The organization of social learning in firms: Should it be formal or informal?*

Alexandra RUFINI†

Abstract

This paper addresses how competences transmission could be organized in firms. Two kinds of social learning process are analyzed and compared. In the first one (formal organization called Mentorship), a mentor is designated among insiders to transmit his competences but must stop working. In the second one (informal organization called Teamwork), mixed teams are composed to encourage an informal learning where on-the-job interactions are likely to occur. Finally, a numerical exploration of the model defines the optimal learning process according to the firm and workers characteristics and surprisingly shows that a formal organization is not always a dominated strategy.

JEL Classification: D23; J24; M54
Keywords: learning, agent's heterogeneity, teams, tournament, incentives.

*I would like to thank Pierre-André Chiappori, Pierre Garrouste, Alan Kirman, Nathalie Lazaric, Patrick Musso and Dominique Torre for their hepful comments and suggestions. Thanks also are due to the participants at the Third Summer School on "Economic analysis of heterogeneity in social organizations", June 2007, Louvain-la-Neuve; and at the International conference on "Incentives and motivations within organizations", February 2008, Paris. All remaining errors are mine.

†University of Nice Sophia-Antipolis - GREDEG(DEMOS) - CNRS, E-mail: rufini@gredeg.cnrs.fr

1 Introduction and Related Works

In a knowledge-based economy learning is seen as enhancing firm performance. Therefore, workers need to continuously acquire the appropriate skills to make them more productive and this may involve, *social learning*[1] processes. Some papers in management science address this precise issue, arguing that firms can encourage peers to work and learn collaboratively and, hence, should facilitate the learning of their members (Senge 1990; Pedler et al. 1991). The economic literature (see Bishop 2000 for an extended abstract) provides several overviews of this concept of learning and training within organizations, but generally without considering the organization of the firm as a driver of learning. One exception is the paper by Garicano (2000), which argues that a "knowledge-based hierarchy is a natural way to organize the acquisition of knowledge"[2]. In his basic model, the question of learning appears when agents are confronted with a problem that they are not able solve. They can either learn on their own, which incurs a cost, and means they have to stop working, or they can seek help from a co-worker, who has to stop working, and this then incurs a cost of transmission. The trade off leads to a specialization process in which some workers produce and learn the most common problems and others confront and overcome the most difficult ones and become specialized in the transmission of knowledge. Garicano and Hubbard (2005) made an extension of this model (*i*) in which agents have heterogeneous learning costs and (*ii*) knowledge has to be acquired on the job. They show that the hierarchical organization is optimal when workers and managers of the same type (according to their learning costs) are matched. Although these papers explore the ways in which firms should organize the acquisition of knowledge through the reorganization of their workforce, the question of the incentives for workers to train others remains clouded. In other words, the reasons that motivate workers to take on this role are not explained.

This paper addresses this question by analysing the ways in which workers are organized and the incentive schemes associated with the generation of knowledge transmission. I focus on the transmission of firm-specific competences. The reason for this choice is to ensure that training would have taken place within the firm. If the choice had been to study general competences, then firms would have the choice to outsource the learning process (via external training programmes, for instance) which would not have an impact on the organization of workers nor on the incentives for the firm to provide training for its workforce. Hence, the vertical heterogeneity[3] of the workers studied includes firm tenure: insiders hold firm-specific competences whereas new hires do not (even if they have achieved tenure in a previous firm)[4]. This kind of heterogeneity is frequent and

[1]When used in other fields than labour economics, the term social learning usually refers to learning by observing the behaviours of others (see for instance Gale 1996 or Ellison and Fudenberg 1993). In my paper, the definition is more restrictive since social learning involves interactions and, hence, expresses more than simple observation of peers.

[2]Garicano (2000), p. 874.

[3]By vertical heterogeneity, I mean that workers have characteristics that permit them to be ranked according to their individual performance (e.g., workers hold low levels of knowledge or high levels of knowledge, which leads respectively to low individual performance and high individual performance). Horizontal heterogeneity, on the other hand, means that individuals have heterogeneous but complementary characteristics (to continue with the example of knowledge: workers hold formal knowledge or informal knowledge, but the level of knowledge and individual performance are the same).

[4]Firm tenure and age are often confused. Indeed, young workers, on average, have lower firm tenure

appears in a particular context: when the firm has hired new workers. Since I assume that the productivity of workers depends on both general and specific competences, it is obvious that new hires are more productive once they have acquired these specific competences. They could learn these competences on their own through a lengthy individual learning process. However, depending on the degree of complexity of the competences involved, some of these could not be achieved without the knowledge or the know how of experienced workers. Hence, the firm's main objective must be to identify strategies to facilitate knowledge transmission from insiders to new hires.

This paper suggests two alternative worker allocations. The first is a *formal organization* called "Mentorship". Mentoring activities and social learning are closely related. Yet, while the management science literature has emphasized the empirical evidence on mentorship for more than 20 years (see Kram 1983; Hunt and Michael 1983 for seminal studies), it has been somewhat neglected in the economic literature. Laband and Lentz (1995) studied the emergence of this mentor-protégé relationship in the legal profession. They conclude that firms use mentoring to (*i*) induce juniors to acquire firm-specific human capital and (*ii*) reduce turnover. This latter issue is not studied in the current paper, which focuses rather on how the former occurs. Some recent economic papers have focused on how mentoring activities impact on the diversity of the workforce (Athey et al. 2000) or improve the promotion process (Arai et al. 2001). But none of these study mentorship as a learning process. Here, I define mentorship as a formal learning process in which the objective of the manager is, on the one hand, to identify the ablest worker (the mentor) to train new hires, and on the other hand to provide sufficient incentives to ensure that the mentor will invest his best efforts. The second learning process analysed occurs through an *informal organization* called "Teamwork". From the papers by Hamilton et al. (2003, 2004), this kind of workers organization seems to be relevant for the study of social learning. Emphasizing the role of the social interactions likely to occur within teams, these author argue that teamwork can generate peer pressure, which limits the incentives to free ride (see also Kandel and Lazear 1992), and claim that "diversity in skill level and ability enhances team productivity if there is significant mutual learning[5] and task coordination within the team"[6]. But, while Hamilton et al. underline mutual learning as a positive consequence of workers' vertical heterogeneity, I explicitly explore the ways *social learning* could take place and be supported by the organization of the firm. Hence, through this informal learning process, I study how managers compose optimal mixed teams of insiders and new hires in order to induce the transmission of competences.

Finally, through these two social learning processes, I analyse workers' organization within firms when the main objective is to facilitate social learning among them. I want to answer the following question: Does a formal organization outperform an informal one?

than older workers. But, firm tenure is related to firm-specific competences whereas age is not (a young worker could hold a high level of firm specific competences if he has been hired for few years and a senior could have a low level if he has just been hired). Firm tenure is thus a relevant type of heterogeneity for a study based on the learning processes likely to occur inside the firm.

[5] This "mutual learning" is the referred to as social learning in the current paper; it consists of transmission of learning from the most skilled to the less skilled workers.

[6] Hamilton et al. (2004), p.15.

The paper is organized as follows. Section 2 describes the structure of the basic model. Sections 3 and 4 present characterize Mentorship and Teamwork respectively, as social learning processes. In Section 5 I discuss the conditions under which one learning process is preferred over another, through a numerical example. Section 6 concludes the theoretical framework.

2 The Basic Structure of the Model

2.1 Workers

Consider a workforce after a recruitment process. There are two categories of workers within the firm: insiders that have firm-specific competences since they have high level of firm tenure, and new hires with no firm-specific competences. Insiders, denoted s, are called *seniors* and new hires, denoted j, *juniors* herafter.

The productivity of a worker x, indexed by y_x, whatever his/her type $x = \{j, s\}$, is related to his/her level of firm tenure and skill level. All seniors have the same level of firm tenure and skill, therefore the same level of productivity y_s. Juniors have the same low level of firm tenure but they are heterogeneous in their level of skill given the recruitment channel they have been hired or their level of education. I assume that the *average* juniors' level of skill is the same than the seniors' one in order to focus on the firm tenure heterogeneity. Hence, as long as juniors have not acquired firm-specific competences the senior's productivity level will always be higher than the average junior's one, *i.e.* $y_s > \bar{y}_j$. This gap in productivity could be reduced through two methods. On the one hand, juniors could learn on their own and acquire firm-specific competences at an average rate $\bar{\gamma}$. On the other hand, seniors could transmit their firm-specific competences to juniors through a level of effort incurring a cost. But, when the levels of competences of workers differ, it is likely that the way that they transmit competences will also differ (Cohendet and Steinmueller 2000). Thus, I assume that seniors have heterogeneous costs of competence transmission. This assumption means that some seniors hold particular attributes, such as patience or pedagogy, which make the effort of transmission less painful and less costly. The parameter α_i denotes the senior's i ability to transmit his/her competences, which would be linked to his/her cost of effort. The lower α_i, the better able is the senior i to train others and the lower is the cost of the transmission effort. I assume that α is uniformly distributed among seniors, such that $[0, \alpha^{\max}]$. In spite of their level of firm tenure, seniors are not able to observe the ability to transmit of others, since social learning processes have not been implemented. In other words, there is no reputation effect on the ability to transmit these competences. However, seniors are able to identify the average value of α, denoted $\bar{\alpha}$, and its distribution, which I suppose is uniform.

2.2 Production, profits and the objective of the manager

In this model, the time horizon of the learning process in firms is finite and begins immediately after the recruitment process. Since workers acquire specific competences, both firm and workers are willing to invest in a relatively long term relationship. Then the firm workforce will be stable during the learning process (there is neither recruitment nor turnover during this period).

The manager pays a fixed wage, w_i, to the worker i negotiated during the hiring process. The payroll supported by the firm is hence $W = \sum_{i=1}^{s+j} w_i$. For this level of wage, worker i always provides the level of production effort required. In fact, the traditional incentive problems related to production have been voluntarily disregarded in order to focus on those related to the learning.

