GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 112

Job Promotion Tournaments and Imperfect Recall

Oliver Gürtler*

May 2006

*Oliver Gürtler, Department of Economics, BWL II, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany. Tel.:+49-228-739214, Fax:+49-228-739210; E-mail: oliver.guertler@uni-bonn.de

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.gesy.uni-mannheim.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Konrad Stahl, Ph.D. · Department of Economics · University of Mannheim · D-68131 Mannheim,
Phone: +49(0621)1812786 · Fax: +49(0621)1812785

Job Promotion Tournaments and Imperfect Recall*

Oliver Gürtler**, *University of Bonn*

Abstract

In this paper, a promotion tournament is considered, where, at the beginning of the tournament, it is unknown how long the tournament lasts. Further, the promotion decision is based on the assessments of a supervisor with imperfect recall. In line with psychological research, the supervisor is assumed to either value early or recent impressions more strongly. It is shown that effort may increase or decrease, as the probability of promotion in a certain period gets higher. The single effects determining the sign of the effort change oftentimes depend on how the supervisor processes information.

Key words: Promotion Tournament, Promotion Probability, Imperfect Recall

JEL classification: J33, M51, M52

*I would like to thank Marion Eberlein, Matthias Kräkel and Anuschka Sieber for helpful comments. Financial support by the Deutsche Forschungsgemeinschaft (DFG), SFB-TR 15 (Governance and the efficiency of economic systems), is gratefully acknowledged.

**Oliver Gürtler, Department of Economics, BWL II, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany. Tel.:+49-228-739214, Fax:+49-228-739210; E-mail: oliver.guertler@uni-bonn.de

1 Introduction

In the economic literature (see e.g. Lazear & Rosen (1981), Green & Stokey (1983), Kräkel (2002), (2003), Chen (2003), or Dubey & Haimanko (2003)), promotion tournaments are usually modeled in the following way: There are several employees (he) who compete for a promotion and thus exert effort. Effort in combination with some random components results in individual output. Then, the employee attaining highest individual output is promoted to a vacant workplace at a higher layer in the hierarchy.

In my view, there are two main reasons, why this approach is inappropriate or at least insufficient to model promotion decisions. First, in contrast to other incentive devices such as e.g. piece-rates or bonus payments, it is ex ante oftentimes unclear, when the promotion takes place and, accordingly, when the wage increase connected with the promotion is realized, i.e. when the winning employee is rewarded for outperforming the other employees. Only rarely do the parties know for sure that a workplace becomes vacant at a certain point in time and so when a promotion actually takes place. More frequently, a workplace should become vacant, if, unexpectedly, an employee leaves the firm. Hence, instead of being sure that a promotion takes place after efforts have been chosen, employees should assign probabilities to workplaces becoming vacant at certain points in time. Second, promotions are rarely linked to verifiable measures such as e.g. individual output. On the contrary, there are usually supervisors (she) deciding about whom to promote. That is, promotion decisions are mainly based on subjective assessments of employees' qualities. In this context, it is probably the case that supervisors have imperfect recall, i.e. when deciding about which employee to promote, they do not recall all their previous impressions of the employees.

I set up a model, where the two points of criticism are taken into account. A multi-period setting is considered, where, at some point in time, a vacancy on a higher layer in the hierarchy may arise. In this case, a supervisor decides about which employee to promote to this vacancy. Thereby, she takes into account impressions of the employees she has obtained in previous periods. Incorporating different ideas from psychology, a case distinction is introduced, when modeling, how the supervisor aggregates the impressions to a single evaluation. First, recent impressions are given higher weights in order to account for the fact that recent impressions are more likely to be recalled. Second, I attach lower weights to recent impressions. This implies that early impressions are most important for the evaluation process. The supervisor might form a view on the employees in early periods and rethink this view, only if she obtains impressions that extremely contradict her view.

In this setting, the determinants of employees' efforts are analyzed. Assuming that the occurrence probability of the vacancy is constant over time, it is focused on how efforts depend on this probability. If the probability changes, promotions after some periods become less likely, while they may become more likely after other periods. Thus, when deciding about their efforts, employees switch attention from the former periods to the latter ones. Then, four effects can be eliminated determining whether or not efforts increase. Interestingly, the sign of these effects oftentimes depends on our assumptions about the supervisor, i.e. on whether or not the supervisor values recent impressions more strongly.

The remainder of the paper is organized as follows: The next section gives some arguments for the behavioral assumptions concerning the supervisor. In Section 3, the basic model is presented. Section 4 contains several model extensions and Section 5 concludes.

2 Information processing in the human brain

In contrast to computers the human brain does not store all the information it receives. Forgetting is a natural task of the brain so that a person's recall is always imperfect. Despite the clearness of this finding, economic models usually consider totally rational decision makers being aware of the complete history of the model, i.e. having perfect recall.[1] I deviate from this assumption and consider a supervisor promoting one of two subordinates on the basis of several impressions received in different periods. While a rational supervisor with perfect recall would make his decision based on the sum of all single impressions, a supervisor with imperfect recall might decide differently. In order to make a more founded prediction of how imperfect recall affects the promotion decision, let us take a look at psychological work dealing with storage of information in the human brain and subjective

[1] There exist a few papers considering the difficulties in transferring results from models with perfect recall to models with imperfect recall. See e.g. Kuhn (1953), Piccione & Rubinstein (1997), or for a textbook treatment of imperfect recall the fourth chapter in Rubinstein (1998).

evaluations of people.

It is clear that, due to the complexity of the topic, only an incomplete overview can be given. I thus restrict attention to two opposing effects that seem to be of particular importance. First, psychologists argue that people are more likely to retrieve recently and frequently used concepts than concepts not recently or frequently activated.[2] Events that occurred a short time ago are often remembered, while earlier events are harder to recall. Applied to our problem, this means that supervisors should be most likely to recall impressions that have been formed not so long ago.[3]

Although this argumentation is quite intuitive, things may work the other way round. In the context of information processing in the human brain, psychologists emphasize the important role of priming.[4] Loosely speaking, priming means that information which is conflicting with or unrelated to previous impressions is less easily stored than information that is in line with previous impressions. A simple example may underline this effect. Consider a woman who is going to meet a man for the first time and is told that the man has a particular habit. During the meeting, the woman stores information fitting in her picture of the man, i.e. being connected with the man's habit, more easily than other information. For our model, this implies that early impressions are of higher importance than later ones, for the supervisor is

[2] See e.g. the textbook by Fiske & Taylor (1984), p. 244.

