Learning by Helping:
A Bounded Rationality Model of Mentoring

Mahmood Arai*, Antoine Billot** and Joseph Lanfranchi* [0]

Abstract

Within an organization, a bounded rational principal organizes a promotion *contest* based on a sequence of tests regarding candidates' relative performances. We assume the principal to suffer from *limited ability to rank* the performances, only identifying the best in each test. Furthermore, he *satisfices* the expected profit from promotion, designing the contest such that expected gains do not decrease with the information generated by additional tests. Then, mentoring is shown to improve the information about candidates' ability when the principal offers help to the current best candidate provided by a manager promoted after a similar contest.

JEL CLASSIFICATION: D10-D83-L22.
KEYWORDS: Mentoring, Selection, Contests, Bounded Rationality.

[0]

* Department of Economics, Stockholm University
** ERMES, Université Panthéon-Assas, Paris.

[0]The authors wish to thank James W. Friedman, Hervé Moulin, Tony E. Smith, Jacques Thisse and John Treble for helpful comments and suggestions. Financial support from the Université Panthéon-Assas, Paris II is gratefully acknowledged. Arai acknowledges financial support by The Swedish Council for Research in the Humanities and Social Sciences (HSFR).

1 Introduction

Mentoring relations involving newcomers and senior managers have never been formally modelled in economic theory even though sociological and management literature have demonstrated their empirical relevance. For example, Collins and Scott (1979), Ochberg, Tischler and Schulberg (1986), and Chao (1997) study various professional environments identifying the extent of relations between mentors and proteges and showing that mentored workers experience greater future rewards or career success than nonmentored ones. This paper presents a model that endogenizes the formation of mentor-protege relations and shows how mentoring can improve the promotion process of workers in organizations.

In a pioneering study, Kram (1985) underlines the career enhancing functions of mentoring. These functions are coaching, sponsorship and teaching, all of which conspire to increase skills, signal ability and prepare the protege for advancement. Accordingly, we interpret mentoring as the set of activities by which a mentor can help workers to increase their productivity.

Laband and Lentz (1995) discuss two rational explanations for the use of mentor-protege relations in organizations. The first is based on the transfer and accumulation of firm specific human capital. The second relates mentoring to job matching theory. To identify the best workers, firms invest time and knowledge of senior managers to cooperate with the promising workers. This cooperation generates information about the quality of job matches and can improve on the matching of workers to jobs. Our contribution clearly belongs to the job-matching interpretation.

We consider a boundedly rational principal who organizes a promotion contest that is, a sequence of tests regarding relative performances of candidates. The tests can be based on the workers' performances during the normal course of production. Though the recorded performances are imperfect signals of ability, they help disclose for each candidate the *likelihood* of being the best. The principal is assumed to have a bounded capacity to process this information because he can only identify the best performance in a given test and he is generally concerned with improvements (and not maximization) of the expected profit associated with the promotion of the best worker. This means that the promotion contest must be designed so that the expected profit does not decrease with additional information. This *bounded rationality* is motivated by the imperfection of the information and the difficulty of mapping this information to expected profits.

We ask, (i) can a procedure with repeated identical tests and aggregated records of performance always be relevant for selection? and, (ii) can the principal improve on learning about workers' ability by adopting a *biased* test design, changing its informative value?

We answer these questions proving that the principal must bias the final test by giving the worker known to have been the best performer in the past, i.e. the leader, some help in the current test. In our framework, this help corresponds to vertical cooperation with a mentor, i.e. a senior worker already promoted.

What is the intuition behind this result? Anticipating a vacancy in the next period, the principal gathers previous information about workers and finally promotes one of them after a final test. Nevertheless, designing the final test in the same way as the previous ones does not always increase the principal's confidence for the best worker, nor does it change the fact that the leader before the final test will be promoted. If such an identical test is irrelevant, changing its informative value by introducing a bias can nevertheless increase the principal's confidence in his selection. Biasing consists of helping the leader, help provided by a senior manager whose mentoring is such that it can alter the informative value of the test. Two outcomes are possible. If the mentored worker wins, he remains the leader and the principal's final confidence about his promotion is enhanced. Otherwise, if a non-mentored worker wins the biased test, this victory is understood as a stronger signal of his ability because the quality of the contestants' field is increased. Thus, the principal always becomes more confident about the winner of the final biased test than he was about the previous leader.

The idea of using a bias to improve the quality of information obtained from imperfect signals has been applied to various contexts (see Calvert (1985), Koh (1994), Billot (1998)). Close to our context, Meyer (1991) considers the problem of a promotion contest between two agents.

Meyer proposes a cardinal bias for the final test which strictly *maximizes* the principal's confidence about the winner. Besides, she argues that a possible way of implementing the bias could be to modify the contents of the leader's job. This interpretation is intuitive and especially attractive in terms of job design. Nevertheless, her framework does not allow for a change in the tasks devoted to the leader. Formally, the modelling of Meyer's bias is perfectly cardinal, operating as if the supervisors were giving a positive handicap to the leader of the competition. This is different from our case as we focus on cooperation between workers and show that the bias can be implemented by vertical cooperation similar to mentoring.

Moreover, in Meyer's model, the principal deviates from standard notions of rationality only in the sense that he has problems in gathering information within the organization. Nevertheless, there is no limitation in his instrumental rationality since he still maximizes the expected output of the promoted worker. In our framework, bounded rationality is instrumental and cognitive since the principal does not maximize expected output nor gather full information.

The remainder of the paper is organized as follows. Section 2 deals with the rules of the promotion contest and the principal's beliefs and behavior. Section 3 describes the way the principal can extract supplementary information about the workers' ability by designing biased contests. Section 4 studies the optimal design of the biased contest, endogenizes mentoring relations and discusses their empirical relevance. Finally, Section 5 summarizes the results and their empirical relevance and concludes with possible extensions of the discrete choice framework within organization theory.

2 The Model

This section introduces a model of an organization managing the optimal assignment of workers within its hierarchy. Our first objective is to describe the process of revelation of the relative abilities of workers along a selection contest. Then, we will explain how the principal in charge of the promotion decision will form his beliefs from the results of this screening process. Finally, his objective and choice behavior will be defined.

2.1 Definitions

We define the organization as a finite set X of agents, ranked in three hierarchical levels: X_1, the junior workers, X_2, the managers who produce and supervise and X_3, the principal. The latter runs the organization as a closed internal labor market in which retiring managers are systematically replaced by workers from the entry level X_1. The principal must thus select the best applicant from the pool of junior workers.

Initially, all junior workers are undifferentiated, in the sense that their relative abilities are totally unknown. Nevertheless, through on-the-job screening, the principal must acquire information about the best applicant for future promotion. We focus on the principal's most efficient learning method in order to finally select the most able worker. The abilities of the applicants are assumed to remain constant over time and to be imperfectly signalled by a periodical performance.

Formally, for each applicant, the performance, denoted $y = y(a, \epsilon)$, depends on his ability a and on ϵ, an idiosyncratic random shock, independently and identically distributed across the contestants.[1,2] This signal can either be a direct measure of the output of a worker at X_1, supposed to be positively correlated with the ability to perform well at X_2, or an indirect evaluation of that applicants potential as a manager. In both cases, a selection rule based on repeated records of performance can solve the assignment problem within the hierarchy. Moreover, we abstract from the moral hazard considerations assuming that each junior worker supplies an identical and perfectly verifiable level of effort or amount of hours of work.