The need for incentives in the case of learning has the same origins as in more traditional settings: lack of information. A natural assumption is to suppose that neither the level of transmission effort nor the individual ability to transmit is observable. The manager only has information on the average value of the ability to transmit $\bar{\alpha}$ and its distribution. The manager can observe the average productivity of workers and the individual learning rate of juniors.

The objective of the manager then is to find the appropriate incentive to encourage seniors to provide high level transmission efforts. But, such high level of efforts may decrease their productivity if too much of their time is spent training juniors. In addition, not all seniors are willing to enter into such a training process. Thus, there will only be a few seniors filling the mentoring role[7] indexed by S such as $S < s$.

Finally, the objective of the manager is to select an organization of the firm's workers'that will enhance the firm's profits by improving the productivity of juniors without losing the productivity of seniors. To estimate this, I consider a centralized reorganization associated with formal and practical learning and a decentralized reorganization associated with informal learning.

These two alternatives devices are described in depth in the following sections.

3 Mentorship

3.1 Mentorship features

One way of facilitating the acquisition of firm-specific competences by juniors is to implement formal on-site learning. A formal on-site learning programme would include courses when high tenure workers could explain the technical, organizational and cultural routines of the firm to juniors. Such formal courses could be complemented by practice, especially when technical competences are concerned. For instance, seniors demonstrate how to perform a task or help in its realization. In this context, juniors are able to learn theoretically and pragmatically since they put into practice what they have learned via the courses. However, the time of those teaching these courses is taken away from their working time.

3.1.1 The role and identification of the mentor

The formal learning process must be rationalized by the firm. Here, I assume that the manager delegates responsibility for formal learning to one senior. To justify this assumption, there are at least three arguments that can be advanced: (*i*) all things being equal, the greater the number of trainers the more that production will be reduced; (*ii*)

[7] These mentoring/training seniors can be chosen by different means: voluntary participation, preselection by the manager, random selection, etc. This determination can be described analytically.

appointing a single mentor avoids the opportunity to free ride; (*iii*) being the only person given this responsibility confers the status of leader, which could be an incentive for the trainer. It can be seen then, that, a formal learning process requires the emergence of a mentor.

However, identifying only one worker implies that the learning process will depend entirely on the mentor's ability to train. Identification of the mentor then is crucial. But, since the manager is not able to observe the seniors' ability to transmit, identification is complex. One might argue that the manager could select the mentor randomly from among seniors, but this is neither a realistic practice nor an optimal process of designation. Hence, since there is nothing that the manager is privy to that will signal who would be the best mentor, an alternative would be to stimulate competition among seniors. Then, a proper and likely process to determine a mentor might be a tournament. In this paper, a tournament is used not to incite workers to provide the highest levels of efforts, but to determine who would make the best trainer, which is an original approach with respect to the traditional economic literature on tournaments.

3.1.2 A specific tournament

A tournament within a firm is a competition among workers. In the labour economics literature, before a tournament, the manager is supposed to announce the "rules of the game" (for instance, the duration of the tournament, the number of workers who will compete and will be promoted, the tasks that will be required of them and the new level of earnings in the case of success). After the tournament, the manager observes the individual workers' production and promotes the worker(s) who has (have) produced the most according to the rules previously announced. Such schemes are often claimed to be efficient incentive processes.

The tournament in this paper is rather different. First, the tournament is used to identify the best trainer; the incentive properties normally associated with the tournament are hence disregarded. Second, after the tournament, seniors are not evaluated by the manager on their levels of production, but on the levels of production achieved by the juniors they have supervised. This assumption is made because the main objective is to determine the most able trainer and not the most productive worker among the firm's seniors.

The specification provided here can be considered as possibly illustrating the case of a tournament where the manager is willing to determine the best mentor. This specification is based on the following assumptions:

- The tournament could be composed of many rounds of training. Each round corresponds to a specific task that juniors are required to perform, denoted k.
- The manager randomly composes as many teams as candidates for mentorship, but includes in each one the same proportion j/S of juniors for a round of training.
- With each round, the composition of the teams changes randomly.
- With each round, the manager randomly matches one single senior to each team.

- During each round of training, each senior, independently and simultaneously, plays the role of the mentor within the team to which he is assigned. For that, each Training Senior (TS) i chooses his/her optimal level of effort denoted e_i^{TS*} which is a continuous control variable such that $e_i^{TS} \in [0,1]$. This level of effort incurs a cost $c_i^{TS}\left[e_i^{TS}, \alpha_i, j, S, y_s, \bar{y}_j\right]$ such that $c_i^{TS}(e_i^{TS}) = 0$ if $e_i^{TS} \leq 0$; $c_i^{TS}\left(e_i^{TS}\right) > 0$ if $e_i^{TS} > 0$ and $\delta c_i^{TS}/\delta e_i^{TS} > 0$; $\partial^2 c_i^{TS}/\partial\left(e_i^{TS}\right)^2 > 0$; $\delta c_i^{TS}/\delta \alpha_i > 0$; $\delta c_i^{TS}/\delta\left(j/S\right) > 0$ and $\delta c_i^{TS}/\delta\left(y_s - \bar{y}_j\right) > 0$.

- Juniors only produce, and seniors only transmit: this combination gives an output associated with each senior for each round.

- At the end of the tournament, the manager observes the amount of output associated with each training seniors, denoted $\sum_{k=1}^{K} Y^k$ (where K is the total number of tasks performed). The winner of the tournament is the senior whose juniors produced the highest output. This senior becomes the mentor and will be awarded a bonus after the mentorship process.

- The manager determines the optimal duration of the tournament denoted t, which is a continuous variable, such as $t \in [0,1]$ and the number of training rounds and the number of associated tasks K, in which the longer the duration the more numerous will be the training rounds.

Since juniors have heterogeneous levels of skills, the average level of skills could differ from one team to another (for instance, one team could include the three highest skilled juniors, while another might include the three least skilled). As a consequence, if there were only one round of training, some seniors could be handicapped by the random draw. On the other hand, the more rounds of training that take place, the greater the opportunity for each senior to train a different team, and the more homogeneous will be the cumulated sample of juniors associated with each senior. Then, for the most able senior (the one whose effort costs are lowest) the chances of being appointed mentor increase. In other words, the more the longer the duration of the tournament, the more the expected value of the mentor's ability to transmit increases, and the more the variance in this ability decreases (the variance however remains different from zero since seniors are heterogeneous).

In the rest of this paper, the tournament duration t will be considered as the quality indicator for the tournament outcome.

3.1.3 The mentorship learning process

Once the tournament process is over, the mentor optimally assigned is put in charge of training all juniors. In all cases, the duration of the tournament is very short compared to the duration of mentorship. For instance, to give an order of magnitude, a tournament could represent a week whereas the mentorship role could correspond to six months. Obviously, in just one week, juniors cannot assimilate the competences that take six months to be acquired.

During the mentorship learning process, juniors begin to learn on their own the firm specific competences (not before because adaptation time is needed). According to the level complexity of the competences, juniors are able to learn on their own completely, partially or not at all, but in the former two cases time is involved. The individual learning rate is cumulative from one period to another but diminishing with time such that: $1 \geq \bar{\gamma}^{\tau} > \bar{\gamma}^{\tau+1} > \ldots > \bar{\gamma}^{\mathrm{T}} \geq 0$ and $\sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^{\tau} \leq 1$. The individual learning reduces the gap between seniors and juniors' productivity (*i.e.*the firm tenure heterogeneity) and therefore induces higher production from the juniors. This outcome is expressed by the term $j\left(y_s - \bar{y}_j\right)\bar{\gamma}^{\tau}$. If the cumulative individual learning rate becomes maximal (equal to unity), the juniors' productivity will equal the seniors'. But this possibility is avoided here in order to make possible the social learning alternative.

Once the mentor is designated, he provides his/her optimal level of effort denoted e_i^{M*} which is the second continuous control variable of training seniors such that $e_i^M \in [0, 1]$. This level of effort incurs a cost $c_i^M[e_i^M, \alpha_i, j, y_s, \bar{y}_j, \bar{\gamma}^2]$, such that $c_i^M(e_i^M) = 0$ if $e_i^M \leq 0$ and $c\left(e_i^M\right) > 0$ if $e_i^M > 0$; respecting the following properties:

$$\frac{\delta c_i^M}{\delta e_i^M} > 0;\ \frac{\partial^2 c_i^M}{\partial\left(e_i^M\right)^2} > 0;\ \frac{\delta c_i^M}{\delta \alpha_i} > 0;\ \frac{\delta c_i^M}{\delta j} > 0;\ \frac{\delta c_i^M}{\delta\left(y_s - \bar{y}_j\right)} > 0 \text{ and } \frac{\delta c_i^M}{\delta \bar{\gamma}^2} < 0$$

Then, the additional outcome produced by juniors given both individual and learning process is expressed by $j\left(y_s - \bar{y}_j\right)\left(\sum \bar{\gamma}^{\tau} + \left(1 - \sum \bar{\gamma}^{\tau}\right)\tilde{e}_m^M\right)$. One can see that the mentor could transmit all of the firm-specific competences if he chooses the maximal level of transmission effort, *i.e.* $e_i^M = 1$. To induce such a high level of effort, the manager must i) be sure that the tournament selection is effecient (for that, he should choose a high level of tournament duration t), ii) find an appropriate remunerative scheme given that only the the results of effort are observable. In line with the conclusions in the standard literature on incentives, the manager could give to the mentor a bonus indexed to the results of his level of effort. Accordingly, the second control variable for the firm (the first being the tournament duration t) is a bonus multiplier, denoted w^M, such as $w^M \in \mathbb{R}^+$.