[3] Note that this effect might be further enforced, if a person's characteristics change over time. In this case, it is very natural to attach a higher weight to recent impressions.

[4] See for experiments highlighting the role of priming e.g. Higgins et al. (1977) or Srull & Wyer (1979).

little likely to reassess her view on the employees. Note that this is very similar to the conservatism effect that has been introduced by Edwards (1968) into the economic literature.

In order to give consideration to both approaches, a supervisor is considered who evaluates the employees according to a weighted sum of the single impressions she receives. Depending on monotonicity properties of the weights, both described ideas can be captured. If the impressions' weights are monotonically decreasing in the temporal distance from the time of evaluation, that is, if impressions from the past are less influential than present impressions, we are in the first case, where supervisors are likely to forget information that was obtained long ago. If, on the other hand, the weights are increasing in the temporal distance from the time of evaluation, the first impression is the most important one and we describe the second scenario. Finally, if the weights are constant, the supervisor is totally rational and takes all relevant information equally into account.

3 The basic model

Consider two risk-neutral employees who, at the same time $t = 1$, begin to work in a department of a firm, which is managed by a risk-neutral supervisor. Overall, each employee works $T > 1$ periods in the firm and discounts future utility at a common factor $\delta < 1$. In an arbitrary period t, employee i, $i = 1, 2$, exerts unobservable effort e_{it} in order to produce output $y_{it} = f(e_{it})$, where f, unless stated otherwise, is an increasing concave function.

Effort entails costs for the employee, which are given by the increasing and convex function $C(e_{it})$. This cost function satisfies $C(0) = 0$, $\frac{\partial C(0)}{\partial e_{it}} = 0$, and $\frac{\partial C(e_{it})}{\partial e_{it}} = \infty$, for $e_{it} \to \infty$, where the last two conditions ensure that, as long as monetary incentives are present, the employees' equilibrium efforts are always strictly positive, but finite.[5] A natural interpretation for this cost function is that an employee voluntarily exerts a certain effort, which is here normalized to zero. Any effort exceeding this voluntary level is costly to the employee. Output is supposed to be unobservable by all parties and so cannot be used to incentivize the employees. Instead, assume that the firm uses promotions as incentive device.[6] That is, the two employees know that, at some point in time, there may arise a vacancy on a higher layer in the hierarchy, which will be filled by one of them. On the higher layer, a wage of $w > 0$ is paid in every period, while the current wage is normalized to zero.[7] Furthermore, suppose that it is important for the firm to fill the vacancy. That is, an unfilled vacancy leads to highly negative profit. Therefore, the firm is always interested in filling the vacancy so that promotions are indeed a credible incentive device. In case the workplace on the higher layer is filled, let the output from this workplace be given by a constant $\bar{y}$. This

[5] In the model, it may sometimes be that monetary incentives vanish. In this case, the employees are indifferent between choosing zero effort and effort slightly above zero. It is then assumed that they decide for the former and choose zero effort.

[6] The optimality of promotions as incentive device is discussed in Section 4.3.

[7] In this section, we assume the wage structure to be exogenously given. This could e.g. be the case, if minimum wages are set by unions and the firm finds it unprofitable to raise wages above their minimum level. Section 4.1 extends the model to encompass endogenous wages.

simply means that this output depends on factors, which are beyond the scope of the model. The emergence of a vacancy on the higher layer follows a simple Markov process. In particular, the probability that, in any period, the vacancy arises and one of the employees is promoted equals λ and does not change from time to time.[8] Further, once the vacancy has been filled, there will never arise a second one.

If a vacancy arises, the supervisor will have to decide, which employee to promote. Her decision is based on her impression of the employees' performances that she has formed in earlier periods. As indicated before, we assume that the supervisor bases her decision on a weighted sum of her obtained impressions. Let the impression the supervisor forms of employee i in period t be given by[9]

$$s_{it} = e_{it} + \varepsilon_{it} \tag{1}$$

ε_{it} denotes an error term, which is independently[10] drawn from a normal distribution with zero mean and variance σ^2. It accounts for the fact that the supervisor cannot perfectly monitor each employee and, thus, does only get an imprecise impression of each employee's performance. Further, we make the assumption that the employees know the realizations of the s_{it}'s.

[8] Section 4.2 deals with varying occurrence probabilities.

[9] In this model, influence activities are neglected. Although subjective evaluation schemes generally suffer from influence activities, they seem to be less important for promotion decisions, as these are usually done more carefully. See, for an argumentation along these lines, e.g. Milgrom and Roberts (1992) p. 407.

[10] The independence assumption affects both, i and t. This means that the realization of the error term is independent of the opponent's error term realization and of realizations of the error terms from prior periods.

They could e.g. monitor each other and get similar impressions or deduce the supervisor's impressions from discussions with her. If the workplace becomes vacant directly after period t, the supervisor will promote the first employee, only if $\sum_{j=1}^{t} \theta_{t-j} (e_{1j} + \varepsilon_{1j} - e_{2j} - \varepsilon_{2j}) > 0$.[11] θ_0 is the weight for the most recent impressions, θ_1 for the impressions that have been obtained one period earlier and so on. Hence, in the model, where recent impressions play the most important role, we have $\theta_0 \geq \theta_1 \geq ... \geq \theta_T \geq 0$, with at least one inequality being strict. If the inequalities are reversed so that $0 \leq \theta_0 \leq \theta_1 \leq ... \leq \theta_T$, we are in the model, where early impressions are most important. Finally, the perfectly rational decision maker is a special case of this model, requiring $\theta_0 = ... = \theta_T > 0$.

Before turning to the model solution, it is convenient to define the distribution function of a normally distributed random variable with zero mean and variance $2\sigma^2 \sum_{j=0}^{m-1} \theta_j^2$ by $F_m(\cdot)$ and the corresponding density function by $f_m(\cdot)$. Further, let $s_{1t} - s_{2t}$, for $t \geq 1$, be given by Δs_t and $\varepsilon_{1t} - \varepsilon_{2t}$ by $\Delta \varepsilon_t$.