The principal can design the selection process of the most able worker as a contest, that is a sequence of production periods before the promotion during which only the identity of the best performer is recorded. Each period is called a *test*. The manager supervising each test is assumed to (i) always single out a winner and (ii) never manipulate the observations when recording and sending them to the principal in order to avoid all kinds of favoritism.

Thus, a *contest*, denoted $\mathbf{C}_A^s$, is defined as a finite sequence of s tests over a subset of workers A included in X_1.[3] The *winner* of any test is the agent w whose performance y_w is superior to any other worker's performance.

From the records of the winners' identity during the repeated tests, the principal constructs a function summarizing the accumulated information for each worker involved in the contest. Consequently, for any contest $\mathbf{C}_A^s$, we define the *reputation* of a worker i as a value $\mathbf{R}^s(i) \geq 0$, increasing with his number of victories after s tests. The reputation is a mapping $\mathbf{R}^s(.)$ defined from 2^{X_1} to $\mathbb{R}^+$ such that, for all $s > 0$, $A \in 2^{X_1}$:

[1] Obviously, we also assume that random shocks are distributed such that a higher level of performance conveys a positive signal about ability.

[2] In Meyer(1991), the information was said to be coarse in the sense that the principal only knows both distributions of the spreads of performance and random schocks and records the rank of the two contestants' performance. In our model, the information is even *coarser* since the principal has no prior about the distribution of relative abilities neither of the random shocks.

[3] Indeed, the contest may include all the junior workers in the entry level or a subset if the promotion process is organized within tenure or age cohorts.

$$\mathbf{R}^s(\cup_{i\in A}) = \sum_{i\in A} \mathbf{R}^s(i) = s \times 1_A \qquad (1)$$

where 1_A corresponds to the *unit value* $\mathbf{R}^1(A)$ of one victory in A, that can be interpreted as the reputation conveyed by organizing one test in A whatever the identity of the winner. The unit value can also be identified as the marginal value of a new test over a given set of candidates.

Furthermore, we assume that when the result of a test is not recorded, it conveys no information. That is, if $A = \emptyset$ then $1_\varnothing = 0$ and if $A \neq \emptyset$ then $1_A > 0$. We also impose a natural requirement of monotonicity: if $A \subset A' \subset X_1$, $1_A < 1_{A'} < 1_{X_1}$. This translates the intuition that winning in a large set of workers does mean more than winning in one of its subsets. Finally, the reputation conveyed by one test is supposed to be completely assigned to the winner of this test. Note that this way of ruling the informative value of each test result clearly recalls the basic features of discrete choice theory in the sense that the response in the model is binary: win or lose.[4]

From (1), $\mathbf{R}^1(A)$ can be decomposed as follows:

$$\mathbf{R}^1(A) = \mathbf{R}^1(w) + \sum_{i\neq w} \mathbf{R}^1(i) = 1_A$$

which means that the supervisor only singles out the winner w and forgets the losers i, $\mathbf{R}^1(w) = 1_A$ and $\mathbf{R}^1(i) = 0$ for all $i \neq w$.

2.2 The Principal's Beliefs

From the reputation function defined above, we can now define how the principal will form and revise his beliefs about the ability of the junior workers involved in the contest. Originally undifferentiated in terms of ability, all workers of X_1 have the same prior reputation before the beginning of the actual tests in the contest. Then, the *prior probability of any worker i to be the best in any subset $A \subset X_1$*, denoted $P_A^0(i)$, can be given by the uniform distribution, $\frac{1}{a}$, for all $i \in A$ with $a = card(A)$. This hypothesis allows identification of the probabilities $P_A^0(.)$ *as if* they were inferred from a number a of previous tests where each worker i of A has won just once. In this way a contest consisting of s tests over a set of contestants A contain

[4]For further details about the relationship between discrete choice theory and selection models, see Lanfranchi, (1994), Arai, Billot, Lanfranchi (1994) and Billot (1998).

$s-a$ actual tests and a initial neutral tests: every contestant is assigned one victory.[5] Thus, for all $i \in A$, $P_A^0(i) = P_A^a(i)$ with $\mathbf{R}^0(i) = \mathbf{R}^a(i) = 1_A$ and $\mathbf{R}^0(A) = \mathbf{R}^a(A) = a \times 1_A$:

$$P_A^0(i) = P_A^a(i) = \frac{\mathbf{R}^0(i)}{\mathbf{R}^0(A)} = \frac{\mathbf{R}^a(i)}{\mathbf{R}^a(A)} = \frac{1_A}{a \times 1_A}.$$

By dynamically extending this formula to any contest, we assume the principal to set up a probability distribution from the results of $\mathbf{C}_A^s$, that were recorded by the supervisor, as follows:

Axiom 1 : For all $i \in A \subset X_1$,

$$P_A^s(i) = \frac{\mathbf{R}^s(i)}{\mathbf{R}^s(A)} = \frac{\mathbf{R}^s(i)}{s \times 1_A} \tag{2}$$

where $P_A^s(i)$ can be viewed as the principal's probability of the event "i is the best contestant after s tests over A".

Since the random shock and ability distributions are assumed to be unknown, the probability given by Axiom 1 is one of the simplest *beliefs* that can be rationally formed from the principal's information.[6]

Throughout this paper, we will follow an example of a promotion contest illustrating the functioning of the model. Here follows an example of the reputation and belief functions from the records of a given contest.

Example 1 : *Consider a contest over $A = \{i, j\}$ such that after $s = 4$ (2 neutral and 2 actual) tests over $A = \{i, j\}$, the reputations are given by:*

$$\mathbf{R}^4(i) = 3 \times 1_A \text{ and } \mathbf{R}^4(j) = 1_A.$$

First, we know that, for i and j:

$$P_A^0(i) = P_A^a(i) = P_A^2(i) = \frac{\mathbf{R}^2(i)}{\mathbf{R}^2(A)} = \frac{1_A}{2 \times 1_A} = \frac{1}{2}.$$

[5] The initial tests constitute the foundations for the principal's beleifs on contestants reputation before the begining of the actual tests.

[6] The form of beliefs given by equation (2) recalls the hypothesis of simple scalability, frequently used in discrete choice models, according to which the choice probability of a given alternative can be expressed as an increasing function of its scale value and a decreasing function of the scale values of the other alternatives.

