JENA ECONOMIC RESEARCH PAPERS

2014 – 028

Taxation and incentives to innovate: a principal-agent approach

by

Diego d'Andria

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Taxation and incentives to innovate: a principal-agent approach

Diego d'Andria*

Abstract

A principal-agent multitasking model is used to explore the effects of different tax schemes on innovation in a pure knowledge economy. Corporate taxes and labor income taxes can affect both the firm owner's and the employee's incentives to commit to innovative tasks, when the former compensates the latter (a manager, technical or R&D employee) by means of variable pay tied to measures of the company's success.

Results point to a complementary role between "patent box" tax incentives and reductions in the tax rate levied on profit sharing schemes. This complementarity holds, albeit with different relative importance for the two tax incentives, also with non-deductible labor costs, with a stochastic innovation value coupled with a risk-averse agent, and with multiple principals competing for talented agents.

JEL Classifications: H2, O31, J33.

Keywords: tax incentives for R&D; patent box; principal-agent models; multitasking models; profit sharing schemes; incentives to innovate.

---

*Friedrich Schiller University of Jena (Germany), DFG Research Training Group "The Economics of Innovative Change". E-mail: diego.dandria@gmail.com
I am very grateful for the helpful comments and suggestions received from Uwe Cantner, Thomas De Haan, Andreas Haufler, Oliver Kirchkamp, Rene Levinski, Silke Uebelmesser, from participants at the Jena Economic Research Workshops, at the CGDE Workshop 2014 in Leipzig, and at the 70th IIPF Congress in Lugano. Of course, sole responsibility for any remaining shortcomings rests with the author. Financial support from the German Science Foundation (DFG RTG 1411) is gratefully acknowledged.

1 Introduction

Innovation is considered as one of the most important factors driving economic growth. Many countries enact policies that aim at supporting innovation and encourage firms to invest in R&D projects. The policy mix varies significantly, and while some countries rely on the use of direct funding, others prefer various forms of tax incentives.[1]

There are several rationales that have been provided for these policy interventions. First, innovation is the outcome of long, risky processes where the investment is highly uncertain. As a consequence, it is hard for investors to properly assess the value of an innovation *ex ante*, and even more to evaluate the chances of actually being successful in going from an initial conception phase to commercialization. This means that firms, especially small young ones, may face severe financial constraints. Second, R&D activities may generate positive spillovers in the form of non-rival (and not entirely excludable) knowledge, thus the private return for the innovator would be lower than the social return. More and better R&D means a country can be more competitive, maintain jobs, and export products with larger value-added.

Our focus is on people producing innovations within knowledge-intensive private companies. The production of knowledge is mainly a human process, conducted by highly skilled people usually possessing some form of advanced tertiary-level education, who are often highly specialized. Workers of this type who are able to produce the "big" innovations are few, and their number does not increase over time, at least not in the short run. While the power of taxation to affect the profitability of R&D projects for firms is well studied (for an example see the work of Hall and Van Reenen 2000), the effects on innovative workers and their incentives are not.

At least since the work of Balkin and Gomez-Mejia (1984) we know that

[1]Nearly all developed countries use some form of tax incentives for R&D, and these are almost exclusively in the form of reduced tax burdens on the side of corporate taxation (Evers et al. 2013). In recent times a specific tax incentive, a so-called "patent box", has gained momentum and has been adopted by 11 European countries. A patent box is nothing more than a reduced corporate tax rate applied to the revenues associated with a patented innovation (in some countries, other forms of intellectual property are also included under the more favorable tax rate). In comparison to other forms of R&D tax incentives, like deductibility for R&D expenditures or tax credits, the advantage of a patent box would be to target the revenue side rather than the cost side. This limits the problem with a possible reclassification of expenditures into R&D (obtained by manipulating accounting rules in order to claim the tax benefit), while reducing the tax distortion on the marginal profits obtained in case of a successful innovation.

companies use compensation schemes which tie part of the pay of R&D personnel and of specific key employees to the company's profits. Profit sharing schemes may include: stock grants, stock options, direct participation to profits, and bonus compensation. Subsequent work has confirmed the use of these schemes by innovative firms as a means to provide incentive to innovative employees for the U.S. (Ittner et al. 2003, Lerner and Wulf 2007, Francis et al. 2011) as well as for Europe (d'Andria and Uebelmesser 2014).

At first this might seem counterintuitive. Previous theories (Lazear 1986) would suggest that in an innovative corporate environment, uncertainty in outcomes and monitoring costs of a worker's effort are large, and consequently risk-averse employees would ask for some fixed or piece-rate compensation scheme tied to some observable input measure (i.e. salaries). Overall we would expect less use of profit sharing in innovative firms, in comparison to traditional firms.

More recent theories (built upon the seminal works of Holmstrom 1989 and Holmstrom and Milgrom 1991) argue that profit sharing schemes can be part of the solution to a problem of asymmetric information in a multitasking setting. This can be the case with R&D workers and managers when the innovation process relies upon exploration of new approaches and directions (Manso 2011), or when there is a tension between application of known practices versus exploitation of new opportunities (Hellmann and Thiele 2011). In such cases the principal can provide the "right" set of incentives to the agent by means of contracts made of a mix of fixed pay, variable pay linked to some observable individual measure, and variable pay tied to the overall performance of the firm. As a successful innovation can lead to a significant increase in a company's value, the latter compensation type becomes a tool to align innovators' incentives to the principal's.

Given that companies use compensation schemes to reward their R&D workers, technical employees and managers, different taxes could affect both the agent's and the principal's incentives to innovate. Few studies deal with the interactions between taxation and variable payment schemes. They are related to the banking sector and to non-innovative firms and they study the effects of bonus taxes (Grossmann et al. 2011, Radulescu 2012, Dietl et al. 2013). To our knowledge, the present study is the first to explicitly link taxation and variable pay, within a principal-agent multitasking setting. By taking into account simultaneously the effects of taxes on the employer and on the employee when profit sharing schemes are used as a means to foster innovation, we are able to analyze possible interactions between labor and corporate income taxation and the expected effects on the feasibility of innovations.

Our goal is to switch the tax debate from the external (financial) constraints that a company faces when innovating, to the internal (incentive) constraints and challenges given by its own workforce. The model we propose is particularly suited to describe a knowledge economy where human capital enters as the main input factor to the production of innovations. A real-world example can be the "New Economy" industry as observed in the last twenty years: the development of new software and Web-based services usually does not require major investments in physical capital in the early stages of development, but its feasibility is strictly dependent upon the availability of the right set of highly specialized human capital coupled with commitment to problem-solving and creative thinking.

To shed light on this matter we build a multitasking principal-agent model. The underlying structure of the model is based on the work of Hellmann and Thiele (2011): the agent is assumed to be able to invest effort in two competing tasks, where one is a standard task, and the other is a task that can produce an innovation. The principal is unable to directly observe the agent's levels of effort. Differently from Hellmann and Thiele (2011) the principal can provide two distinct forms of variable pay for the two tasks. Also differently, the policymaker can levy four tax rates on the income generated by the agent or the principal, and from the standard or the innovative task.

We show that under the assumption that the principal's gross profit from the standard task is not much larger than the related cost of labor for the same task and considering very valuable innovations, a tax incentive on profit sharing schemes can be a more powerful tool to foster innovation (not necessarily related to formal R&D activities) than a patent box incentive. When the expected value generated by the innovative task is not very high (we call these "marginal" innovations), the two policies appear to be complementary in making such innovations feasible. In an extension of the base model with risk-averse agents, the effects of a change in the labor tax are modified and can either be reduced, or when risk aversion is sufficiently large the relation between the tax and the incentives to innovate can become non-monotonic. A multi-principal setting where principals compete to attract highly-skilled agents is also discussed.

The structure of the paper is as follows. Section 2 presents the basic form of the model and derives a number of propositions. Section 3 discusses some implications of the model that are relevant for policy-makers. Section 4 extends the model by changing a number of key assumptions. Section 5 concludes.

2 The Base Model

The following model is based on the work by Hellmann and Thiele (2011). We extend and modify the original principal-agent model in two ways. In Hellmann and Thiele (2011) the division of the innovation value between the principal and the agent was assumed to be exogenous and stemming from different degrees of appropriability of the innovation value. In our model, on the contrary, we consider two distinct variable pay schemes that the principal can use in order to drive the agent's effort investments and which endogenously determine how the value of the innovation is divided. The focus of the analysis is also different, as we are interested in how taxation may shape the incentives to innovate, enforcing or alleviating the incentives stemming from the payment schemes. To this end we add four types of taxes introducing fiscal wedges for both the agent and the principal.

The principal is a profit-maximizing owner of a firm, who provides compensation to the agent in the form of a contract with two distinct types of pay, the amount of which depends upon variables that the principal cannot observe *ex ante*.

The agent is a worker who can invest effort into two distinct tasks, and we can interpret the agent as either an employee in R&D or technical functions, or as a manager: in all cases these are workers who are able to commit effort either in a known job task (exploiting knowledge generated in the past), or in an attempt to create new knowledge (in the form of new products, new processes, or improvements over existing products and processes), with some degree of freedom over such choice. The first task is "traditional" or standard and produces a binary signal $S = \{0, 1\}$ that the principal can observe *ex post*. The probability of getting $S = 1$ is linearly increasing in the amount of effort e_S invested by the agent in the standard task. When $S = 1$ the agent obtains a pay equal to β. The second task is innovative and produces, if successful, an innovation of value Y. The probability of success for the innovative task is a linear function of effort e_I invested by the agent in this task, scaled down by a factor $0 < n < 1$.[2] In case of success the agent receives γY and the principal retains $(1 - \gamma)Y$, while in case of failure both get zero. In this setting, β and γ represent the two elements of the compensation contract.[3] Both the principal and the agent are assumed to be risk-neutral

[2]This ensures that the probability of success for the innovation is always lower than the probability of getting signal $S = 1$ for the standard task, for equal levels of effort put in the task.