In the tournament framework, notification of the rules of the game is a prerequisite. The manager, then, has to choose the duration t and the bonus multiplier w^M prior to the tournament and, hence, without knowing the identity of the mentor. Such uncertainty about the tournament's outcome adds complexity to the manager's strategic choice. Indeed, the manager is able to determine the level of effort of the "average mentor" from the program of seniors since he knows the average cost of effort $\bar{\alpha}$. But he would not be able to determine the exact level of effort of the future mentor i which will depend on his own level of cost of effort α_i. In this context, the manager makes an expectation on the mentor's level of effort. This expectation is rational and, thus, can only be formulated according to the information held by the manager. More precisely, the manager uses the average mentor's level of effort $\bar{e}^M$ which depends on w^M but not on the other control variable t. Indeed, the optimal level of effort chosen by the "average mentor" during the mentorship process comes from the traditional trade off between the costs and the benefits, but does not take account of the uncertainty of the tournament. Workers consider this uncertainty only when they choose their level of effort during the tournament process: $\bar{e}^{TS}$ would be a function of (w^M, t) but $\bar{e}^M$ would only be a function of (w^M). Yet, the level of t would impact on the ability to transmit of the future mentor.

The uncertainty around the determination of the mentor leads the manager to use an ex ante expected relation between the two control variables (w^M, t). Therefore, the mentor's level of effort expected by the manager could be written as $\tilde{e}_m^M = f[w^M, t, \bar{e}^M]$, where the function f satisfies the following properties and conditions on the limit value of the mentor's effort expectation:

$$\frac{\delta f}{\delta w^M} > 0; \frac{\delta f}{\delta t} > 0 \text{ and } \frac{\delta f}{\delta \bar{e}^M} > 0;$$

$$\forall t, f[w^M = 0] = 0; \forall w^M, f[t = 0] > 0 \text{ and } \lim_{w^M \to +\infty} f[w^M, t = 1] = 1$$

From these conditions, one can deduce that $\forall w^M$ and t, $0 \leq f[w^M, t] \leq 1$, *i.e.* that the expected effort $\tilde{e}_m^M$ is consistent with the domain of definition of e_i^M.

Given the role of the expectation function $\tilde{e}_m^M$ of the manager in the ongoing determination of the optimal solution in w^M and t, it is necessary to introduce a condition of consistency that will ensure that $\tilde{e}_m^M$ provides an accurate forecast of the result. Assume that the firm can expect this result from the selection capacity of the tournament and from the subsequent efforts of the selected mentor. An adequate formulation could be one of the following: (i) whatever the pair (w^M, t) chosen by the manager, the expected level of effort $\tilde{e}_m^M$ is the same as the average observed level of effort obtained in previous experiments when levels of (w^M, t) were chosen (weakly rational expectation) or (ii) whatever the pair (w^M, t) chosen by the manager, the expected level of effort $\tilde{e}_m^M$ is the expected value of the mentor's effort, given the relevant model based on the type of tournament used by the manager (strongly rational expectation).

3.2 The program of the workers and the manager

3.2.1 The program of the workers

On the assumption that the behaviour of seniors that do not participate in the transmission process and the behaviour of juniors does not impact on the strategic choice of the manager, only the training seniors' behaviour is studied.

The objective of any training senior i is to choose the level of e_i^{TS*} during the tournament and e_i^{M*} during thementorship process in order to maximize the expected utility $u_i\,(e_i)$ described by equation (1).

$$\begin{aligned}
u_i^* =& \max_{e_i^{TS}, e_i^M} u_i \\
u_i =& w_i - c_i^{TS}\left[e_i^{TS}\right] + \left(\frac{1}{1+r}\right)\left[\Pr(M)\left(w_i - c_i^M\left[e_i^M\right]\right) + (1 - \Pr(M))\,w_i\right] \\
&+ \left(\frac{1}{1+r}\right)^2 \left[\Pr(M)(w_i + w^M j\,(y_S - \bar{y}_J)\left(1 - \bar{\gamma}^2\right) e_i^M) + (1 - \Pr(M))\,w_i\right] \\
&+ \ldots + \left(\frac{1}{1+r}\right)^{\mathrm{T}-1} \left[\Pr(M)(w_i + w^M j\,(y_S - \bar{y}_J)\,(1 - \sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau) e_i^M)\right. \\
&\left. + (1 - \Pr(M))\,w_i\right]
\end{aligned} \quad (1)$$

where $\Pr(M) = p_i\left[e_i^{TS}, \alpha_i, \bar{\alpha}, S, t\right]$, $0 \leq e_i^{TS} \leq 1$ and $0 \leq e_i^M \leq 1$.

The probability of being mentor $\Pr(M)$ depends on the information held by each training senior i: his own level of effort during the tournament e_i^{TS} such that $\delta \Pr(M)/\delta e_i^{TS} > 0$; $\partial^2 \Pr(M)/\partial\left(e_i^{TS}\right)^2 < 0$, his cost of effort α_i such that $\delta \Pr(M)/\delta\alpha_i < 0$, the average cost of effort $\bar{\alpha}$ such that $\delta \Pr(M)/\delta\bar{\alpha} > 0$, the number of training seniors S such that $\delta \Pr(M)/\delta S < 0$ and the tournament duration t such that $\delta \Pr(M)/\delta t \leq 0$ if $\alpha_i \geq \bar{\alpha}$ and $\delta \Pr(M)/\delta t > 0$ otherwise.

Since the time horizon of the learning process is finite, equation (1) expresses intertemporal utility function. Each period τ (with $\tau \in [1, \mathrm{T}]$) is associated with a discount rate where r is the interest rate.

The training senior i's expected utility is composed in each period by the wage received w_i. At time 1, the tournament takes place. It is depicted by the cost of transmission effort the training senior i incurs $c_i^{TS}\left[e_i^{TS}, \alpha_i, j, S, y_S, \bar{y}_J\right]$. At time 2, the mentorship process takes place. The training senior i becomes the mentor with the probability $\Pr(M)$ and incurs in this case the cost of transmission effort $c_i^M\left[e_i^M, \alpha_i, j, S, y_S, \bar{y}_J, \bar{\gamma}^2\right]$. Otherwise, with the inverse propbability $1 - \Pr(M)$, the training senior i does not support any cost. From time 3, the mentorship process ends. The training senior i receives the bonus w^M indexed to the result of his level of effort still with the probability $\Pr(M)$ and does not receive any bonus with the inverse probability $1 - \Pr(M)$. The result of the mentor's level of effort is the additional outcome produced by juniors due to the social learning depicted by the term $j\left(y_s - \bar{y}_j\right)\left(1 - \sum \bar{\gamma}^\tau\right)\tilde{e}_m^M$.

3.2.2 The program of the manager

The ultimate objective of the manager is to choose the level of the pair (w^{M*}, t^*) in order to maximize the expected profit $\pi(w^M, t)$ with:

$$\pi^* = \sup[\pi^{M*}, \pi^r] \tag{2}$$

where π^r is the reservation profit such as:

$$\begin{aligned}
\pi^r = \; & [sy_s + j\bar{y}_j - W] + \left(\frac{1}{1+r}\right)\left[sy_s + j\bar{y}_j + j\left(y_s - \bar{y}_j\right)\bar{\gamma}^2 - W\right] \\
& + \left(\frac{1}{1+r}\right)^2 \left[sy_s + j\bar{y}_j + j\left(y_s - \bar{y}_j\right)\left(\bar{\gamma}^2 + \bar{\gamma}^3\right) - W\right] \\
& + \ldots + \left(\frac{1}{1+r}\right)^{\mathrm{T}-1}\left[s\bar{y}_s + j\bar{y}_j + j\left(y_s - \bar{y}_j\right)\left(\sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau\right) - W\right]
\end{aligned} \tag{3}$$

where $\pi^r \geq 0$ and $\pi^M\left(w^M, t\right)$ is the mentorship expected profit such that:

$$
\begin{aligned}
\pi^{M*} = \quad & \max_{w^M, t} \pi^M \\
\pi^M = \quad & s y_s + j \bar{y}_j - W - c\,[t] + \left(\frac{1}{1+r}\right) \Big[(s-1)\, y_s + j\bar{y}_j - W \\
& + j\,(y_s - \bar{y}_j)\left(\bar{\gamma}^2 + (1-\bar{\gamma}^2)\,\tilde{e}_m^M\right)\Big] + \left(\frac{1}{1+r}\right)^2 \Big[s y_s + j\bar{y}_j \\
& - W + j\,(y_s - \bar{y}_j)\left(\bar{\gamma}^2 + \bar{\gamma}^3 + (1-\bar{\gamma}^2-\bar{\gamma}^3)\tilde{e}_m^M\right) \\
& - w^M j\,(y_s - \bar{y}_j)\,(1-\bar{\gamma}^2)\tilde{e}_m^M\Big] + \ldots + \left(\frac{1}{1+r}\right)^{\mathrm{T}-1} \Big[s y_s \\
& + j\bar{y}_j - W + j\,(y_s - \bar{y}_j)\left(\sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau + \tilde{e}_m^M \left(1 - \sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau\right)\right) \\
& - w^M j\,(y_s - \bar{y}_j)\left(1 - \sum_{\tau=2}^{\mathrm{T}-1} \bar{\gamma}^\tau\right)\tilde{e}_m^M\Big]
\end{aligned} \tag{4}
$$

where $\tilde{e}_m^M = f\left[w^M, t, \bar{e}^M\right]$, $0 \leq w^M$ and $0 \leq t \leq 1$.

The reservation profit π^r is composed at each period by the wage costs W and the total production, where sy_s is the senior's production and $j\bar{y}_j$ the junior's. Individual learning becomes effective from time 2 and the additional production of workers is depicted by $j\,(y_s - \bar{y}_j)\,\bar{\gamma}^\tau$.