Then, $\mathbf{R}^4(i) = 3$ *and* $\mathbf{R}^4(j) = 1$ *means that* i *wins* 2 *times and* j *never. Hence, the probabilities for the two candidates to be the best after* 4 *tests correspond to:*

$$P_A^4(i) = \frac{\mathbf{R}^4(i)}{\mathbf{R}^4(A)} = \frac{\mathbf{R}^2(A)+1}{\mathbf{R}^4(A)} = \frac{2+1}{4} = 3/4$$

and

$$P_A^4(j) = \frac{\mathbf{R}^4(j)}{\mathbf{R}^4(A)} = \frac{\mathbf{R}^2(j)+1}{\mathbf{R}^4(A)} = \frac{0+1}{4} = 1/4.$$

2.3 The Principal's Behavior

The principal is here assumed to be *boundedly rational* for the following two reasons:

- (i) The problem of the promotion decision in the contest arises from the imperfect nature of the information conveyed by the recording of workers' performances. This imperfection is not only due to the uncertainty about workers' ability but also stems from the principal's *incapability to cardinally measure or totally rank* workers' performances. Since the principal only identifies the winner and promotion contests may include more than 2 contestants, the principal's ability to rank is obviously *bounded*.

- (ii) The principal only tries to *increase* (and not to maximize) the expected profit after the promotion. From the assumption of positive correlation between the current performance of workers in X_1 and their future performance in X_2, looking for a higher expected profit is perfectly equivalent to *increase* (and not to maximize) the final confidence in the promoted worker, i.e. the leader at the end of the contest.

After each test, the probabilistic beliefs of the principal are revised according to the results. In test after test, the principal tries to learn about the relative ability of all contestants in order to limit the risk of inefficient assignment of workers. Hence, it is natural to consider that the principal must design the contest in order to extract the relevant information for the promotion selection.

3 Information within Contests

In this section, we show how the principal can extract information from the results of a sequence of identical tests. Moreover, we demonstrate that organizing supplementary identical tests is not *relevant* for the purpose of efficient selection of workers when the beliefs of the principal about the abilities of the workers are not uniform. Finally, we introduce the notion of *faith contests* conveying the same amount of information while organized over different sets of applicants and for a different number of tests.

3.1 Expected reputations from virtual tests

After a given contest $\mathbf{C}_A^s$, the principal can extract information from a supplementary identical test, designed with the same rules within the same set of workers A. Before this supplementary test, the principal is able to compute the expected reputation for each contestant after the $(s+1)$th test. If we denote $\mathbf{R}^{s+1}(i \mid i = w)$ the conditional reputation of the worker i when he is the winner of the supplementary test, the conditional reputation of any worker i can a priori take the two following values:

$$\begin{cases} \mathbf{R}^{s+1}(i \mid i = w) = \mathbf{R}^s(i) + \ 1_A, \\ \mathbf{R}^{s+1}(i \mid j = w) = \mathbf{R}^s(i) \text{ for } i \neq j. \end{cases}$$

Furthermore, the principal knows the probabilities $P_A^s(.)$. The probabilities for the workers to be chosen are used to compute the associated expected reputation. Denoting $V^{s+1}(i)$ the expected reputation of i in $s+1$, we obtain:

$$V^{s+1}(i) = \mathbf{R}^{s+1}(i \mid i = w) \times P_A^s(i) + \mathbf{R}^{s+1}(i \mid j = w) \times (1 - P_A^s(i)),$$

then,

$$V^{s+1}(i) = (\mathbf{R}^s(i) + \ 1_A) \times P_A^s(i) + \mathbf{R}^s(i) \times (1 - P_A^s(i)) = \frac{(s+1)}{s}\mathbf{R}^s(i). \quad (3)$$

Basically, this result can be generalized over more than a supplementary virtual test. For a given contest $\mathbf{C}_A^s$, the expected reputation of a worker i after n virtual supplementary tests can be shown to be as follows:

$$V^{s+n}(i) = \frac{(s+n)}{s}\mathbf{R}^s(i). \quad (4)$$

This result in terms of expected utility can also be reinterpreted in terms of probabilities in the following proposition.

Proposition 2 : *Consider any contest $\mathbf{C}_A^s$. Then, for all $A \subset X_1$, $i \in A$, and for all $s' > s$,*

$$P_A^{s'}(i) = P_A^s(i).$$

Proof: From (4), we know for all $i \in A$:

$$V^{s'}(i) = \frac{s'}{s}\mathbf{R}^s(i),$$

i.e. from axiom 1, $P_A^{s'}(i) = P_A^s(i)$.□

According to this proposition, a virtual contest $\mathbf{C}_A^{s'}$ can be said to convey the same information about the selection of a contestant within A, i.e. the same ranking as in a real contest $\mathbf{C}_A^s$. This establishes that the principal can reinterpret the records obtained after s actual tests over A in a longer sequence of s' tests over the same set without any change in the probability for each worker to be promoted. This means that the contestants' ranking remains invariant under virtual replication of the number of tests by expectation.[7]

3.2 Relevant Tests

At the end of a contest, we know that the principal should rationally promote the best candidate, that is the one with the highest probability of being the best. We define the *leader* after s tests over A as the worker w_A^s with the highest reputation and the *follower* after s tests as the one j_A^s with the lowest reputation. Let us suppose that after any given sequence of s identical tests, the principal knows that he must replace a manager by a junior worker in the next period. The principal then needs to run a $(s+1)$th test which allows him to increase his confidence regarding the best contestant. We define such a test as *relevant* (for the confidence).

Formally, for any given contest $\mathbf{C}_A^s$, a $(s+1)$th test over A is said to be *relevant* (*for the confidence*) if and only if:

$$\max_A P_A^{s+1}(i) = P_A^{s+1}\left(w_A^{s+1}\right) \geq \max_A P_A^s(i) = P_A^s(w_A^s). \quad (5)$$

The following example will give us some intuition about the possible design of this last test.

[7] In fact, the probability $P_A^{s'}(i)$ can also be computed and interpreted as the expected probability for i to be the best worker after s' tests for a given actual contest C_A^s.

Example 3 : *Consider a contest where after* 4 *tests over* $A = \{i, j\}$, *the reputations are given by:*

$$\mathbf{R}^4(i) = 3 \times 1_A \text{ and } \mathbf{R}^4(j) = 1_A.$$

Hence, the probabilities for the two candidates to be the best correspond to:

$$P_A^4(i) = 3/4 \text{ and } P_A^4(j) = 1/4.$$

Therefore, a $5th$ *test over* A *can only lead to the following results:*

-(i) $P_A^5(i) = 3/5 < 3/4$ *and* $P_A^5(j) = 2/5 < 3/4$ *which confirms that* i *is the best candidate even if* j*'s victory decreases the principal's confidence regarding* i *as the best or*

-(ii) $P_A^5(i) = 4/5$ *and* $P_A^5(j) = 1/5$ *which also confirms that* i *is the best candidate.*

A $5th$ *test over* A *would be relevant if and only if* $P_A^5(w_A^5) \geq 3/4$ *which is not the case if* $w_A^5 = j$. *After running this test, the probability attached to the previous leader* i *can decrease and the additional test is thus irrelevant.*

Note that this test is useless for the selection since the principal's opinion, i.e. the final ranking on the workers, does not alter by the results of this test. The worker who should finally be promoted is i, regardless of who wins the $5th$ *test.*

This example highlights the fact that running a supplementary identical test is not always relevant in terms of confidence for the leader of the contest. On the contrary, in the case of identical reputations of all workers after s tests, the problem of selection is trivial because a supplementary identical test will be relevant. So, by assumption, we now restrict ourselves only to nondegenerated situations for which, after any given $\mathbf{C}_A^s$, the reputations are not uniform, i.e. $\mathbf{R}^s(w_A^s) > \mathbf{R}^s(j_A^s)$.

