On the joint production of research and training*

António Freitas[†]and Inés Macho-Stadler[‡]

Universitat Autònoma de Barcelona

February 9, 2011

Abstract

Universities and research institutions have the responsibility to produce science and to provide training to new generations of researchers. In this paper, we propose a model to analyze the determinants of a senior scientist's decisions about allocating time between these tasks. The results of this decision depend upon the characteristics of the research project, the senior scientist's concern for training and the expected innate ability of the junior scientist involved. We analyze the role that a regulator can play in defining both the value of scientific projects and the future population of independent scientists.

Keywords: Allocation of time between tasks; research and training; senior and junior scientists

*We are grateful to David Pérez-Castrillo, Pau Olivella and Pedro Rey-Biel for their insightful comments. We would also like to thank the participants in the presentations at Jornadas de Economía Industrial (Madrid, 2010), Universitat Autònoma de Barcelona, Universidade do Porto and Universitat Rovira i Virgili. We would like to thank ECO2009-7616, Consolider-Ingenio-CSD2006-16, 2009SGR-169, and Fundação Ciência e Tecnologia (Grant SFRH/BD/40182/2007) for financial support. Inés Macho-Stadler is a research fellow of MOVE (Markets, Organizations and Votes in Economics) and Barcelona Economics.

†Department of Economics. Universitat Autonoma de Barcelona. E-mail: antonioluis.cardoso@uab.es

‡Corresponding author. Department of Economics. Universitat Autonoma de Barcelona. Fac CCEE Edifici B. 08193 Bellaterra (Barcelona). Spain. Phone: 34 93 5811812. Fax: 34 93 5813767. E-mail: ines.macho@uab.es

1 Introduction

It is widely accepted that universities and research institutions have the responsibility to produce science. However, there is another task of great importance to our society's advancement of knowledge: training the new generations of researchers. In this paper, we consider senior scientists to be involved both in doing research and in providing training to junior scientists, as in a system of apprenticeship. Understanding the allocation of time among the two activities is of great interest because the training of junior researchers needs to be performed by the people who know how to do research, and this is crucial in assuring a high-quality research workforce for the future. However, there are voices that point out that our research institutions may be failing in this dimension, meaning that there is a shortage of time devoted to training scientists able to perform outstanding independent research in the future.[1] In this paper, we propose a model to address this problem, to discuss the allocation of time between research and training the next generation of researchers and to discuss the problems that may arise.

Our motivation comes from two facts. On the one hand, it is a documented fact that the most prominent candidates who attain an independent research status are PhDs and postdoctoral researchers who, thrive through more experience and skills in science, either in academics or in industry (Cech and Bond, 2004). The literature on higher education, human resource management and mentoring extensively recognizes the effects of training by senior staff in the competence, productivity, career development and independent skills of young professionals, both in industry and academia. The *student-supervisor* relationship is the most critical issue affecting the quality of the PhD training (which affects both eventual job placement and success in obtaining a degree). In this process, it is natural that doctoral students hold expectations with respect to the role of their supervisors. The most important expectations are guidance in the early days of obtaining a PhD, knowledge about the area they are working in, and most importantly involvement with their work (Pole *et al.*, 1997).[2] On a postdoctoral level, Vogel (1999) reports the experience

[1] Obviously, an alternative to training one's own researchers is to attract researchers trained elsewere. While this is an interesting idea, we choose to ignore this topic in this paper.

[2] Murray (2004) illustrates this issue on a study on academic scientists working for the industry in the biosciences. The author identifies one of the sources of social capital to be the scientist's laboratory

of a principal investigator (PI) supervising postdocs in an internationally appraised lab.[3] She states that the PI's key to producing successful and high-quality junior scientists is to provide them with original ideas and orientation, to encourage strong participation in the projects, and to listen to them to assess their skills, motivations and ambitions.

On the other hand, training problems persist on a global scale. Student doctoral attrition remains a common problem in PhD training, and this is estimated at approximately 50% on the U.S. (Lovitts, 2001). In a case study about former students who spent at least two years in a PhD program Golde (2000) identifies a lack of support and guidance from supervisors as one of the causes of attrition. The author is also able to identify characteristics for good supervision: the amount of time spent, the quality of interactions between student and supervisor, and an interest in the student's work are important to guarantee training success. Accessibility seems to be an important issue as well.[4] Training problems also occur at the postdoctoral level. Puljak (2006) reports that the most common complaint in postdoctoral training is ironically, a the lack of postdoctoral training. Postdocs join a research lab and, shortly thereafter, many realize that they are on their own. It has also been identified that some advisors tend to take over the design of experiments, making postdocs feel like they are overeducated technicians.[5]

We study this issue by building a multitask model that examines the incentives of a senior scientist to provide training to a junior scientist. The senior scientist chooses the time to allocate to her own research and the time to train to the junior scientist under her supervision. We then evaluate the impact of time allocation on the level of research

network, which includes his former Ph.D. advisor, post-doctoral mentor, graduate student colleagues and his own graduate student, resident, and fellow advisees.

[3] A PI is a head researcher and author who supervises doctoral students, conceives ideas and conducts projects that may include collaborating with research assistants (Armbruster, 2008).

[4] It seems that there can be a mismatch in the perceptions of the supervisor and of doctoral students with respect to accessibility. In a study on the provisions of PhD training in biomedical research PhD programs, virtually all supervisors reported meeting frequently with their students, whereas 1/4 of the students reported problems in accessing their supervisor (Frame and Allen, 2002).

[5] Nerad and Cerny (1999) also survey the perspective on postdoctoral employment in the U.S. and report that there exists a generalized discontent on behalf of postdoctoral researchers. The length of postdoctoral appointments has increased and these appointments are increasingly being seen as 'holding base', rather than being an important step in a young researcher's career.

and the final skills of the junior scientists as a researcher. The junior scientist is not an active player in our model (he does not make any decisions), but has a productive role in the project (he contributes to its final value). In addition, we assume that the junior scientist's final capabilities are not only affected by the training received from the senior scientist but are also affected by his innate ability. On this respect, we abstract from information features: senior and junior scientists have the same information about the junior scientist innate ability.[6]

Not surprisingly, when we analyze the senior scientist's allocation of time between research and training, we find that when she has more time available this results in more time allocated to both tasks. Also, we show that when there is an increase in the innate ability of the junior scientist, an increase in the importance that training has on the senior scientist's utility function, or a decrease in the productivity of the senior scientist in the project then there is a tendency to increase the time allocated to training. We also discuss the final capability of the junior scientist and the final value of the scientific project in different scenarios. Most interestingly, we show that ignorance about the true innate ability of the junior scientist may lead to more training for less able junior scientists, while there is a tendency toward an under-investment in training for the most talented ones.

As a robustness check, we consider two extensions. First, we consider the case where the senior scientist can also spend time selecting a better junior scientist. Second we examine the case where the senior scientist chooses the total amount of time she will work (and allocate to research and training), that is, the total amount of time spent on both tasks.

We also discuss possible policy instruments for a regulator who is concerned with maximizing the value of projects and attaining highly qualified scientists in a desired proportion. We highlight that the implementation of training programs in earlier education

[6] This does not mean that the junior scientist's innate ability is public information. It is ex-ante unknown by both participants. Even though a student is selected to participate in a graduate programme or in a lab according to a GRE score, and other internal admission criteria of a department, significant uncertainty remains in predicting if a student has the potential to become a successful independent researcher (Lovitts, 2005).

and tougher selection processes to attract high-ability junior scientists for research under supervision, as well as attractive training conditions, can be effective measures to attain it.

Following the work by Holmstrom and Milgrom (1991), where the authors propose a principal-agent model where the principal wants the agent to perform multiple tasks, several papers have considered the incentives for scientists to perform different tasks. For example, Lacetera and Zirulia (2008), in a context of corporate science with a great deal of competition, propose a model to explain the optimal choice of an effort to do applied research and an effort to do basic research. They analyze the strength of incentives in the effort allocation decision of the scientist and the effects of different levels of competition. In Banal-Estañol and Macho-Stadler's work (2010), the authors present a model of incentives of a researcher who can choose to either allocate time between undertaking a new research idea or developing an existing one that will deliver immediate commercial benefits. In the same branch of the literature, Walckiers (2008) argues about whether it is more attractive for a university to produce both research and teaching. The author conducts his analysis in a contractual setting between the university (principal) and the academic/scientist (agent) and studies the incentives for university scientists to perform either one of the tasks or both of them. In contrast to Walckiers (2008), where the agent does teaching at the undergraduate level, we consider training at the graduate level which implies that there are complementarities among the two tasks. Walckiers (2008) uses an adverse selection framework, where researchers differ on their preference for both tasks[7] and he shows that it can be optimal to produce research and teaching in the same institution (bundling the two tasks).

This paper is organized as follows. Section 2 describes the model and analyzes the equilibrium allocation of time to the two tasks: research and training. It also provides the comparative statics of the equilibrium efforts with respect to the parameters of the model. In Section 3, we evaluate and draw the patterns that the project's expected value and the junior scientist's final capability follow. We also present the ex-post ability of the junior scientist and the role of imperfect information in the distortions with respect to

[7] In our model, we could also discuss the researcher preferences, but this is not the main aspect of the analysis.

the full information and efficient outcomes. In Section 4 we perform a robustness check by considering two possibilities. First, we consider that the senior scientist can choose the total amount of time to exert in both tasks. Second, we consider the incentives for a senior scientist to spend time in previous activities that allow her to know more about the innate ability of the junior scientist. In Section 5 we discuss some policy instruments that may change the time allocation. In Section 6 we conclude. All proofs are remitted in the Appendix.

2 Basic Model

We consider a senior scientist who is in charge of a research project and allocates her time between research and the training of a junior scientist under her supervision. We denote the research effort by e_R and the training (guidance or education) effort by e_G and assume that the senior scientist has limited time t to allocate to these tasks. Formally, the senior scientist's time constraint is written as:

$$e_R + e_G = t.$$

In Section 4.1, we study how the available time t that the senior scientist works is determined. For now, we assume that $t > 0$ is exogenously given.

