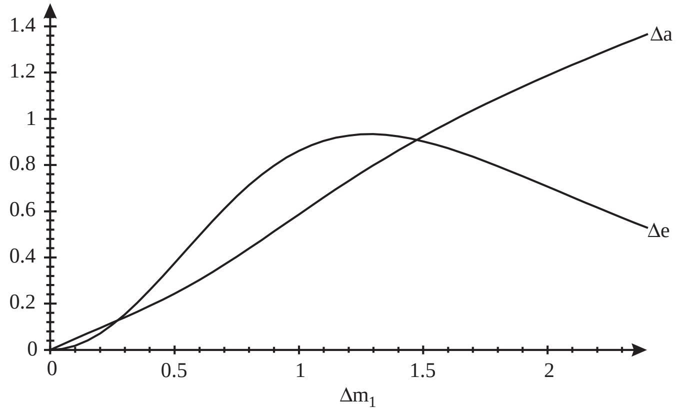


Fig. 2. Effort and ability effect.

gain Δu smaller than some threshold, it is never optimal to dismiss the manager. (ii) *For larger values of Δu, there exists at least one interval $[\Delta\underline{m}_1, \Delta\bar{m}_1]$, $\Delta\underline{m}_1 > 0$, $\Delta\bar{m}_1 < \infty$, such that dismissing the manager is optimal if $\Delta m_1 \in [\Delta\underline{m}_1, \Delta\bar{m}_1]$*. (iii) *For Δu large enough, there is a unique interval such that dismissing the manager is optimal if and only if Δm_1 is inside this interval.*

Proof. See the appendix.

In Fig. 2 the effort and the ability effect are plotted for an example where again the existence of the equilibrium is ensured. Dismissing the manager is optimal if and only if Δm_1 is in the interval where the graph of Δe lies above the graph of Δa.

The nonmonotonicity can be explained as follows: If Δm_1 is very large we know that the ability effect will be large: The owner expects that one of the two workers is far worse than the other and when employing a new manager she takes the higher risk that this worse type will be selected. At the same time we know from Lemma 3 that for large values of Δm_1 the effort effect will be small.

In contrast, for very small values of Δm_1 the ability effect will be small. For $\Delta m_1 = 0$ the old manager's optimal selection is exactly the same as the new manager's: As he has gained no information from the first period, he also selects the worker with the higher second period performance. The effort effect is zero for the same reason in that case. However, the ability effect is a first-order effect with respect to Δm_1 whereas the effort effect is of second order. Therefore, the ability effect dominates the effort effect for small values of Δm_1.[14] Only for intermediate values of Δm_1, dismissing the manager increases expected profits. Here the ability effect is not yet too large and the effort effect from intensifying the competition by installing an uninformed new manager is large enough to overcompensate the ability loss.

[14] Compare for a similar result in biased tournaments Meyer (1992, p. 174).

Given this policy, the workers will choose the optimal first period effort levels, which in turn all players will use to calculate the expected ability difference Δm_1. These levels will be smaller than those exerted if the owner was not able to dismiss the manager [15] as they have a lower incentive to impress a superior who might not be in charge anymore tomorrow.

5. Conclusion

Our model does not yield a rule like "Dismiss the manager if performance is low" which might seem plausible at first glance. Rather, the owner should use her information on workers' heterogeneity to judge whether competition among the subordinates can be strengthened sufficiently with a new manager. If it does, dismissing the manager can be beneficial, but this can be the case with high as well as low profits.

Our informational assumptions can be weakened in two respects: If the owner has no information on performance, it can be shown, that she will dismiss the manager in period 2 if Δu is large enough. But even if the owner observes only aggregate profits and can pass this information to a new manager, the manager will be dismissed if Δu is sufficiently large and the realization of the first period profit is not extreme. [16] We assumed that performance contingent contracts are infeasible. Our results continue to hold qualitatively if position contingent payments can be made: As incentives are provided only by the worker's competition for Δu, the principal can only increase this gain by paying a higher wage for the favorable position.