More generally, the following result establishes the impossibility of running a relevant $(s+1)$th test over A after any given contest $\mathbf{C}_A^s$.

Proposition 4 : *After any given contest* $\mathbf{C}_A^s$, *any* $(s+1)th$ *test over* A *is irrelevant.*

Proof: Suppose the $(s+1)$th test over A to be relevant. Then, by (5), for all w_A^{s+1}, we must have:

$$P_A^{s+1}\left(w_A^{s+1}\right) \geq P_A^s\left(w_A^s\right)$$

which means

$$\frac{V^{s+1}\left(w_A^{s+1}\right)}{(s+1)\times 1_A} \geq \frac{\mathbf{R}^s\left(w_A^s\right)}{s\times 1_A}$$

i.e.

$$\frac{s}{s+1}V^{s+1}\left(w_A^{s+1}\right) \geq \mathbf{R}^s\left(w_A^s\right).$$

Due to equation (4), this expression becomes:

$$\mathbf{R}^s\left(w_A^{s+1}\right) \geq \mathbf{R}^s\left(w_A^s\right).$$

Now, by definition of w_A^s, it is only consistent with $\mathbf{R}^s\left(w_A^{s+1}\right) = \mathbf{R}^s\left(w_A^s\right)$which is not true if $w_A^{s+1} = j_A^s$.$\square$

This proposition means that the informative value of any new test **over** A is not sufficient for the principal to increase the expected profits. Hence, this impossibility is clearly related to the set of candidates. Intuitively, relevant information for the purpose of selection can only be obtained from a relevant test with a higher unit value, that is involving a different number of contestants. The next subsection deals with *biased tests* that convey supplementary information about the relative ability of the workers.

3.3 Biased Tests

The following definition allows to focus on a particular subset of contests within the organization, that are comparable in terms of information. Two contests $\mathbf{C}_A^s$, $\mathbf{C}_{A'}^{s'}$, with $s' > 0$ and $A' \subset X$, are said to be *faith* when

$$\mathbf{R}^s(A) = \mathbf{R}^{s'}(A'). \tag{6}$$

i.e. by the definition of the reputation: $s \times 1_A = s' \times 1_{A'}$

By monotonicity of the unit values 1_A and $1_{A'}$, we have necessarily $s < s'$ if $A' \subset A$ and symmetrically $s' < s$ if $A \subset A'$. More generally, (6) means that designing a small number of tests in a large set of candidates can be interpreted as faithfully informative as designing a large number of tests in a smaller set of candidates.

When after s tests, the principal needs to promote a worker in the next period, he already knows from proposition 2 that an identical test over A is irrelevant since he cannot always increase his confidence regarding the best candidate. Hence, he should modify the informative value of the $(s+1)$th test in order to transform it in such a way that it becomes relevant. Using Meyer's language, 'changing the informative value of a test' means '*biasing* a test'.

Actually, from equation (6), we know that the principal can design the $(s+1)$th test over a larger set $A^{'} \supset A$. In this way, he obtains the same information as the one which should be generated by $(n+1)$ more tests over A if $\mathbf{C}_{A'}^{1}$ is a faith contest of $\mathbf{C}_{A}^{n+1}$, i.e.e:

$$\mathbf{R}^{n+1}(A) = \mathbf{R}^{1}(A^{'}).$$

The principal will organize one test over $A^{'}$ instead of a $(s+1)$th over A and obtains information that is faithful to $n+1$ supplementary tests over A. More particularly, we define the agent(s) $(A'-A)$ as the *bias* introduced for the $(s+1)$th test. Then, the principal just has to add the results of this test to those previously recorded from $\mathbf{C}_{A}^{s}$. Finally, everything goes *as if* he sets up a global actual contest $\mathbf{C}_{A}^{s+n+1}$.

Note that in designing a biased test $\mathbf{C}_{A'}^{1}$ in place of an irrelevant $(s+1)$th test over A, we distinguish one actual victory and n expected ones in such a way that this single biased test over A' is equivalently informative as $(n+1)$ more tests over A. In that sense, a biased test is relevant (for the confidence) if and only if:

$$\max_{A} \left[P_{A}^{s+n+1}(i)\right] \geq P_{A}^{s}(w_{A}^{s}). \tag{7}$$

Consider the problem of how to precisely modify the set A. From the definition of two faith contests, we immediately infer that A' must be larger than A. However, we have to evaluate the relation between the size of the bias $(A'-A)$ and the corresponding number $(n+1)$ of virtual tests over A resulting from the actual test over A'. The following proposition establishes the reputation of the bias.

Proposition 5 : *Consider two faith contests $\mathbf{C}_{A}^{n+1}$, $\mathbf{C}_{A'}^{1}$. Then, the reputation of the bias $\mathbf{R}^{1}(A'-A)$ is defined as:*

$$\mathbf{R}^{1}(A'-A) = 1_{A'} - 1_{A} = n \times 1_{A}. \tag{8}$$

Proof: The reputation of the bias $\mathbf{R}^{1}(A'-A)$ can be decomposed as:

$$\mathbf{R}^{1}(A'-A) = \mathbf{R}^{1}(A') - \mathbf{R}^{1}(A).$$

or, by definition of the reputation:

$$\mathbf{R}^1(A' - A) = 1_{A'} - 1_A.$$

Now, focus on the special family of virtual contests $\mathbf{C}_A^{n+1}$ equivalent to $\mathbf{C}_{A'}^1$, i.e.:

$$1_{A'} = (n+1) \times 1_A$$

which means that *one* victory in A', namely in the contestants set A enlarged with the bias $(A' - A)$, is equivalently informative as $(n+1)$ victories in the actual set A. Then, we have:

$$\mathbf{R}^1(A' - A) = n \times 1_A.\square$$

In enlarging the set A towards A', the principal defines a set union between the previous set A and the bias $(A' - A)$, i.e. a subset of virtual contestants belonging to the organization such that their reputation modifies the informative value of the $(s+1)$th test.

4 Mentoring

We already know 'how to precisely modify the set A' due to the introduction of the bias $(A' - A)$ whose reputation is defined by $n \times 1_A$. Now, we have to search for the best 'implementation' of the bias or the efficient design of the rules of the biased test. First, for consistency of the promotion process, the principal must be sure that the bias $(A' - A)$ is not sufficiently involved in the contest $\mathbf{C}_{A'}^1$ to win it.