In our model, the junior scientist does not make any decisions. He is endowed with an innate ability $\hat{a}$, *ex-ante unknown* by all the players. We assume that there is a population of junior scientists with different innate abilities. The innate ability of the junior scientist who works with the senior takes a value in the interval $[\underline{a}, \bar{a}]$, with $\bar{a} > \underline{a}$, and the expected innate ability is $E(a)$.[8]

The senior scientist's vector of efforts affects two outcomes: the quality (the value) of the research project and the final capability of the junior she is training.

The final capability of the junior scientist depends on his innate ability and on the senior scientist's educational effort. Our view is that education provided by the senior is

[8] In our model, there is always symmetric information about the junior scientist's innate ability. Under complete information senior and junior know that his innate ability is $\hat{a}$; under ignorance they expect it to be $E(a)$.

a necessary input to develop the junior's scientific capability. The junior scientist's final capability, denoted by q, is a function of his *true* innate ability $\hat{a} \in [\underline{a}, \bar{a}]$ and the training he receives e_G, and it is defined as follows:

$$q = \hat{a} e_G. \tag{1}$$

Without training, even the most gifted junior scientist will not be able to acquire the capability to work on the research project in a profitable way (and maybe run a research project in the future).

The project's value depends on the direct research effort exerted by the senior scientist and on the junior scientist's final capabilities. The scientific value of the senior's project is given by:

$$v = \alpha e_R + \delta q e_R \tag{2}$$

where α is the productivity of the senior's research time e_R, and δ captures the synergies of working together with a junior scientist of capability q. In this sense, senior and junior scientists provide complementary inputs to the research project. Researchers may have different projects defined by (α, δ) and we will discuss this further on. The level v may represent the publications obtained, patents achieved or other results of the discoveries. Note that by following this functional form, no value will be produced from the project if the senior scientist does not provide any research effort.

We assume that the senior scientist's utility function combines the project's value and the junior's final capability. Formally,

$$u(v, \beta, q) = v + \beta q.$$

The project's value is included in the senior's preferences because it is a verifiable outcome that determines the senior scientist's payoff. It is also a proxy for the usual argument of peer recognition and the "puzzle joy" (Stephan and Levin, 1992). The junior scientist's final capability enters the senior's utility in a proportion β, which represents the relative appreciation of the training outcome. We may also interpret this second term of the senior's utility as a concern with reputation associated to having a network and disciples who excel in the profession.[9]

[9] Our model aims to encompass the fact that both senior and junior scientists benefit from the train-

Given the parameters $(\alpha, \delta, \beta, t)$ and the ex-ante expectation about the junior scientist's innate ability a with $a \in \{\hat{a}, E(a)\}$, the senior scientist chooses the optimal allocation of time between e_R and e_G that maximizes the ex-ante (expected) value of her utility:[10]

$$\underset{e_R, e_G}{Max} \; \{\alpha e_R + \delta a e_G e_R + \beta a e_G\}$$
$$s.t. \qquad e_R + e_G = t$$
$$e_R \in [0, t]$$
$$e_G \in [0, t]$$

From the solution to this problem we obtain the result that follows.

Lemma 1 *Given $(\alpha, \delta, \beta, t)$ and the innate ability of the junior scientist a, with $a \in \{\hat{a}, E(a)\}$, the senior scientist's allocation of time among the tasks of research and training is:*

a) When $\beta > \delta t$ and $a > \frac{\alpha}{\beta - \delta t}$,

$$e_R^* = 0 \text{ and } e_G^* = t$$

b) When $a < \frac{\alpha}{\beta + \delta t}$,

$$e_R^* = t \text{ and } e_G^* = 0$$

c) Otherwise,

$$e_R^* = \frac{t}{2} + \frac{\alpha - \beta a}{2a\delta} \text{ and } e_G^* = \frac{t}{2} - \frac{\alpha - \beta a}{2a\delta}$$

ing relationship. For the senior scientist, training increases visibility and reputation when the young professional is a productive member. Therefore, she earns more respect from the organization by developing the trainee (Kram, 1983). This model also includes the trainer's inner satisfaction in passing along knowledge (Levinson et al., 1978). For the junior scientist, the benefits include learning technical aspects of the profession, developing writing and critical skills, defining career perspectives, performing research collaborations (Kram, 1983) and receiving an important push toward building networks (Kram and Isabella, 1985).

[10] Note that v is a linear function of a that allows us to use a simplification where we write the ex-ante value of the project as a function of a, $a \in \{\hat{a}, E(a)\}$. This allows us to consider at the same time the cases where the information about the junior scientist's innate ability is perfect or imperfect.

The senior scientist's allocation of time for the interior solution depicted in Lemma 1

$$\left(e_R^* = \frac{t}{2} + \frac{\alpha - \beta a}{2a\delta}, e_G^* = \frac{t}{2} - \frac{\alpha - \beta a}{2a\delta} \right)$$

depends on all the relevant parameters. It shows the deviation from the half-half distribution of time t as a function of the senior scientist's effectiveness in the research process (α), the complementarity among the senior's and junior scientist's participation (δ), the expected innate ability of the junior (a) and the senior's concern about the junior's training (β). Lemma 1 also shows that when the junior has a low expected ability he does not receive any training. It also shows that when the senior scientist's concern about the junior's training is high as compared to the time available and the complementarity ($\beta > \delta t$) then the possibility exists that she decides ,only to perform training. Note for example, that if $\delta = 0$ (and $a \neq \frac{\alpha}{\beta},$), i.e., there is no effect of the junior scientist toward the result of the research project, then time will only be allocated to training.

Corollary 2 *For the combination of parameters satisfying $\alpha \in [a(\beta - \delta t), a(\beta + \delta t)]$ (region c) in Lemma 1), the static comparative of the efforts is presented in Table 1:*

	α	β	δ	t	a
e_R^*	$+$	$-$	$-$ *iff* $\alpha > a\beta$	$+$	$-$
e_G^*	$-$	$+$	$+$ *iff* $\alpha > a\beta$	$+$	$+$

Table 1

As expected, the senior scientist's research effort (resp., training effort) increases (resp., decreases) when α increases, β decreases or a decreases. Both efforts go in different directions when δ increases, and the direction of the change depends on the sign of $\alpha - a\beta$. In addition, both efforts increase with the time t the senior scientist has to work.

The effect of a in equilibrium is in accordance with stylized facts. In a 5 $\frac{1}{2}$ year longitudinal investigation with 233 PhD students (Paglis, Green, and Bauer, 2006, which was an extension of a similar study by Green and Bauer, 1995), the effects of supervisory mentoring of advisors on PhD students in the applied sciences were analyzed. The results show that supervisory mentoring increases the productivity and the self-efficacy of PhD students. Most importantly, a positive relationship between student potential (ability, experience and commitment to the training) and the extent to which training functions are provided by the faculty advisor is identified. That is, students who show more promise to be productive researchers receive more supervisory training and mentoring from the advisor.

If we consider, ceteris paribus, a constellation of projects defined by the pair (α, δ), for the projects whose success is heavily based on the time the senior scientist spends on it, training will be weak.[11] Also, if the direct participation of the senior scientist in the project is very important (α) and the complementarity (δ) among junior and senior participation increases, then the allocation of time to training will increase.

With respect to β, when $\beta = \frac{\alpha}{a}$, the senior scientist will exert an equal effort toward both tasks, $\frac{t}{2}$. For a higher β, the senior is relatively more concerned about training and the effect on her reputation, so she exerts a greater amount of effort to training in detriment to research. Because the efforts are linear in this parameter, we observe a constant rate of substitution. Note that even if $\beta = 0$ (in which case we will have $\delta t \geq \beta$), the senior scientist may be interested in allocating time to training. To make this point clear, let us take the example $\delta = 1$. Then the important comparison is in between α and a. If the productivity of the senior scientist is high, $\frac{\alpha}{t} > a$, she will just allocate time to research and the project will be run exclusively on her effort. If α is low when compared to a then she will allocate time to training because the complementarity of her research effort with the junior scientist will be the motor of the project.

In Figure 1 we represent the senior's effort equilibrium allocation to the two tasks in the space (a, t) by representing the iso-effort curves (the dotted lines). Keeping constant the remaining parameters, any combination (a, t) provides an optimal combination of efforts. A low amount of time tends to concentrate the senior scientist's attention on one of the tasks: if a is also low the task will be research, if a is high the task will be training. In the interior solution, the shape of the iso-effort curves shows the conflicting effects between t and a in the optimal allocation of time. As seen in Corollary 2, both parameters are aligned in their effect on training, but induce different behaviors in their effect on research, which translates in the different slopes of the iso-effort curves. For

[11] Note that in our model any project can provide the same quality of training since we assume that training may be field-specific but not project-specific.

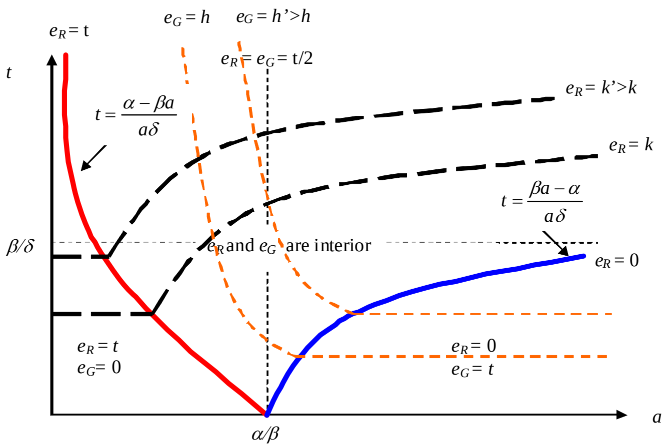

Figure 1: Optimal allocation of time

example, as a increases and t decreases, less time is devoted to research but training may decrease or increase. If a and t increase then training also increases but research may decrease or increase depending on the relative change of both parameters. Finally, note that as δ decreases, $\frac{\alpha-\beta a}{a\delta}$ and $\frac{\beta a-\alpha}{a\delta}$ increase and there are more corner solutions where the senior scientist allocates her time only to one of the tasks.