The model is solved by backward induction. In period T, the employees have no incentive to choose higher effort than their voluntary effort, hence $e_{1T} = e_{2T} = 0$.[12] This is also true in each period after the promotion tour-

[11] I implicitly assume that the supervisor always proposes the employee for promotion that she assesses as the better one. This assumption is relatively innocuous. If the decision rule was changed, effort incentives would break down and performance in the supervisor's department would decline. This would clearly be problematic for the supervisor.

[12] One might wonder, why the employees are not fired in the final period of their career. A possible reason is closely related to our assumption that employees voluntarily choose a

nament. That is, if a vacancy has arisen and one of the employees has been promoted, effort incentives disappear and optimal efforts are zero. Turn now to period $T-1$ and suppose that both employees are still on the lower layer of the hierarchy. The two employees' maximization problems are then, respectively, given by

$$\underset{e_{1T-1}}{Max} \quad F_1(\theta_0\,(e_{1T-1}-e_{2T-1})+\sum_{t=1}^{T-2}\theta_{T-1-t}\Delta s_t)\lambda w\delta - C(e_{1T-1}) \tag{2}$$

$$\underset{e_{2T-1}}{Max} \quad \left(1-F_1(\theta_0\,(e_{1T-1}-e_{2T-1})+\sum_{t=1}^{T-2}\theta_{T-1-t}\Delta s_t)\right)\lambda w\delta \tag{3}$$
$$-C(e_{2T-1})$$

Each employee maximizes the discounted payment in case of being promoted times the probability of being promoted minus costs entailed by effort. The first-order conditions to the two maximization problems are[13]

$$f_1(\theta_0\,(e_{1T-1}-e_{2T-1})+\sum_{t=1}^{T-2}\theta_{T-1-t}\Delta s_t)\lambda w\delta\theta_0 - \frac{\partial C(e_{1T-1})}{\partial e_{1T-1}} = 0 \tag{4}$$

$$f_1(\theta_0\,(e_{1T-1}-e_{2T-1})+\sum_{t=1}^{T-2}\theta_{T-1-t}\Delta s_t)\lambda w\delta\theta_0 - \frac{\partial C(e_{2T-1})}{\partial e_{2T-1}} = 0 \tag{5}$$

It is straightforward to see that the equilibrium is symmetric, i.e. both

certain effort level. This effort may be high enough to justify an employment in the final period. Formally, if we assume $f(0) \geq 0$, there is no need to fire the employees, if they choose zero effort. Note, however, that, normalizing the employees' reservation utilities to zero, the results remain unchanged, even if the employees are fired in the final period of their career.

[13] As shown by e.g. Lazear & Rosen (1981), the second-order conditions will only hold, if the density function $f_1(\cdot)$ is sufficiently flat or the effort cost function sufficiently steep. In what follows, it is assumed that this is the case so that the first-order conditions indeed characterize a maximum.

employees choose same effort $e_{1T-1} = e_{2T-1} =: e_{T-1}$. This effort is implicitly given by $\frac{\partial C(e_{T-1})}{\partial e_{T-1}} = f_1(\sum_{t=1}^{T-2} \theta_{T-1-t} \Delta s_t) \lambda w \delta \theta_0 =: Y$. As the effort cost function is convex, effort increases in the right-hand-side of the equality. For instance, $f_1(\sum_{t=1}^{T-2} \theta_{T-1-t} \Delta s_t)$ gets lower the more $\sum_{t=1}^{T-2} \theta_{T-1-t} \Delta s_t$ deviates from zero. The absolute value of this sum measures the head start of the currently leading employee before period $T-1$. Hence, if it increases, competition in period $T-1$ will become less intense and effort will decline. The following proposition states a corresponding result on the relation between e_{T-1} and λ.

Proposition 1 *If, in period $T-1$, both employees are still on the lower layer of the hierarchy, this period effort is strictly increasing in λ.*

Proof. As $C(e_t)$ is convex, $\frac{\partial e_{T-1}}{\partial \lambda} > 0 \Leftrightarrow \frac{\partial Y}{\partial \lambda} > 0$. This condition is equivalent to $f_1(\sum_{t=1}^{T-2} \theta_{T-1-t} \Delta s_t) w \delta \theta_0 > 0$, which is clearly satisfied. ∎

Proposition 1 is very intuitive. It says that the more likely it is that a promotion takes place, the higher is the effort chosen by the employees. This is not very surprising, as, in period $T-1$, efforts will be worthless, if no vacancy arises. In other words, in period $T-1$, efforts depend on the expected tournament prize λw so that higher λ makes it more worthwhile to exert effort.

Consider now period $T-2$ in order to see, whether or not the results from Proposition 1 will continue to hold, if we allow efforts to affect the promotion probabilities in different periods. The two employees' maximization problems are, in this case, given by (6) and (7), respectively.

$$\underset{e_{1T-2}}{Max} \quad F_1(\theta_0\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\lambda w\left(\delta+\delta^2\right) \tag{6}$$
$$+F_2\left(\theta_1\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)(1-\lambda)\lambda w\delta^2-C(e_{1T-2})$$
$$\underset{e_{2T-2}}{Max} \quad \left(1-F_1(\theta_0\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\right)\lambda w\left(\delta+\delta^2\right) \tag{7}$$
$$+\left(1-F_2\left(\theta_1\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)\right)(1-\lambda)\lambda w\delta^2-C(e_{2T-2})$$

Effort here affects two probabilities. It affects the promotion probability in the current period $(F_1(\theta_0\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\lambda)$ and in the following period $(F_2\left(\theta_1\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)\lambda(1-\lambda))$. Differentiating with respect to effort, yields the following first-order conditions:

$$\begin{aligned}0 &= f_1(\theta_0\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\lambda w\left(\delta+\delta^2\right)\theta_0 \qquad (8)\\ &\quad +f_2\left(\theta_1\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)(1-\lambda)\lambda w\delta^2\theta_1-\frac{\partial C(e_{1T-2})}{\partial e_{1T-2}}\\ 0 &= f_1(\theta_0\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\lambda w\left(\delta+\delta^2\right)\theta_0 \qquad (9)\\ &\quad +f_2\left(\theta_1\left(e_{1T-2}-e_{2T-2}\right)+\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)(1-\lambda)\lambda w\delta^2\theta_1-\frac{\partial C(e_{2T-2})}{\partial e_{2T-2}}\end{aligned}$$