Proposition 6 : *Consider two faith contests $\mathbf{C}_A^{n+1}$, $\mathbf{C}_{A'}^1$, such that there exists a bias $(A' - A)$ whose reputation is defined by $\mathbf{R}^1(A' - A) = n \times 1_A$. Then, the winner w of the $(s+1)$th biased test $\mathbf{C}_{A'}^1$ is different from the bias $(A' - A)$.*

Proof: If $\mathbf{C}_A^{n+1}$ is a faith contest of $\mathbf{C}_{A'}^1$ then, one victory in A', denoted $1_{A'}$, is faith to $(n+1)$ victories in A, each denoted 1_A, i.e.:

$$1_{A'} = (n+1) \times 1_A.$$

Now, suppose that $(A' - A)$ wins the test over A', i.e. the $(n+1)$ tests over A, then:

$$\mathbf{R}^1 (A' - A) = 1_{A'} = (n+1) \times 1_A.$$

Hence, since $\mathbf{R}^1 (A' - A) = n \times 1_A$ by Proposition 3, for $n \geq 1$, it implies

$$(n+1) \times 1_A = n \times 1_A,$$

i.e. $1_A = 0$ which is wrong for all $A \neq \emptyset$.□

This proposition establishes that the design of the biased test is such that the bias $(A' - A)$ cannot win it and, consequently, does not change the set of actual candidates for promotion, namely A. Being involved in the contest without any possibility to win, implies *mentoring* one of the contestants of A where *mentoring* means that (i) the bias *helps* a candidate, i.e. does not really participate in the contest and (ii) belongs to an *upper* hierarchical level.

4.1 The Mentor Theorem

This theorem deals with identifying the bias in a contest as a mentor. The proof is in four steps:

First step : It is easy to remark that, in fact, A' can always be decomposed into $A \cup (A' - A)$. Hence, we have:

$$A' = \cup_{j \in A - \{i\}} \{j\} \cup \{i\} \cup (A' - A).$$

Denote $i_m = \{i\} \cup (A' - A)$, the union between the agent i and the bias $(A' - A) = m$ (m for mentor). This set union reflects the fact that a contestant i will be helped during the production period by an agent (or a group of agents) m. So, for a given contest $\mathbf{C}^s_A$, we define a biased test as a contest $\mathbf{C}^1_{A'}$ faith to $\mathbf{C}^{n+1}_A$, where one agent $i \in A$ is helped by $m = (A' - A)$ whose reputation is equal to $\mathbf{R}^1(m) = n \times 1_{A.}$

The performance of the aggregated agent i_m corresponds to a *joint performance union*,[8] such that it is impossible for the principal to precisely distinguish the respective contribution of i and m within the

[8] See Arai, Billot & Lanfranchi for further details about the concept of 'joint performance union'.

global performance of i_m. This interpretation is consistent with the following Proposition which defines the reputation of the winner of the biased test:

Proposition 7 : *Consider a contest* $\mathbf{C}_A^s$*, the reputation of the winner* w *of a biased test* $\mathbf{C}_{A'}^1$ *faith to* $\mathbf{C}_A^{n+1}$ *is the following:*

$$\begin{cases} \text{- (i) } \textit{the reputation } \mathbf{R}^1(i) = 1_A \textit{ if } i_m = w, \\ \text{- (ii) } \textit{the reputation } \mathbf{R}^1(j) = (n+1) \times 1_A \textit{ if } j \neq i_m \textit{ and } j = w. \end{cases}$$

Proof: (i) Straightforward, since by Proposition 3 and the definition of the reputation, $\mathbf{R}^1(i) = \mathbf{R}^1(i_m) - \mathbf{R}^1(m) = 1_{A'} - (n \times 1_A) = 1_A$.

(ii) By definition of the reputation, $\mathbf{R}^1(j) = 1_{A'} = (n+1) \times 1_A$.□

This proposition establishes that since i does not win the biased test in A' by his own, the principal cannot see him as the true winner and does not credit him with the reputation $1_{A'}$, neither with the $(n+1)$ victories in A. On the contrary, if the helped leader loses this biased test, then the winner j, who is not helped, wins by his own. Consequently, in that case, j will benefit from the whole reputation of his victory in A'.

Second step : We need to know how the principal can in every possible situation, translate the record of a biased test over A' into equivalent records of $(n+1)$ tests over A.

From proposition 5, it is clear that the information acquired from the biased test is qualitatively different from the one that could be acquired from $(n+1)$ supplementary tests over A with a unique winner. In fact, when i_m wins the biased test, the sum of the reputation of the contestants in A only represents the reputation value of one real victory in A because $\mathbf{R}^1(i) = 1_A$ while, for all the other workers j in A, $\mathbf{R}^1(j) = 0$, just like if only one of the $(n+1)$ tests had been recorded. If the principal wants to infer from this unique record the workers' reputation corresponding to $(n+1)$ tests over A, he must consider the remaining n tests as virtual ones for which Proposition 1 applies. In the opposite case, when i_m loses the biased test, the winner j obtains a reputation equal to $(n+1) \times 1_A$. The next proposition expresses the rules to calculate the reputation of the workers over A after the biased test whoever the winner is:

Proposition 8 : *Consider two faith contests* $\mathbf{C}_A^{n+1}$, $\mathbf{C}_{A'}^1$, *such that there exists a mentor* $m = (A' - A)$ *whose reputation is* $\mathbf{R}^1(m) = n \times 1_A$. *Then, after any contest* $\mathbf{C}_A^s$, *for any* $(s+1)$*th biased test* $\mathbf{C}_{A'}^1$, *the reputations* $\mathbf{R}^{s+n+1}(.)$ *are defined by:*

$$\begin{cases} \mathbf{R}^{s+n+1}(i) & = \mathbf{R}^s(i) + (n+1) \times 1_A \text{ if } i \text{ wins and } i \text{ is not helped} \\ \mathbf{R}^{s+n+1}(j) & = \mathbf{R}^s(j) \text{ otherwise,} \\ or & \\ \mathbf{R}^{s+n+1}(i) & = \frac{s+n}{s}\mathbf{R}^s(i) + 1_A \text{ if } i \text{ wins and } i \text{ is helped} \\ \mathbf{R}^{s+n+1}(j) & = \frac{s+n}{s}\mathbf{R}^s(j) \text{ otherwise.} \end{cases}$$

Proof: Two cases occur. First, i wins without help: $\mathbf{R}^1(A') = \mathbf{R}^{n+1}(A) = \mathbf{R}^{n+1}(i) = 1_{A'}$. Hence,

$$\mathbf{R}^{s+n+1}(i) = \mathbf{R}^s(i) + \mathbf{R}^{n+1}(i)$$

which is equal to

$$\mathbf{R}^{s+n+1}(i) = \mathbf{R}^s(i) + 1_{A'} = \mathbf{R}^s(i) + (n+1) \times 1_A,$$

while the other agents j of A keeps the same records:

$$\mathbf{R}^{s+n+1}(j) = \mathbf{R}^s(j).$$

Those results correspond to the situation where i can claim all the victories since he is not helped.