3 Expected Project Value and Junior Scientist Capability at the Optimal Time Allocation

Let us now focus on the effect that the optimal time allocation to both tasks has on the expected project value, v, as well as on the expected junior scientist capability, q.

Using the results of Lemma 1, we compute the *ex-ante* (expected) equilibrium levels of v and q. We find that:

a) When $\beta > \delta t$ and $a > \frac{\alpha}{\beta - \delta t}$,

$$v^*(a) = 0 \text{ and } q^*(a) = at$$

b) When $a < \frac{\alpha}{\beta + \delta t}$,

$$v^*(a) = \alpha t \text{ and } q^*(a) = 0$$

c) Otherwise,

$$v^*(a) = \frac{(a\delta t + \alpha)^2 - (\beta a)^2}{4a\delta} \text{ and } q^*(a) = \frac{ta}{2} - \frac{\alpha - \beta a}{2\delta}$$

When the parameters satisfy $\alpha \in [a(\beta - \delta t), a(\beta + \delta t)]$ (region c) in Lemma 1) for which a strictly positive amount of time is allocated to both tasks, the static comparative of $v^*(a)$ and $q^*(a)$ is the one summarized in Table 2.

	α	β	δ	t	a
$v^*(a)$	$+$	$-$	$+$ if and only if $a^2(\beta^2 + t^2\delta^2) > \alpha^2$	$+$	$+$ if and only if $a^2(t^2\delta^2 - \beta^2) > \alpha^2$
$q^*(a)$	$-$	$+$	$+$ if and only if $\alpha > a\beta$	$+$	$+$

Table 2

As shown in Table 2, the value of the project $v^*(a)$ is increasing in α and decreasing in β. It only increases in δ for high levels of a, i.e, for $a^2 > \frac{\alpha^2}{\beta^2+t^2\delta^2}$. This threshold decreases with the complementarity among senior and junior scientists in the project. Also, $v^*(a)$ is decreasing in a when $t\delta < \beta$. When $t\delta \geq \beta$, it decreases in a if a is smaller than a certain cut-off, $a > \frac{\alpha}{\sqrt{(\beta+\delta t)(\delta t-\beta)}}$ and increases otherwise. We finally note that $v^*(a)$ is convex in a. As to $q^*(a)$, we can see that the expected capability is linear in a and the static comparative provides us the predicted intuitive results.

We would now consider different projects defined by the pair (α, δ). For research projects that differ on α, those projects whose success is heavily based on the time the senior scientist directly spends on them, will have a better final quality. This finding is true because the senior scientist will dedicate more of her time to the project, allocating more effort to research and leaving junior scientist's guidance, and hence his final capabilities are at low levels. If we compare two research projects that differ only by δ, there are cases where in the project with higher δ, the senior scientist can produce better innovative results and more capable junior scientists. Senior researchers that have projects that differ in both dimensions (α, δ) are more difficult to compare because the effects go in opposite dimensions.

3.1 Ex-post Project Value and the Junior Scientist Final Capability

After focusing on the ex-ante equilibrium values, we now compute the *ex-post* project value and the *ex-post* junior scientist's capability that are obtained in equilibrium. We use the case of perfect information about the innate ability of the junior as a benchmark to measure the difference between ex-ante and ex-post efficiency under ignorance.[12] When the senior scientist decides on the time allocation of efforts under ignorance, the allocation of time depends on her beliefs $a = E(a)$. However, the ex-post project value (resp., the junior capability) depends on the true innate ability of the junior scientist $\hat{a}$. Hence, we denote the ex-post levels as $v^o(\hat{a}, E(a))$ and $q^o(\hat{a}, E(a))$.

As a function of the parameters, the ex-post outcomes are:

a) When $\beta > \delta t$ and $a > \frac{\alpha}{\beta - \delta t}$,

$$v^o(\hat{a}, E(a)) = 0 \text{ and } q^o(\hat{a}, E(a)) = \hat{a}t$$

b) When $a < \frac{\alpha}{\beta + \delta t}$,

$$v^o(\hat{a}, E(a)) = \alpha t \text{ and } q^o(\hat{a}, E(a)) = 0$$

c) Otherwise, using $x \equiv \frac{E(a)(\delta t - \beta) + \alpha}{2E(a)\delta}$ and $y \equiv \frac{E(a)(\delta t + \beta) - \alpha}{2E(a)}$,

$$v^o(\hat{a}, E(a)) = v^o(\hat{a}, E(a)) = \alpha x + \hat{a}xy \text{ and } q^o(\hat{a}, E(a)) = \hat{a}\left(\frac{t}{2} - \frac{\alpha - \beta E(a)}{2\delta E(a)}\right)$$

We comment first on the ex-post value of the project. In the interior case, $v^o(\hat{a}, E(a))$ is linear and increasing in $\hat{a}$.[13] To measure the distortion ex-post on the value of the project, i.e., we compute $v^*(\hat{a})$ and $v^o(\hat{a}, E(a))$ and depict the difference. The ex-post value of the project may be higher or lower under ignorance than under full information. Intuitively, under ignorance the value will be higher when expectations on the innate ability of the junior are high and its true level low. Figure 2 represents the comparison

[12] With complete information about the junior scientist's innate ability $\hat{a} = E(a)$, and the ex-post and ex-ante values of the project (resp., the junior capability) coincide. The corresponding expression is obtained by substituting a for $\hat{a}$ in the expression of the above mentioned subsection.

[13] Because $\text{sign}(xy) = \text{sign}(E(a)(\delta t - \beta) + \alpha))(E(a)(\delta t + \beta) - \alpha)$ is positive in this region.

for different levels of $E(a)$ when the concern for training is low enough, $\delta t \geq \beta$. Under ignorance, the project value is higher than under perfect information for low levels of the true innate ability and the opposite occurs when it is high.[14]

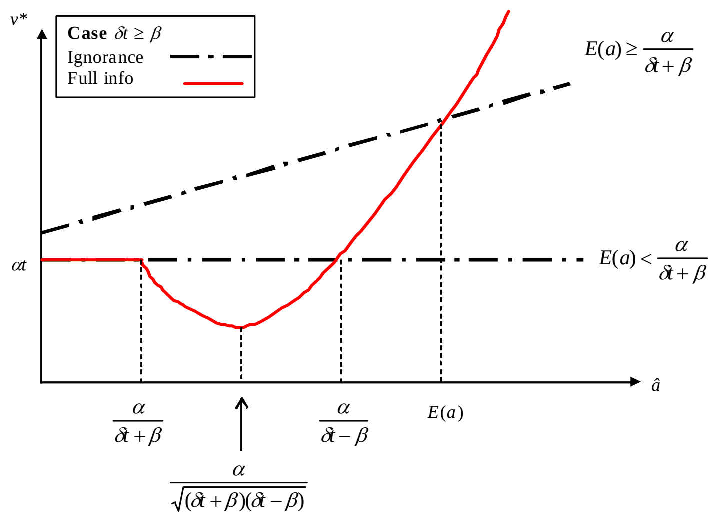

Figure 2: $v^*(\hat{a})$ and $v^o(\hat{a}, E(a))$ when $\delta t \geq \beta$

In the case where there is a high enough concern for training, $\delta t < \beta$, this distortion is more extreme. As Figure 3 shows, the value under ignorance keeps increasing (dotted lines) but the value under full information is always non-increasing in $\hat{a}$ (full lines). Also, for intermediate values of the expected innate ability, the senior scientist is always overinvesting in the training task.

A more interesting situation is to analyze how the ex-post project value under imperfect information changes with the senior scientist's prior belief about the ability of the

[14] Note that this comparison is between the decision of the senior scientist with full information and the decision under ignorance, but none of these decisions may be in accordance with the social optimum. The distortion at the bottom that may lead senior scientists to train more juniors under ignorance may be good from a social point of view. The distortion at the top may indicate more of a concern for a society interested in guaranteeing that excellent projects and high-potential junior scientists are better identified. We discuss these issues in Section 5.

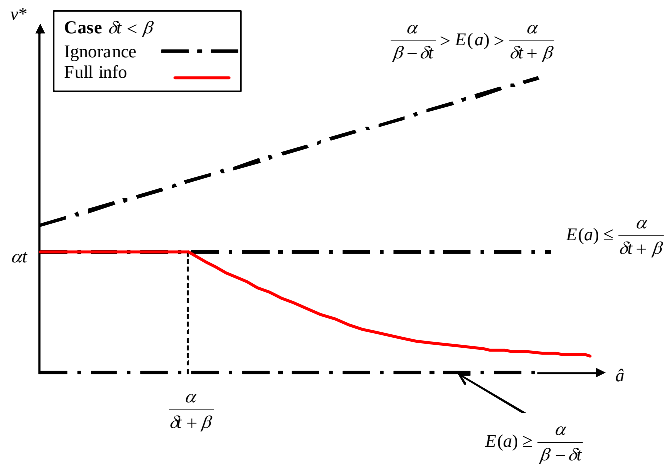

Figure 3: $v^*(\hat{a})$ and $v^o(\hat{a}, E(a))$ when $\delta t < \beta$

junior scientist, $E(a)$. Given $\hat{a}$, when $E(a)$ is low enough,[15] an increase in $E(a)$ leads to an increase in $v^o(\hat{a}, E(a))$. Otherwise it will decrease. One can also read this result the other way around. Given $E(a)$, for high values of $\hat{a}$ a higher prior will increase $v^o(\hat{a}, E(a))$, and for low $\hat{a}$ a higher prior will decrease $v^o(\hat{a}, E(a))$. This means that a higher prior in the case of Figure 2, increases the distortion between $v^*(\hat{a})$ and $v^o(\hat{a}, E(a))$ for intermediate values of $\hat{a}$ but decreases the distortion for very low and very high values of $\hat{a}$. So, when the junior scientist is indeed very good, this inefficiency becomes smaller as $E(a)$ is closer to $\hat{a}$. In the case of $\delta t < \beta$ (Figure 3), the effect of an increase in $E(a)$ will decrease the distortion for low levels of $\hat{a}$ and increase it onwards.