[3]If the profit sharing pay γ is interpreted as stock-based, then the product of the stan-

and gifted with perfect foresight about the expected value of Y.[4] Only the agent can observe a signal about the opportunity to produce an innovation of value Y, while the principal only knows about said opportunity *ex post* in case the innovation is successfully developed.

For analytical convenience we do not include any fixed pay. We also assume the agent has no reservation utility. This assumption does not change any of the following results, but of course it is not realistic as optimally-designed contracts should satisfy the additional constraint of a minimum utility level obtained by the agent. We assume, instead, that the agent has an exit option: if the principal offers him or her an undesired type of contract that pays for innovation, we assume the agent always has the possibility to leave the current principal for a competing firm who offers the desired type of contract paying for the standard task only. This equates to assume that there are some firms that cannot innovate, and just produce the standard good by means of the same production function as the firms that can innovate.

The novelty of the model is found in the inclusion of different types of taxes affecting directly either the agent or the principal. A tax on labor income is levied on the agent's earnings, which is made of two distinct tax rates, t_S and t_I, levied on the income earned respectively from the standard and the innovative tasks. Also, the principal's profits are subject to a tax rate, which again is made of differentiated rates τ_S and τ_I levied on the profits generated, respectively, by the standard and the innovative task.

A tax scheme featuring $t_S < t_I$ means there is some bonus tax in place that burdens profit sharing schemes more than normal labor compensations. On the contrary a tax scheme where $t_S > t_I$ resembles tax benefits that exist in some countries and that allow some tax reductions for specific types of profit sharing schemes (i.e. ISO stock options in the US, HMRC-approved schemes in the U.K.). With regard to corporate taxation, a tax scheme where $\tau_S > \tau_I$

dard task would also be captured by γ. But within the boundaries set by our assumptions, it is always possible for the principal to initially offer a contract where β is equal to the market wage and γ is just zero, in order to measure the average output from the standard task. Then in subsequent periods, an optimal contract with $\gamma > 0$ could be employed and applied to any profit generated above said average output. Analytically this would require to slightly modify the model so that Y would not represent the entire expected value of the innovation, but rather the mean of the distribution of the additional profits made on top of the average profits from standard tasks. This modification, though, would not affect the results in any meaningful way and therefore is omitted for the following discussion.

[4] As an alternative setting, the value Y represents the expected value drawn from some distribution. If both the agent and the principal share the same expectations and information about this distribution, because of their risk neutrality this reduces to the equivalent problem with a certain value Y discussed in the text.

is interpreted as a patent box type of R&D tax incentive.

The model assumes a single period of time divided into three stages. In the first stage the policymaker sets the four tax rates. In the second stage, the principal observes the tax rates, the value of the innovation that could potentially be produced, and then offers a single contract (β, γ) to the agent which maximizes expected profits. In the third stage, the agent has full knowledge about all the relevant variables. She maximizes her own utility by deciding about the amounts of effort e_S and e_I to invest. We proceed by backward induction to derive the optimal solutions for the agent, and then for the principal.

We do not include a formal maximization problem for the policymaker. We assume that the policymaker is maximizing welfare and is constrained in having to raise an exogenously given amount of total tax revenues by means of both a labor income tax and a corporate income tax in order to finance some public expenditure. We also assume that a market failure exists which leads private agents to produce less innovation in comparison to the social optimum (i.e. positive knowledge spillovers), or alternatively that taxation is itself distortionary along some other choice dimension of firms and workers, which again leads to a suboptimal amount of innovation. Therefore the policymaker introduces a tax incentive on the side of corporate taxation to foster aggregate innovation.

Our analysis assumes as a starting point a preexisting tax incentive for R&D on corporate income taxation in the form of a patent box. We look at small changes in tax rates where the patent box incentive is reduced, while a new tax incentive on profit sharing schemes is introduced so that total expected tax revenues stay unchanged. The latter tax incentive comes in the form of a reduced tax rate levied on personal income obtained by a worker by means of profit sharing compensation. The aim of the analysis is to assess the general direction of the change in the incentives to innovate, and consequently of the change in the expected production of innovation, therefore providing insight about the relative efficiency of said tax incentives.

Sections 2.1 and 2.2 present the optimization problems faced by the agent and the principal, respectively. The incentives they face, as well as the direct effects of taxation on the decisions about effort and contracts, will be discussed. Section 2.3 studies the global effects of taxation in the case where the expected innovation value is large. Some relevant interactions between the two tax incentives will be highlighted and discussed. Section 2.4 analyzes the same effects for the case of not-so-valuable ("marginal") innovations.

2.1 The Agent

The agent solves the following maximization problem by taking both the tax system and the compensation scheme as given, optimally choosing the effort levels:

$$\max_{e_S, e_I} U(.) = (1 - t_S)\beta e_S + (1 - t_I)\gamma Y n e_I - \frac{(e_S + e_I)^2}{2} \tag{1}$$

where a quadratic function for the private effort costs is employed, following the Principal-Agent literature (see for instance Grossmann et al. 2011). The probability of getting signal $S = 1$ is equal to e_S, and the probability of getting Y is equal to ne_I.

The first order conditions (FOCs) for problem (1) are the following:

$$e_S = (1 - t_S)\beta \tag{2}$$

$$e_I = (1 - t_I)\gamma Y n \tag{3}$$

The assumption of linearity in e_S and e_I for the probabilities of getting $S = 1$ and Y, respectively, provides us with a nice simplification, namely that the agent specializes either in the standard task, or in the innovative task. We can then define a threshold value Y_A such that for $Y < Y_A$ the agent prefers to specialize in the standard task, and for $Y > Y_A$ the agent specializes in the innovative task. The value of that threshold Y_A is defined as the value of Y for which the agent is indifferent between committing to either the standard or the innovative task. Y_A must therefore satisfy the following equality:

$$\beta(1 - t_S)e_S^* - \frac{(e_S^*)^2}{2} = \gamma Y_A n(1 - t_I)e_I^* - \frac{(e_I^*)^2}{2} \tag{4}$$

where e_S^* and e_I^* denote the optimal levels of effort chosen by the agent. Note that e_S^* does not depend on the value of Y when $Y < Y_A$, and when $Y > Y_A$, $e_S^* = 0$.

The solutions derived from these FOCs can be divided based on whether $Y > Y_A$, or $Y < Y_A$:

$$\text{if } Y > Y_A \Rightarrow e_S^* = 0;\ e_I^* = (1 - t_I)\gamma Y n \tag{5}$$

$$\text{if } Y < Y_A \Rightarrow e_S^* = (1 - t_S)\beta;\ e_I^* = 0 \tag{6}$$

Eq. (6) also expresses the optimal amount of effort for the case when the innovation signal observed by the agent is zero, so that no opportunity for innovation occurs. It easily follows from these FOCs that, if β and γ do not change with the tax rates:

$$\frac{\partial e_I^*}{\partial \gamma} > 0; \; \frac{\partial e_I^*}{\partial t_I} < 0 \tag{7}$$

$$\frac{\partial e_S^*}{\partial \beta} > 0; \; \frac{\partial e_S^*}{\partial t_S} < 0 \tag{8}$$

meaning that the effort invested in a task is increasing with the compensation (β or γ) offered for that task, and decreasing in the tax levied.

The following proposition defines how the threshold Y_A is affected by changes in the compensation contracts and in labor taxation. All proofs are provided in the Appendix.

Proposition 1: The following relations hold true:

$$i) \; \frac{\partial Y_A}{\partial \gamma} < 0; \; \frac{\partial Y_A}{\partial t_I} > 0$$

$$ii) \; \frac{\partial Y_A}{\partial \beta} > 0; \; \frac{\partial Y_A}{\partial t_S} < 0$$

2.2 The Principal

The principal maximizes expected profits by offering a compensation contract (β, γ) to the agent, taking the tax system as given:

$$\max_{\beta,\gamma} \; \mathit{\Pi}(.) = (1 - \tau_S)(V(e_S) - \beta e_S) + (1 - \tau_I)(1 - \gamma)Y n e_I \tag{9}$$

where $V(e_S)$ denotes the value obtained from the standard task, with $\frac{\partial V}{\partial e_S} > 0$ and $\frac{\partial^2 V}{\partial e_S^2} < 0$. The principal, similarly to what was previously discussed for the agent for the threshold Y_A, prefers the agent to innovate if Y is above a threshold value Y_P. The value of Y_P is defined by the equality:

$$(1 - \tau_S)(V(e_S^*) - \beta^* e_S^*) = (1 - \tau_I)(1 - \gamma^*)Y_P n e_I^* \tag{10}$$

where β^* and γ^* are the unconstrained optimal payment schemes obtained by solving problem (9).

The term "unconstrained" means that these contracts are obtained from a maximization problem where the agent could commit effort only in either the standard task or in the innovative task. In other words these contracts maximize expected profits obtained from each individual task, but they are together optimal for the principal only as long as, when offered to the agent, the latter opts for the principal's preferred task. Once an unconstrained optimal contract γ^* (respectively, β^*) is sufficient to make the agent choose the innovative (respectively, the standard) task, then the other compensation element β (respectively, γ) can be set to zero by the principal without any effect on the agent's decision (assuming the principal would be able to observe *ex ante* the same signal observed by the agent about the opportunity for an innovation). In a subsequent section more general solutions will be provided, and a constrained optimal contract with (β^*, γ^{**}) where $\gamma^{**} > \gamma^*$ will be shown to exist in cases where the distance between Y and Y_A is not large.