Second, i_m wins: $\mathbf{R}^1(i_m) = 1_{A'}$, but because i is helped, $\mathbf{R}^1(i) = 1_A$ and $\mathbf{R}^1(i_m) = \mathbf{R}^{n+1}(A) > \mathbf{R}^{n+1}(i)$. In such a case, the n first virtual tests do not change the probabilities obtained from $\mathbf{C}_A^s$, while the $(n+1)$th is considered as won by i. Hence, from Proposition 2, we can write:

$$\mathbf{R}^{s+n+1}(i) = \mathbf{R}^s(i) + \mathbf{R}^{n+1}(i) = \mathbf{R}^s(i) + V^n(i) + \mathbf{R}^1(i)$$

i.e., from equation (4):

$$\mathbf{R}^{s+n+1}(i) = \frac{s+n}{s}\mathbf{R}^s(i) + \mathbf{R}^1(i).$$

Then,

$$\mathbf{R}^{s+n+1}(i) = \frac{s+n}{s}\mathbf{R}^s(i) + 1_A.$$

Moreover, in order to satisfy $\mathbf{R}^{s+n+1}(A) = (s+n+1) \times 1_A$, when i_m wins, we have for all $j \neq i_m$:

$$\mathbf{R}^{s+n+1}(j) = \frac{s+n}{s}\mathbf{R}^s(i).\square$$

When the winner of the biased test has been helped, he cannot be considered as fully responsible for the victory. In that case, the principal increases his reputation with one true victory and assign the value of the bias to all the contestants proportionally to their actual reputation after s tests, due to equation (4). On the contrary, if the winner is not helped, his reputation is increased by the equivalent in A of the unit value of the test over A'.

Third Step : Another question concerns the right contestant to help. The biased test must be designed in order to be relevant for promotion. In Example 2 above, it is straightforward that if the principal wants the 5th test to be relevant, he must help the leader after 4 tests. Actually, in any biased test where the leader is helped and wins, the principal learns that the leader remains the best candidate but with a higher probability. On the contrary, if a worker wins against the helped leader, the principal could then conclude that j is the agent whose probability of being the best is the highest if the biased test is relevant. In this way, a victory of any of the workers who do not receive help against a stronger field of contestants because of the mentor's intervention, is a sufficient signal that he is more likely to be the most able contestant. Moreover, the quantity measuring the *advantage of the leader(lead)*, that is $d\mathbf{R}^s = V^{s+n+1}(w_A^s) - \mathbf{R}^s(j_A^s)$ where w_A^s is the leader and j_A^s the follower, is decisive for the relevance of the biased test:

Theorem 9 (of relevance) : *Consider two faith contests $\mathbf{C}_A^{n+1}$, $\mathbf{C}_{A'}^1$ such that there exists a mentor $m = (A' - A)$ whose reputation $\mathbf{R}^1(m) \geq d\mathbf{R}^s$. Then, after any contest $\mathbf{C}_A^s$, the $(s+1)$th biased test $\mathbf{C}_{A'}^1$ is relevant iff $i_m = \{w_A^s \cup m\}$.*

Proof: (i) Assume any other contestant $i \neq w_A^s$ to be helped. Then, the $(s+1)$th biased test is relevant iff:

$$\max_A \left[P_A^{s+n+1}(i)\,; i_m \neq \{w_A^s \cup m\}\right] \geq P_A^s(w_A^s).$$

Now, consider that $i \neq w_A^s$ wins, i.e.:

$$\max_A \left[P_A^{s+n+1}(i)\,; i_m \neq \{w_A^s \cup m\}\right] = \frac{\max[\frac{s+n}{s}\mathbf{R}^s(w_A^s), \frac{s+n}{s}\mathbf{R}^s(i) + 1_A]}{(s+n+1)\times 1_A}.$$

Then, for the relevance, we need:

$$\frac{\max[\frac{s+n}{s}\mathbf{R}^s(w_A^s), \frac{s+n}{s}\mathbf{R}^s(i) + 1_A]}{(s+n+1)\times 1_A} \geq \frac{\mathbf{R}^s(w_A^s)}{s\times 1_A}$$

which is true if and only if, for all $i \in A$:

$$\frac{s+n}{s}\mathbf{R}^s(w_A^s) < \frac{s+n}{s}\mathbf{R}^s(i) + 1_A$$

i.e.

$$\mathbf{R}^s(w_A^s) - \mathbf{R}^s(j) < \frac{s}{s+n}1_A \leq 1_A \text{ for any } n \geq 0.$$

Then, since $\mathbf{R}^s(w_A^s) - \mathbf{R}^s(i) \geq 1_A$ by the assumption of non uniformity of the reputation, it is inconsistent with $\mathbf{R}^s(w_A^s) - \mathbf{R}^s(i) < 1_A$ and the test is then irrelevant.

(ii) Assume the leader w_A^s to be helped. Then, the $(s+1)$th biased test is relevant iff:

$$\max_A \left[P_A^{s+n+1}(i)\,; i_m = \{w_A^s \cup m\}\right] \geq P_A^s(w_A^s).$$

Two cases occur: (1) The leader w_A^s wins, i.e., by Proposition 7:

$$\max_A \left[P_A^{s+n+1}(i)\,; i_m = \{w_A^s \cup m\}\right] = \frac{\frac{s+n}{s}\mathbf{R}^s(w_A^s) + 1_A}{(s+n+1)\times 1_A}.$$

Then, we always have:

$$\frac{\frac{s+n}{s}\mathbf{R}^s(w_A^s) + 1_A}{(s+n+1)\times 1_A} \geq \frac{\mathbf{R}^s(w_A^s)}{s\times 1_A}$$

since $\mathbf{R}^s(w_A^s) \leq s \times 1_A$ and the biased test is always relevant.

(2) Another contestant $j \neq w_A^s$ wins, i.e.:

$$\max_A \left[P_A^{s+n+1}(i)\,; i_m \neq \{w_A^s \cup m\}\right] = \frac{\max\left[\mathbf{R}^s(w_A^s), \mathbf{R}^s(j) + (n+1)\times 1_A\right]}{(s+n+1)\times 1_A}. \tag{9}$$

Then, for the relevance, we must have:

$$\frac{\max\left[\mathbf{R}^s(w_A^s), \mathbf{R}^s(j) + (n+1) \times 1_A\right]}{(s+n+1) \times 1_A} \geq \frac{\mathbf{R}^s(w_A^s)}{s \times 1_A}$$

which is true if and only if for all $j \in A$ and then for the follower j_A^s with the lowest reputation:

$$\frac{s+n+1}{s}\mathbf{R}^s(w_A^s) < \mathbf{R}^s(j_A^s) + (n+1) \times 1_A$$

i.e.

$$d\mathbf{R}^s = V^{s+n+1}(w_A^s) - \mathbf{R}^s(j_A^s) \leq \mathbf{R}^1(m) = n \times 1_A.$$

Hence, the $(s+1)$th test is relevant if and only if the leader w_A^s is helped by m such that $\mathbf{R}^1(m) \geq d\mathbf{R}^s$.$\square$

This theorem is an ordinal extension to the case of n contestants of Meyer's result under which the most efficient contestant in the past must be given a positive handicap.[9]

Such a result could seem to be linked with *favoritism*. Prendergast and Topel, 1995, define favoritism from a supervisor in a three-tier hierarchy as a bias in his preferences towards one of the supervised workers. Consequently, such tastes can lead to inefficient assignment of workers within the hierarchy. In our framework, the mentoring relation cannot be mixed up with favoritism for two reasons. (i) At the beginning of the selection, the principal has no prior preference for any of the contestants. The objective results of the contest reveal the relative ability of the workers and then, the principal chooses to bias the final test and to mentor the leader according to the relevance criterion. In other words, mentor-protege relations help to promote the best worker, whoever he is, while favoritism leads to promotion of a particular worker, whatever his relative ability (otherwise, it would not be a favor). (ii) Our selection process implying that the leader is helped during the biased test, can nonetheless not be viewed as favoring the leader. Actually, if the last test is still designed as the previous ones, the leader is sure to be promoted (see Example 2, Proposition 2). Even though mentoring him increases his performance during the last test, it does not fully protect him against the

[9] When after a given sequence of s tests, the principal identifies multiple leaders with equal reputations, each of them should receive the same help. Nevertheless, the reputation of the mentor remains the same, each leader sharing equally his help.

risk of finally being outdone by another contestant. *Hence, if we ask this leader to choose the design of the final test, he should rationally prefer not to be helped*. Furthermore, from the viewpoint of the other contestants, the relevant informative value of the biased test yields a last chance of promotion for all of the contestants.