We now consider the distortions of the junior scientist's final capability. Note that for the interior solution, $q^o(\hat{a}, E(a))$ is linear and increasing in $\hat{a}$ and increasing and concave in $E(a)$. Comparing the ex-post junior's training under full information $q^*(\hat{a})$ and under ignorance $q^o(\hat{a}, E(a))$, we obtain the results depicted in Figures 4 and 5. Figure 4 represents the case when the concern for training is low, $\delta t \geq \beta$. We can see from the two curves that there is over-training of low-ability and under-training of high-ability juniors. The further away the senior scientist's prior is from the true ability of the junior, either from above or below, the higher is the distortion in the ex-post formation, and

[15] For $E(a) \leq \left(\frac{\hat{a}\alpha}{2(\alpha+\hat{a}\beta)}\right)^{\frac{1}{2}}$.

the distortion increases proportionately. The effect of a higher prior $E(a)$ over the junior scientist's ability is a higher slope of the "ignorance" ex-post curve, which means that the distortion increases for lower values of $\hat{a}$ and will decrease for higher values. More accurate training is provided to the population of junior scientists with more potential.

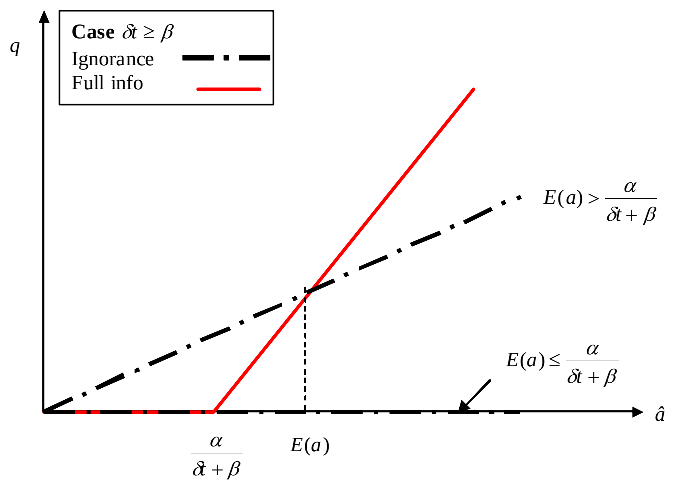

Figure 4: $q^*(\hat{a})$ and $q^o(\hat{a}, E(a))$ when $\delta t \geq \beta$

For $\delta t < \beta$, as illustrated in Figure 5, the distortion and the effects of a higher $E(a)$ are very similar.

As in the case with the value of the project, the sign of the distortion may be aligned or not aligned with social interest. We will discuss this aspect in Section 5. Note also that if the interval $[\underline{a}, \bar{a}]$ is smaller or if, for a given interval the expected innate ability of the junior scientist, $E(a)$ is higher, then it may be the case that more education is provided under ignorance and the distortion is not too big.

4 Extensions

Let us consider two natural questions that may come to mind that we present here in two independent extensions. In the first one, we allow for the senior scientist to choose the amount of time that she will work. In other words, t is not exogenous but her choice. In the second extension we assume that the time available is exogenously given, but we

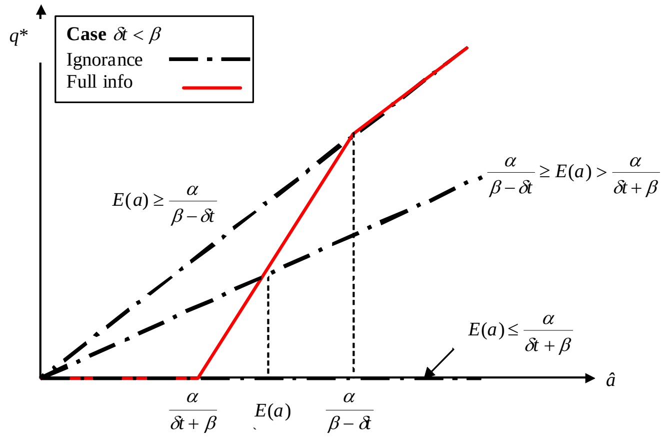

Figure 5: $q^*(\hat{a})$ and $q^o(\hat{a}, E(a))$ when $\delta t < \beta$

allow the senior scientist to have access to a better pool of junior scientists at the cost of some of her time.

Both extensions can be viewed as a sequential decision problem where, once either the time allocated to work or the ability of the junior scientist is determined, the analysis of the previous sections tells us the result in terms of research and training. Hence, it is useful to use the optimal allocation of time as a function of t and a to write the utility of the senior in terms of these variables. Using Lemma 1 we see that:

a) When $\beta > \delta t$ and $a > \frac{\alpha}{\beta - \delta t}$,

$$u^*(a,t) = \beta a t$$

b) When $a < \frac{\alpha}{\beta + \delta t}$,

$$u^*(a) = \alpha t$$

c) Otherwise,

$$u^*(a,t) = \frac{(a\delta t + \alpha)^2 - (\beta a)^2}{4a\delta} + \beta\left(\frac{ta}{2} - \frac{\alpha - \beta a}{2\delta}\right)$$

4.1 Total Time Worked

Until now we have considered that the time t the senior scientist works is fixed. Let us now consider that, as in the line of more traditional moral hazard models, the senior scientist may decide the total amount of time t she will devote to work (the total effort). To determine this total working time t, the senior scientist maximizes her expected utility net of the cost of the working time. We will assume that the cost of working time c is high enough. More precisely, we assume $c \geq \frac{a\delta}{2}$.[16] Hence, the senior solves

$$\underset{t}{Max} \left\{ u^*(a,t) - \frac{c}{2}t^2 \right\}.$$

From this problem, we obtain the following result:[17]

Lemma 3 *The senior scientist's total time as a function of the parameters is*

a) When $\alpha - a\beta \geq 0$ *and* $c \geq \frac{a\delta\alpha}{\alpha - a\beta}$,

$$t^* = \frac{\alpha}{c}$$

b) When $\alpha - a\beta \geq 0$ *and* $\frac{a\delta}{2} < c \leq \frac{a\delta\alpha}{\alpha - a\beta}$ *or* $\alpha - a\beta < 0$ *and* $\frac{\beta a^2 \delta}{a\beta - \alpha} > c$,

$$t^* = \frac{\alpha + a\beta}{2c - a\delta}$$

c) When $\alpha - a\beta < 0$ *and* $\frac{a\delta}{2} < c \leq \frac{\beta a^2\delta}{a\beta - \alpha}$,

$$t^* = \frac{a\beta}{c}$$

Lemma 3 is depicted in Figure 6 in the space $(c, (\alpha - a\beta))$.

From Lemma 3 we conclude that, as expected, the time the senior scientist works t^* is a non-decreasing function of a, α, β, δ and is decreasing in c. The comparative statics

[16] If the cost c is smaller that $a\delta/2$ the optimal time goes to infinite. In this case it would be natural to include a maximum time limit T. We will comment on this assumption later, but we will concentrate on the case where c is high for the sake of simplicity.

[17] Note that for $\alpha - a\beta \geq 0$, we have $\frac{a\delta}{2} \leq \frac{a\delta\alpha}{\alpha - a\beta}$, and for $\alpha - a\beta < 0$, we have $\frac{a\delta}{2} < \frac{\beta a^2\delta}{a\beta - \alpha}$. Hence, the regions of Lemma 3 are well defined.

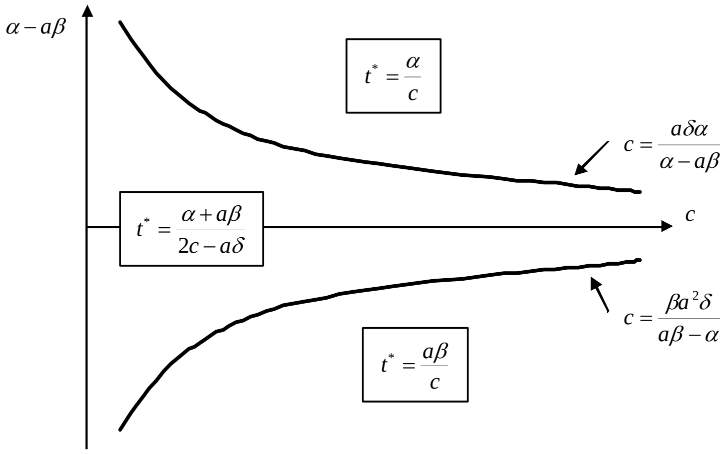

Figure 6: Optimal t^* in the space $(c, (\alpha - a\beta))$

are summarized in Table 3.

	α	β	δ	c	a
$t^* = \frac{\alpha}{c}$	$+$	0	0	$-$	0
$t^* = \frac{\alpha+a\beta}{2c-a\delta}$	$+$	$+$	$+$	$-$	$+$
$t^* = \frac{a\beta}{c}$	0	$+$	0	$-$	$+$

Table 3

Given the optimal t^*, we compute the time allocated to each task. As expected, the time allocated to both tasks is decreasing in c (that now plays a role similar to a decrease in t in the previous sections). Effort e_R is non-increasing and e_G is non-decreasing in the concern for training, β. More precisely, for $t^* = \frac{\alpha}{c}$, the optimal allocation of time is $\left(e_R = \frac{\alpha}{c}, e_G = 0\right)$ and research increases with α but nothing is affected by β or a. At the other extreme, for $t^* = \frac{a\beta}{c}$ the optimal allocation of time is $\left(e_R = 0, e_G = \frac{a\beta}{c}\right)$. For the case $t^* = \frac{\alpha+a\beta}{2c-a\delta}$ the time allocated to both tasks deserves some attention, and we perform comparative statics given in Corollary 4.