The FOCs for problem (9) are the following:

$$\beta = \frac{\partial V}{\partial e_S}\frac{\partial e_S^*}{\partial \beta} + \frac{(1-\tau_I)}{(1-\tau_S)}(1-\gamma)Yn\frac{\partial e_I^*}{\partial \beta} \tag{11}$$

$$\gamma = \frac{1-\tau_S}{1-\tau_I}\frac{\frac{\partial V}{\partial e_S^*} - \beta}{Yn} - \frac{e_I^*}{\frac{\partial e_I^*}{\partial \gamma}} + 1 \tag{12}$$

Similarly to the agent's optimization problem and because the agent chooses to invest in one type of task only, the principal's optimal contract is specialized based on the value of Y relative to Y_P. The unconstrained optimal contract (β^*, γ^*) is obtained by substituting (5) and (6) in place of e_S^* and e_I^*:

$$\text{if } Y > Y_P \;\Rightarrow\; \beta = 0; \gamma^* = \frac{1}{2} \tag{13}$$

$$\text{if } Y < Y_P \;\Rightarrow\; \beta^* = \frac{1}{2}\frac{\partial V(e_S^*)}{\partial e_S^*}; \gamma = 0 \tag{14}$$

If the principal knew *ex ante* that the innovation signal was positive, then he could just set $\beta = 0$ in case he expects $Y > Y_P$. But because the principal cannot know *ex ante* whether an opportunity for an innovation arose and was observed by the agent, the standard pay offered to the agent will be β^* (and not $\beta = 0$) also in the case $Y > Y_P$. The optimal unconstrained innovation contract is therefore (β^*, γ^*).

If we are in a scenario where Y is either much larger or much smaller than both Y_A and Y_P, it is immediate to verify that the optimal values set by the principal for β and γ are not affected by the value of the tax rates τ_S and τ_I.

Proposition 2: With non-confiscatory tax rates ($\tau_S < 1$ and $\tau_I < 1$), there exists a value $Y = Y_{high}$, where $Y_{high} >> Y_A$ and $Y_{high} >> Y_P$, for which the principal's optimal contract is $\left[\beta = \beta^*, \gamma = \frac{1}{2}\right]$ and it is invariant to profits taxation. There is a value $Y_{low} << Y_A$ and $Y_{low} << Y_P$ for which the principal's optimal contract is $\left[\beta = \frac{1}{2}\frac{\partial V(e_S^*)}{\partial e_S^*}, \gamma = 0\right]$ and it is invariant to profits taxation.

In the scenario where Y is extremely high, therefore, the principal and the agent both commit to innovation irrespectively of the level of the tax burden on profits.[5] An analogous reasoning applies when Y is extremely low, and no innovation happens regardless of taxation.

It remains to see how tax rates τ_S and τ_I affect Y_P in the more general case where the distance between Y and Y_P is not extremely large. The following proposition summarizes the results.

Proposition 3: The following relations hold true:

$$i)\ \frac{\partial Y_P}{\partial \tau_I} > 0;$$

$$ii)\ \frac{\partial Y_P}{\partial \tau_S} < 0;$$

$$iii)\ \text{If } \tau = \tau_S = \tau_I,\ \text{then } \frac{\partial Y_P}{\partial \tau} = 0$$

.

A higher tax rate on innovation profits, τ_I, thus reduces the willingness for the principal to invest in innovation, while a higher tax rate on standard profits, τ_S, on the contrary provides stronger incentives to commit to innovation. As stated in Proposition 3.iii, a flat-rate tax on profits does not modify the principal's willingness to invest in innovation.

[5]Note that although any positive value of t_I cannot make the agent switch to the standard task when $Y >> Y_A$, this tax still affects the degree of effort put into the innovative task. To see this, note that because $\frac{\partial \gamma^*}{\partial t_I} = 0$ (from eq. (13)), then it is straightforward to verify $\frac{\partial e_I^*}{\partial t_I} = -\gamma^* Y n < 0$.

2.3 Interactions between labor taxation and the principal's incentives to innovate when the innovation is non-marginal

We now turn to a set of scenarios where the value Y of the innovation is high enough to make, absent any taxation, the principal commit to provide incentives to innovate to the agent. We define such highly valuable innovations as "non-marginal" and, as stated by Proposition 2, such large expected value makes the principal offer an unconstrained contract $\left[\beta = \beta^*, \gamma = \frac{1}{2}\right]$. We will look at the case where it is verified that $Y_A(\gamma) > Y_P(\gamma)$. The latter is probably the most interesting case to study: the principal has an interest in driving the agent toward the innovative task as long as $Y > Y_P$. With an optimally designed contract, the agent's threshold Y_A can be driven below Y, so that innovation occurs.

We know already (from Hellmann and Thiele 2011, lemma 4) that starting from $\gamma = 0$, $\frac{dY_P}{d\gamma}$ is first negative and then positive as γ approaches 1, or equivalently, $Y_P(\gamma)$ follows a U-shaped curve, with its minimum at the unconstrained optimal value γ^*. This can be seen by decomposing $\frac{\partial Y_P}{\partial \gamma}$ similarly to what was done in eq. (24), and observing that because at the unconstrained optimal compensation γ^* the value of Π^I is maximal while Π^S does not vary with γ, then $Y_P(\gamma^*)$ must necessarily be a minimum. Intuitively, the idea is that at some point, increasing γ would make the loss in expected profits greater than the gains due to additional effort.

On the contrary $\frac{dY_A}{d\gamma}$ is always negative so that $Y_A(\gamma)$ is a monotonically decreasing function. This can be seen immediately from eq. (4), as a larger share of the innovation value always increases the value of the innovative task for the agent, relative to the value of the standard task. Overall by increasing the share of the innovation value given to the agent above γ^*, the principal makes the innovative task more appealing for the agent but less for herself.

In Hellmann and Thiele (2011) a proof is provided (see Proposition 2 in said paper) that $Y_A(\gamma)$ and $Y_P(\gamma)$ cross once. We assume in the following discussion that the return to the standard task for the principal at e_S^* is close in value to the return expected by the agent, so that it is always verified that $\frac{V(e_S^*(\gamma^*))}{\beta^* e_S^*(\gamma^*))} < 2$. In this way the crossing point between $Y_A(\gamma)$ and $Y_P(\gamma)$ lies on the right side of γ^*, which makes the model more interesting while not changing the main arguments here discussed.[6] See the curves for $Y_A(\gamma)$ and

[6]If one obtains $Y_A = \frac{\beta^* e_S^*(\gamma^*)}{\gamma^* n e_I^*(\gamma^*)}$ from eq. (4) following a similar reasoning to the expla-

$Y_P(\gamma)$ drawn in Figure 1 for an illustration.

<Figure 1 goes about here>

In previous sections we discussed some effects of taxation on Y_A and Y_P. From Proposition 1 we know the direction of the change in Y_A caused by a modification of the tax rates on labor income. From Propositions 2 and 3, we learned the effect of a change in profit taxation on Y_P. We have seen that profit taxation has no direct effect on the optimal choice of the agent, but what about the effects of labor taxation on Y_P? The following proposition explores this question.

Proposition 4: With a payment scheme optimally set to the unconstrained optimum γ^*, the following relations hold true:

$$i)\ \frac{\partial Y_P(\gamma^*)}{\partial t_S} < 0;$$

$$ii)\ \frac{\partial Y_P(\gamma^*)}{\partial t_I} > 0;$$

Proposition 4 highlights an asymmetry that exists between the effects of tax incentives on labor income and on profits. While a reduction in τ_I was shown to affect Y_P only, a reduction in t_I affects both Y_A and Y_P. From the perspective of the policymaker, this means that in principle τ_I could be increased and t_I could be lowered, in a way that makes the incentives threshold represented by Y_P for the principal stay the same, while lowering the incentives threshold Y_A for the agent.

nation in the proof of Proposition 1, and $Y_P = \frac{V(e_S^*)-\beta^* e_S^*(\gamma^*)}{(1-\gamma^*)ne_I^*(\gamma^*)}$ from eq. (10), and then compares them at $\gamma^* = \frac{1}{2}$, it is immediate to see that with no taxation $Y_A(\gamma^*) < Y_P(\gamma^*)$ *iff* $\frac{V(e_S^*(\gamma^*))}{\beta^* e_S^*(\gamma^*))} > 2$. This means that without taxes the curve $Y_A(\gamma^*)$ lies below $Y_P(\gamma^*)$ only when the profits earned by the principal from the standard task (net of all other costs, as we did not include other production factors in the model), are more than double the labor costs paid for it. It is a possible scenario, but it is quite rare to see this in the real world, as the standard task is meant to represent some well-known (mature) product or process. Our focus on the case where $Y_A(\gamma^*) > Y_P(\gamma^*)$ is likely to represent the one most often met in real enterprises.

2.4 Marginal innovations

We now focus on "marginal" innovations, which we define as innovations whose expected value is lower than $Y_A(\gamma^*)$. We therefore drop the assumption previously employed in Proposition 2 ($Y >> Y_A$ and $Y >> Y_P$), and consider innovations which are still highly valuable, but not as much.

Marginal innovations, as we are going to discuss in the present section, may lead the principal to offer a different contract where the pay for the standard task is again β^*, but the performance-related pay is some γ^{**} with $\gamma^* < \gamma^{**} < 1$. Differently from the previous case where the expected innovation value was high enough so that the optimal contract always featured $\gamma^* = 0.5$ irrespective of tax rates, with marginal innovations different tax rates can affect how γ^{**} is set by the principal.