Again, the equilibrium is symmetric, hence $e_{1T-2}=e_{2T-2}=: e_{T-2}$. The first-order conditions thus simplify to $\frac{\partial C(e_{T-2})}{\partial e_{T-2}}=f_1(\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\lambda w\left(\delta+\delta^2\right)\theta_0+$ $f_2\left(\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)(1-\lambda)\lambda w\delta^2\theta_1 =: Z$. From the last condition, it is easy to derive the next proposition:

Proposition 2 *Suppose that, in period $T-2$, both employees are still on the* *lower layer of the hierarchy. Then,* $\frac{\partial e_{T-2}}{\partial \lambda} > 0$, *if and only if* $f_1(\sum_{t=1}^{T-3} \theta_{T-2-t} \Delta s_t) \left(\delta + \delta^2\right) \theta_0 > (2\lambda - 1) f_2 \left(\sum_{t=1}^{T-3} \theta_{T-1-t} \Delta s_t\right) \delta^2 \theta_1$.

Proof. As $C(e_t)$ is convex, $\frac{\partial e_{T-2}}{\partial \lambda} > 0 \Leftrightarrow \frac{\partial Z}{\partial \lambda} > 0$. The latter condition is given by $f_1(\sum_{t=1}^{T-3} \theta_{T-2-t} \Delta s_t) w \left(\delta + \delta^2\right) \theta_0 - (2\lambda - 1) f_2 \left(\sum_{t=1}^{T-3} \theta_{T-1-t} \Delta s_t\right) w \delta^2 \theta_1 > 0 \Leftrightarrow f_1(\sum_{t=1}^{T-3} \theta_{T-2-t} \Delta s_t) \left(\delta + \delta^2\right) \theta_0 > (2\lambda - 1) f_2 \left(\sum_{t=1}^{T-3} \theta_{T-1-t} \Delta s_t\right) \delta^2 \theta_1$.
■

Proposition 2 offers the rather surprising result that effort may be decreasing in λ. To explain this, recall that, in period $T-2$, effort affects the promotion probability after two periods, period $T-2$ and period $T-1$. The two events, promotion after period $T-2$ or after period $T-1$, are attached weights according to their probabilities of occurrence. If λ decreases, the first event becomes less likely, whereas the second event becomes more likely, iff $\lambda > 0.5$. If a decrease in λ makes no event more likely, i.e. if $\lambda \leq 0.5$, effort will certainly decrease. This is similar to the effect in Proposition 1. If $\lambda \leq 0.5$, it becomes more likely that no promotion takes place at all and effort is completely worthless. Hence, the expected reward for performing best decreases and so do optimal efforts. Things become more complicated in case $\lambda > 0.5$. Then, the employees switch their attention, while deciding about how much effort to provide, from period $T-2$ to period $T-1$. Four effects influencing the provision of effort must, in this case, be distinguished:

First, there is a simple time effect. The later an employee is promoted, the less periods remain, where the employee receives the higher wage. Therefore, a promotion after period $T-2$ is assigned a weight $\delta + \delta^2$, whereas the corresponding weight for a promotion after period $T-1$ is δ^2. Hence, this effect entails a decrease in effort, when attention is switched from period $T-2$ to period $T-1$.

Second, there is a variance effect. An evaluation after period $T-2$ is connected with less exogenous noise than an evaluation after period $T-1$. This may either enhance or decrease effort. The first case arises, if, at the beginning of period $T-2$, one employee has a high head start. In this case, competition is very low and will be intensified, if the impact of random components on the promotion decision gets higher. With more noise, the initial head start is less decisive for the promotion decision and it is more worthwhile to exert effort. If, on the other hand, no employee has a very high head start at the beginning of period $T-2$, more noise will decrease efforts, as, in this case, more noise reduces the effect of effort on the promotion decision. Graphically, this effect can be nicely demonstrated by plotting the density functions of two normal distributions with zero mean and variances $\sigma_1^2 = 2\sigma^2\theta_0^2$, $\sigma_2^2 = 2\sigma^2(\theta_0^2 + \theta_1^2)$ into one diagram. One can then see that, for values close to zero, the density function of the first distribution (the distribution with the lower variance) lies above the other density function, while, for values far from zero, the relationship is the other way round.

Third, there is a judgment effect. Effort exerted in period $T-2$ may affect the supervisor's judgment after periods $T-2$ and $T-1$ differently. If a promotion takes place after period $T-2$, the productivity of e_{T-2} in influencing the supervisor is given by θ_0. In contrast, the respective productivity in case a promotion takes place after period $T-1$ is θ_1. Hence, the change in effort due to the judgment effect is positive (negative, zero), if $\theta_1 > (<,=)\theta_0$. In words, if the supervisor values past impressions more (less, equally) strongly than present ones, a switch in attention from period $T-2$ to $T-1$ has a positive (negative, neutral) effect on e_{T-2}, as e_{T-2} is more (less, equally) productive in influencing the supervisor's judgment after period $T-1$ than after period $T-2$.