Corollary 4 *When $t^* = \frac{\alpha+a\beta}{2c-a\delta}$, region b) in Lemma 3, we have that the allocation of time* *to the tasks is*

$$\begin{aligned} e_R^* &= \frac{1}{2}\left(\frac{\alpha + a\beta}{2c - a\delta} + \frac{\alpha - a\beta}{a\delta}\right) \\ e_G^* &= \frac{1}{2}\left(\frac{\alpha + a\beta}{2c - a\delta} - \frac{\alpha - a\beta}{a\delta}\right) \end{aligned}$$

and the comparative statics are presented in Table 4:

	α	β	a	δ
$e_R^*(a, \alpha, \beta, \delta)$	$+$	$+$ *iff* $c < a\delta$	$+$ *iff* $c < \frac{\delta a}{2}\frac{(2\alpha + a\beta)}{\alpha}$	$+$ *iff* $c < \frac{\delta a}{2}\left(\frac{(\alpha+a\beta)^{1/2}}{(\alpha-a\beta)^{1/2}} - 1\right)$
$e_G^*(a, \alpha, \beta, \delta)$	$+$ *iff* $c < a\delta$	$+$	$+$	$+$

Table 4

For completeness let us remark that in a version of the model where c is not constrained from below, and there is a maximum amount of time T that the senior scientist has available, the results will be similar, except for low costs c $\left(c \leq \frac{a\delta}{2}\right)$ and/or T small enough $\left(T < \frac{|\alpha - a\beta|}{a\delta}\right)$. In these cases, the senior chooses to work for all the available time and she allocates all the time T either to research ($\alpha > a\beta$) or to training ($\alpha < a\beta$) (except if $\alpha = a\beta$, case where she is indifferent). Changes in α or in β do not affect the total time allocated to work and only discrete changes may affect to which task this time T is allocated. In these cases, only changes of the time available for these activities may have an effect on the senior scientist's behavior.

4.2 When Expected Ability and Time Are Related

In Section 2 we analyzed the decisions of a senior scientist that is (randomly) matched with a junior scientist of innate ability a. However, one may wonder about what happens if the junior's expected innate ability a depends on some previous activity that the senior performs and that consumes time. This may correspond to a selection process that tries to identify a better population of junior scientists, an advertising or investment procedure that aims to attract a junior scientist with a higher expected innate ability, or an undergraduate system that provides better skills and better information about the juniors' abilities. Here, we model the relationship between the time invested in increasing the innate ability of the junior she works with and the remaining time available for research and training.

Let us assume that T is the maximum amount of time available for the three tasks and A the innate ability of the junior scientist if no effort is made to improve it. Let us denote by $g(T-t)$, with $g > 0$, the improvement of the innate ability of the junior scientist that the senior can obtain by using an amount of time $(T-t)$ to improve the quality of the junior with whom she works, in such a way that she will have t to allocate to the tasks of research and formation. Hence, the ability of the junior scientist with whom the senior will work with is $a = A + g(T-t)$.

In this case, the senior scientist chooses (a, t) by maximizing her utility function, taking into account the constraints $a = A + g(T-t)$. To simplify presentation and to avoid cumbersome calculations for different regions of parameters, we just present an example where we assume that the senior has no appreciation of the junior scientist's final capability ($\beta = 0$) and that the complementarity effect is 1 ($\delta = 1$). This implies that we will be in Region $\delta t \geq \beta$ (that, in this case, is reduced to $t \geq 0$). Under this parameter combination, as a function of t, the senior's allocation of time depends on whether $a \geq \frac{\alpha}{\delta t}$ (and her time will be allocated to research and formation) or $a \leq \frac{\alpha}{\delta t}$ (and she will only do research).

Lemma 5 *Assuming $\beta = 0$, $\delta = 1$ and the relation between time and ability given as $a = A + g(T-t)$, the senior scientist's decision on the optimal innate ability and on the optimal amount of time spent in previous activities is:*

a) When $(gT+A)^2 - 12\alpha g \geq 0$,

$$a^* = \frac{gT + A + \sqrt{(gT+A)^2 - 12\alpha g}}{6} \text{ and } t^* = \frac{5(gT+A) - \sqrt{(gT+A)^2 - 12\alpha g}}{6g}$$

b) When $(gT+A)^2 - 12\alpha g < 0$,

$$a^* = A \text{ and } t^* = T$$

Lemma 5 illustrates that in projects where the productivity of the senior scientist's direct research (α) is low enough, the senior is willing to spend time in selection activities that allow to work with a junior scientist of higher expected innate ability. This finding is because being a less productive scientist, the senior will want to increase the prospects of working with a more talented junior scientist. When α is high enough, then she will choose not to spend any time in activities to retrieve more information about the junior's innate ability, leaving it at level A. This way, she chooses to allocate all of the time resources to training and research only. Note also that for a given combination of the other parameters (g, α), when T or A are small it is more often the case that the senior's optimal decision is not spend time in improving the innate ability of the junior (while this does not mean that she will not allocate some time to formation).

5 Welfare Analysis

We would like to consider here a situation with a social planner who is concerned about the level of research that the senior scientist achieves and the final capacity of the junior, that he interprets as a measure of the potential of the next generation of researchers. We consider first that this social planner has the welfare function:

$$W = E(v^o(\hat{a}, E(a))) + \lambda E(q^o(\hat{a}, E(a))),$$

where λ can be interpreted as the society's relative concern about the capability of the next generation of researchers.

If λ coincides with β, then the decision of the senior and the aims of the society concur. If λ and β do not coincide, the social planner may be tempted to intervene. To discuss this possibility, we take as a starting point our basic model presented in Section 2, where we assume that there is no moral hazard problem, just a decision about the allocation of time. An alternative way of looking at the comparative static in Table 1 (and the discussion after it) is to consider how society may induce changes in some parameters to affect the senior scientist's allocation of time to research and formation. For this purpose, the social planner must affect the senior's utility $u = \alpha e_R + \delta a e_G e_R + \beta a e_G$, possibly using (α, δ) and β as instruments.[18]

[18]Obviously, the social planner can also change the time available for these tasks t (for example, by reducing the senior's involvement in other time-consuming tasks, such as administrative ones).

If the outcome of research v and the outcome of training q are verifiable, the regulator can manipulate the decision of the senior scientist by changing her awareness about these two variables. The planner can increase the senior scientist's utility from the project's value (that is, increasing α and δ in the same proportion or, equivalently decreasing β) or to increase the senior's payoff as a function of the quality of the junior she mentors (increasing β) via the definition of a successful career or the allocation of research funds that weight this aspect of the academic career. Note that increasing both perceptions is useless when the aim is to change the allocation of total time, because total time is fixed. Also, if only publications (and other measures of the senior scientist project results) are verifiable, the social planner can only encourage more time to research (through tenure tack rules, opportunities to travel and access to research funds, or peer esteem, which in our model corresponds to decrease β) but he cannot increase it above the natural inclination of the senior scientist. Only by discouraging research can the time allocated to training be increased.

The social planner can change the junior scientist's innate quality a (for example, by having an attractive and selective program of fellowships) that allows the attraction of better students. Indeed, several European expert institutions (e.g., EURAB, ESF) have given priority to the training of scientists and developed actions so that postdoctoral researchers ascend to PIs in recent years. These actions involve providing access to special grants, as well as promoting free and secure mobility.

When total time is fixed, these instruments have a positive effect on one task but a negative effect on the other. Both efforts only increase simultaneously by inducing a higher t, as already mentioned. If a moral hazard situation exists, and the senior scientist decides how much time to work, the previous discussion of the instruments to use holds partially. In this case, incentivizing the results for both research and training may be optimal because these instruments affect not only time allocation but also how much time the senior decides to work. As shown in Corollary 4, if the cost of the effort or the quality of the junior is high enough, then increases in (α, δ), which correspond to a higher utility associated to the value of the research project, or increases in β induce more research and more training (because they induce more incentives to work). If juniors are gifted enough, both instruments (increasing the utility the senior scientist receives from research or from training) have positive effects on the senior scientist's dedication to both tasks. In a society where the population of juniors is of low expected ability, the instruments have positive effects on one task and negative on the other and encouraging one activity crowds out the effort on the other one. This emphasizes the importance of attracting a good population of junior scientists. This comment connects with the analysis conducted in Section 4.2 where Lemma 5 draws attention to the possibility that the population of juniors can be linked to the time allocated to select them. Note however, that measures that increase a^* will decreases t^*. This may lead to an increase in the senior's dedication to a task but may trigger a decrease in the time allocate to the other task unless the cost of obtaining better pools of junior scientists, g, decreases.

Another point of view is to consider that the social planner is not just concerned about the expected level of research and training. His concern may be to reach high enough outcomes in both tasks. In other words, it can be the case that the social planner is only interested in excellence and in achieving the highest innovation level (project quality) and the highest level of ex-post capability.[19] Imagine a social planner considers research to be valuable only if $v \geq V$ and wants junior scientists to be endowed with a minimum final capability $q \geq Q$ to be considered good independent researchers. In this framework, the social planner cares about $\tilde{W}$,

$$\tilde{W} = E_{v \geq V}(v^o(\hat{a}, E(a))) + \lambda E_{q \geq Q}(q^o(\hat{a}, E(a))),$$

where the minimum requirements (V, Q) are given by the social planner a level of exigency.

We use now the results presented in Section 3.1 based on the model where time t is given. To have projects and young researchers above the cutoffs (V, Q) with high probability (or a high proportion) the social planner may use the available instruments β, a, t. We have seen in Figures 5 and 6 that the level of final capability under ignorance, $q^o(\hat{a}, E(a))$, increases with the senior scientist's priors with respect to the innate ability of the junior.

To help the discussion along, in Figure 7 (using the information conveyed in Table 2)

[19] The reason may be that the society may not consider results below a minimum requirement on both outcomes as an achievement: the society may value only "good enough" discoveries to be patented or to improve knowledge, and only "capable enough" junior scientists may be considered good researchers.

we represent in the space (a, t) the expected value of the project in equilibrium, as well as the iso-project value curves and the iso-final capability curves of the junior scientist, keeping constant other parameters (the dotted lines). This figure shows that a higher total amount of time always induces a higher expected project value as well as a higher junior scientist capability. However, increasing only a does not have the same effect. If the social planner wants project values to have at least value V' and the junior capability Q' then it has, on one hand, to procure a higher total time t available to the senior scientist, and on the other hand induce as much as possible a selected junior scientist with enough potential.