The unconstrained optimal innovation contract we derived earlier is ($\beta^* = \frac{1}{2}\frac{\partial V(e_S^*)}{\partial e_S^*}, \gamma^* = \frac{1}{2}$). This contract was obtained in a context where for such a value of γ, $Y > Y_A$ and $Y > Y_P$. Let us now consider the case where at $\gamma = \frac{1}{2}$, we have $Y_A > Y > Y_P$. This would imply that the agent chooses the standard task or, if $\beta < \beta^*$, leaves the principal for a competitor offering the optimal contract β^*. As long as there is room for it, the principal could then forgo an additional share of the expected profits from the innovation, in order to lower Y_A. If there is some level of $\gamma > \frac{1}{2}$ for which $Y = Y_A \geq Y_P$, then the new "constrained" optimal contract will be characterized by this value, which we label γ^{**}. From eq. (4) by substituting e_I^* and e_S^* from eqs. (5) and (6), we can write the value of the constrained-optimal contract γ^{**} as:

$$\gamma^{**} = \frac{1 - t_S}{1 - t_I}\frac{\beta^*}{Yn} \tag{15}$$

where, as before, β^* indicates the unconstrained optimal compensation scheme for the standard task. γ^{**} is decreasing in Y, and it increases with the difference $t_I - t_S$. In the case of a constrained-optimal innovation contract therefore, differently from the unconstrained case discussed earlier, labor tax rates directly modify γ^{**}.

In a setting where $Y_A > Y > Y_P$ at γ^*, remembering that the agent is assumed to always be able to leave for a competitor where she can obtain a contract β^*, a constrained innovation contract γ^{**} is feasible for the principal as long as the profits from innovation exceed the profits from the standard task:

$$(1 - \tau_2)(1 - \gamma^{**})\gamma^{**}Y^2n^2 \geqslant (1 - \tau_1)[V(\beta^*(1 - t_S)) - (\beta^*)^2(1 - t_S)] \tag{16}$$

where eq. (16) is derived by substituting the values of e_S^* and e_I^* from eqs. (5) and (6) into eq. (10).

From a policymaker's perspective it may be important to see how different taxes affect both Y_A and Y_P. The reason is that for marginal innovations, the distance $Y_A - Y_P$ can affect the feasibility of a constrained optimal innovation contract. The following Lemma extends Proposition 4 to the constrained optimal contract.

Lemma 1: With a payment scheme optimally set to a constrained optimum γ^{**}, the following relations hold true:

$$i)\ \frac{\partial Y_P(\gamma^{**})}{\partial t_S} < 0;$$
$$ii)\ \frac{\partial Y_P(\gamma^{**})}{\partial t_I} > 0;$$

Lemma 1 shows that the asymmetric effect of a change in t_I or τ_I holds also for the constrained optimal contract γ^{**}. Again, a reduction of t_I lowers both thresholds Y_P and Y_A, while a reduction of τ_I only reduces Y_P.

3 Policy implications

We are now ready to derive some policy implications. In the following sections we will discuss three distinct cases: one with a highly valuable innovation, a second case with a marginal innovation, and a third case where the expected innovation value is very low.

As an underlying framework we assume that the policymaker aims at collecting a fixed exogenous tax revenue target which is used to finance a public consumption good entering additively into the agent's and the principal's payoff functions. The policymaker also wants to increase aggregate innovation produced by the economy to overcome some market failure that makes it sub-optimal. To this end, a patent box incentive was introduced. The main research question is whether it is possible to obtain the same tax revenues and equal or more innovation with a policy reform substituting some preexisting patent box incentive with a tax incentive on profit sharing schemes.

Section 3.1 will address the effects of the reform on innovation. Section 3.2 discusses the effects on tax revenues. It will be shown that it is always possible to substitute some profit tax incentive with some profit sharing tax incentive while keeping the same goal for tax revenues to be collected. Section 3.3 summarizes the implications for policymaking and discusses some additional elements that are not explicitly represented in the formal model.

3.1 Effects of the reform on innovation

Consider as a starting point a situation without any labor tax incentive. We assume that preexisting taxation is such that a positive flat-rate labor tax $t_S = t_I$ and a positive corporate tax τ_S plus some incentive rate for R&D $\tau_I < \tau_S$ are levied in order to obtain some exogenous revenue target. This is the policy mix most often encountered in real countries. Figure 1 depicts the incentives Y_A and Y_P before introducing any tax incentive on the side of labor taxation. If we now apply a tax incentive in the form of a reduced tax rate t_I, from Propositions 1 and 4 we expect both the curves $Y_A(\gamma)$ and $Y_P(\gamma)$ to shift downward. This scenario is the one represented in Figure 2.

If the value Y of the innovation is very high and $Y > Y_A$, innovation would occur without any tax incentive, notwithstanding whether it is in the form of a reduced t_I or τ_I. The level of effort e_I^* invested by the agent (from eq. (7)) would be lower than in the case with the labor tax incentive. Therefore, for highly valuable innovations it can be efficient to slightly increase τ_I and to reduce t_I in order to increase the probability that an innovation is successfully produced.

If the value Y is lower than (and close to) Y_A at γ^*, with no labor tax incentive the principal chooses a constrained-optimum compensation scheme $\gamma^{**} > \gamma^*$. In the latter scenario innovation still occurs, but the principal obtains a smaller share of its value if successful. By reducing t_I the intensity of the change in e_I^* in comparison to the scenario with $t_S = t_I$ is generally ambiguous: on the one hand, a reduction in t_I would increase effort; but the consequent reduction in γ from γ^{**} to γ^* provides smaller incentive to the agent to commit effort to the innovative task, therefore counteracting to some extent the reduction in effort. With the functional forms used here, in this case e_I^* does not change with t_I: this can be seen by substituting eq. (15) into the optimal effort e_I^* given by eq. (5), and noticing that $(1 - t_I)$ at numerator and denominator cancel out.

<Figure 2 goes about here>

Therefore for marginal innovations the two incentives, on tax rate t_I and on τ_I, are perfect substitutes as long as changes in profit sharing schemes (that is, in the value of γ^{**}) have no further implications for the policymaker[7], and as long as the reduction in t_I is small enough so that the potential increase in effort is fully offset by a reduction in compensation γ^{**}. But, if the reduction in t_I is so large that the offered contract is set to the unconstrained optimum γ^*, then further reductions in t_I will again make the agent commit more effort to the task.

In the case depicted in Figure 2, the tax rate t_I could be slightly reduced and substituted by some larger tax rate on profits τ_I. If the innovation is socially valuable, then such a reform may be desirable, as it could allow to obtain the same total tax revenues and the same amount of innovation, while reducing distortions brought on other dimensions (not included in this model, but which could in principle exist) by the profit tax incentive.[8]

It might be the case that the value Y is so low that some marginal innovation does not take place at all without tax incentives. This is represented in Figure 3, where for any level of γ that is high enough to make the agent commit to the innovative task (at any point where $Y > Y_A$), the crossing point where $Y = Y_A$ lies below the dashed curve $Y_P(\gamma)$.

<Figure 3 goes about here>

In the case represented in Figure 3, in order to make the innovation feasible, again the policy-maker can reduce t_I. Alternatively, given the assumption that a preexisting profit tax $\tau_S > \tau_I$ is also levied and the outcome is the one depicted in Figure 3, the policy-maker could further reduce τ_I. In both cases the constrained-optimal contract γ^{**} is going to be above the unconstrained optimal γ^* (as long as the labor tax incentive is not so large that it drives the offered compensation scheme to γ^*). A reduction of t_I, though, is generally more effective than a reduction of τ_I: it reduces both Y_A and Y_P, while a reduction of τ_I would only affect Y_P.

[7] As a side note, it may be interesting to highlight that changes in the value of γ would not affect social well-being, if one measures social welfare as a sum of individual utilities and therefore without any distributional concern. This is because different values of γ represent only how the principal and the agent divide the produced value among themselves.

[8] The latter observation rests on the assumption that the deadweight burden of taxation increases quadratically with the tax rate, as per traditional public finance arguments. If this is the case, then it may be more efficient to substitute some reduction in τ_I with some reduction in t_I even if the latter is in itself distortionary along some different behavioral dimension.

3.2 Effects of the reform on tax revenues

In order to better understand the revenue effects of a reform substituting some patent box incentive with some tax incentive on profit sharing schemes, the following equation represents total tax revenues G expected to be raised by the policy maker in the case with very high Y:

$$G = Yn[t_I\gamma^* + \tau_I(1-\gamma^*)]e_I^* \tag{17}$$

Here the innovation contract is constant and set to $\gamma^* = 0.5$, and a small reduction of the patent box incentive (which translates into an increase of τ_I) will raise expected tax revenues. The introduction of a profit sharing scheme incentive (a reduction of t_I) on the one hand reduces revenues through the term $t_I\gamma^*$, but it also increases revenues by making optimal effort e_I larger. It is then straightforward to see that for any small increase in τ_I, some reduction in t_I exists that keeps G constant. As $\frac{\partial G}{\partial t_I} = (Yn\gamma^*)^2(-t_I^2 - t_I + 1)$, the loss in revenues is marginally decreasing.

With marginal innovations eq. (17) changes to:

$$G = Yn[t_I\gamma^{**} + \tau_I(1-\gamma^{**})]e_I^* \tag{18}$$

where the contract γ^{**} is now affected by a modification of t_I. While an increase of τ_I has the same effect as in the previous eq. (17), a reduction of t_I now also reduces the optimal contract γ^{**}. As stated before, effort in this case remains unchanged. This means that although it is still always feasible to exchange some increase of τ_I with some decrease of t_I while keeping G constant, the marginal loss in revenues due to a reduction of t_I can be larger than the case in eq. (17) if $t_I > \tau_I$ (or lower if $t_I < \tau_I$), because part of the taxable value of the innovation will shift to profit taxation from the (now more favorable) labor tax.

Figure 4 plots the couples of values of t_I and τ_I that would obtain the same expected tax revenues raised in an initial benchmark case with a patent box set at $\tau_I = 0.10$, a profit tax $\tau_S = 0.27$, a tax on labor $t_S = 0.40$, a scaledown parameter $n = 0.5$, and no incentive on profit sharing schemes. The expected value Y of the innovation is set starting from very high numbers (top-left graph) to marginal values (bottom-right graph).