Finally, there is a weight effect. If a period passes, the weights the supervisor attaches to different impressions may change. This implies that unequal starting positions in period $T-2$ might be transformed into less or even more unequal starting positions in period $T-1$. To assess this effect on a more formal level compare the two sums in Proposition 2 and assume without loss of generality that $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t > 0$. We want to determine, when the first sum exceeds the absolute value of the second sum. Suppose, for the moment, that the single weights are interconnected in that $\theta_t = k\theta_{t-1}$, for $t = 1, ..., T$, with $k > 0$. Then, we get the following Lemma:

Lemma 3 *Let $\theta_t = k\theta_{t-1}$, for $t = 1, ..., T$, and $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t > 0$. Then,*

$$\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t \left\{\begin{array}{c} > \\ = \\ < \end{array}\right\} \left|\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right| \Longleftrightarrow k \left\{\begin{array}{c} < \\ = \\ > \end{array}\right\} 1 \tag{10}$$

Proof. We first show that $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t > 0$ together with $\theta_t = k\theta_{t-1}$ implies that $\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t > 0$. From $\theta_t = k\theta_{t-1}$, it follows that $\theta_t = k^t\theta_0$. Then, $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t$ can be rewritten as $\sum_{t=1}^{T-3}k^{T-2-t}\theta_0\Delta s_t = \frac{1}{k}\sum_{t=1}^{T-3}k^{T-1-t}\theta_0\Delta s_t = \frac{1}{k}\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t$. It directly follows that $\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t$ is positive. To prove Lemma 3, it remains to be shown when $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t > \sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t$. Using $\sum_{t=1}^{T-3}k^{T-2-t}\theta_0\Delta s_t = \frac{1}{k}\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t$, this condition simplifies to $k<1$. Similarly, $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t < (=) \sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t \Leftrightarrow k > (=)1$. ■

Again, there is a simple intuition behind this Lemma. Employee 1, i.e. the initially leading employee (recall the assumption $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t > 0$), must have reached a better impression than employee 2 in at least one period. On the other hand, employee 2 may or may not have reached better impressions than employee 1 in some periods. If the supervisor attaches a higher weight to recent (early) impressions, these better impressions in previous periods are given higher (lower) weight in period $T-2$ than in period $T-1$. Hence, with $\theta_t = k\theta_{t-1}$, the head start of the initially leading employee is more (less) likely to be outweighed in period $T-1$ than in period $T-2$. A switch in attention from period $T-2$ to period $T-1$ then yields more (less) intense competition and thus higher (lower) efforts.

Note that this argumentation does not hold for arbitrary weight sequences $\theta_{(t)}$. For example, under an arbitrary sequence, the weights of the currently losing employee's better impressions might decrease stronger than the other weights and this may entail lower competition. To demonstrate this, suppose that we consider the fourth period[14] and that $\Delta s_1 = 0$, $\Delta s_2 = -1$, $\Delta s_3 = 1$, $\theta_0 = \theta_1 = 1$, $\theta_2 = 0.9$, $\theta_3 = \theta_4 = ... = 0$. In this case, $\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t = 0.1 > 0$ and $\sum_{t=1}^{T-3}\left(\theta_{T-2-t} - \theta_{T-1-t}\right)\Delta s_t = -0.8 < 0$. Hence, although recent impressions are attached highest weights, the starting positions become more uneven, as time passes.

[14] Note that this implies that $T = 6$.

Despite this counter example, it is oftentimes the case that, even for arbitrary weight sequences, the results from Lemma 3 remain valid. If time passes and the supervisor attaches highest value to recent impressions, unequal starting positions tend to become more even. Further, notice that, when all Δs_t have same sign, the results from Lemma 1 definitely hold, even for arbitrary weight sequences.

Summarizing, we have seen that the signs of the third and the fourth effect depend on how the supervisor processes information. If early impressions are of particular relevance, the third effect is positive and the fourth effect likely to be negative. Similarly, if the supervisor attaches a higher weight to recent impressions (acts as a totally rational decision maker), the third effect is negative (neutral) and the fourth effect probably positive (neutral).

Continuing backward induction, one can see that the results derived in Proposition 2 combined with the given intuition are quite robust. In a given period, a change in λ alters the occurrence probabilities of the events "promotion after the current period", "promotion after the next period" and so on. Effort is chosen such that the new occurrence probabilities are properly taken into account. Some events become more likely and are attached a higher weight, while others become less likely and are attached a lower weight. To decide, whether or not effort increases, one has to trade-off effort incentives concentrating on the events that become more likely with incentives concentrating on the events that become less likely. The effects determining this trade-off are the same effects as described above. That is, determining the sign of $\frac{\partial e_t}{\partial \lambda}$, for $t = 1, ..., T-3$ may become more difficult, but the effects affecting the sign do not change.

Consider some arbitrary period t. With a similar reasoning as before, one can show that optimal efforts are again symmetric and given by

$$\frac{\partial C(e_t)}{\partial e_t} = \frac{\lambda w}{1-\delta} \sum_{k=1}^{T-t} f_k \left(\sum_{j=1}^{t-1} \theta_{t-j+k-1} \Delta s_j \right) (1-\lambda)^{k-1} \theta_{k-1} (\delta^k - \delta^{T-t+1}) \quad (11)$$

It is then very easy to show that $\frac{\partial e_t}{\partial \lambda} > 0$ is equivalent to

$$\begin{aligned} & \sum_{k=1}^{T-t} f_k \left(\sum_{j=1}^{t-1} \theta_{t-j+k-1} \Delta s_j \right) (1-\lambda)^{k-1} \theta_{k-1} (\delta^k - \delta^{T-t+1}) \quad (12) \\ > \; & \lambda \sum_{k=1}^{T-t} f_k \left(\sum_{j=1}^{t-1} \theta_{t-j+k-1} \Delta s_j \right) (k-1)(1-\lambda)^{k-2} \theta_{k-1} (\delta^k - \delta^{T-t+1}) \end{aligned}$$

, which can be interpreted as explained before.

4 Extensions

In this section, the model assumptions are relaxed in several ways. First (in Section 4.1), I discuss, whether or not the model results will change, if we endogenize the wages. Thereafter (in Section 4.2), I let the probability that a vacancy arises differ from period to period. Finally, in Section 4.3, I discuss the optimality of the used incentive scheme under the current model assumptions.