Corollary 10 : *Consider two faith contests* $\mathbf{C}_A^{n+1}$, $\mathbf{C}_{A'}^1$ *such that there exists a mentor* $m = (A' - A)$ *whose reputation* $\mathbf{R}^1(m) \geq d\mathbf{R}^s$*. Then, after any contest* $\mathbf{C}_A^s$*, a relevant* $(s+1)$*th biased test* $\mathbf{C}_{A'}^1$ *allows the follower* j_A^s *to win the contest* $\mathbf{C}_A^{s+n+1}$*.*

Proof: Straightforward since by relevance, we have:

$$d\mathbf{R}^s < (n+1) \times 1_A$$

i.e.

$$\mathbf{R}^s(j_A^s) + (n+1) \times 1_A > \frac{s+n+1}{s}\mathbf{R}^s(w_A^s).$$

Now, if the follower j_A^s wins the biased test $\mathbf{C}_{A'}^1$, it follows from Proposition 6:

$$\mathbf{R}^{s+n+1}(j_A^s) = \mathbf{R}^s(j_A^s) + (n+1) \times 1_A > \frac{s+n+1}{s}\mathbf{R}^s(w_A^s) > \mathbf{R}^s(w_A^s)$$

and the result holds.□

The qualitative nature of these results, that is mentoring the leader of the contest while authorizing the follower to be promoted after the last test, only depends on the relevance criterion. Hence, designing a biased test introducing a joint performance union between the leader and a mentor does not require that the principal acts as a probability maximizer. A satisficing behavior is clearly sufficient.

Fourth Step : The following theorem allows identification of the mentor as a previously promoted worker, i.e. belonging to X_2 or X_3. Actually, we know that the principal promotes the worker with the highest reputation, that is with the highest probability of being the best. Then, we show that each winner of a biased contest, that is a contest in which some test is biased, can bias afterwards any other contest where the advantage of the leader is inferior to his final reputation.

Theorem 11 (of the mentor) : *Consider two faith contests* $\mathbf{C}_A^{n+1}$, $\mathbf{C}_{A'}^1$ *such that there exists a mentor* $m \in X$, *whose reputation is* $\mathbf{R}^1(m) \geq d\mathbf{R}^s$. *Then, the winner* w *of the biased contest* $\mathbf{C}_A^{s+n+1}$ *can bias a* $(\tilde{s}+1)$*th test of any contest* $\mathbf{C}_{\tilde{A}}^{\tilde{s}}$ *such that* $d\mathbf{R}^s \geq d\mathbf{R}^{\tilde{s}}$.

Proof: A winner w of a biased contest $\mathbf{C}_A^{s+n+1}$ can always bias a $(\tilde{s}+1)$th test of any contest $\mathbf{C}_{\tilde{A}}^{\tilde{s}}$ when his reputation $\mathbf{R}^{s+n+1}(w)$ is greater or equal to $d\mathbf{R}^{\tilde{s}}$, as shown by Theorem 1.

By Theorem 1, a relevant biased $(s+1)$th test is such that $i_m = \{w_A^s \cup m\}$. Then, two possibilities emerge: (i) A contestant $w \neq w_A^s$ wins. In this last case, Proposition 6 establishes that:

$$\mathbf{R}^{s+n+1}(w) = \mathbf{R}^s(w) + (n+1) \times 1_A > d\mathbf{R}^s.$$

(ii) The leader w_A^s wins, i.e. $w = w_A^s$:

$$\mathbf{R}^{s+n+1}(w) = \frac{s+n}{s}\mathbf{R}^s(w_A^s) + 1_A > d\mathbf{R}^s.$$

Thus, we know, for any relevant biased contest $\mathbf{C}_A^{s+n+1}$, that the promoted worker's reputation $\mathbf{R}^{s+n+1}(w)$ is always greater than $d\mathbf{R}^s$, whoever he is. Now, consider a contest $\mathbf{C}_{\tilde{A}}^{\tilde{s}}$ such that $d\mathbf{R}^s \geq d\mathbf{R}^{\tilde{s}}$. Then, it is always true that:

$$\mathbf{R}^{s+n+1}(w) \geq d\mathbf{R}^{\tilde{s}}.\square$$

The following final example summarizes the design and results of a relevant biased test

Example 12 : *Consider as, in Examples 1 and 2, a contest where after* 4 *previous tests over* $A = \{i, j\}$, *the reputations are given by:*

$$\mathbf{R}^4(i) = 3 \times 1_A \text{ and } \mathbf{R}^4(j) = 1_A.$$

Hence, the probabilities for the two candidates to be the best correspond to:

$$P_A^4(i) = 3/4 \text{ and } P_A^4(j) = 1/4.$$

We know any 5*th test to be irrelevant. Hence, introduce a mentor* m *such that* $\mathbf{R}^1(m) = n \times 1_A$ *with* $d\mathbf{R}^4 \leq n \times 1_A$, *since* $i = w_A^s$ *and* $j = j_A^s$. *We can choose* $n = 11$ *for instance, i.e. the necessary minimum to bias since:*

$$d\mathbf{R}^4 = \left(\frac{(5+n)}{4}3 - 1\right) \leq n.$$

Then, we can consider the actual contest $\mathbf{C}^1_{A\cup\{m\}}$ *as equivalently informative as a virtual contest* C_A^{11+1}. *By Theorem 1, we know that it is relevant to help the leader, that is to form a joint performance union* i_m. *Now, according to Proposition 6, two possibilities emerge:*

-(i) i_m *wins and* $\mathbf{R}^{16}(i) = \frac{4+12}{4}\mathbf{R}^4(i) + 1_A = (4 \times 3 + 1) \times 1_A = 13 \times 1_A$. *Hence,* $P_A^{16}(i) = \frac{13}{16} \times 1_A > 3/4$,

-(ii) j *wins and* $\mathbf{R}^{16}(j) = \mathbf{R}^4(j) + 12 \times 1_A = 13 \times 1_A$. *Hence,* $P_A^{16}(j) > 3/4$.

In both cases it is clear that the principal increases the quality of his choice measured as the confidence attached to the candidate who is finally promoted.