<Figure 4 goes about here>

As discussed in the text, above a given innovation value the innovation contract is set by the principal to the unconstrained γ^*, and revenue-neutral changes in the tax rates must be the same irrespective of the value of Y. This is represented in the two top graphs of Figure 4, which are just identical (even though the expected tax revenues in the first graph are larger). When the innovation is marginal (this is represented in the two bottom graphs of Figure 4), the set of possible revenue-neutral pairs of rates (t_I, τ_I) gets smaller as the innovation value is lower, and further reductions in t_I would require a bonus tax on profits $\tau_I > \tau_S$.

3.3 Further considerations

The present discussion makes a case for a reduction of labor taxation burdening profit sharing scheme compensations. While a "patent box" type of tax incentive is helpful in making room for marginal innovations, a reduction of t_I can be at least as effective and also induce higher commitment of effort for the very valuable, non marginal innovations. A substitution of some patent box incentives with said incentives can be revenue-neutral. Tax incentives on profit sharing schemes can at the same time increase the expected profitability of more innovative projects for firms (as traditional tax incentives for R&D would do), while making innovative employees more committed to pursue valuable innovations.

Similar arguments can be employed against the use of "bonus taxes" that apply a tax rate $t_I > t_S$, for example in cases where the incentive pay γ is provided as bonus pay, stock options, or direct participation to profits. A bonus tax would reduce agent's effort for the most valuable innovations and make some marginal innovation not being developed at all. Moreover, if the intent of a bonus tax is to reduce the employment of such a kind of compensation schemes, our model shows that the opposite may be true, as the principal is pushed toward using a larger γ^{**} in order to motivate the agent to commit effort in the innovative task when the innovation value is marginal.

As a concluding remark, we would like to spend some few words on the design of tax incentives for R&D. R&D tax incentives granted to firms necessitate strict definitions in order to limit reclassifications of accounts and, generally, to limit tax avoidance. Incentives are almost always conditioned upon severe limitations on the expenses which can be listed as related to R&D. Patent box incentives require that the innovation is one that can be protected by an intellectual property right (IPR). But not all innovations fulfill these prerequisites, and those that do not are basically not affected by existing tax incentives.

On the contrary, incentives described here on the side of profit sharing schemes are in principle able to capture also those innovations that cannot be covered by IPRs, or that cannot be associated with any formal R&D expenditure. Organizational innovations, or process innovations that are not technologically viable for a patent grant, would still improve the profitability of the firm and, therefore, innovative employees have incentives to commit effort to their production, if profit sharing schemes are offered to them. This means that, as long as the assumption $Y_A(\gamma^*) > Y_P(\gamma^*)$ is met, a reduction in t_I would make some of these marginal innovations that cannot be affected by R&D incentives on profits, feasible. While this kind of distinction between types of innovations is not formally made in the model presented here, nevertheless it is worth stressing that its existence may make a reduction of labor taxation burdening incentive pay on the innovative task a better tool (and not a substitute for) in comparison to reductions in the corporate income tax.

4 Model extensions

The following sections modify or extend the base model in distinct ways. Some critical assumptions are changed in order to see how this would affect the effectiveness of different tax incentives.

4.1 Non-deductible labor costs from the corporate tax base

We assume here that profit taxation reduces the return to innovation also for the agent. This could be closer to a situation where the agent's reward from the innovative task is not deductible from the corporate tax base.[9]

The agent maximizes the following:

$$\max_{e_S, e_I} U(.) = (1 - t_S)\beta e_S + (1 - t_I)(1 - \tau_I)\gamma Y n e_I - \frac{(e_S + e_I)^2}{2} \tag{19}$$

where the compensation obtained for the innovative task is now reduced also by $(1 - \tau_I)$ in addition to $(1 - t_I)$. This changes the optimal level of effort

[9]This might apply, for instance, in case the agent's reward is provided as stocks, or as a direct participation to the revenues generated by patent licensing.

in case the innovative task is chosen to: $e_I^* = \gamma Y n(1 - t_I)(1 - \tau_I)$, and it is straightforward to verify that $\frac{\partial e_I^*}{\partial \tau_I} < 0$ and $\frac{\partial Y_A}{\partial \tau_I} > 0$.

The following proposition summarizes the effect of a change of τ_I on Y_P.

Proposition 5: With non-deductible labor costs for the innovative task, it is true that $\frac{\partial Y_P}{\partial \tau_I} > 0$.

By comparing the value of $\frac{d\Pi^I}{d\tau_I}$ in Propositions 3 and 5 and substituting $\frac{\partial e_I^*}{\partial \tau_I}$, we see that the effect of the rate τ_I is stronger with non-deductible labor costs by an addendum equal to $-\gamma^*(1-\gamma^*)Y^2 n^2 (1-t_I)(1-\tau_I)$. A patent box policy is more effective in this scenario than the case represented in the base model, because it can have a larger effect on the principal's incentive to innovate, and also a positive effect on the optimal effort choice of the agent (because as we saw before, $\frac{\partial e_I^*}{\partial \tau_I} < 0$). The asymmetry found in the base model where only modifications in t_I could affect both Y_A and Y_P, does not hold anymore when labor costs are non-deductible.

4.2 Stochastic innovation and risk aversion

Previous models assume risk-neutrality for both the principal and the agent following the prevalent literature. In reality, though, often the agent is unable to diversify the investments and is risk averse.

To see how risk aversion might influence previous results, the value of the innovation is now stochastically determined, and the agent is assumed risk-averse. To model this scenario, we substitute Y with a function $\phi(\bar{Y}, \sigma)$ where the possible values for Y are distributed according to a Rectified Gaussian distribution (where negative values of the corresponding non-rectified Normal Distribution are set to zero) with mean $\bar{Y} = E(Y)$ and standard deviation σ.

In order to introduce risk-aversion for the agent, we rewrite her maximization problem as follows:

$$\max_{e_S, e_I} U(.) = (1 - t_S)\beta e_S + (1 - t_I)\gamma \bar{Y} n e_I - \delta(.) - \frac{(e_S + e_I)^2}{2} \tag{20}$$

where we introduced some function $-\delta(.)$ that is equal to zero for any σ if the agent is risk-neutral, while it takes increasingly negative values as her risk-aversion and σ get larger. The assumption taken here is that the agent and the principal possess perfect knowledge about $\bar{Y}$ and σ, and about the value that function $\delta(.)$ obtains for the agent for the specific σ related to the innovative task. The value of $-\delta(.)$ is interpreted as the risk premium that the agent would demand in order to be indifferent between choosing to invest effort in an innovative task with value given by a density function $\phi(\bar{Y}, \sigma)$ or in a non-stochastic innovation with certain value $Y = \bar{Y}$.

An example for the function $\delta(.)$ is the following one, which is derived by assuming a constant absolute risk-aversion (CARA) utility function (see Bolton and Dewatripont 2005 and Grossmann et al. 2011):

$$\delta(.) = \frac{r}{2}(1 - t_I)^2 \gamma^2 n^2 e_I^2 \sigma^2 \tag{21}$$

where $r > 0$ is a parameter reflecting the agent's absolute risk aversion. We will employ this functional form in order to bring forward the present discussion.

Some effects of the current changes in comparison to the base model are immediate to see. First, in a no-taxation setting the risk-averse agent will face lower incentives to innovate. Second, the pro-innovation effects of a reduction in the tax rate t_I are now offset by an increase in the term δ, and the latter term changes quadratically. This makes the effects of a change in t_I more difficult to assess without providing a value to the risk-aversion parameter r, because from some point onward the derivatives $\frac{\partial Y_A}{\partial t_I}$ and $\frac{\partial Y_P}{\partial t_I}$ could even switch sign from being positive to negative. Only as long as the inequality $r < \frac{Y}{(1-t_I)\gamma n e_I^* \sigma^2}$ holds at Y_A, the relation between Y_A and t_I remains positive and monotone. In very general terms we can assert that a reduction in t_I is now, in comparison to the base model, less effective as a tool to foster innovation.

With risk-aversion a flat-rate tax $t = t_S = t_I$ can also distort the agent's incentives, making the innovation task more attractive. This immediately descents from eq. (4), by noticing that for a risk-neutral agent an increase in t does not affect Y_A, because if for some value of t it is true that $\beta(1 - t) < \gamma Y_A n(1 - t)$ (a similar argument holds for the opposite inequality sign), then any other value of t will not change the inequality sign. That is, for each possible level of effort, the task with the highest expected return remains the same one. On the contrary with risk-aversion, any modification of t affects the left-hand side of eq. (4) linearly, while affecting the right-hand side quadratically. This leads to the following lemma:

Lemma 2: With risk-aversion expressed in the form of eq. (21) and a flat-rate labor tax t, it is verified that $\frac{\partial Y_A}{\partial t} < 0$

On the side of the principal, agent's risk aversion reduces the effectiveness of the innovation contracts, so that for any level of γ^* or γ^{**} and all other things the same, the amount of effort e_I^* will be lower and the threshold Y_A higher in comparison to the base model. As γ^{**} gets larger than γ^*, it will cross with a flatter curve $Y_A(\gamma)$ which makes it less likely that a marginal innovation will be feasible for some γ^{**}. The effectiveness of a patent box policy is in this case the same as for the base model, with the exception that some marginal innovations which could be made feasible in the base model by a patent box taxation, will not be feasible with risk aversion.

Changes in the labor tax rates can affect the principal's incentives to innovate differently with risk aversion than in the base model. Proposition 4 shows that $\frac{\partial Y_P}{\partial t_I} > 0$ as long as it is true that $\frac{\partial e_I^*}{\partial t_I} < 0$ and $\frac{\partial \gamma^*}{\partial t_I} = 0$. With a large enough risk aversion parameter r, the derivative $\frac{\partial e_I^*}{\partial t_I}$ is first negative and then positive as t_I is increased above zero. This would translate in an inverse U-shaped relation between t_I and Y_P.