A similar effect to ours may be found in a different set-up if the new manager has the same information as the old one but values different characteristics. In this case the dismissal of the old manager should lead to higher incentives as the subordinates might want to convince a new manager that they do have those characteristics. [17]

It might be interesting to think about how the theory can be tested empirically. There are some studies on the effects of the dismissal of coaches in football as data are readily available. Breuer and Singer (1996) have investigated dismissals of football managers in the German Bundesliga. They find some evidence for a short-run impact of the dismissal of a coach: Teams whose coach has been dismissed performed significantly better in the 4 games after the dismissal than teams that kept their coaches ranking one place higher at the time of the dismissal. [18] No significant effect appeared afterwards. Note that our model does not yield a simple prediction such as: Dismissing the manager leads to a higher team performance. A dismissal rather has positive and negative effects. One interpretation of those results would be that the dismissal yields a performance push, but this push is only transitory as the new coach learns more and more on the player's abilities. A testable implication of our model is that the effect should be weak for either very homogenous or very heterogenous teams.

[15] Although we do not solve the game back to the first period explicitly, the manager is always dismissed with positive probability if an interval for Δm_1 exists where a dismissal is optimal.

[16] Formal proofs of both statements can be obtained from the authors.

[17] We thank an anonymous referee for pointing this out.

[18] However, for other control groups no significant difference was detected.

Similar effects should in principle be observable within firms or other organizations. For instance, our model might suggest that subordinates' absence rates are lower or the number of proposals for process innovations higher for some time after a manager is replaced.[19] An empirical investigation of the performance effects of management changes within organizations by testing similar hypotheses would be an interesting project for future research.

Acknowledgements

We are grateful to James Dow, Tore Ellingsen, Matthias Kräkel, Georg Nöldeke, Stefan Reichelstein, P. Schütz, Patrick Schmitz, Urs Schweizer, Roland Singer and Thomas Tröger as well as the Editor Klaus Schmidt and two anonymous referees for many helpful and encouraging comments or discussions. Financial support from Deutsche Forschungsgemeinschaft, SFB 303 at the University of Bonn is gratefully acknowledged.

Appendix A.

Proof of Lemma 1. Let $\hat{e}_2^i$ be the equilibrium efforts. We have that

$$P\{m_2^i > m_2^j | I_1, \hat{e}_2^j\} = P\{\tfrac{\alpha_2}{1-\alpha_2}\Delta m_1^i + e_2^i - \hat{e}_2^i > \eta^j - \eta^i + \varepsilon_2^j - \varepsilon_2^i | I_1\}$$

with $\Delta m_1^i := m_1^i - m_1^j$. Define $\delta^i := \eta^j - \eta^i + \varepsilon_2^j - \varepsilon_2^i$ with $g^i(\delta^i | I_1)$ as conditional density. The first-order condition of i's maximization problem is

$$g^i\left(\frac{\alpha_2}{1-\alpha_2}\Delta m_1^i + e_2^i - \hat{e}_2^i | I_1\right)\Delta u = c'(e_2^i). \tag{A.1}$$

As $\delta^a = -\delta^b$ and $g^a(x|I_1) = g^b(-x|I_1)$, we have $e_2^a = e_2^b$. Furthermore,[20]

$$\mathrm{E}[\delta^i | I_1] = \mathrm{E}[\eta^j - \eta^i + \varepsilon_2^j - \varepsilon_2^i | \eta^i + \varepsilon_1^i, \eta^j + \varepsilon_1^j] = -\Delta m_1^i, \tag{A.2}$$

$$V[\delta^i | I_1] = V[\eta^j - \eta^i | \eta^i + \varepsilon_1^i, \eta^j + \varepsilon_1^j] + V[\varepsilon_2^i - \varepsilon_2^j] = \frac{2h_0 + 4h_\varepsilon}{(h_0 + h_\varepsilon)h_\varepsilon}. \tag{A.3}$$

Using this, (4) is equivalent to (A.1). □

[19] There are some findings indicating that career concerns may matter for absence rates. Barmby et al. (2002) find that absence rates increase with tenure even when age is controlled for and are lower for jobs with more responsibility (where career prospects may be more important). Ichino and Riphahn (2001) discuss whether career concerns may explain a tenure effect. We thank an anonymous referee for pointing out those references.