4.2 Empirical Relevance

The bias in our model, namely help provided by a senior manager, sounds natural according to previous empirical results about the positive effects of mentors on the career of proteges in different types of productive organizations. Mentors' influences on access to partnership within law firms are documented by Laband and Lentz. Similar patterns are reported by Ochberg, Tischler and Schulberg (1986) for the medical sector. Various surveys about young managers and professional workers arrive at similar conclusions (see e.g. Roche (1979), Collins and Scott 1979, Chao). Whytely, Dougherty and Dreher (1991) for example, show that the success of workers in their early career is positively related to mentoring relationships, arguing that successful newcomers in firms have received frequent help, advice or exposure from senior managers.

Moreover, in our model, the previous performance records of the promoted workers serve as a specific asset for the organization in contrast to external entrants in ranks X_2 or X_3 who are not associated with such a previous reputation. In this way, the organization can be seen as a stock of reputation accumulated from its previous workers' selection policy. Choosing to promote someone already in the organization, can be seen as formally equivalent to an investment insuring that future quality of promotion will be improved. Nevertheless, the supply of reputation is naturally limited by the size of the organization and more particularly by the limited number of senior managers who are able to help the junior workers. This restriction has an appealing consequence for the design of the biased contests. Because the extent of the bias is increasing with the ability lead of a leading worker, the more quickly a candidate signals his productive advantage, the more difficult it will be to find a mentor whose reputation can match the lead. Hence, if an agent seems definitively better, the principal needs to stop the contest earlier, thus offering a faster internal career for this agent. This corresponds to a well-known stylized fact, that early signalling of ability by workers and their later achievements in terms of internal career are positively correlated (see Rosenbaum (1984), Forbes (1987), and Bruderl, Diekmann and Preisendorfer (1991).

5 Final Remarks

The model we propose in this paper is to our knowledge, the first theoretic approach formally establishing that mentor-protege relations can be beneficial to the organization.

Two remarks can be made in relation to the related literature:

(i) The main result, as an extension of the Meyer's idea, allows us to interpret the learning behavior of the principal in the context of the job design literature. Earlier literature studying the optimal choice of the job contents as a part of the motivation system, show that cooperation and coordination between workers of the same hierarchical rank can complement incentive compensation policies (See e.g. Holmstrom and Milgrom (1991) and Valsecchi (1996)). Analogous to this 'incentive' argument for cooperation, our contribution establishes the importance of work organization for the selection scheme.

(ii) Mentoring as an integral part of a selection process according to our model, serves also as a means of creating long term employment relations between the workers and their firm in line with the human capital theory. The implementation of vertical cooperation, as learning by doing, naturally increases the productive knowledge of the worker. Moreover, after the promotion, a worker is regarded as a specific asset for the organization since these workers reputation can be used to optimally bias future promotion contests. Such a specific human capital reinforces the link between senior workers and the organization. In some sense, expected gains from efficient screening with mentoring are similar to the use of Old Boys Networks by firms to proceed in external hiring (see Simon & Warner (1992)).

From a methodological viewpoint, a new feature of our contribution is the modelling of the selection problem in a discrete choice style. For further research, this family of probabilistic models seem to be useful for describing imperfect decision making within the context of organization theory. A natural extension of this paper is to model the promotion process when supervisors record the results of repeated *quota*-like tests in order to analyze the properties and relative advantage of the two systems of selection: the tournaments and quotas (see Lazear and Rosen (1981) for tournaments and Drago and Turnbull (1991) for quotas). An appropriate tool for that purpose could be the Tversky's selection model of discrete alternatives (1972), based on binary experiments about specific attributes.

References

Arai, Mahmood, Billot, Antoine and Joseph Lanfranchi, 1994. Efficient selection of agents under limited ability to rank, WP-ERMES 94-10, Université Panthéon-Assas, Paris 2.

Billot, Antoine, 1998, Autobiased choice theory, *Annals of Operations Research*, 80, 85-103.

Billot, Antoine and Tony E. Smith, 1998, Informative contests and the efficient selection of agents, In: Martin J. Beckmann and Roland Thord, (Eds.), *Knowledge and networks in a dynamical economy*, Essays in honor of Ake Andersson. Springer-Verlag, New York-Heidelberg, pp 273-287.

Bruderl, Josef, Diekmann, Andreas and Peter Presendorfer, 1991. Patterns of intraorganizational mobility: Tournament models, path dependency, and early promotion effects, *Social science Research*, 20, 197-216.

Calvert, Randall L., 1985. The value of biased information: a rational choice model of political advice, *Journal of Politics*, 47, 530-555.

Chao,G., 1997. Mentoring phases and outcomes, *Journal of Vocational Behavior*, 51, 15-28.

Collins, Eliza and Patricia Scott, 1978. Everyone who makes it has a mentor, *Harvard Business Review*, July-August, 89-101.

Drago, Robert and Geoffrey K. Turnbull, 1991. Competition and cooperation in the workplace, *Journal of Economic Behavior and Organization*, 15, 347-364.

Forbes, John B., 1987. Early intraorganizational mobility: patterns and influence, *Academy of Management Journal,* 30, 57-71.

Holmstrom, Bengt and Paul Milgrom, 1991. Multi-task principal-agent relations: incentive contracts, asset ownership and job design, *Journal of Law, Economics and Organization,* 5 Supp, 24-52.

Koh, Winston T.H., 1994. Fallibility and sequential decision-making, *Journal of Institutional and Theoretical Economics,* 150, 362-374.

Kram, Kathy K., 1985, *Mentoring at work: Developmental relationships in organizational life*, (Scott, Foresman: Glenview, Illinois).

Lanfranchi, Joseph, 1994, *Sélection, promotion et tournois,* Ph.D-Thesis, Université Panthéon-Assas, Paris.

Laband, David N. and Bernard F. Lentz, 1995. Workplace mentoring in the legal profession, *Southern Economic Journal*, 61, 783-802.

Lazear, Edward P. and Sherwin Rosen, 1981. Rank-Order tournaments as optimal labor contracts, *Journal of Political Economy*, 89, 841-864.

Meyer, Margareth, 1991. Learning from coarse information: biased contests and career profiles, *Review of Economic Studies,* 58, 15-41.

Ochberg, Richard L., Tischler, Gary L. and Herbert C. Schulberg, 1986. Mentoring relationships in the careers of mental health administrators, *Hospital and Community Psychiatry,* 12, 939-940.

Prendergast, Canice and Robert Topel, 1995. Favoritism in organizations, *Journal of Political Economy,* 103, 958-978.

Roche, Gerard F., 1979. Much ado about mentors, *Harvard Business Review,* January-February, 14-28.

Rosenbaum, James., 1984, *Career mobility in a corporate hierarchy*, (Acadmic Press, Orlando).

Simon, Curtis J. and John T. Warner, 1992. Matchmaker: the effects of old boys networks on job match quality, earnings and tenure, *Journal of Labor Economics,* 12, 306-330.

Simon, Herbert A., 1955, A behavioral model of rational choice, *Quarterly Journal of Economics*, Vol 69, pp 99-118.

Tversky, Amos, 1972. Choice by elimination, *Journal of Mathematical Psychology,* 9, 341-367.

Whitely, William, Dougherty, Thomas W. and George F. Dreher, 1991. Relationship of career mentoring and socio-economic origin to managers and professionals early career progress, *Academy of Management Journal,* 34, 331-351.