In the case when the optimal innovation contract is γ^{**}, Lemma 1 shows that $\frac{\partial Y_P}{\partial t_I} > 0$ as long as $\frac{\partial e_I^*}{\partial t_I} < 0$ and $\frac{\partial \gamma^{**}}{\partial t_I} > 0$. Again, for some values of the parameter r, the derivative $\frac{\partial e_I^*}{\partial t_I}$ could switch sign. Moreover, the effects of risk-aversion could be so strong, that an increase in t_I could make the reduction of δ larger than the reduction in the expected gain in case of success from the innovative task, thus making the derivative $\frac{\partial \gamma^{**}}{\partial t_I}$ negative.

Taking stock from this discussion, the general policy-relevant effect of risk aversion is that a reduction in t_I is either affecting effort e_I^* in a weaker way (in comparison to the benchmark case given by the base model); or, if risk aversion is strong enough to produce non-monotonic relationships between t_I and e_I^*, risk aversion requires to levy either a very low t_I, or a very high t_I with the aim to exploit the fact that the tax now provides positive incentives to innovate to both the agent and the principal, thanks to its variance-reduction effect.

4.3 Opportunity signals with information about the value of the innovation

In the base model a signal was assumed that tells the agent whether an opportunity for an innovation is available to him. But it can be argued that often the agent is in a better position and can assess the approximate value of the innovation, before the principal could do the same. This situation is more likely to arise when the agent is an expert in his field, so that *ex ante* evaluation of the value is made according to private knowledge that is not fully transferable to the principal.

To inquire how such a scenario would modify previous propositions, we now assume the agent observes a more informative signal A, which is not binary anymore but provides, with 100% accuracy, the value of the innovation to be produced. Signal A takes value $A = 0$ if no opportunity for an innovation is present, and $A = y$ when an opportunity to produce an innovation of value y is observed. The value y is again drawn from a distribution $\phi(\bar{Y}, \sigma)$ known to both the principal and the agent.

The agent maximizes the following:

$$\max_{e_S, e_I} U(.) = (1 - t_S)\beta e_S + (1 - t_I)\gamma y n e_I - \frac{(e_S + e_I)^2}{2} \tag{22}$$

where y is now an exact value for the innovation, given by signal A.

The information set available to the principal is the same as in the base model. The difference lies in the fact that an occurrence of a value $Y_P < y < Y_A$ would drive the agent's choice toward the standard task, while the principal would still prefer the innovative task if $\bar{Y} > Y_P$. If $\bar{Y} > Y_P$ at γ^*, the principal can be better-off by increasing compensation γ above the optimal values derived in the base model, up to the point where:

$$\max_{\beta,\gamma} \mathit{\Pi}_S(\beta) \int_0^{Y_A(\gamma)} \phi(y; \bar{Y}, \sigma) dy + \int_{Y_A(\gamma)}^{+\infty} \phi(y; \bar{Y}, \sigma) \mathit{\Pi}_I(\gamma, y) dy \tag{23}$$

where $\mathit{\Pi}_S(\beta) = (1 - \tau_S)(V(e_S^*) - \beta e_S^*)$ and $\mathit{\Pi}_I(\gamma, y) = (1 - \tau_I)(1 - \gamma) y n e_I^*$.

Generally the solution to problem (23) is a contract made of a wage β^* that is the same as the one derived in the base model, and of a compensation γ that is no lower than γ^*. This is immediate to see by substituting (from eq. (5)) $e_I^* = (1 - t_I)\gamma y n$ into the right-hand side of eq. (23) and rewriting it as:

$$\gamma(1 - \gamma) n^2 (1 - t_I)(1 - \tau_I) \int_{Y_A(\gamma)}^{+\infty} \phi(y; \bar{Y}, \sigma) y^2 dy$$

and noticing that the part to the left of the integral is decreasing as γ gets larger (or lower) than the unconstrained optimal γ^*, while the quantity $\int_{Y_A(\gamma)}^{+\infty} \phi(y; \bar{Y}, \sigma) y^2 dy$ is always increasing in γ.

In this setting, a reduction of the tax rate τ_I lowers Y_P as in the base model by increasing Π_I for each value of y. A reduction of t_I increases effort invested by the agent in the innovative task for each possible value y, thus lowering both Y_A and Y_P as in the base model.

4.4 Mobile agents and competition over hiring

The base model is built on the assumption that there is one agent for one principal, and the agent can always leave the current principal for a competitor offering the unconstrained-optimal compensation β^*. But there might be multiple principals, all capable to innovate, who compete in order to attract agents from a limited pool of candidates. Workers who are able to innovate are few and their supply is usually limited in the short run, and frequent changes in technology can shift the demand from firms so that skill shortage occurs. A market for talented workers could in principle modify results from the base model.

To model a simplified labor market for agents we start by considering a scenario where:

1. There are multiple homogeneous principals;
2. There are many agents all having the same endowment of human capital;
3. The principals can observe agents' human capital with 100% accuracy;
4. A higher human capital makes the agent more productive in the standard task (that is, function $V(e_S)$ takes larger values for equal level of effort e_S), and also more likely to produce innovations for equal levels of effort e_I;
5. Agents can freely change employer without incurring in any costs.

In this symmetric market, absent any collusive behavior, it is straightforward to see that for any principal it is always optimal to increase compensation above the optimal values β^* and γ^* (or γ^{**}) discussed before. By raising compensation just above the highest compensation offered by competitors a principal is able to attract all agents, thus the loss in expected per-agent profits is more than compensated by the added profits generated by having more employees. With symmetric principals, this leads to an equilibrium (by standard arguments) where the contract offered by all principals is ($\beta = \frac{V(e_S^*)}{e_S^*}, \gamma = 1$), and therefore agents appropriate all the surplus (as in "super managers" models like Baranchuk et al. 2011).

In this polar scenario, profit taxation would not produce any revenues and would not affect principals' decisions. Incentives on the side of labor taxation fully apply their effects, so that a reduction in t_I again makes agents invest additional effort in the innovative task and lowers Y_A and Y_P. Differently from the base model, with a marginal innovation value a reduced t_I also makes agents invest more effort because the principals are now unable to compensate this by reducing γ closer to γ^*. This suggests that having mobile agents across principals may make the labor tax incentives even more effective in fostering aggregate innovation.

We now introduce heterogeneity in agents' human capital. If human capital endowments were perfectly observable and contracts could be conditioned on them, then we would end up with a number of different markets (one for each possible level of human capital), each one characterized by a compensation at the equilibrium given by $\gamma = 1$, and by $\beta = \frac{V(e_S^*)}{e_S^*}$ where the value of function $V(.)$ would be increasing in human capital (because of the assumption that a higher human capital makes agents also more productive in the standard task). The same considerations as before with regard to taxation would apply.

Consider a newly introduced constraint: principals now cannot observe human capital endowments at all. This is equivalent to state that contracts must be the same for all agents, irrespective of their level of human capital. This different assumption can better fit the case of a market characterized by a very rapid turnover of the workforce (as an example, think of New Economy firms in the U.S. during the second half of 1990s), where agents do not stay employed long enough with the same employer to be properly evaluated (as modeled in Acharya et al. 2013). Principals cannot observe skill directly, but they are assumed still to be able to know about the distribution of human capital among the population of agents. The expected equilibrium contract is again the one that, on average, makes agents appropriate of all the surplus. This will be again ($\beta = \frac{V(e_S^*)}{e_S^*}, \gamma = 1$), where function $V(.)$ is computed based on a proper mean value across the population of agents. As in the scenario with homogeneous agents, any tax incentive on the side of profit taxation is totally ineffective. A reduction of t_I has the effect of inducing more effort by the agents who chose a principal specialized in innovation, and a second effect is to bring marginal agents to switch from the standard task to the innovative task.

5 Concluding remarks

We have shown how a reduction in labor income taxation levied on profit sharing compensation schemes can be complementary to a tax incentive on corporate profits in increasing the feasibility of innovations. Our results build on the fundamental idea that innovation is mainly the outcome of personal effort, and that a tension between competing tasks exists.

Some econometric studies (Lerner and Wulf 2007, Sauermann and Cohen 2010, Francis et al. 2011, Azoulay et al. 2011) as well as laboratory experiments (Ederer and Manso 2013) supply evidence that providing the right monetary (and non-monetary) incentives to key employees to make them invest in innovation is an important challenge for R&D-intensive firms and industries. A patent box policy alone can foster aggregate innovation as long as external constraints faced by firms are the sole obstacle. If interior constraints are also relevant, then our results advocate to choose a tax mix that provides tax reductions both to companies and to key employees by means of reduced tax rates levied on profit sharing schemes.

The present research is a first step in this direction, and several extensions of the model could be implemented in order to test its predictions in a richer setting. Also, the empirical literature on the interactions between taxation, innovation, and variable payment schemes is still underdeveloped. A prediction of our model is that firms operating in countries where tax benefits are allowed for profit sharing schemes should be observed, *ceteris paribus*, to innovate more. Other methodological approaches (laboratory experiments, agent-based simulations) could also be employed to shed more light on the aggregate effects of different tax schemes when individuals or firms are heterogeneous with respect to some relevant characteristics (i.e. skill, risk-aversion, technological capabilities). Finally, our model explores the use of profit sharing schemes as a means to provide incentives, but such forms of compensation can also be offered by firms in order to attract and retain highly skilled workers. An interesting future extension would be to introduce a formal labor market in order to jointly examine these two motives and the way they may be shaped by tax policy.