For the social planner, the senior scientist's priors is a possible instrument to obtain a superior outcome in the training component because a higher prior induces more time allocated to training. Besides the quantity effect, which is the fact that more of the population reaches an independent research status (attains q above Q), there is a quality effect on junior researchers since they are better prepared. Analyzing v^o here, we conclude that the project value under ignorance increases with the senior scientist's priors, but only until a certain point. For very high values of the expected ability the equilibrium value of the project starts to decrease. Hence, when fixing the level of $E(a)$ both effects must be taken into account. Increasing the expected ability of the juniors population, $E(a)$, can be performed by implementing or increasing subsidies to a tougher selection of scientists eligible to perform research under supervision.[20] Implementing good programs in earlier education can also cause this shift. Also, offering more attractive conditions in programs for PhD and postdocs may attract better candidates for the task, who are otherwise drawn to more attractive careers in other sectors. These conditions mean not only better stipends, but also better lab equipment accessible to junior scientists. Implementing such measures will shift the population to higher levels of innate ability, first order stochastically dominating the initial population distribution or even an increase in

[20] One can assume that the senior sceintist or ither department are in a better position to assess the ability of a junior scientist, but it also seems reasonable to think that higher resources allocated to the selection processes may help. Any selection process would include the past education of the junior sceintist, as well as considering the university of origin and inviting the juniors for an interview. If these resources are not available, their own students may be less risky that outsiders.

the lower bound $\underline{a}$ of the distribution on the innate ability of junior scientists (which can correspond to a higher A in the analysis of Section 4.2).

The productivity of the senior scientist with respect to her own research, α, can also be used to either promote higher project value, and the concern that the senior has over formation, β, to promote the higher final capability of junior scientists. The social planner will obviously face a trade-off here as well. However, some outcomes are unreachable controlling only the parameter β; if the regulator wants to have simultaneously a high enough expected project value (above V) and a high enough expected q^* (above Q), this objective may not be reached only by manipulating β. If these cutoffs V and Q are very demanding, they can be obtained only for some combination of the remaining parameters (a, t, α, δ).

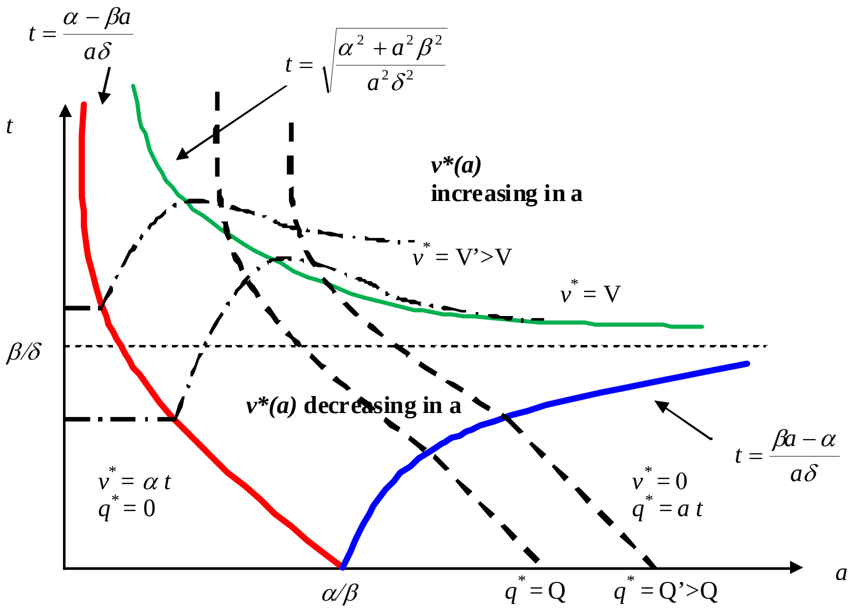

Figure 7: v^* and q^* as function of (a, t)

6 Conclusion

In this paper we provide a model where not only research results of scientists are important but also their involvement and effect in training young scientists. To this aim we propose a multitask model where the training of a junior scientist depends on the incentives that a senior scientist has to allocate time to training when this is at the cost of spending less time doing research. Our model proposes the (testable) predictions that researchers are more inclined to be involved in training the more time they have to allocate to these tasks, the higher the expected innate ability of the junior scientist, and the less important is her own involvement for the success of the research project is as compared to the complementarities of working with a competent junior.

We study the possibility that a senior scientist engages in previous activities to obtain more information about the innate ability of the junior scientist. This possibility is used when her own productivity on the project is not too high and investing in this activity provides better chances of working with a good junior. We also analyze the distortions that arise in the value of a research project and in the final capability of the junior scientist between imperfect information (ex-ante efficiency) and full information (ex-post efficiency). The patterns of the distortions will vary with the relevant parameters, and more or less training may be provided when we compare the decision of the senior under ignorance or full information. This aspect is also important when a regulator considers policies to induce more training and decrease these distortions. We also discuss that, in some cases, if a regulator implements attractive, recognized training programs in earlier education (leading to a higher proportion of good junior scientists), as well as a tougher selection process for research under supervision and attractive conditions to appointments, the proportion and the quality level of independent scientists in the population will be positively affected.

This model is still a first approach to the problem. We think that it may help to provide a simple framework to conduct some empirical studies and to think about policy issues. This model has the potential to be generalized toward the study of additional problems where the matching among a senior scientist and junior scientist, the senior's choice of a research project, or the team component of some research processes are included.

7 APPENDIX

7.1 Proof of Lemma 1

The program can be rewritten as

$$\underset{e_R, e_G}{Max}\ \alpha e_R + a\delta\left(t - e_R\right) e_R + \beta a\left(t - e_R\right)$$

$$s.t. \qquad e_R \geq 0 \qquad \text{and} \qquad e_R \leq t$$

The Lagrangian of this program is $\mathcal{L} = \alpha e_R + a\delta\left(t - e_R\right) e_R + \beta a\left(t - e_R\right) + \lambda e_R + \mu\left(t - e_R\right)$. Its FOC is

$$\alpha + a\delta\left(t - 2e_R\right) - \beta a + \lambda - \mu = 0.$$

Then the possibilities are (1) $\lambda = 0$, $\mu = 0$ and $e_R = \frac{a\delta t + \alpha - \beta a}{2a\delta}$, which is a candidate when $\frac{a\delta t + \alpha - \beta a}{2a\delta} \geq 0$ and $\frac{a\delta t + \alpha - \beta a}{2a\delta} \leq t$, or equivalent, when $\alpha + a\delta t - \beta a \geq 0$ and $\alpha - a\delta t - \beta a \leq 0$, (2) $\lambda > 0$, $\mu = 0$ and $e_R = 0$, which is a candidate when $\alpha + a\delta t - \beta a \leq 0$ (3) $\lambda = 0$, $\mu > 0$ and $e_R = t$, which asks for $\alpha - a\delta t - \beta a \geq 0$. Finally, the combination of these cases and the definition of $e_G = t - e_R$ give the results

Solution e_R^*	Solution e_G	When $\delta t \geq \beta$	When $\delta t < \beta$
$e_R^* = 0$	$e_G = t$		$\frac{\alpha}{\beta - \delta t} < a$
$e_R^* = \frac{t}{2} + \frac{\alpha - \beta a}{2a\delta}$	$e_G = \frac{t}{2} - \frac{\alpha - \beta a}{2a\delta}$	$\frac{\alpha}{\beta + \delta t} \leq a$	$\frac{\alpha}{\beta + \delta t} \leq a \leq \frac{\alpha}{\beta - \delta t}$
$e_R^* = t$	$e_G = 0$	$a < \frac{\alpha}{\beta + \delta t}$	$a < \frac{\alpha}{\beta + \delta t}$

presented in the Lemma.

7.2 Proof of Lemma 3

We proceed for simplicity by regions of parameters:

Case 1: $\alpha - a\beta > 0$

In this case:

$$u(t) = \frac{\left(a\delta t + \alpha\right)^2 - \left(\beta a\right)^2}{4a\delta} + \frac{2\beta a\left(a\delta t - \alpha + \beta a\right)}{4a\delta} \text{ if } t \geq \frac{\alpha - a\beta}{a\delta} \tag{3}$$

$$u(t) = \alpha t \text{ if } t \leq \frac{\alpha - a\beta}{a\delta} \tag{4}$$

Step (i) Let's first consider the function $u(t) = \alpha t$. The maximization problem is

$$\underset{t}{Max} \left\{ \alpha t - \frac{c}{2} t^2 \right\}$$
$$s.t. \qquad 0 \leq t \leq \frac{\alpha - a\beta}{a\delta}$$

The Lagrangian function is $\mathcal{L} = \alpha t - \frac{c}{2}t^2 + \lambda(\frac{\alpha - a\beta}{a\delta} - t) + \mu t$. The FOC is

$$\alpha - ct - \lambda + \mu = 0.$$

Then the possibilities are (i.1) $\lambda = 0$ and $\mu = 0$, then $t = \frac{\alpha}{c}$, which is a candidate when $\frac{\alpha}{c} \leq \frac{\alpha - a\beta}{a\delta} \iff \frac{a\delta\alpha}{\alpha - a\beta} \leq c$. (ii.2) $\lambda > 0$ and $\mu = 0$, then $t = \frac{\alpha - a\beta}{a\delta}$, which is a candidate when $\alpha - c(\frac{\alpha - a\beta}{a\delta}) > 0$, i.e., $\frac{\alpha}{c} > \frac{\alpha - a\beta}{a\delta} \iff \frac{a\delta\alpha}{\alpha - a\beta} > c$. (ii.3) $\lambda = 0$ and $\mu > 0$, then $t = 0$, which is a candidate if $\alpha < 0$ which is never the case.