[20] For the conditional variance see e.g. DeGroot (1970, p. 167).

Proof of Lemma 2. If $e_2^i = e_2^j$ a new manager chooses agent i in $t=3$ if $\mathrm{E}[\eta^i|y_2^i, y_2^j] > \mathrm{E}[\eta^j|y_2^i, y_2^j] \Leftrightarrow y_2^i > y_2^j$. Worker i exerts e_2^i to optimize $P\{y_2^i > y_2^j|I_1, \hat{e}_2^j\}\Delta u - c(e_2^i)$. We then proceed as in Lemma 1 now investigating $g^i(0|I_1)$. □

Proof of Lemma 3. (i) If $\Delta m_1 = 0$, we have that $e_2^{\mathrm{N}} = e_2^{\mathrm{O}}$. By implicit differentiation of e_2^{N} and e_2^{O} we get $\partial\Delta e/\partial\Delta m_1|_{\Delta m_1=0} = 0$. Furthermore, $e_2^{\mathrm{N}}, e_2^{\mathrm{O}} \to 0$ for $\Delta m_1 \to \infty$.

(ii) With the quadratic cost function the effort effect can be calculated as

$$\Delta e = \frac{\Delta u}{\sqrt{2\pi\sigma_\delta^2 k}}\left(\exp\left(-\frac{\Delta m_1^2}{2\sigma_\delta^2}\right) - \exp\left(-\frac{1}{2\sigma_\delta^2}\left(\frac{h_0+2h_\varepsilon}{h_\varepsilon}\Delta m_1\right)^2\right)\right). \tag{A.4}$$

To see that $\Delta e(\Delta m_1)$ has a unique maximum note that

$$\frac{\partial\Delta e}{\partial\Delta m_1} > 0 \quad\Leftrightarrow\quad \exp\left(\frac{\Delta m_1^2}{2\sigma_\delta^2}\left(1-\left(\frac{h_0+2h_\varepsilon}{h_\varepsilon}\right)^2\right)\right) > \left(\frac{h_\varepsilon}{h_0+2h_\varepsilon}\right)^2.$$

The right-hand side is strictly between zero and one. The inequality is met for small values of Δm_1 as the left-hand side gets arbitrarily close to 1 if Δm_1 small enough. It is not met if Δm_1 bigger than a threshold, as the left-hand side strictly decreases in Δm_1 and tends to 0 for $\Delta m_1 \to \infty$. □

A.1. Derivation of A_3^{old} and A_3^{new}

To derive A_3^{old} we reformulate (7) yielding

$$A_3^{\mathrm{old}} = \sum_{i=a,b}\left(P\left\{\eta^i > \eta^{-i} + \varepsilon_2^{-i} - \varepsilon_2^i - \frac{h_0+h_\varepsilon}{h_\varepsilon}\Delta m_1|I_1\right\}\right.$$
$$\left.\times\mathrm{E}\left[\eta^i|\eta^i > \eta^{-i} + \varepsilon_2^{-i} - \varepsilon_2^i - \frac{h_0+h_\varepsilon}{h_\varepsilon}\Delta m_1; I_1\right]\right).$$

From result R.187ii in Gourieroux and Monfort (1989, p. 528) we derive

$$\mathrm{E}[X|X>Y]\cdot P\{X>Y\}$$
$$= m_X\Phi\left(\frac{m_X-m_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}}\right) + \frac{\sigma_X^2}{\sqrt{\sigma_X^2+\sigma_Y^2}}\phi\left(\frac{m_X-m_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}}\right). \tag{A.5}$$

Using that[21]

$$\mathrm{E}\left[\eta^b + \varepsilon_2^b - \varepsilon_2^a - \frac{h_0+h_\varepsilon}{h_\varepsilon}\Delta m_1|I_1\right] = m_1^b - (h_0+h_\varepsilon)/h_\varepsilon\Delta m_1,$$

$$V[\eta_a|I_1] = 1/(h_0+h_\varepsilon)$$

[21] For the conditional variance of η^a see again DeGroot (1970, p. 167).