The mentorship expected profit π^M is described by the same terms but obviously also takes into account the benefits and costs incurred by the actions of the mentor and his designation. The term $c\,[t]$ is the cost of the tournament. Even if training seniors enhance juniors' production during the tournament, there is ultimately a cost because a part of seniors' effort is diverted from producing. This cost increases at a growing rate with the duration of the tournament such that $\delta c/\delta t > 0$ and $\partial^2 c/\partial t^2 > 0$. At time 2, the mentorship process begins. The mentor's level of effort expected by the manager, accelerates the individual learning process and increases the juniors' production. This latter outcome is described by the term $j\,(y_s - \bar{y}_j)\left(\sum \bar{\gamma}^\tau + (1 - \sum \bar{\gamma}^\tau)\,\tilde{e}_m^M\right)$. From time 3, the mentorship process ends but the manager pays a bonus to the mentor indexed to the result of his expected level of effort, $w^M j\,(y_s - \bar{y}_j)\,(1 - \sum \bar{\gamma}^\tau)\,\tilde{e}_m^M$, where w^M is the bonus multiplier.

Finally, the main difference between the two profit functions is the mentor's lack of production emphasized by the term $(s-1)\,y_s$ in the mentorship expected profit. This difference matters since it ensures that the mentorship process is not inevitably a dominating strategy. Indeed, if the optimal choice of the manager is $\left(w^{M*}, t^*\right) = (0,0)$, the reservation profit would always be greater than the expected mentorship profit since the mentor does not produce. Outside of this extreme case, the choice is not trivial and it is not possible to identify *a priori* the best strategy.

From the definition of $\tilde{e}_m^M = f[w^M, t, \bar{e}^M]$ and from the properties of the function employed, one can prove the following lemma:

Lemma 1. *Whatever the values of the parameters $\bar{\alpha}, j, s, S, y_s, \bar{y}_j, \bar{\gamma}, r, W$, there exists an upper level of w^M ensuring that the mentorship expected profit is positive.*

Proof. Given that $\lim_{w^M \to +\infty} \bar{e}^M = 1$ and $\lim_{w^M \to +\infty} f\left[w^M, t = 1, \bar{e}^M = 1\right] = 1$, then $\forall t, \lim_{w^M \to +\infty} w^M j\,(y_s - \bar{y}_j)\,\left(1 - \sum \bar{\gamma}^\tau\right)\, f\left[w^M, t, \bar{e}^M\right] = +\infty$ and therefore $\lim_{w^M \to +\infty}$ $\pi^M\left[w^M, t, \bar{e}^M\right] = -\infty$. From the continuity of $\pi^M\left[w^M, t, \bar{e}^M\right]$ in w^M, one can deduce that there exists a finite value $w^M_{\pi^M=0}[\forall t, \bar{e}^M]$ of w^M, such that $\forall w^{M\prime}, w^{M\prime} > w^M_{\pi^M=0}[t, \bar{e}^M]$, $\pi^M[w^{M\prime}, t, \bar{e}^M] < 0$ ■

3.3 The sequence of the model

Suppose that the manager and the training seniors are rational, actions can be depicted by a two stage Stackelberg game in which the manager plays leader.

Time 0: The decisions of the manager
The manager rationally expects the mentor's level of effort $\tilde{e}_m^M$ and simultaneously determines the optimal values of the bonus multiplier w^M and the tournament duration t in order to maximize the mentorship profit. If the expected mentorship profit is greater than the reservation profit, the tournament takes place. Then the manager announces to the training seniors the level of the pair $\left(w^{M*}, t^*\right)$ and that the mentor's bonus will be paid after the mentorship process, according the surplus of production observed.

Time 1: Tournament
Training seniors are made acquainted with the rules of the tournament (tournament duration, type of remuneration in case of success, level of the bonus multiplier, number of competitors, tasks to be performed). Each training senior determines his optimal level of effort during the tournament e_i^{TS*} and, if he is appointed mentor, his optimal future level of effort during the mentorship process e_i^{M*}.

Between Time 1 and Time 2: Identification of the mentor
The manager observes the amount of output associated with the S training seniors $[\sum_{k=1}^K Y_1^k, \sum_{k=1}^K Y_2^k, \ldots, \sum_{k=1}^K Y_S^k]$. The one associated with the highest level of output is assigned mentor.

Time 2: Mentorship process
The mentor provides the level of effort e^{M*} already determined in time 1 in respect to the trade off between the bonus and the cost of his effort. This level of effort may not match with that expected by the manager $\tilde{e}_m^M$.

After Time 2: Payment of the bonus
The manager pays the bonus corresponding to the observed surplus of production $w^{M*}j$ $(y_s - \bar{y}_j)(1 - \sum \bar{\gamma}^\tau)e^{M*}$ until the end of the mentorship learning process. This amount of bonus could not match with the level expected $w^{M*}j(y_s - \bar{y}_j)(1 - \sum \bar{\gamma}^\tau)\tilde{e}_m^M$.

3.4 Results

From the program defined in equation (4), the existence and the generic uniqueness of the pair $\left(w^M, t\right)$ associated with a maximal expected profit for the manager, can be derived.

Proposition 1. *Whatever the values of the parameters associated with the average transmission cost ($\bar{\alpha}$), the number of workers (j, s, S), the workers' productivity ($y_s, \bar{y}_j$), the individual learning rate ($\bar{\gamma}$), the interest rate (r) and the payroll (W), the program given by expression (2) has at least one pair of solutions* $\left(w^{M^*}, t^*\right)$.

Proof. see Appendix 1.

Proposition 2. *The maximum solution of expression (2) is generically unique.*

Proof. see Appendix 2.

4 Teamwork

4.1 Teamwork features

An alternative way to facilitate the acquisition of firm-specific competences by juniors is to encourage informal on-site learning. Informal on-site learning is practical learning acquired within the firm. For instance, in the course of production, seniors show juniors how to perform certain tasks. At the same time, juniors perform these tasks by observing seniors and asking for advice when they are unable to be successful on their own. Hence, both juniors and seniors work during these informal exchanges.

4.1.1 The role and the composition of the team

To be efficient, the informal process should be organized by the firm. Indeed, the firm can promote relationship between juniors and seniors by forming mixed teams, since interactions are more likely to occur in this context. Moreover, working in mixed teams can be an incentive for seniors to transmit knowledge if there is a reward associated with team output.

In a teamwork context, the organization and, especially, the composition of the team(s) matter. If informal learning and working were incompatible, the number of teams and the number of seniors per teams would be obvious. In this case (that is, formal learning described in the previous section), the most efficient organization would be one team associated with one senior (which amounts to mentorship).

But here, since informal learning and working are compatible, I emphasize the existence of a trade off between the positive effects associated with large teams and those associated with small teams. On the one hand, assuming that the seniors' transmission level of effort is cumulative within a team, the more seniors there are in a team the higher will be the transmission effort within the team. Moreover, a large team may incur lower cost of transmission since seniors are more available to help juniors when they experience problems in performing a task. On the other hand, it has been pointed out that the free rider problem may occur in large teams. This implies that the more seniors there are in a team, the less the individual level of effort matters and, hence, the lower is the joint level of effort.

The composition of the team is therefore crucial and needs to be optimally determined by the manager.

4.1.2 The teamwork learning process

Since teamwork is not the default organizational form of the firm, the manager would delegate to a team only seniors who are training seniors, and juniors. Such mixed teams do not impact on the joint production of the team (the production function remains separable), but facilitate the transmission of competences due to the interactions that occur among team members. The other seniors carry on producing independently.

The composition of the mixed team is assumed to be the same from one team to another. In other words, the ratio of juniors by team denoted j^T and training seniors by team denoted S^T will be the same for every team. Obviously, these ratios j^T and S^T are related to the number of teams denoted N such as $S^T = S/N$ and $j^T = j/N$. Therefore, since the number of juniors j and training seniors S are given, the manager has only to optimally determine one of these three proportions and the two others will be given[8]. Without any loss of generality, I assume that the number of training seniors by team S^T is the control variable for the manager, such as $1 \leq S^T \leq S$. Finally, seniors are assumed to be distributed randomly in each team since the manager does not know their individual abilities to transmit.

Within each team, each senior i chooses his/her optimal level of effort denoted e_i^* which is a continuous variable, such as $e_i \in [0, 1]$. As in the mentorship process, this level of effort incurs a cost $c_i[e_i, \alpha_i, j, S^T, y_s, \bar{y}_j, \bar{\gamma}^2]$, such as $c(e_i) = 0$ when $e_i \leq 0$; $c(e_i) > 0$ when $e_i > 0$ and $\bar{\alpha}_i > 0$; respecting the following properties:

$$\frac{\delta c_i}{\delta e_i} > 0; \ \frac{\partial^2 c_i}{\partial (e_i)^2} > 0; \ \frac{\delta c_i}{\delta \alpha_i} > 0; \ \frac{\delta c_i}{\delta (j/S^T)} > 0; \ \frac{\delta c_i}{\delta (y_s - \bar{y}_j)} > 0 \text{ and } \frac{\delta c_i}{\delta \bar{\gamma}^2} < 0$$

The level of effort optimally chosen by each senior i gives the optimal level of effort of the team, denoted e^T. This team level of effort has to be seen as an overall effort depending on the seniors' individual levels of effort in the team. But seniors do not know the identity of their peers and cannot observe their individual levels of effort and thus do not know precisely the utlimate team level of effort. Senior i has therefore to choose his level of effort e_i taking account of the seniors' expected average level of effort denoted $E(\bar{e})$ such as $E(\bar{e}) \in [0, 1]$ where $\bar{e}$ is the average level of effort, $\bar{e} = \frac{1}{S}\sum_{i=1}^{S} e_i$. Hence the team level of effort expected by agent i denoted $\tilde{e}_i^T$ is represented by $h_i[e_i, E(\bar{e}), S^T]$, where the function h_i satisfies the following properties:

$$\frac{\delta h_i}{\delta e_i} > 0; \ \frac{\delta h_i}{\delta E(\bar{e})} > 0 \text{ and } \frac{\delta h_i}{\delta S^T} \lesseqqgtr 0$$

This last condition means that the higher the number of seniors in a team, the higher the relevancy of the seniors' expected average level of effort $E(\bar{e})$ and the less the individual level of effort e_i matters. Thus, the number of seniors does not affect the team level of effort if $E(\bar{e}) = e_i$, and increases with the team level of effort if $E(\bar{e}) > e_i$ and decreases otherwise). Finally assume that: $\forall S^T$, $h_i[e_i = 0, E(\bar{e}) = 0] = 0$ and $h_i[e_i = 1, E(\bar{e}) = 1] = 1$. From these conditions, it can be deduced that $\forall e_i, E(\bar{e}), S^T$, $0 \leq h_i[e_i, E(\bar{e}), S^T] \leq 1$, *i.e.* that the team level of effort $\tilde{e}_i^T$ is consistent with the domain of the definition of e_i and $E(\bar{e})$.