APPENDIX

Proof of Proposition 1: The proof for the first inequalities in i) and ii) is obtained from eq. (4) by noticing that, because the marginal utility obtained by additional effort on both sides of the equation is constant due to the linearity of such functions (that is, the quantities $\beta(1-t_S)$ and $\gamma Y n(1-t_I)$ do not change with effort), and the marginal effort cost is the same for the two tasks, to keep the equality sign in $\beta(1-t_S) = \gamma Y_A n(1-t_I)$ (thus making the agent indifferent between the two tasks) it is required that $Y_A = \frac{\beta(1-t_S)}{\gamma n(1-t_I)}$. The derivatives of the latter equality with respect to β and γ obtain $\frac{\partial Y_A}{\partial \beta} > 0$ and $\frac{\partial Y_A}{\partial \gamma} < 0$, which completes the proof for the first inequalities in i) and ii). The second inequalities in i) and ii) are immediately derived from the previous function $Y_A = \frac{\beta(1-t_S)}{\gamma n(1-t_I)}$, simply calculating the first-derivatives $\frac{\partial Y_A}{\partial t_S} < 0$ and $\frac{\partial Y_A}{\partial t_I} > 0$. $\square$

Proof of Proposition 2: The derivatives of e_S^* and e_I^* with respect to τ_S and τ_I, from eqs. (5) and (6), are both equal to zero. Also the derivatives of the unconstrained optimal compensation schemes β^* and γ^* with respect to τ_S and τ_I are equal to zero from either eqs. (13) or (14). As long as Y is high enough (respectively: low enough) to make the inequalities $U^S < U^I$ ($U^S > U^I$) and $\Pi^S < \Pi^I$ ($\Pi^S > \Pi^I$) not change sign, any non-confiscatory profit tax rate cannot increase Y_P above Y_{high} (respectively, decrease Y_P below Y_{low}). $\square$

Proof of Proposition 3: To prove i) and ii) we can implicitly differentiate eq. (10) in the following way:

$$\frac{dY_P}{d\tau} = \frac{\frac{d\Pi^S}{d\tau} - \frac{d\Pi^I}{d\tau}}{\frac{d\Pi^I}{dY}} \tag{24}$$

where the denominator $\frac{d\Pi^I}{dY}$ is evaluated at Y_P.
By applying the envelope theorem to function $\Pi(.)$, we know from eq. (10)

that:

$$\frac{d\Pi^S}{d\tau_S} = -(V(e_S^*) - \beta^* e_S^*) < 0 \tag{25}$$

$$\frac{d\Pi^S}{d\tau_I} = 0 \tag{26}$$

$$\frac{d\Pi^I}{d\tau_S} = 0 \tag{27}$$

$$\frac{d\Pi^I}{d\tau_I} = -(1-\gamma^*)Y_P n e_I^* < 0 \tag{28}$$

$$\frac{d\Pi^I}{dY} = (1-\tau)(1-\gamma^*)ne_I^* + (1-\tau)(1-\gamma^*)n\frac{\partial e_I^*}{\partial \gamma} > 0 \tag{29}$$

the last one being positive because $\frac{\partial e_I^*}{\partial \gamma} > 0$.
The proofs for i) and ii) are then obtained from (24), by noting that it is either $\frac{d\Pi^S}{d\tau_I} = 0$ and $-\frac{d\Pi^I}{d\tau_I} > 0$ which leads to $\frac{\partial Y_P}{\partial \tau_I} > 0$; or $\frac{d\Pi^S}{d\tau_S} < 0$ (because the difference $V(e_S^*) - \beta^* e_S^*$ is necessarily positive as the optimal β^* was chosen to maximize profits from the standard task, which are given by $(1-\tau_S)V(e_S^*) - \beta^* e_S^*$) and $\frac{d\Pi^I}{d\tau_S} = 0$ which leads to $\frac{\partial Y_P}{\partial \tau_S} < 0$.
The proof for iii) is immediately obtained from the fact that Y_P is implicitly defined by the equality $\Pi^S = \Pi^I$, and dividing it by $(1-\tau)$ we can infer that the value of Y_P is independent from τ as long as $\tau < 1$. □

Proof of Proposition 4: Proceeding as for Proposition 3, we see that $\frac{dY_P}{dt_S} < 0$ because $\frac{d\Pi_S}{dt_S} < 0$ as from eq. (8) we know that $\frac{\partial e_S^*}{\partial t_S} < 0$. Similarly, we see that $\frac{dY_P}{dt_I} > 0$ because $\frac{d\Pi_I}{dt_I} < 0$ as we know (see eq. (7)) that $\frac{\partial e_I^*}{\partial t_I} < 0$ and $\frac{\partial \gamma^*}{\partial t_I} = 0$. □

Proof of Lemma 1: The proof for i) is the same as in Proposition 4.
To prove ii), we see that $\frac{dY_P}{dt_I} > 0$ because $\frac{d\Pi_I}{dt_I} < 0$ as we know that $\frac{\partial e_I^*}{\partial t_I} < 0$ and, from eq. (15), that $\frac{\partial \gamma^{**}}{\partial t_I} > 0$. □

Proof of Proposition 5: We proceed as for Proposition 3. Implicitly differentiating like in eq. (24) we obtain the following:

$$\frac{d\Pi^S}{d\tau_I} = -(V(e_S^*) - \beta^* e_S^*) = 0$$
$$\frac{d\Pi^I}{d\tau_I} = -(1-\gamma^*)Yne_I^* + (1-\gamma^*)Yn(1-\tau_I)\frac{\partial e_I^*}{\partial \tau_I} < 0$$
$$\frac{d\Pi^I}{dY} = (1-\tau_I)(1-\gamma^*)ne_I^* + (1-\tau_I)(1-\gamma^*)n\frac{\partial e_I^*}{\partial \gamma} > 0$$

where $\frac{d\Pi^I}{d\tau_I} < 0$ because $\frac{\partial e_I^*}{\partial \tau_I} < 0$. Consequently from (24), $\frac{dY_P}{d\tau_I} > 0$ which concludes the proof. □

Proof of Lemma 2: With risk-neutrality ($\delta = 0$), it is straightforward to verify from eq. (4) that $\frac{\partial Y_A}{\partial t} = 0$. With risk-aversion in order to keep the equality sign in eq. (4), Y_A has to change in order to compensate for the risk premium δ while $\frac{\partial U^S}{\partial t}$ is still the same as in the risk-neutral case. As from eqs. (20) and (21) by applying the envelope theorem it is verified that $\frac{\partial \delta}{\partial t} < 0$, increasing t will increase $-\delta$ closer to zero, thus a lower Y_A is now needed for equality between U^S and U^I to be maintained. □

References

Acharya, Viral V., Marco Pagano, and Paolo Volpin (2013). *Seeking Alpha: Excess Risk Taking and Competition for Managerial Talent*. Working Paper 18891. National Bureau of Economic Research.

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso (2011). “Incentives and creativity: evidence from the academic life sciences”. *The RAND Journal of Economics* 42(3), pp. 527–554.

Balkin, David B. and Luis R. Gomez-Mejia (1984). “Determinants of R and D Compensation Strategies in the High Tech Industry”. *Personnel Psychology* 37(4), pp. 635–650.

Baranchuk, Nina, Glenn MacDonald, and Jun Yang (2011). “The Economics of Super Managers”. *Review of Financial Studies* 24(10), pp. 3321–3368.

Bolton, Patrick and Mathias Dewatripont (2005). *Contract Theory*. MIT Press Books. The MIT Press.

d’Andria, Diego and Silke Uebelmesser (2014). *The relationship between R&D intensity and profit-sharing pay: Evidence from Europe*. Conference Paper. International Schumpeter Society Conference, 27-30 July in Jena.

Dietl, Helmut M. et al. (2013). "Incentive effects of bonus taxes in a principal-agent model". *Journal of Economic Behavior & Organization* 89, pp. 93–104.

Ederer, F. and G. Manso (2013). "Is Pay for Performance Detrimental to Innovation?" *Management Science*.

Evers, Lisa, Helen Miller, and Christoph Spengel (2013). *Intellectual property box regimes: Effective tax rates and tax policy considerations*. 13-070. ZEW Discussion Papers.

Francis, Bill B., Iftekhar Hasan, and Zenu Sharma (2011). *Incentives and Innovation: Evidence from CEO Compensation Contracts*. SSRN Scholarly Paper ID 1762621. Rochester, NY: Social Science Research Network.

Grossmann, Martin, Markus Lang, and Helmut M. Dietl (2011). *Why Taxing Executives' Bonuses Can Foster Risk-Taking Behavior*. SSRN Scholarly Paper ID 1952729. Rochester, NY: Social Science Research Network.

Hall, Bronwyn and John Van Reenen (2000). "How effective are fiscal incentives for R&D? A review of the evidence". *Research Policy* 29(4), pp. 449–469.

Hellmann, Thomas and Veikko Thiele (2011). "Incentives and Innovation: A Multitasking Approach". *American Economic Journal: Microeconomics* 3(1), pp. 78–128.

Holmstrom, Bengt (1989). "Agency costs and innovation". *Journal of Economic Behavior & Organization* 12(3), pp. 305–327.

Holmstrom, Bengt and Paul Milgrom (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design". *Journal of Law, Economics & Organization* 7, p. 24.

Ittner, Christopher D., Richard A. Lambert, and David F. Larcker (2003). "The structure and performance consequences of equity grants to employees of new economy firms". *Journal of Accounting and Economics* 34(1), pp. 89–127.

Lazear, Edward P. (1986). "Salaries and Piece Rates". *The Journal of Business* 59(3), pp. 405–431.

Lerner, Josh and Julie Wulf (2007). "Innovation and Incentives: Evidence from Corporate R&D". *Review of Economics and Statistics* 89(4), pp. 634–644.

Manso, G. (2011). "Motivating Innovation". *The Journal of Finance* 66(5), pp. 1823–1860.

Radulescu, Doina Maria (2012). "The Effects of a Bonus Tax on Manager Compensation and Welfare". *FinanzArchiv: Public Finance Analysis* 68(1), pp. 1–16.

Sauermann, Henry and Wesley M. Cohen (2010). “What Makes Them Tick? Employee Motives and Firm Innovation”. *Management Science* 56(12), pp. 2134–2153.

Figure 1: Plot of functions $Y_A(\gamma)$ (red line) and $Y_P(\gamma)$ (blue line), for $0 \leq \gamma \leq 1$.