4.1 Endogenous Wages

In this section, the supervisor endogenously determines the two employees' wages on the two layers of the hierarchy. In this context, assume that the supervisor wishes to maximize the present value of firm profits. Alternatively, one could introduce a firm manager into the model determining the employees' wages and suppose that the supervisor's only task is to monitor the employees. Two assumptions are made facilitating the analysis. First, renegotiation of the original contract is supposed to be prohibitively costly, hence the parties do never change the initial contract terms. Renegotiation may, for example, prevent the employees from working on their tasks so that the gains from renegotiation are outweighed by the loss in production. Second, the employees are assumed to be limitedly liable implying that the wages on both layers have to be non-negative. This allows us to neglect participation constraints. Under limited liability, the employees always achieve a strictly positive payoff. This is easily demonstrated. By choosing zero effort, an employee occurs no effort costs and nevertheless may win the tournament and become promoted. With non-negative prizes and at least one strictly positive prize, the employee can ensure a positive expected payoff by choosing zero effort. Hence, if deviating and choosing strictly positive effort, the employee must do better so that expected payoffs are strictly positive. Normalizing the reservation utilities to zero (as I do in this paper), participation of the employees is always guaranteed.[15]

From the preceding discussion, it is clear that the supervisor sets the

[15] Note that the cancellation of either assumption would make the model intractable. If renegotiation were possible, the supervisor would adjust the wage structure to the actual standing in the promotion tournament. Then, by choosing effort the employees could not only affect their promotion probability, but also the wage structure. This would extremely complicate the analysis. Similarly, unlimited liability would make the model intractable, as participation not only in the first tournament, but in all rounds needed to be ensured.

wage on the lower layer of the hierarchy equal to zero. A higher wage would entail higher monetary costs and decrease incentives for the employees and so cannot be optimal. In order to keep the exposition as simple as possible, consider the case $T = 4$ and assume that the supervisor discounts future profits at the same factor δ as the employees. The supervisor's optimization problem is then given by

$$\begin{aligned} & \underset{w}{Max}\ 2f\left(e_1\right)+\delta\left[\left(1-\lambda\right)2E\left[f\left(e_2\right)\right]-\lambda w\right] \\ & +\delta^2\left[\left(1-\lambda\right)^2 2E\left[f\left(e_3\right)\right]-\lambda\left(2-\lambda\right)w\right]-\delta^3\lambda\left(3-3\lambda+\lambda^2\right)w \\ s.t.\ \frac{\partial C(e_3)}{\partial e_3} &= f_1(\theta_1\Delta s_2+\theta_2\Delta s_1)\lambda w\delta\theta_0, \\ \frac{\partial C(e_2)}{\partial e_2} &= f_1(\theta_1\Delta s_1)\lambda w\left(\delta+\delta^2\right)\theta_0+f_2\left(\theta_2\Delta s_1\right)\left(1-\lambda\right)\lambda w\delta^2\theta_1, \\ \frac{\partial C(e_1)}{\partial e_1} &= f_1(0)\lambda w\left(\delta+\delta^2+\delta^3\right)\theta_0+f_2\left(0\right)\left(1-\lambda\right)\lambda w\left(\delta^2+\delta^3\right)\theta_1 \\ & +f_3\left(0\right)\left(1-\lambda\right)^2\lambda w\delta^3\theta_2 \end{aligned} \tag{13}$$

, with $E\left[\cdot\right]$ denoting the expectations operator. It is implicitly assumed that the wage contracts to be derived only hold for the two competing employees. Wages of employees who are already on the higher layer of the hierarchy are unaffected. Further, in maximization problem (13), the output to be realized from a filled workplace on the higher layer is neglected. As this output is constant, this negligence is unproblematic. The following proposition contains the solution to problem (13):

Proposition 4 *In case $T = 4$, the optimal wage w^* on the higher layer of* *the hierarchy is implicitly given by*

$$\begin{aligned} 0 &= 2\frac{\partial f(X(t^*))}{\partial X}\frac{\partial X\left(t^*\right)}{\partial t}\frac{t^*}{w^*} + \delta\left[(1-\lambda)\,2\int\limits_{-\infty}^{+\infty}\frac{\partial f(X(u^*))}{\partial X}\frac{\partial X\left(u^*\right)}{\partial u}\frac{u^*}{w^*}g(\Delta\varepsilon_1)d\Delta\varepsilon_1 - \lambda\right] \\ &\quad +\delta^2\left[(1-\lambda)^2\,2\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\frac{\partial f(X(z^*))}{\partial X}\frac{\partial X\left(z^*\right)}{\partial z}\frac{z^*}{w^*}g(\Delta\varepsilon_1)g\left(\Delta\varepsilon_2\right)d\Delta\varepsilon_1 d\Delta\varepsilon_2 - \lambda\left(2-\lambda\right)\right] \\ &\quad -\delta^3\lambda\left(3-3\lambda+\lambda^2\right) \end{aligned}$$

, with $X = \left(\frac{\partial C(e)}{\partial e}\right)^{-1}$, $t^* = f_1(0)\lambda w^*\left(\delta+\delta^2+\delta^3\right)\theta_0 + f_2\left(0\right)\left(1-\lambda\right)\lambda w^*\left(\delta^2+\delta^3\right)\theta_1 +$ $f_3\left(0\right)\left(1-\lambda\right)^2\lambda w^*\delta^3\theta_2$, $u^* = f_1(\theta_1\Delta\varepsilon_1)\lambda w^*\left(\delta+\delta^2\right)\theta_0 + f_2\left(\theta_2\Delta\varepsilon_1\right)\left(1-\lambda\right)\lambda w^*\delta^2\theta_1$ *,* $z^* = f_1(\theta_1\Delta\varepsilon_2+\theta_2\Delta\varepsilon_1)\lambda w^*\delta\theta_0$ *and* $g(\cdot)$ *denoting the density function of a normally distributed random variable with zero mean and variance* $2\sigma^2$*.*

Proof. Defining $X := \left(\frac{\partial C(e)}{\partial e}\right)^{-1}$, it follows that

$$\begin{aligned} E\left[f\left(e_2\right)\right] &= \int\limits_{-\infty}^{+\infty} f\left(X\left(f_1(\theta_1\Delta\varepsilon_1)\lambda w\left(\delta+\delta^2\right)\theta_0 + f_2\left(\theta_2\Delta\varepsilon_1\right)\left(1-\lambda\right)\lambda w\delta^2\theta_1\right)\right) g(\Delta\varepsilon_1)d\Delta\varepsilon_1, \\ E\left[f\left(e_3\right)\right] &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f\left(X\left(f_1(\theta_1\Delta\varepsilon_2+\theta_2\Delta\varepsilon_1)\lambda w\delta\theta_0\right)\right) g(\Delta\varepsilon_1)g\left(\Delta\varepsilon_2\right)d\Delta\varepsilon_1 d\Delta\varepsilon_2, \\ e_1 &= X(f_1(0)\lambda w\left(\delta+\delta^2+\delta^3\right)\theta_0 + f_2\left(0\right)\left(1-\lambda\right)\lambda w\left(\delta^2+\delta^3\right)\theta_1 \\ &\quad + f_3\left(0\right)\left(1-\lambda\right)^2\lambda w\delta^3\theta_2) \end{aligned}$$