[8] For example, if $j = 36$, $S = 12$ and if the manager chooses $S^T = 3$, then $N = 4$ and $j^T = 9$.

Seniors can transmit all of the firm-specific competences if they collectively reach the maximal team level of transmission effort. In order to incite seniors to provide such a high team level of effort, I assume that the manager uses the same remuneration scheme as in the mentorship process. Each training senior will receive a bonus indexed to the result of the team level of effort. Accordingly, the second control variable for the manager (the first being the number of training seniors by team S^T) is the bonus multiplier, $w^T \in \mathbb{R}^+$.

The optimal individual level of effort e_i^* chosen by the senior i can be determined by the manager from the program of the seniors. According to e_i^*, which depends on the control variables $\left(w^T, S^T\right)$, and according to the average ability to transmit $\bar{\alpha}$, the manager can deduce the expected average level of effort $E\left(\bar{e}\right)$, which then also depends on $\left(w^T, S^T\right)$[9]. Here, with the average level of effort $E\left(\bar{e}\right)$, the manager is able to anticipate the team level of effort $\tilde{e}_m^T$ and then to choose the bonus multiplier w^T and the number of seniors by team S^T optimally.

The expected team level of effort could be written as $\tilde{e}_m^T = g[E\left(\bar{e}\right), S^T]$, where the function g satisfies the following properties: $\delta g/\delta E\left(\bar{e}\right) > 0$ and $\delta g/\delta S^T \gtrless 0$. This last double inequality means that the team level of effort expected by the manager increases or decreases with the number of seniors per team, according to the dominating effect. If the positive (negative) effects associated with a large team dominate the negative (positive) ones, then the expected team level of effort would increase (decrease) with the number of seniors by team. The positive effects of a large team come from: i) the accumulation of the levels of effort (expressed in the function g); ii) the lower cost of transmission associated with it (expressed by inside $E\left(\bar{e}\right)$ - remembering that $E\left(\bar{e}\right)$ depends on w^T and S^T). The negative effects are related to the occurrence of free riding in large teams (expressed by $E\left(\bar{e}\right)$). I introduce additional conditions on the value of the team level of effort expected by the manager for the lower bound of S^T and on the limit value of the function g:

$$\forall E\left(\bar{e}\right), g[S^T = 1] = E\left(\bar{e}\right) \ ; \forall S^T, g[E\left(\bar{e}\right) = 0] = 0 \ ; \forall S^T, g[E\left(\bar{e}\right) = 1] = 1$$

From these conditions, it can be deduced that $\forall E\left(\bar{e}\right)$ and S^T, $0 \leq g\left[E\left(\bar{e}\right), S^T\right] \leq 1$, *i.e.* that the expected effort $\tilde{e}_m^T$ is consistent with the domain of definition of e_i, $E\left(\bar{e}\right)$ and $\tilde{e}_i^T$.

4.2 The program of the workers and the manager

In order to enable comparison with the mentorship process, the time horizon of the teamwork process is the same and the juniors' learning occurs at the same time, *i.e.* time 2.

4.2.1 The program of the workers

Within each team, the objective of any training senior i is to choose the level of e_i^* in order to maximize the expected utility $u_i\left(e_i\right)$ by equation (5).

[9]Note that it makes a difference with the mentorship process where the uncertainty about the tournament involves that the average mentor's level of effort does not depend on the two control variables but only on w^M.

$$
\begin{aligned}
u_i^* = & \max_{e_i} u_i \\
u_i = & \ w_i + \left(\frac{1}{1+r}\right) \left[w_i - c_i \left[e_i, \alpha_i, j, S^T, y_s, \bar{y}_j, \bar{\gamma}^2\right]\right] \\
& + \ldots + \left(\frac{1}{1+r}\right)^{\mathrm{T}-1} \left[w_i + w^T j \left(y_s - \bar{y}_j\right) \left(1 - \sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau\right) \tilde{e}_i^T\right]
\end{aligned} \tag{5}
$$

with $\tilde{e}_i^T = h_i \left[e_i, E\left(\bar{e}\right), S^T\right]$ and $0 \leq e_i \leq 1$.

The senior i's expected utility is composed in each period by the wage received w_i. At time 2, teams are formed and the teamwork process takes place. It is depicted by the cost of transmission effort the senior i incurs $c_i \left[e_i, \alpha_i, j, S^T, y_s, \bar{y}_j, \bar{\gamma}^2\right]$. From time 3, the senior i receives the bonus indexed to the result of the team level of effort he expects $\tilde{e}_i^T$.

4.2.2 The program of the manager

The objective of the manager is to choose the level of pair $\left(w^{T*}, S^{T*}\right)$ in order to maximize the expected profit $\pi\left(w^T, S^T\right)$ with:

$$
\pi^* = \sup\left[\pi^T, \pi^r\right] \tag{6}
$$

where π^r is the reservation profit depicted in equation (3) and $\pi^T\left(w^T, S^T\right)$ is the teamwork expected profit depicted in equation (7)

$$
\begin{aligned}
\pi^{T*} = & \max_{w^T, S^T} \pi^T \\
\pi^T = & \ s y_s + j \bar{y}_j - W + \left(\frac{1}{1+r}\right) \left[s y_s + j \bar{y}_j - W\right. \\
& \left. + j\left(y_s - \bar{y}_j\right)\left(\bar{\gamma}^2 + \tilde{e}_m^T (1 - \bar{\gamma}^2)\right)\right] + \left(\frac{1}{1+r}\right)^2 \left[s y_s + j \bar{y}_j\right. \\
& - W + j\left(y_s - \bar{y}_j\right)\left(\bar{\gamma}^2 + \bar{\gamma}^3 + \tilde{e}_m^T \left(1 - \bar{\gamma}^2 - \bar{\gamma}^3\right)\right) \\
& \left. - S w^T j\left(y_s - \bar{y}_j\right)(1 - \bar{\gamma}^2)\tilde{e}_m^T\right] + \ldots + \left(\frac{1}{1+r}\right)^{\mathrm{T}-1} \left[s y_s\right. \\
& + j \bar{y}_j - W + j\left(y_s - \bar{y}_j\right)\left(\sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau + \tilde{e}_m^T \left(1 - \sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^\tau\right)\right) \\
& \left. - S w^T j\left(y_s - \bar{y}_j\right)\left(1 - \sum_{\tau=2}^{\mathrm{T}-1} \bar{\gamma}^\tau\right) \tilde{e}_m^T\right]
\end{aligned} \tag{7}
$$

where $\tilde{e}_m^T = g\left[E\left(\bar{e}\right), S^T\right]$, $0 \leq w^T$ and $1 \leq S^T \leq S$.

The teamwork expected profit π^T is described by the same terms as the reservation profit, but also takes account of the benefits and costs arising from the training seniors' decisions. At time 2, the manager designates the mixed teams and the teamwork process takes place. The team level of effort expected by the manager accelerates the individual learning process and thus increases the juniors' production. This latter outcome is depicted by the term $j\left(y_s - \bar{y}_j\right)\left(\sum \bar{\gamma}^\tau + \left(1 - \sum \bar{\gamma}^\tau\right)\tilde{e}_m^T\right)$. From time 3, the manager pays a bonus to the S training seniors indexed to the result of the expected team level of effort, such as $Sw^T j\left(y_s - \bar{y}_j\right)\left(1 - \sum \bar{\gamma}^\tau\right)\tilde{e}_m^T$.

Finally, given the definition of the two profit functions π^r and $\pi^T\left(w^T, S^T\right)$, the teamwork expected profit amounts to the reservation profit if the optimal value of the bonus multiplier is $w^{T^*} = 0$. Indeed, this value of w^{T^*} would discourage seniors from training juniors. Outside this extreme case, the teamwork process is always a dominating strategy. This would seem to be consistent with an informal learning process which allows workers to acquire and transmit some competences without it stopping them from working.

From the definition of $E\left(\bar{e}\right)$ and $\tilde{e}_m^T$ and from the properties of the utility function and the teamwork expected profit, one can prove the following lemma:

Lemma 2. *Whatever the values of the parameters $\bar{\alpha}, j, s, S, y_s, \bar{y}_j, \bar{\gamma}^\tau, r, W$, there exists an upper level of w^T ensuring that the teamwork expected profit is positive.*

Proof. Given that $\lim_{w^T \to +\infty} E\left(\bar{e}\right) = 1$, implying $\forall S^T, \lim_{w^T \to +\infty} g\left[E\left(\bar{e}\right), S^T\right] = 1$, and that $\forall S^T, \lim_{w^T \to +\infty}\left(Sw^T j\left(y_s - \bar{y}_j\right)\left(1 - \sum \bar{\gamma}^\tau\right)\tilde{e}_m^T\right) = +\infty$, then $\lim_{w^T \to +\infty} \pi^T\left[w^T, S^T\right] = -\infty$. From the continuity of $\pi^T\left[w^T, S^T\right]$ in w^T, it can be concluded that there is a finite value $w^T_{\pi^T=0}[\forall S^T]$ of w^T, such that $\forall w^{T\prime}, w^{T\prime} > w^T_{\pi^T=0}[S^T]$, $\pi^T[w^{T\prime}, S^T] < 0$ ■

4.3 The sequence of the model

As in the mentorship process, actions can be depicted by a two stage Stackelberg game in which the manager plays leader.