Step (ii) Let's now consider the function $u(t) = \frac{(\alpha + at\delta)^2 - (a\beta)^2}{4a\delta} + \frac{2a\beta(a\delta t - \alpha + a\beta)}{4a\delta}$. The maximization problem is

$$\underset{t}{Max} \left\{ \frac{(\alpha + at\delta)^2 - (a\beta)^2}{4a\delta} + \frac{2a\beta(a\delta t - \alpha + a\beta)}{4a\delta} - \frac{c}{2}t^2 \right\}$$
$$s.t. \qquad t \geq \frac{\alpha - a\beta}{a\delta}$$

The Lagrangian is $\mathcal{L} = \frac{(\alpha + at\delta)^2 - (a\beta)^2}{4a\delta} + \frac{2a\beta(a\delta t - \alpha + a\beta)}{4a\delta} - \frac{c}{2}t^2 + \lambda(t - \frac{\alpha - a\beta}{a\delta})$. The FOC is

$$\frac{\alpha + at\delta + a\beta}{2} - ct + \lambda = 0.$$

Then, the possibilities are: (ii.1) $\lambda = 0$ and $t = \frac{\alpha + a\beta}{2c - a\delta}$, which is a candidate when $\frac{\alpha + a\beta}{2c - a\delta} > \frac{\alpha - a\beta}{a\delta}$ or, equivalently, when $\frac{a\delta\alpha}{\alpha - a\beta} \geq c$. (ii.2) $\lambda > 0$ and $t = \frac{\alpha - a\beta}{a\delta}$, which asks for $\frac{\alpha + a\delta(\frac{\alpha - a\beta}{a\delta})}{2} - c(\frac{\alpha - a\beta}{a\delta}) < 0$, or equivalently $\frac{a\delta\alpha}{\alpha - a\beta} < c$.

Step (iii) Comparing the results from the previews steps, and realizing that the candidate for solution in one of cases is a possible solution in the other case we obtain the result presented in the Lemma.

Case 2: $\alpha - a\beta < 0$

Here we have:

$$u(t) = \frac{(a\delta t + \alpha)^2 - (\beta a)^2}{4a\delta} + \frac{2\beta a\,(a\delta t - \alpha + \beta a)}{4a\delta} \text{ if } t \geq \frac{a\beta - \alpha}{a\delta} \tag{5}$$
$$u(t) = a\beta t \text{ if } t \leq \frac{a\beta - \alpha}{a\delta} \tag{6}$$

Step (i) Let's consider the function $u(t) = a\beta t$. The maximization problem is

$$\underset{e_R,e_G}{Max}\ \left\{a\beta t - \frac{c}{2}t^2\right\}$$
$$s.t. \qquad 0 \leq t \leq \frac{a\beta - \alpha}{a\delta}$$

The Lagrangian function is $\mathcal{L} = a\beta t - \frac{c}{2}t^2 + \lambda(\frac{a\beta-\alpha}{a\delta} - t) + \mu t$. The FOC is

$$a\beta - ct - \lambda + \mu = 0.$$

Then the possibilities are (i.1) $\lambda = 0$ and $\mu = 0$ which gives $t = \frac{a\beta}{c}$, which is a candidate when $\frac{a\beta}{c} \leq \frac{a\beta-\alpha}{a\delta} \iff \frac{\beta a^2\delta}{a\beta-\alpha} \leq c$. (ii.2) $\lambda > 0$ and $\mu = 0$, and $t = \frac{a\beta-\alpha}{a\delta}$, which is a candidate when $\alpha - c\left(\frac{a\beta-\alpha}{a\delta}\right) > 0$, i.e., $\frac{\alpha}{c} > \frac{a\beta-\alpha}{a\delta} \iff \frac{\beta a^2\delta}{a\beta-\alpha} > c$. (ii.2) $\lambda = 0$ and $\mu > 0$, and $t = 0$, which is a candidate when $a\beta < 0$, which is never the case.

Step (ii) Let's now consider the function $u(t) = \frac{(\alpha+at\delta)^2-(a\beta)^2}{4a\delta} + \frac{2a\beta(a\delta t-\alpha+a\beta)}{4a\delta}$. The maximization problem is

$$\underset{t}{Max}\ \left\{\frac{(\alpha + at\delta)^2 - (a\beta)^2}{4a\delta} + \frac{2a\beta(a\delta t - \alpha + a\beta)}{4a\delta} - \frac{c}{2}t^2\right\}$$
$$s.t. \qquad t \geq \frac{a\beta - \alpha}{a\delta}$$

The Lagrangian function is $\mathcal{L} = \frac{(\alpha+at\delta)^2-(a\beta)^2}{4a\delta} + \frac{2a\beta(a\delta t-\alpha+a\beta)}{4a\delta} - \frac{c}{2}t^2 + \lambda(t - \frac{a\beta-\alpha}{a\delta})$. The FOC is

$$\frac{\alpha + at\delta + a\beta}{2} - ct + \lambda = 0.$$

Then, the possibilities are: (ii.1) $\lambda = 0$ and $t = \frac{\alpha+a\beta}{2c-a\delta}$, which is a candidate when $\frac{\alpha+a\beta}{2c-a\delta} > \frac{a\beta-\alpha}{a\delta}$ or, equivalently, when $\frac{\beta a^2\delta}{a\beta-\alpha} > c$. Note that $\frac{\beta a^2\delta}{a\beta-\alpha} < \frac{a\delta}{2}$ (ii.2) $\lambda > 0$ and $t = \frac{a\beta-\alpha}{a\delta}$, which asks for $\frac{\alpha+a\delta(-\frac{\alpha-a\beta}{a\delta})+a\beta}{2} - c(\frac{a\beta-\alpha}{a\delta}) < 0$, or equivalently $\frac{\beta a^2\delta}{a\beta-\alpha} < c$.

Step (iii). Comparing the results from the previews steps, and realizing that the candidate for solution in one of cases is a possible solution in the other case we obtain the result presented in the Lemma.

Case 3: $\alpha - a\beta = 0$

In this case:

$$u(t) = t\frac{2\left(\alpha + \beta a\right) + a\delta t}{4} \text{if } t \geq 0$$

Then the maximization problem of the senior is:

$$\underset{t}{Max}\left\{\frac{2t\left(\alpha+\beta a\right)+a\delta t^{2}}{4}-\frac{c}{2}t^{2}\right\}$$
$$s.t. \qquad t\geq 0$$

The Lagrangian function is $\mathcal{L}=\frac{2t(\alpha+\beta a)+a\delta t^{2}}{4}-\frac{c}{2}t^{2}+\lambda t$. The FOC is

$$\frac{\alpha+a\beta+at\delta}{2}-ct+\lambda=0$$

Then, the possibilities are: (i) $\lambda=0$ and $t=\frac{\alpha+a\beta}{2c-a\delta}$, which is a candidate when $\frac{\alpha+a\beta}{2c-a\delta}\geq 0$ or, equivalently, when $c>\frac{a\delta}{2}$. (ii) $\lambda>0$ and $t=0$, which asks for $\frac{\alpha+a\beta}{2}<0$, which is never the case.

7.3 Proof of Lemma 5

For the case $\beta=0$ and $\delta=1$, we consider the candidates that satisfy $a^{2}-(gT+A)a+\alpha g\geq 0$ and $a^{2}-(gT+A)a+\alpha g<0$ sequentially and then we provide the solution as function of the parameters.

Step 1: $a^{2}-(gT+A)a+\alpha g\geq 0$

The utility function to be considered is $u^{*}(t)=\alpha\left(T-\frac{a-A}{g}\right)$.

Depending on the parameters, the values of a such that $a^{2}-(gT+A)a+\alpha g=0$ (they only exist if $(gT+A)^{2}-4\alpha g\geq 0$) will lie, or not, in the interval where the junior's ability is defined, $[A,gT+A]$. When $a=A$, $a^{2}-(gT+A)a+\alpha g$ has a positive value if $AT\leq\alpha$, and a negative value otherwise. When $a=gT+A$, the value for $a^{2}-(gT+A)a+\alpha g$ is always positive. Hence we can define 3 possible regions to attain a solution: a) when $AT\leq\alpha$ and $(gT+A)^{2}-4\alpha g\geq 0$; b) when $AT\leq\alpha$ and $(gT+A)^{2}-4\alpha g<0$; and c) when $AT>\alpha$ and $(gT+A)^{2}-4\alpha g\geq 0$. The One last case could be $AT>\alpha$ and $(gT+A)^{2}-4\alpha g<0$, however it is not possible since this would mean that the function never has negative values and, simultaneously, has a negative value when $a=A$.

Case a) $AT\leq\alpha$ and $(gT+A)^{2}-4\alpha g\geq 0$

In this case the optimal ability lies in either one of the two following regions: $A\leq a\leq\frac{gT+A-\sqrt{(gT+A)^{2}-4\alpha g}}{2}$ or $\frac{gT+A+\sqrt[2]{(gT+A)^{2}-4\alpha g}}{2}\leq a\leq gT+A$. We first formalize the problem with respect to the first region:

$$\begin{array}{rrcl} & & & \underset{a}{Max} \left\{ \alpha \left(T - \frac{a-A}{g} \right) \right\} \\ s.t & a & \geq & A \\ & a & \leq & \frac{gT + A - \sqrt{(gT+A)^2 - 4\alpha g}}{2} \end{array}$$

The lagrangian is $\mathcal{L} = \alpha\left(T - \frac{a-A}{g}\right) + \lambda(a - A) + \mu(\frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2} - a)$. The FOC is $-\frac{\alpha}{g} + \lambda - \mu = 0$

There are two possible cases for the lagrangian multipliers:

1) $\lambda > 0$, $\mu = 0$. $a = A$ is a candidate.

2) $\lambda > 0$, $\mu > 0$. This holds when $A = \frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2} \Leftrightarrow AT = \alpha$ which is a particular case of 1). Hence $a = A$ is again a candidate.

Formalizing the problem with respect to the second region, the FOC is:

$$-\frac{\alpha}{g} - \lambda + \mu = 0$$

One possible case exists for the lagrange multiplier:

1) $\lambda = 0$, $\mu > 0$. In this case, $a = \frac{gT+A+\sqrt{(gT+A)^2-4\alpha g}}{2}$ is a candidate.

Since the utility function is decreasing in a, $a = A$ is a candidate for the optimal ability.

Case b) $AT \leq \alpha$ and $(gT + A)^2 - 4\alpha g < 0$

The region to work with is $A \leq a \leq gT + A$, since the function always has positive values in this case. Since the utility function of the senior is decreasing in a, $a = A$ is a candidate for the optimal ability.

Case c) $AT > \alpha$ and $(gT + A)^2 - 4\alpha g \geq 0$

This is a particular case of case 1.a), where we only consider the second region, hence $a = \frac{gT+A+\sqrt{(gT+A)^2-4\alpha g}}{2}$ is a candidate for the optimal ability.