The condition in Proposition 4 is then obtained by inserting these values into the profit formula and calculating the first-order condition. ■

When determining the wage on the higher layer of the hierarchy, the supervisor trades-off the gains from increasing w with the corresponding costs. Increasing w leads to higher effort (or, alternatively, to higher expected effort), and consequently, to higher output, but also to higher costs in terms of wage payments. These gains and costs are weighted according to their temporal occurrence and their respective occurrence probabilities. For instance, a wage payment in period 4 is attached a weight $\delta^3\lambda\left(3-3\lambda+\lambda^2\right)$, as it is discounted by the factor δ^3 and occurs with probability $\lambda\left(3-3\lambda+\lambda^2\right) = \lambda + (1-\lambda)\lambda + (1-\lambda)^2\lambda$.

In order to assess, whether or not the results from Section 3 continue to hold, it is convenient to simplify the model by assuming specific forms for $f(e)$ and $C(e)$. Let, in this context,

$$f(e) = \begin{cases} \ln e, \text{ for } e \geq 1 \\ 0, \text{ for } 0 \leq e < 1 \end{cases} \tag{14}$$

and

$$C(e) = 0.5ce^2,\ c > 0 \tag{15}$$

Further, assume that c is so small that, in equilibrium, $e > 1$ always holds. The condition in Proposition 4 then simplifies to

$$w^* = 2\frac{1+\delta(1-\lambda)+\delta^2(1-\lambda)^2}{\delta\lambda+\delta^2\lambda(2-\lambda)+\delta^3\lambda\left(3-3\lambda+\lambda^2\right)} \tag{16}$$

It is easy to see that $\frac{\partial w^*}{\partial\lambda} < 0$. A higher λ makes a payment of w^* in all periods more likely. So, the expected wage costs increase, leading to a decrease in w^*. A change in λ has thus a further effect on effort. To recognize this, consider the total derivative $\frac{de_t}{d\lambda}$. This derivative can be split up into $\frac{de_t(\lambda,w)}{d\lambda} = \frac{\partial e_t(\lambda,w)}{\partial\lambda} + \frac{\partial e_t(\lambda,w)}{\partial w}\frac{\partial w}{\partial\lambda}$. The second summand is strictly negative. Hence, with endogenous wages, the result that effort may be decreasing in λ, is not offset, but even enforced. Besides the effects on effort illustrated in Section 3, an increase in λ yields a decrease in w^* which induces effort to decline.

4.2 Consideration of varying occurrence probabilities

Let us return to the case of exogenous wages. Up to this point, the occurrence probability of a vacancy was assumed to remain constant over time. In this section, we depart from that assumption and assume that, in period h, a vacancy arises with the following probability:

$$\text{Prob}\left\{\text{vac in h}\right\} = \begin{cases} Min\left\{q^{h-1}\lambda, 1\right\}, \text{ if no vacancy has arisen in } t = 1, ..., h-1 \\ 0, \text{ otherwise} \end{cases} \tag{17}$$

Assume that $q > 0$. Further, let q either exceed 1 or be smaller than 1. In the former case, q^h is monotonically increasing in h, in the latter case, it is monotonically decreasing.[16] For $q > 1$, it becomes more likely that a promotion actually takes place, as time passes, while, for $q < 1$, the reverse is true.

Again, the main results can already be obtained by considering effort choices in periods $T-1$ and $T-2$. In period $T-1$, effort is implicitly given by $\frac{\partial C(e_{T-1})}{\partial e_{T-1}} = f_1(\sum_{t=1}^{T-2} \theta_{T-1-t}\Delta s_t)q^{T-2}\lambda w\delta\theta_0$. It is thus still increasing in λ. This is a simple extension of Proposition 1. In period $T-1$, the employees base their effort decisions on the expected winner-prize, hence efforts increase in λ.

[16]Notice that, in the former case, the model parameters are assumed to be such that $q^h\lambda \leq 1$ always holds on the equilibrium path.

In period $T-2$, one can show that optimal efforts are implicitly given by

$$\begin{aligned}\frac{\partial C(e_{T-2})}{\partial e_{T-2}} &= f_1(\sum_{t=1}^{T-3} \theta_{T-2-t}\Delta s_t) q^{T-3}\lambda w \left(\delta+\delta^2\right)\theta_0 \\ &+ f_2\left(\sum_{t=1}^{T-3} \theta_{T-1-t}\Delta s_t\right)\left(1-q^{T-3}\lambda\right) q^{T-2}\lambda w \delta^2\theta_1\end{aligned} \tag{18}$$

The condition in Proposition 2 guaranteeing that e_{T-2} increases in λ is then replaced by $f_1(\sum_{t=1}^{T-3}\theta_{T-2-t}\Delta s_t)\left(\delta+\delta^2\right)\theta_0 > \left(2q^{T-2}\lambda - q\right) f_2\left(\sum_{t=1}^{T-3}\theta_{T-1-t}\Delta s_t\right)\delta^2\theta_1$. Its interpretation is qualitatively the same as in Section 3. The only difference is an additional effect due to the varying occurrence probabilities. This effect alters the first factor on the right-hand side of the inequality from $(2\lambda-1)$ to $\left(2q^{T-2}\lambda - q\right)$. Comparing these factors, one can show that the second factor exceeds the first one, only if $2\lambda > 1$ and $q > 1$, or $2\lambda < 1$ and $q < 1$. This means that, for $q > 1$, the effects of a change in λ on the promotion probability at the end of period $T-1$ are enforced, while, for $q < 1$, they are weakened. If e.g. a marginal decrease in λ makes a promotion after period $T-1$ more likely, i.e. if $2\lambda > 1$, this positive effect on the promotion probability is enforced for $q > 1$ so that the switch in attention from period $T-2$ to period $T-1$ is even more distinctive.