Time 1: The decisions of the firm

The manager expects a certain team level of effort according to which he chooses the level of the bonus multiplier w^{T^*} and the number of training seniors by teams S^{T^*} (and the corresponding number of teams N) and announces them publicly.

Time 2: Teamwork and the decisions of the training seniors

The manager designates the mixed teams and the teamwork process takes place. Within each team, each training senior chooses his optimal level of effort e_i^* and trains the juniors in his/her team.

After Time 2: Payment of the bonus

The manager pays the bonus corresponding to the surplus of production observed to the S training seniors until the end of the teamwork learning process.

4.4 Equilibrium conditions and properties

4.4.1 Equilibrium concepts and conditions

The equilibrium concepts and conditions can be considered separately for the two steps of the game.

i) At step 2, the level of bonus w^T and the number of seniors per teams S^T are known, and each training senior has to determine his/her level of effort, given the expected average level of effort of all training seniors. When confronted and determined by the same expected average level of effort $E(\bar{e})$, the S optimal individual levels of effort of the training seniors provide the effective average level. This process can be illustrated as follows:

$$E(\bar{e}) \begin{array}{c} \nearrow \\ \nearrow \\ \cdots \\ \searrow \end{array} \left. \begin{array}{c} e_1^*\left(E(\bar{e}), w^T, S^T\right) \\ e_2^*\left(E(\bar{e}), w^T, S^T\right) \\ \\ \\ e_S^*\left(E(\bar{e}), w^T, S^T\right) \end{array} \right\} \rightarrow \bar{e}\left(E(\bar{e}), w^T, S^T\right)$$

The equilibrium of the second step is then the average level of effort of the training seniors $\bar{e}^*$, such that all training seniors should expect this average level, *i.e.* $E(\bar{e}) = \bar{e}^*$, then it is effectively observed across the teams as resulting in the optimal individual decisions of the training seniors.

ii) At step 1, the manager chooses the optimal level of bonus w^{T^*} and the optimal number of training seniors per team S^{T^*}, in order to maximize the expected profit, using his or her own expectation of the average level of effort of the training seniors.

An equilibrium of the full game is then a triplet $\left(\bar{e}^*, w^{T^*}, S^{T^*}\right)$ such that:

- $\bar{e}^*$ is an equilibrium of the second step sub-game when the bonus is w^{T^*} and the number of seniors per team is S^{T^*};

- $\left(w^{T^*}, S^{T^*}\right)$ is the optimal choice of the manager at the first step sub-game when he expects that the average training effort will be $\bar{e}^*$.

4.4.2 Results

The general properties of the teamwork training can be considered by the following formulation of the results:

Lemma 3. *Whatever the values of the first step variables w^T and S^T and those of the parameters $\bar{\alpha}, j, s, S, y_s, \bar{y}_j, \bar{\gamma}^\tau, r, W$, at least one equilibrium exists $\bar{e}^*$ for the second step sub-game.*

Proof. see Appendix 3.

Lemma 4. *According to the form of the function $h_i[e_i, E(\bar{e}), S^T]$, the equilibrium of the second step sub-game can be unique or not.*

Proof. see Appendix 4.

The consequence of Lemma 4 is that, according to the content of the individual function $h_i[e_i, E(\bar{e}), S^T]$, the possibility that coordination failures exist among the training seniors can not be excluded. They may occur when $\bar{e}$ is non-decreasing in $E(\bar{e})$.

Proposition 3. *If the second step solution is unique, whatever the values of the parameters associated with the average transmission cost ($\bar{\alpha}$), the number of workers (j, s, S),* *the workers' productivity* $(y_s, \bar{y}_j)$*, the individual learning rate* $(\bar{\gamma})$*, the interest rate* (r) *and the payroll* (W)*, there is always at least one solution* $\left(\bar{e}^*, w^{T^*}, S^{T^*}\right)$ *to the full game.*

Proof. see Appendix 5.

Proposition 4. *If the second step solution is unique, the maximum solution of the full game is also unique.*

Proof. see Appendix 6.

5 Mentorship vs Teamwork

5.1 A specification of the model

The expected profit functions presented above can be specified in order to enable comparison. I propose to use particular functions consistent with the derivative properties and limit value conditions defined above.

For the Mentorship learning process, consider the following specified functions: the senior i's cost of effort would be $c_i^{TS}[e_i^{TS}, \alpha_i, j, S, y_s, \bar{y}_j] = j(y_s - \bar{y}_j)\alpha_i(e_i^{TS})^2/2S$ during the tournament and $c_i^M[e_i^M, \alpha_i, j, y_s, \bar{y}_j, \bar{\gamma}^2] = j(y_s - \bar{y}_j)(1-\bar{\gamma}^2)\alpha_i(e_i^M)^2/2$ during the mentorship process. The tournament cost can be expressed as $c[t] = t^2/2$ and the expected mentor's level of effort as $f[w^M, t, \bar{e}^M] = \inf[1, w^M t + \bar{e}^M(t(1-\varepsilon-t)+\varepsilon)]$[10] where ε is a constant such as $0 < \varepsilon < 1$.

For Teamwork, let the senior i's cost of effort be described by the same type of cost function $c_i[e_i, \alpha_i, j, S^T, y_s, \bar{y}_j, \bar{\gamma}^2] = j(y_s - \bar{y}_j)(1-\bar{\gamma}^2)\alpha_i(e_i)^2/2S^T$, the team level of effort expected by the training senior i be $h_i[e_i, E(\bar{e}), S^T] = ((S - S^T + 1)e_i + (S^T - 1)E(\bar{e}))/S$ and the one expected by the manager be $g[E(\bar{e}), S^T] = E(\bar{e}) + (S^T - 1)(1 - E(\bar{e}))E(\bar{e})/S$.

Finally, consider that the time horizon of each learning process represents four periods, such as T = 4.

Based on the above specifications, it becomes possible to solve analytically the program of the manager when he decides to implement mentorship or teamwork. But, the analytical solution to profit maximization is rather complicated to calculate and express with general parameters, in the case of teamwork. Indeed, the complex[11] formulation of $g\left[E(\bar{e}), S^T\right]$ depending on the two control variables w^T and S^T and its combination with w^T lead to the resolution of a high degree polynomial with discontinuity outside the set of definitions of w^T and S^T. The resolution ultimately yields both complex and real solutions, but with only one maximum corresponding to the set of definitions of w^T and S^T.

[10]Note that the second right hand term of the specific function f means that the average level of effort $\bar{e}^M$ matters especially for the manager's expectation when the duration tournament t is not maximal. Indeed, a maximal level of t ($t = 1$) increases the chances of the most able senior to be mentor, whatever the value of $\bar{e}^M$. On the other hand, if t is not maximal, the tournament outcome becomes uncertain and the value of $\bar{e}^M$ matters: a high (low) level of $\bar{e}^M$ increases (decreases) the chances that the mentor's level effort will be high.

[11]Note that the formulation of g seems to be the simplest according to the derivative properties and limit value conditions.

This complexity of the analytical expression of the solution (w^{T^*}, S^{T^*}) prevents general description and discussion. In any case, such a discussion is not relevant to the general parameters since it is not intuitive to conclude whether an optimum strategy exists, or to discuss in which cases one learning process is preferred over another. However, this could be done through numerical calculation. The study of numerical examples presented below tends to illustrate the most relevant situations.

5.2 Numerical comparison

The parameters of the model presented above could be divided into two categories. The first type of parameters $(W, \varepsilon, r, s, S)$ has an impact only on the value of firm profit. The second set of parameters has an impact on firm profit and also on the decision to implement a social learning process, since they indicate the "training supply" $(\bar{\alpha})$ and "training demand" $(j, y_s, \bar{y}_j, \bar{\gamma}^2, \bar{\gamma}^3, \bar{\gamma}^4)$ of the firm. More precisely, this training demand which can be understood as the need for training within the firm would be high if: (i) firm tenure heterogeneity $(y_s - \bar{y}_j)$ is high; (ii) number of new hires (j) is high; or (iii) the individual learning process is incomplete such that $\sum_{\tau=2}^{\mathrm{T}} \bar{\gamma}^{\tau} < 1$.

Consider, for instance, the following parameters: $W = 50\,;\ \varepsilon = 0.1\,;\ r = 0.01\,;\ s = 80\,;\ S = 12$ and $\bar{\alpha} = 5\,;\ j = 20\,; \bar{y}_s = 0.65\,;\ \bar{y}_j = 0.4\,;\ \bar{\gamma}^2 = 0.4\,;\ \bar{\gamma}^3 = 0.35\,;\ \bar{\gamma}^4 = 0.2$.
In the figures 1 to 5, I maximize firm profit for each social learning process within the parameters described above and note in a table which learning process is associated with the higher profit, when one parameter, linked to the training supply or demand, varies, the others remaining the same. The plots associated describe the mentorship profit (the blue curve) and the teamwork profit (the purple one) according to the value of the parameter studied. The dashed line depicts the value of the paramater from which one learning process is prefered to another.