We now summarize the candidates for step 1:

$$\begin{array}{rcl} a & = & A \quad \text{if} \quad AT \leq \alpha \\ a & = & \frac{gT + A + \sqrt{(gT+A)^2 - 4\alpha g}}{2} \quad \text{if} \quad AT > \alpha \text{ and } (gT+A)^2 - 4\alpha g \geq 0 \end{array}$$

Step 2: $a^2 - (gT + A)a + \alpha g \leq 0$

The utility function to be considered in this case is $u^*(a,t) = \frac{(at+\alpha)^2}{4a}$.

Following the same logic as in step 1, we have 2 subcases to solve this problem: 2.a) when $AT \geq \alpha$ and $(gT+A)^2 - 4\alpha g \geq 0$; 2.b) when $AT \leq \alpha$ and $(gT+A)^2 - 4\alpha g \geq 0$. The other two subcases are impossible, since $(gT+A)^2 - 4\alpha g < 0$ means the function always has positive values. Hence, in this step we take as given that $(gT+A)^2 - 4\alpha g \geq 0$.

Case a) $AT \geq \alpha$

There is one region for a to work with: $A \leq a \leq \frac{gT+A+\sqrt[2]{(gT+A)^2-4\alpha g}}{2}$. The maximization problem is:

$$\underset{a,t}{Max} \left\{ \frac{1}{4a} \left(a(T - \frac{a-A}{g}) + \alpha \right)^2 \right\}$$

$$s.t. \quad a \quad \geq \quad A \qquad \text{and} \qquad a \leq \frac{gT+A+\sqrt{(gT+A)^2-4\alpha g}}{2}$$

The FOC is:

$$\frac{1}{a^2} \left(a(T - \frac{a-A}{g}) + \alpha \right) \left(a(T - \frac{3a-A}{g}) - \alpha \right) + \lambda - \mu = 0$$

There are 4 possible cases for the lagrange multipliers:

1) $\lambda > 0$, $\mu = 0$. Looking at the FOC, it must be that $a(T - \frac{3a-A}{g}) - \alpha < 0$, that is, $a \in (\frac{gT+A-\sqrt{(gT+A)^2-12\alpha g}}{6}, \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6})$, which is the case when $a = A$.

2) $\lambda = 0$, $\mu > 0$. In this case, $a = \frac{gT+A+\sqrt[2]{(gT+A)^2-4\alpha g}}{2}$ is a candidate if $a \in (\frac{gT+A-\sqrt{(gT+A)^2-12\alpha g}}{6}, \frac{gT+A+\sqrt[2]{(gT+A)^2-12\alpha g}}{6})$. We check that indeed a belongs to this interval. This holds only if $(gT+A)^2 - 12\alpha g \geq 0$.

3) $\lambda = 0$, $\mu = 0$. $a = \frac{gT+A\pm\sqrt[2]{(gT+A)^2-12\alpha g}}{6}$ are candidates, provided $(gT+A)^2 - 12\alpha g \geq 0$.

4) $\lambda > 0$, $\mu > 0$. This holds when $A = \frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2} \Leftrightarrow AT = \alpha$. Hence $a = A$ is a candidate again.

The utility function is decreasing from A until $\frac{gT+A-\sqrt{(gT+A)^2-12\alpha g}}{6}$, increasing from then on until $\frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6}$, and decreasing onwards. Hence, when $(gT+A)^2 - 12\alpha g \geq 0$ $a = \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6}$ is candidate for the optimal ability and when $(gT+A)^2 - 12\alpha g < 0$, $a = A$.

Case b) $AT \leq \alpha$

There is one region for a to work with: $\frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2} \leq a \leq \frac{gT+A+\sqrt{(gT+A)^2-4\alpha g}}{2}$. The FOC is the same as in case a), hence the possible cases for the lagrangean multipliers are:

1) $\lambda > 0, \mu = 0$. $a = \frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2}$ and it must be that $a \in (\frac{gT+A-\sqrt{(gT+A)^2-12\alpha g}}{6}, \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6})$. Since $\frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2} < \frac{gT+A-\sqrt{(gT+A)^2-12\alpha g}}{6}$, indeed it is a candidate.

2) $\lambda = 0, \mu > 0$. $a = \frac{gT+A+\sqrt{(gT+A)^2-4\alpha g}}{2}$ and it must be that $a \in (\frac{gT+A-\sqrt{(gT+A)^2-12\alpha g}}{6}, \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6})$. It is straightforward to check that a indeed belongs to this interval, so it is a candidate, provided that $(gT+A)^2 - 12\alpha g \geq 0$.

3) $\lambda = 0, \mu = 0$. In this case, $a = \frac{gT+A\pm\sqrt{(gT+A)^2-12\alpha g}}{6}$ provided that $(gT+A)^2-12\alpha g \geq 0$.

4) $\lambda > 0, \mu > 0$. In this case, $a = \frac{gT+A}{6}$, which happens when $(gT+A)^2 - 12\alpha g = 0$, a particular case of 3).

Analyzing all the candidates and the behavior of the utility function, $a = \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6}$ is a candidate for the optimal ability when $(gT+A)^2-12\alpha g \geq 0$ and $a = \frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2}$ when $(gT+A)^2 - 12\alpha g < 0$.

We now summarize all candidates for case 2:

$$
\begin{aligned}
a &= \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6} \qquad \text{if} \qquad (gT+A)^2 - 12\alpha g \geq 0 \\
a &= A \qquad \text{if} \qquad AT \geq \alpha \text{ and } (gT+A)^2 - 12\alpha g < 0 \\
a &= \frac{gT+A-\sqrt{(gT+A)^2-4\alpha g}}{2} \qquad \text{if} \qquad AT < \alpha \text{ and } (gT+A)^2 - 12\alpha g < 0
\end{aligned}
$$

As function of the parameters, we evaluate and compare the utility of the solutions attained in step 1 and step 2. The final solutions for a^* and t^* (attained recursively) are:

$(gT+A)^2 - 12\alpha g \geq 0$	$a^* = \frac{gT+A+\sqrt{(gT+A)^2-12\alpha g}}{6}$	$t^* = \frac{5(gT+A)-\sqrt{(gT+A)^2-12\alpha g}}{6g}$
$(gT+A)^2 - 12\alpha g < 0$	$a^* = A$	$t^* = T$

References

[1] Armbruster, L. (2008): "The Rise of The Post-Doc as Principal Investigator? How PHDs May Advance in Their Career and Knowledge Claims in the New Europe of Knowledge", *Policy Futures in Education* 6(4), 409-423.

[2] Banal-Estañol, A. and Macho-Stadler, I. (2010): "Scientific and Commercial Incentives in R&D: Research vs. Development", *Journal of Economics and Management Strategy* 19(1), 185-221.

[3] Cech, T. and Bond, E. (2004): "Managing Your Own Lab", *Science* June 18, Vol. 304, No. 5678, 1717.

[4] European Commission (2004): "*Increasing Human Resources for Science and Technology in Europe*", Report chaired by J.M. Gago, High Level Group (HLG) Human Resources for Science and Technology in Europe.

[5] Frame, I. and Allen, L (2002): "A Flexible Approach to PhD Research Training", *Quality Assurance in Education* 12(2), 98-103.

[6] Golde, C. (2000): "Should I stay or should I go? Student descriptions of the doctoral attrition process", *The Review of Higher Education* 23(2), 199-227.

[7] Green, S. and Bauer, T. (1995): "Supervisory Mentoring by Advisers: Relationships With Doctoral Student Potential, Productivity and Commitment", *Personnel Psychology* 48(3), 537-561.

[8] Kram, K. (1983): "Phases of the Mentor Relationship", *Academy of Management Journal* 26, 608-625.

[9] Kram, K. and Isabella, L. (1985): "Mentoring Alternatives: The Role of Peer Relationships in Career Development", *Academy of Management Journal* 28(1), 110-132.

[10] Holmstrom, B. and Milgrom, P. (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics and Organization* 7, 24-51.

[11] Levinson, D. (1978): "*The season's of a man's life*", New York: Knopf.

[12] Lacetera, N. and Zirulia, L. (2008): "Knowledge Spillovers, Competition, and Taste for Science in a Model of R&D Incentive Provision", Universita' di Bologna, Working Paper..

[13] Lovitts, B. E. (2001): "*Leaving the ivory tower: the causes and consequences of departure from doctoral study*", Lanham, MA, Rowman & Littlefield.

[14] Lovitts, B. E. (2005): "Being a good course-taker is not enough: a Theoretical Perspective on the Transition to Independent Research", *Studies in Higher Education* 30(2), 137-154.

[15] Murray, F. (2004), 'The role of academic inventors in entrepreneurial firms: sharing the laboratory life,' *Research Policy* 33(4), 643–659.

[16] Nerad, M. and Cerny, J. (1999): "Postdoctoral Patterns, Career Advancement, and Problems", *Science* September 3, Vol. 285. No. 5433, 1533-1535.

[17] Paglis, L., Green, S. and Bauer T. (2006): "Does Advisor Mentoring Add Value? A Longitudinal Study of Mentoring and Doctoral Student Outcomes", *Research in Higher Education* 47(5), 451-476.

[18] Pole, C., Sprokkereef, A., Burgess, Robert G., and Lakin, E. (1997): "Supervision of Doctoral Students in the Natural Sciences: expectations and experiences", *Assessment & Evaluation in Higher Education* 22(1), 49-63.

[19] Puljak, L. (2006): "Career Blocker: Bad Advisors" in http://sciencecareers.sciencemag.org.

[20] Stephan, P. E. and S.G. Levin, 1992, Striking the Mother Lode in Science: The Importance of Age, Place, and Time, New York: Oxford University Press.

[21] Vogel, G. (1999): "A Day in the Life of a Topflight Lab", *Science* September 3, Vol. 285, 1531-1532.

[22] Walckiers, A. (2008): "Multidimensional Contracts with Task-Specific Productivities: An Application to Universities", *International Tax and Public Finance* 15, 165-198.