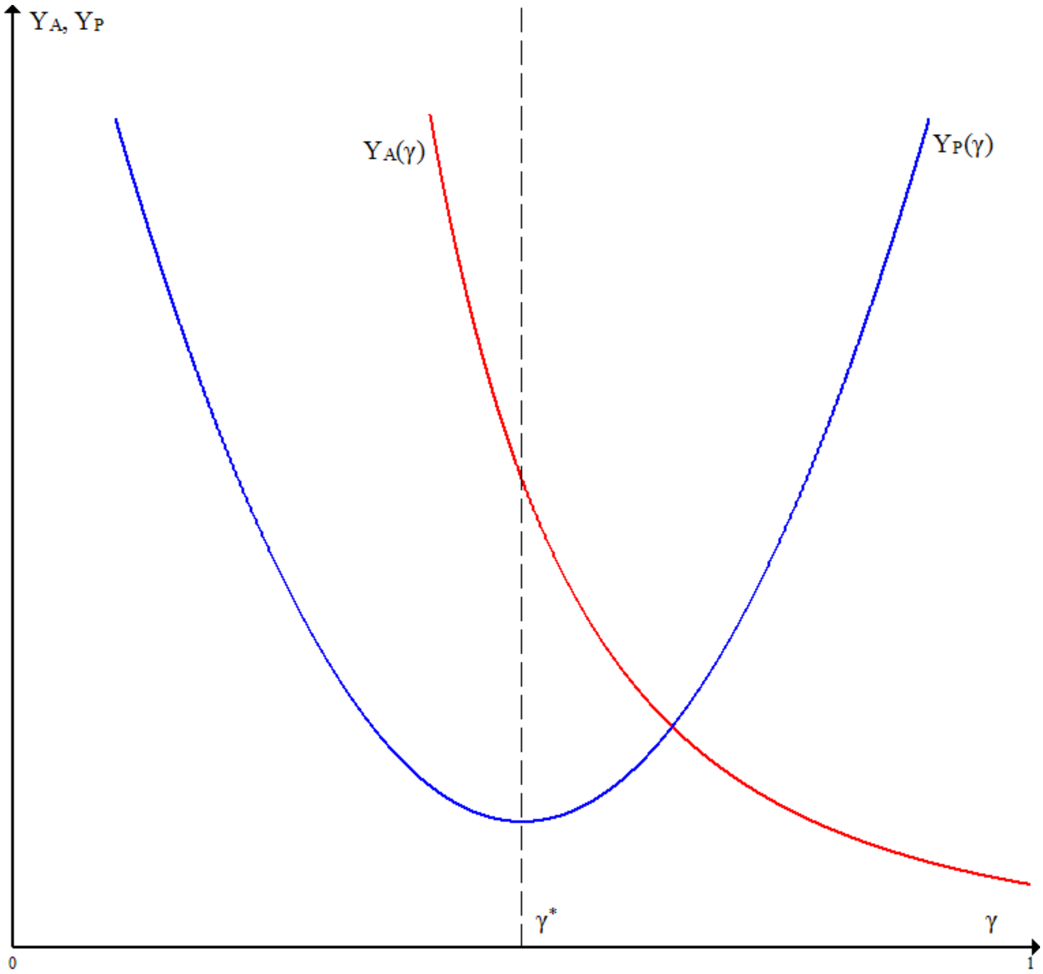

Figure 2: Plot of functions $Y_A(\gamma)$ and $Y_P(\gamma)$, without labor tax incentives (dashed lines) and with a reduced tax rate t_I (continuous lines). Without labor tax incentives, the optimal contract is γ^{**}.

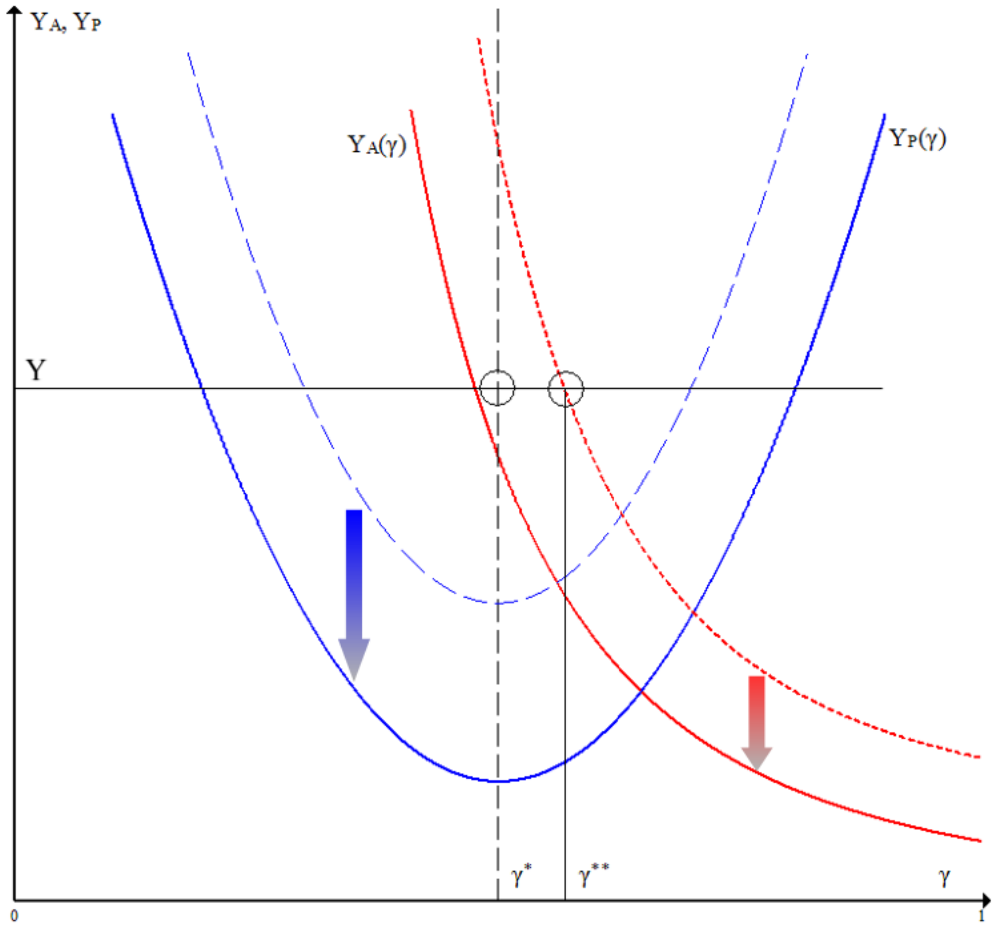

Figure 3: Plot of functions $Y_A(\gamma)$ and $Y_P(\gamma)$, without labor tax incentives (dashed lines) and with a reduced tax rate t_I (continuous lines). A reduction of the tax t_I lowers Y_A and Y_P so much that innovation occurs.

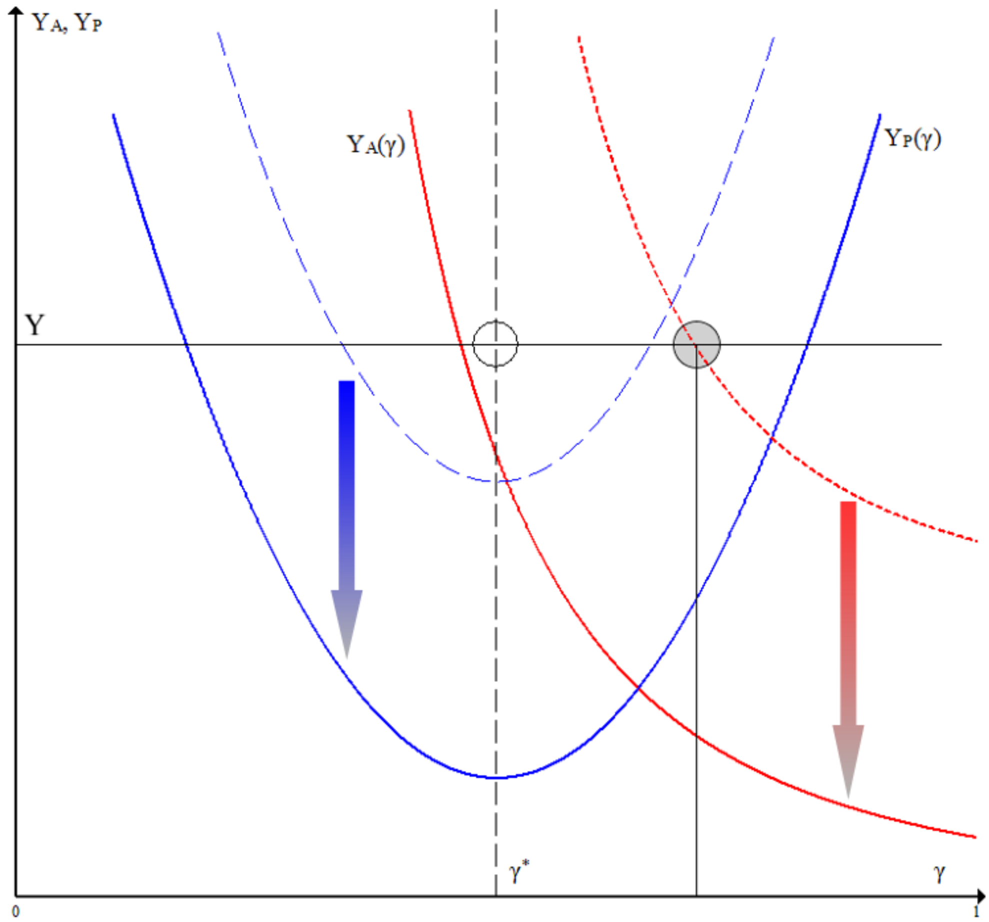

Figure 4: Graphical representation of different revenue-neutral tax reforms. The graphs are ordered from the highest expected innovation value (top-left graph) to the lowest (bottom-right graph).