All things being equal	$j = 20\,;\ y_s = 0.65\,;\ \bar{y}_j = 0.4\,;$ $\bar{\gamma}^2 = 0.4\,;\ \bar{\gamma}^3 = 0.35\,;\ \bar{\gamma}^4 = 0.2$
$2.89 \leq \bar{\alpha} \leq 5$	Teamwork
$\bar{\alpha} < 2.89$ or $\bar{\alpha} > 5$	Mentorship

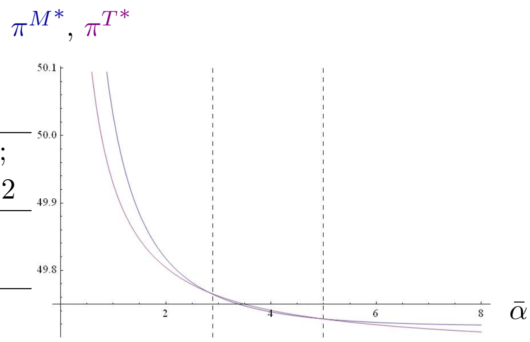


Figure 1: Social learning process associated with the highest profit according to the value of the average cost of transmission effort $\bar{\alpha}$

Figure 1 shows that mentorship is the best strategy when the value of the average cost of transmission effort is rather high or rather low, *i.e.* when the training supply is not intermediate. On the one hand, a very high value of the average cost of transmission effort means a low value of the team level of effort in the teamwork process. Indeed, the less that the training seniors are able to transmit on average, the lower are the individual and team levels of effort and, hence, the lower is the teamwork profit. In contrast, a high value of the average cost of transmission effort does not damage the mentorship profits if the firm invests sufficiently in the tournament to identify the best trainer. On the other hand, a very low value of the average cost of transmission effort means that training seniors transmit efficiently their competences in average. This leads the firm to invest in the tournament to determine the best mentor who is able to transmit all the firm specific competences.

All things being equal	$y_s = 0.65\,;\ \bar{y}_j = 0.4\,;\ \bar{\alpha} = 5\,;$ $\bar{\gamma}^2 = 0.4\,;\ \bar{\gamma}^3 = 0.35\,;\ \bar{\gamma}^4 = 0.2$
$j \leq 20$	Teamwork
$j > 20$	Mentorship

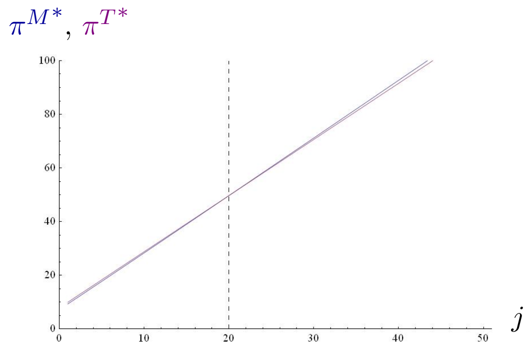


Figure 2: Social learning process associated with the highest profit according to the number of juniors j

All things being equal	$j = 20\,;\ y_s = 0.65\,;\ \bar{\alpha} = 5\,;$ $\bar{\gamma}^2 = 0.4\,;\ \bar{\gamma}^3 = 0.35\,;\ \bar{\gamma}^4 = 0.2$
$0.4 \leq \bar{y}_j \leq 0.65$	Teamwork
$\bar{y}_j < 0.4$	Mentorship

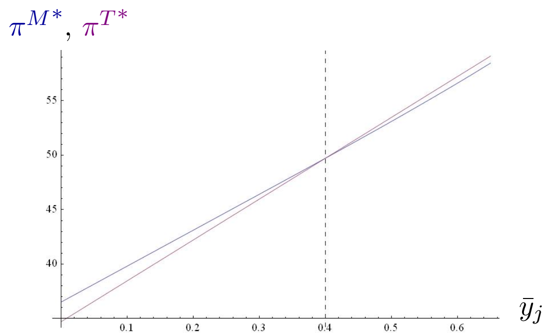


Figure 3: Social learning process associated with the highest profit according to the value of the average productivity of juniors $\bar{y}_j$

All things being equal	$j = 20\,;\ y_s = 0.65\,;\ \bar{y}_j = 0.4\,;$ $\bar{\alpha} = 5\,;\ \bar{\gamma}^3 = 0.35\,;\ \bar{\gamma}^4 = 0.2$
$0.4 \leq \bar{\gamma}^2 \leq 0.45$	Teamwork
$0.35 \leq \bar{\gamma}^2 < 0.4$	Mentorship

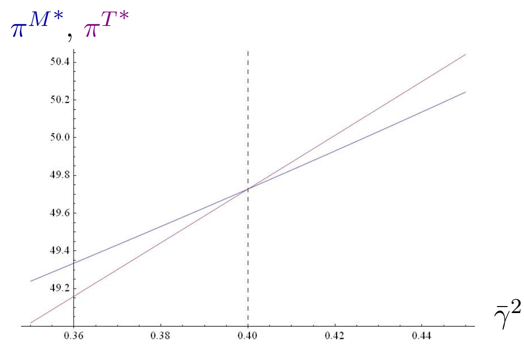


Figure 4: Social learning process associated with the highest profit according to the value of the average individual learning rate at time 2 $\bar{\gamma}^2$

Figures 2 to 5 show that mentorship is the dominated strategy when the training demand is low. In such cases, for instance if there are few new hires (Figure 2), or if there

All things being equal	$j = 20\,;\ y_s = 0.65\,;\ \bar{y}_j = 0.4\,;$ $\bar{\alpha} = 5\,;\ \bar{\gamma}^2 = 0.4\,;\ \bar{\gamma}^4 = 0.2$
$0.35 \leq \bar{\gamma}^3 \leq 0.4$	Teamwork
$0.2 \leq \bar{\gamma}^3 < 0.35$	Mentorship

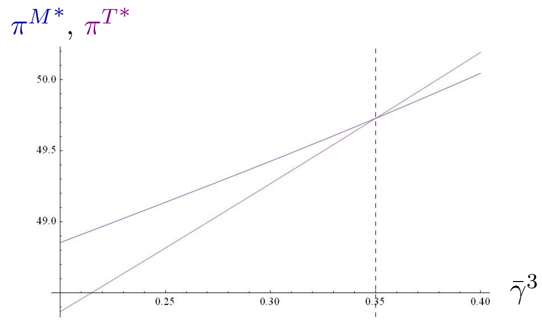


Figure 5: Social learning process associated with the highest profit according to the value of the average individual learning rate at time 3 $\bar{\gamma}^3$

are only a few firm-specific competences to acquire (Figure 3), or if the individual learning process is quite complete (Figures 4 and 5), the benefits associated with competence transmission are low in relation to the costs incurred by the determination of the mentor and his/her lack of production. As a consequence, encouraging informal training such as a teamwork process, would be more appropriate.

Finally, all these tables emphasize lack of superiority of one learning process over another.

6 Concluding Remarks

This paper has attempted to analyse the learning processes likely to occur inside firms after new workers have been hired. Even if new hires could acquire some firm-specific competences on their own, some could not be achieved without the know-how of experienced workers. This paper has studied how optimally to accelerate and complete this individual learning process and suggests that managers should use at least two strategies for the transmission of competences from experienced workers to new hires: mentorship and teamwork.

The general model in this paper shows the existence of an optimal solution for each social learning process, which is unique in the mentorship case and may be unique in the teamwork case.

Since these two learning processes involve different costs and benefits, the conditions under which one learning process might be preferred over another have been discussed. The main conclusion from the numerical observations emphasizes that neither learning process is superior. The mentorship learning process would be the best strategy if : i) there is a need for training within the firm, emphasized by characteristics such as firm tenure heterogeneity, number of new hires to train, individual learning efficiency; ii) experienced workers, on average, have either very low or very high ability to transmit. Under these conditions, the manager would invest massively in the tournament and provide to the mentor a motivating reward. If one of these two conditions is not satisfied, forming mixed teams would be a suitable organization of workers. In other words, an informal learning process seems to appropriate in most cases. These findings are consistent with a French study (Courault, Bourlier, and Trouvé 2004) on seniors and the competence transmission. These authors emphasized that workers are more often casual trainers (70%)

than official mentors (17%).

Finally, the paper highlights the importance of identifying the best mentor among insiders (since the manager has to invest massively in the tournament when the formal learning process is chosen). This conclusion is consistent with that of Tonidandel et al. (2007), which suggests that “more emphasis should be placed on the capabilities of the mentor” in a paper which analyses the impact of mentoring on juniors’ performance. Since identifying the mentor seems to be the cornerstone of any formal learning process, the method used to do so perhaps needs more research. In this paper, the mentor is identified through a tournament involving insiders. But, this tournament has not been modelled and is only represented technically by a tournament cost. Perhaps, opening the black box of the tournament achieved in this paper, then analysing other means of designation, and making a comparison could be an interesting way of research.

Appendix

Appendix 1: Existence of at least one pair of solution $\left(w^{M^*}, t^*\right)$

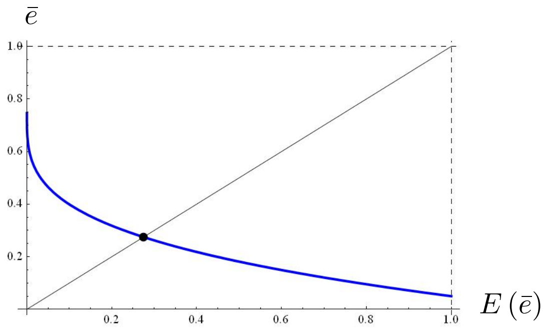
Consider the program given by equation (2) to (4) and the definition of $\tilde{e}_m^M$. Given these equations, expression of the profit $\pi\left(w^M, t\right)$ is continuous in the two control variables $\left(w^M, t\right)$, the set of definitions of which being the following compact subset of $\mathbb{R}^+$, $\left\{\left(0, \sup_{t\in[0,1]} w^M_{\pi^M=0}[t]\right), (0,1)\right\}$. The function $\pi\left(w^M, t\right)$ then reaches at least a maximum and a minimum value in this subset.