and

$$V\left[\eta^b + \varepsilon_2^b - \varepsilon_2^a - \frac{h_0 + h_\varepsilon}{h_\varepsilon}\Delta m_1 | I_1\right] = (2h_0 + 3h_\varepsilon)/((h_0 + h_\varepsilon)h_\varepsilon)$$

and rearranging terms we obtain

$$A_3^{\text{old}} = m_1^b + \Delta m_1 \Phi\left(\Delta m_1 \frac{h_0 + 2h_\varepsilon}{\sigma_\delta h_\varepsilon}\right) + \frac{2}{\sigma_\delta(h_0 + h_\varepsilon)}\phi\left(\Delta m_1 \frac{h_0 + 2h_\varepsilon}{h_\varepsilon \sigma_\delta}\right). \tag{A.6}$$

To derive A_3^{new} we apply again (A.5), now using $V[\eta^b + \varepsilon_2^a + \varepsilon_2^b | I_1] = 2h_0 + 3h_\varepsilon/(h_0 + h_\varepsilon)h_\varepsilon$:

$$A_3^{\text{new}} = m_1^b + \Delta m_1 \Phi\left(\frac{\Delta m_1}{\sigma_\delta}\right) + \frac{2}{(h_0 + h_\varepsilon)\sigma_\delta}\phi\left(\frac{\Delta m_1}{\sigma_\delta}\right). \tag{A.7}$$

Hence,

$$\begin{aligned}\Delta a &= A_2^{\text{old}} + A_3^{\text{old}} - A_2^{\text{new}} - A_3^{\text{new}} \\ &= \frac{\Delta m_1}{2} + \Delta m_1\left(\Phi\left(\Delta m_1 \frac{h_0 + 2h_\varepsilon}{\sigma_\delta h_\varepsilon}\right) - \Phi\left(\frac{\Delta m_1}{\sigma_\delta}\right)\right) \\ &\quad + \frac{2}{(h_0 + h_\varepsilon)\sigma_\delta}\left(\phi\left(\Delta m_1 \frac{h_0 + 2h_\varepsilon}{\sigma_\delta h_\varepsilon}\right) - \phi\left(\frac{\Delta m_1}{\sigma_\delta}\right)\right).\end{aligned} \tag{A.8}$$

Proof of Proposition 2. The ability loss is given by (A.8), which we rewrite as

$$\begin{aligned}\Delta a = &\frac{\Delta m_1}{2} + \Delta m_1\left[\Phi\left(\frac{\gamma \Delta m_1}{\sigma_\delta}\right) - \Phi\left(\frac{\Delta m_1}{\sigma_\delta}\right)\right] \\ &+ \frac{\sigma_\delta}{\gamma}\left[\phi\left(\frac{\gamma\Delta m_1}{\sigma_\delta}\right) - \phi\left(\frac{\Delta m_1}{\sigma_\delta}\right)\right],\end{aligned}$$

where again $\Phi(x)$ is the distribution function, and $\phi(x)$ the density of a standard normal distribution and $\gamma := (h_0 + 2h_\varepsilon)/h_\varepsilon > 1$. Taking the derivative with respect to the ability difference Δm_1 and rearranging terms we get

$$\frac{\partial \Delta a}{\partial \Delta m_1} = \frac{1}{2} + \underbrace{\Phi\left(\frac{\gamma\Delta m_1}{\sigma_\delta}\right) - \Phi\left(\frac{\Delta m_1}{\sigma_\delta}\right)}_{A(\Delta m_1)} - \underbrace{\phi\left(\frac{\Delta m_1}{\sigma_\delta}\right)\frac{\Delta m_1}{\sigma_\delta}\frac{\gamma - 1}{\gamma}}_{B(\Delta m_1)} > 0.$$