4.3 Optimality of promotions as incentive device

In the model, the firm was not able to offer the employees explicit incentives, as no contractible performance measure exists. However, even under this assumption promotions are not the only feasible incentive device. The firm could for example pay the employees a bonus in every period depending on the (relative) assessments of the supervisor. In this way, it might perform far better than under the promotion tournament. Yet, in practice promotions are often argued to be an extremely important incentive device. There are two main problems of subjective bonus payments that are absent under promotions.

First, empirical studies have demonstrated that bonus schemes suffer from a compression of ratings.[17] Supervisors tend to distort evaluations by not sufficiently differentiating between employees. The two most relevant forms of compression are "centrality bias" and "leniency bias". The former bias refers to a situation, where supervisors evaluate all employees similar to a certain norm. The latter bias says that supervisors overstate the evaluations of bad employees. Solutions to these problems as e.g. forced distribution systems are often unsatisfactory as well. First, the forced distribution may not match the actual distribution so that the supervisor is forced to choose incorrect evaluations. Further, the supervisor may achieve a compression of ratings by evaluating employees as good in one period and as bad in another one. Under a promotion tournament, the supervisor is unable to achieve a compression of ratings, so these problems do not arise.

Second, there are fairness considerations. Behavioral economists argue that employees compare themselves with their peers. Employees that come off badly may feel envy and employees that come off well may feel compassion.[18] These feelings might affect the work environment in an unfavorable way. Envy e.g. might lead to mobbing and this may decrease performance.

[17] See, for an overview, e.g. Prendergast (1999).

[18] See, for example, Fehr & Schmidt (1999) or Grund & Sliwka (2005).

Under a bonus scheme, these problems should be more severe, as, after payment of bonuses, employees still work in the same department and so belong to the same peer. Hence, feelings should be stronger under the bonus scheme, for there is inequality between employees belonging to the same peer. Further, mobbing as a consequence of these feelings is little likely to occur under promotions, where the employees are separated after payments have been made.

5 Concluding remarks

In this paper, I tried to model promotions more accurately and added two ingredients to the existing literature on rank-order tournaments that I think to be of particular importance. First, employees can rarely be sure that a workplace becomes vacant and so a promotion actually takes place. As a result, the employees should assign probabilities to a workplace becoming vacant at certain points in time. Second, promotions are usually based on assessments of supervisors rather than on contractible performance measures. It is thus important to understand, how supervisors aggregate information to form a view on the employees.

I explore, in this context, two possibilities. On the one hand, early impressions may be most relevant, when assessing an employee. On the other hand, the supervisor may only recall recent impressions so that these impressions are the more important ones. These two kinds of information processing are of importance, when analyzing the effort change resulting from a change in the probability that a vacancy arises in a certain period.

Finally, the model might be a starting point for an explanation of hierarchies. Suppose that each member of the firm quits his job with an exogenously given probability ρ. The firm may then want to determine the relative number of employees on a layer in the hierarchy to the number of employees on the next lower layer such that the promotion probability λ is set optimally, i.e. to induce optimal incentives on the lower layer. This would complement existing ideas concerning hierarchies (see e.g. Williamson (1967), Calvo & Wellisz (1978), (1979) or Qian (1994)) that emphasize the relevance of loss in control for the optimal organization of hierarchies.

References

Calvo, G., Wellisz, S., 1978, Supervision, Loss of Control and the Optimal Size of the Firm. Journal of Political Economy, 86, 943-952.

Calvo, G., Wellisz, S., 1979, Hierarchy, Ability, and Income Distribution. Journal of Political Economy, 87, 991-1010.

Chen, K.P., 2003, Sabotage in Promotion Tournaments. Journal of Law, Economics, & Organization, 19, 119-140.

Dubey, P., Haimanko, O., 2003, Optimal scrutiny in multi-period promotion tournaments. Games and Economic Behavior, 42, 1-24.

Edwards, W., 1968, Conservatism in Human Information Processing. In: Formal Representation of Human Judgement, New York, 17-52.

Fehr, E., Schmidt, K., 1999, A Theory of Fairness, Competition, and. Co-operation. Quarterly Journal of Economics, 114, 817-868.

Fiske, S.T., Taylor, S.E., 1984, Social Cognition. Addison-Wesley Publishing Company, Reading, Massachusetts.

Green, J.R., Stokey, N.L., 1983, A Comparison of Tournaments and Contracts. Journal of Political Economy, 91, 349-364.

Grund, C., Sliwka, D., 2005, Envy and Compassion in Tournaments. Journal of Economics and Management Strategy, 14, 187-207.

Higgins, E.T., Rholes, W.S., Jones, C.R., 1977, Category accessibility and impression formation. Journal of Experimental Social Psychology, 13, 141-154.

Kräkel, M., 2002, U-Type versus J-Type Tournaments. Journal of Institutional and Theoretical Economics, 158, 614-637.

Kräkel, M., 2003, U-Type versus J-Type tournaments as alternative solutions to the unverifiability problem. Labour Economics, 10, 359-380.

Kuhn, H.W., 1953, Extensive Games and the Problem of Information. In: Contributions to the Theory of Games II, Princeton University Press, Princeton, NJ, 193-216.

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

Milgrom, P., Roberts, J., 1992, Economics, Organization and Management. Prentice Hall, Englewood Cliffs, New Jersey.

Piccione, M., Rubinstein, A., 1997, On the interpretation of Decision Problems with Imperfect Recall. Games and Economic Behavior, 20, 3-24.

Prendergast, C., 1999, The Provision of Incentives in Firms. Journal of Economic Literature, 37, 7-63.

Qian, Y., 1994, Incentives and Loss of Control in an Optimal Hierarchy. Review of Economic Studies, 61, 527-544.

Rubinstein, A., 1998, Modeling Bounded Rationality. The MIT Press, Cambridge, Massachusetts.

Srull, T.K., Wyer, R.S., Jr., 1979, The role of category accessibility in the interpretation of information about persons: Some determinants and implications. Journal of Personality and Social Psychology, 37, 1660-1672.

Williamson, O., 1967, Hierarchical Control and Optimum Firm Size. Journal of Political Economy, 75, 123-138.