Appendix 2: Uniqueness of the solution $\left(w^{M^*}, t^*\right)$

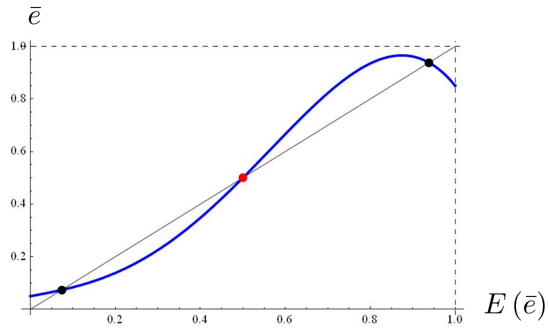
Given the continuity of the function $\pi(w^M, t)$ in the two control variables, the uniqueness of (w^{M^*}, t^*) is a consequence of the form of the function. When $\sup[\pi^{M*}, \pi^r] = \pi^r$, the solution is always unique and equal to $(0, 0)$. When $\sup[\pi^{M*}, \pi^r] = \pi^{M*}$, given the definition of $\pi^M(w^M, t)$ and $\tilde{e}_m^M$, $\pi^M(w^M, t)$ cannot be bi-linear. As a consequence, if there is a multiple solutions, these solutions can only be isolated pairs (w^{M^*}, t^*), $(w^{M'^*}, t'^*)\ldots$ such that $\pi^M(w^{M^*}, t^*) = \pi^M(w^{M'^*}, t'^*) = \ldots$ Even if there are several local maxima in $\pi^M\left(w^M, t\right)$ inside the compact $\left\{\left(0, \sup_{t\in[0,1]} w^M_{\pi^M=0}[t]\right), (0,1)\right\}$, those local maxima will have few chances of providing the same value for $\pi^M\left(w^M, t\right)$.

Appendix 3: Existence of at least one solution $(\bar{e}^*)$

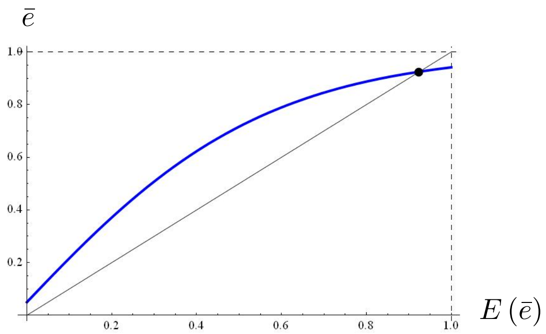
Consider the program given by equation (5) and the definition of $\tilde{e}_i^T$. Given those equations, the expression of the utility $u_i(e_i)$ is continuous in the control variable e_i, the set of definitions of which being the simplex $[0, 1]$. The function $u_i(e_i)$ then reaches at least a maximum and a minimum value in this subset. Hence, an optimal level of effort exists, whatever the senior and the level of the first step variables. Consider, then, the equilibrium condition. The function $\tilde{e}_i^T = h_i\left[e_i, E(\bar{e}), S^T\right]$ being continuous and derivable in $E(\bar{e})$ on the interval $[0, 1]$, the optimal solution $e_i^* = \varphi_{\alpha_i}(E(\bar{e}))$ of u_i is a continuous function of $E(\bar{e})$ depending on the parameters of the model and on the first step variables $\left(w^T, S^T\right)$. The effective average level of effort of the training seniors is then also a continuous function of $E(\bar{e})$ having the form $\bar{e} = \varphi_{\bar{\alpha}}(E(\bar{e}))$. An equilibrium can be considered



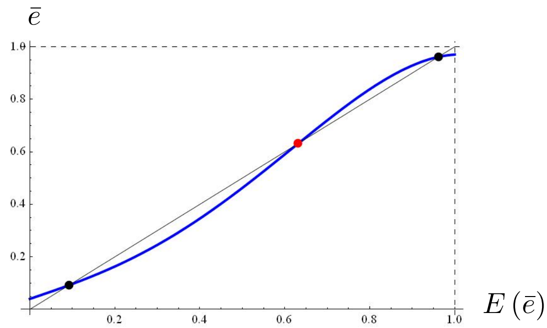
Case with 1 equilibrium when $\bar{e}$ is a decreasing function of $E(\bar{e})$



Case with 3 equilibria when $\bar{e}$ is a non monotonic function of $E(\bar{e})$



Case with 1 equilibrium when $\bar{e}$ is an increasing function of $E(\bar{e})$



Case with 3 equilibria when $\bar{e}$ is an increasing function of $E(\bar{e})$

Figure 6: Examples of unique or multiple equilibria

as a fixed point in the transformation $\bar{e} = \varphi_{\bar{\alpha}}(\bar{e})$. Since $\varphi_{\bar{\alpha}}(\bar{e})$ is continuous and maps from the compact $[0,1]$ into itself, at least one fixed point is associated to $\varphi_{\bar{\alpha}}(\bar{e})$ inside the interval $[0,1]$, as an usual application of Brouwer Theorem.

Appendix 4: Number of equilibria of the second step sub-game

Given the assumptions on the form of $h_i\left[e_i, E(\bar{e}), S^T\right]$, it is not excluded that $e_i^* = \varphi_{\alpha_i}(E(\bar{e})), \forall i$ and $\bar{e} = \varphi_{\bar{\alpha}}(E(\bar{e}))$ could be increasing or non-monotonic functions of $E(\bar{e})$ on the interval $[0,1]$. In these cases the occurrence of multiple equilibria is possible. Four possibilities, among others, are illustrated in Figure 6, the first corresponding to the case where $\bar{e}$ is a decreasing function of $E(\bar{e})$, the second to the case where $\bar{e}$ is a non-monotonic function of $E(\bar{e})$ and the last two corresponding to the case where $\bar{e}$ is an increasing of $E(\bar{e})$.

Appendix 5: Existence of an equilibrium for the full game (w^{T*}, S^{T*})

When the second step full game is unique, one single average level of effort $\bar{e}$ exists, associated to each pair $\left(w^T, S^T\right)$ chosen by the firm. From the properties of $g\left[E(\bar{e}), S^T\right]$ and Lemma 1, one can verify that the profit function $\pi^* = \sup\left[\pi^T, \pi^r\right]$ is continuous in $\left(w^T, S^T\right)$ and that those variables are defined in the program given by equation (6) to

(7) and the definition of $E(\bar{e})$ and $\tilde{e}_m^T$. Given these equations, the expression of the profit $\pi^T\left(w^T, S^T\right)$ is continuous in the two control variables $\left(w^T, S^T\right)$, the set of definition of which being the compact subset of R^2, $\left\{\left(0, \sup_{S^T\in[1,S]} w^T_{\pi^T=0}[S^T]\right), (1, S)\right\}$. The function $\pi^T\left(w^T, S^T\right)$ then reaches at least a maximum and a minimum value in this subset.

Appendix 6: Uniqueness of the solution (w^{T^*}, S^{T^*})

Similarly to the proof on the uniqueness of the solution in the mentorship process, the uniqueness of (w^{T^*}, S^{T^*}) can be demonstrated, given the continuity of the function $\pi^T(w^T, S^T)$ in the two control variables and the definition of $\pi^T(w^T, S^T)$ and $\tilde{e}_T$. Once again, if there are multiple solutions, those solutions can only be isolated pairs (w^{T^*}, S^{T^*}), $(w^{T'^*}, S^{T'^*})\ldots$ such that $\pi^T(w^{T^*}, S^{T^*}) = \pi^T(w^{T'^*}, S^{T'^*}) = \ldots$ Even if there are several local maxima in $\pi^T(w^T, S^T)$ inside the compact $\left\{\left(0, \sup_{S^T\in[1,S]} w^T_{\pi^T=0}[S^T]\right), (1, S)\right\}$, these local maxima will have few chances of providing the same value to $\pi^T(w^T, S^T)$.

References

Arai M., Billot A. and Lafranchi J. (2001). "Learning by helping: a bounded rationality model of mentoring", *Journal of Economic Behavior and Organization*, vol. 45, n°2, pp. 113-131.

Athey S., Avery C.N. and Zemsky P. (2000). "Mentoring and diversity", *American Economic Review*, vol. 90, n°4, pp. 765-786.

Bishop J.H. (1997). "What we know about employer-provided training: a review of the literature", *Research in Labor Economics*, vol. 16, pp. 19-87.

Cohendet P. and Steinmueller E.W. (2000). "The codification of knowledge: a conceptual and empirical exploration", *Industrial and Corporate Change*, vol. 9, n°2, pp. 195-209.

Courault B., Bourlier E. and Trouvé P. (2004). "Les seniors et les transferts de compétences dans les TPE et PME d'Auvergne : un état des lieux", *Rapport de recherche du CEE*, n°14.

Ellison G. and Fudenberg D. (1993). "Rules of thumb for social learning", *Journal of Political Economy*, vol. 101, n°4, pp. 612-643.

Gale D. (1996). "What have we learned from social learning?", *European Economic Review*, vol. 40, n°3, pp. 617-628.

Garicano L. (2000). "Hierarchies and the organization of knowledge in production", *Journal of Political Economy*, vol. 108, n°5, pp. 874-904.

Garicano L. and Hubbard T.N. (2005). "Hierarchical sorting and learning costs: Theory and evidence from the law", *Journal of Economic Behavior and Organization*, vol. 58, n°2, pp. 349-369.

Hamilton B.H., Nickerson J. and Owan H. (2003). "Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation", *Journal of Political Economy*, vol. 111, n°3, pp. 465-497.

Hamilton B.H., Nickerson J. and Owan H. (2004). "Diversity and productivity in production teams", Washington University Olin School of Business, Mimeo.

Hunt D.M. and Michael C. (1983). "Mentorship: A career training and development tool", *Academy of Management Review*, vol. 8, n°3, pp. 475-485.

Kandel E. and Lazear E.P. (1992). "Peer pressure and partnerships", *Journal of Political Economy*, vol. 100, n°4, pp. 801-817.

Kram K.E. (1983). "Phases of the mentor relationship", *Academy of Management Journal*, vol. 26, n°4, pp. 608-625.

Pedler M., Burgoyne J. and Boydell T. (1991). *The Learning Company, a Strategy for Sustainable Development*, London: McGraw-Hill.

Senge P.M. (1990). *The Fifth Discipline: The Art and Practice of the Learning Organization*, New York: Doubleday/Currency.

Tonidandel S., Avery D.R. and Phillips M.G. (2007). "Maximizing Returns on Mentoring: Factors Affecting Subsequent Protégé Performance", *Journal of Organizational Behavior*, vol. 28, n°1, pp. 89-110.