The term $A(\Delta m_1)$ is strictly positive as $\gamma > 1$, and so is $B(\Delta m_1)$, but $B(\Delta m_1)$ strictly smaller than $\frac{1}{2}$: To see this check that $B(\Delta m_1)$ is maximized at $\Delta m_1^* = \sigma_\delta$. Thus

$$\max_{\Delta m_1} B(\Delta m_1) = \frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{2})\frac{\gamma - 1}{\gamma} \approx 0.096\frac{\gamma - 1}{\gamma} < 0.5$$

which proves that indeed $\partial\Delta a/\partial\Delta m_1 > 0$. Furthermore, it is easy to check that $\Delta a \to \infty$ for $\Delta m_1 \to \infty$. To see that Δa is always positive just verify that $\Delta a = 0$ for $\Delta m_1 = 0$. □

Proof of Proposition 3. The total gain of dismissing the manager depends only on the expected ability difference Δm_1 and is given by $\Delta e(\Delta m_1) - \Delta a(\Delta m_1)$. As $\partial\Delta a/\partial\Delta m_1|_{m_1=0} = \frac{1}{2}$ larger than $\partial\Delta e/\partial\Delta m_1|_{m_1=0} = 0$ for small values of Δm_1 the total effect is negative. $\Delta a(\Delta m_1)$ is always positive: it starts at 0 for $\Delta m_1 = 0$ and is strictly increasing. $\Delta e(0)$ is also equal to 0 and is strictly positive for $\Delta m_1 > 0$, furthermore it is strictly quasiconcave and tends to 0 for $\Delta m_1 \to \infty$, hence for large Δm_1 again the total effect is negative. For Δu large enough there exist a lowest and highest intersection between $\Delta a(\Delta m_1)$ and $\Delta e(\Delta m_1)$, intermediate intersections can be ruled out for Δu sufficiently large. □

References

Barmby, T., Ercolani, M., Treble, J., 2002. Sickness absence: An international comparison. The Economic Journal 112, 315–331.

Bhattacharya, S., Guasch, J., 1988. Heterogeneity, tournaments, and hierarchies. Journal of Political Economy 96, 867–881.

Bolton, P., 1992. The theory of organizations: Discussion of Harris and Raviv, and Tirole. In: Laffont, J. (Ed.), Advances in Economic Theory: Proceedings of the Sixth World Congress of the Econometric Society. Cambridge University Press, Cambridge.

Breuer, C., Singer, R., 1996. Trainerwechsel im Laufe der Spielsaison and ihr Einfluss auf den Mannschaftserfolg. Leistungssport 4, 41–46.

Carmichael, L., MacLeod, B., 1993. Multiskilling, technical change and the Japanese firm. The Economic Journal 103, 142–160.

Crémer, J., 1995. Arm's length relationships. Quarterly Journal of Economics 110, 275–295.

DeGroot, M., 1970. Optimal Statistical Decisions. McGraw-Hill, New York.

Fairburn, J., Malcomson, J., 1994. Rewarding performance by promotion to a different job. European Economic Review 38, 683–690.

Gourieroux, C., Monfort, A., 1989. Statistique et Modèles Économétriques, Vol. 2. Economica, Paris.

Green, J., Stokey, N., 1983. A comparison of tournaments and contracts. Journal of Political Economy 91, 349–364.

Ichino, A., Riphahn, R.T., 2001. The effect of employment protection on worker effort. A comparison of absenteeism during and after probation. IZA Discussion Paper 385.

Ickes, B., Samuelson, L., 1987. Job transfers and incentives in complex organizations: Thwarting the ratchet effect. Rand Journal of Economics 18, 275–286.

Lazear, E., Rosen, S., 1981. Rank order tournaments as optimum labour contracts. Journal of Political Economy 89, 841–864.

Meyer, M., 1992. Biased contests and moral hazard: Implications for career profiles. Annales D'Économie et de Statistique 25/26, 165–187.

Meyer, M.A., Vickers, J., 1997. Performance comparisons and dynamic incentives. Journal of Political Economy 105, 547–581.

Nalebuff, B., Stiglitz, J., 1983. Prizes and incentives: Towards a general theory of compensation and competition. Bell Journal of Economics 3, 21–43.

Prendergast, C.J., 1993. The role of promotion in inducing specific human capital acquisition. Quarterly Journal of Economics 108, 523–534.

Waldman, M., 1990. Up-or-out contracts: A signaling perspective. Journal of Labor Economics 8, 230–250.