

JENA ECONOMIC RESEARCH PAPERS



2015 – 004

Motivating innovation in a knowledge economy with tax incentives

by

**Diego d'Andria
Ivan Savin**

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University 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

© by the author.

Motivating innovation in a knowledge economy with tax incentives*

D. d'Andria^{†‡} and I. Savin[†]

Abstract

In the past decades the role of profit sharing schemes (PSS) as a way to foster innovation in a principal-agent context, and more generally of innovation in economic growth, have been widely acknowledged and studied. However, surprisingly little has been done to analyze the interactions between tax policy, PSS and innovative activity, not least because of severe data limitations. In this study we propose an agent-based model to explore the effects of two distinct tax policies on innovation in a pure knowledge economy: a ‘patent box’ incentive and a tax incentive on compensation earned by agents as PSS. A distinct feature of this paper is that in contrast to the conventional assumption that firms (principals) decide on whether to innovate or not, we propose that this decision is actually taken by their employees (agents). We compare the two tax incentives under several distinct specifications and find that the tax incentive on PSS is more efficient than a ‘patent box’ incentive when the role of capital investments in R&D is negligible. With R&D investments in the form of a capacity constraint, both tax incentives are found to play a role in fostering innovation. In addition we find important effects on the incentives’ relative efficacy due to labor mobility and due to the ability of firms to benefit from knowledge spillovers.

Keywords: *agent-based model; innovation; knowledge economy; profit sharing schemes; tax incentives for R&D*

JEL Classification: H2, O3, J33

*We would like to thank Uwe Cantner, Holger Graf, Johannes Herrmann, Martin Kalthaus, Oliver Kirchkamp, Silke Übelmesser, and participants at the Jena Economic Research workshops, the GENED Workshop at the Kiel Institute for the World Economy and the 15th International Schumpeter Society Conference for very helpful comments and suggestions. Financial support from the German Science Foundation (DFG RTG 1411) is gratefully acknowledged.

[†]Corresponding author, Bachstrasse 18k Room 5/B, D-07743 Jena, Germany. Email: diego.dandria@gmail.com

[‡]Graduate College ‘Economics of Innovative Change’, Friedrich Schiller University and Max Planck Institute of Economics, Jena, Germany.

1 Introduction

Innovation has been playing an increasingly prominent role in shaping the competitive advantage of industries and countries. ‘Knowledge workers’ and other highly specialized individuals are the main driver: innovation means new knowledge is generated, incorporated into products or processes and finally commercialized, and this process requires past knowledge held by experts as one of its main input factors.

In an attempt to foster innovation, several policies have been proposed, ranging from regulatory reforms of the intellectual property rights system (Boldrin and Levine, 2008) to tax incentives designed to obtain the maximum boost in innovation with the least loss in tax revenues. In recent times, ‘patent box’ incentives have gained momentum and have been adopted in several European countries (Evers *et al.*, 2013). All these policies have in common their focus on choices with respect to what investments to start, and what amount of research and development (R&D) activities to pursue. Such a kind of choice is made by managers within private and public firms, or by entrepreneurs facing the choice whether to commence a new enterprise.

In a recent contribution d’Andria (2014) proposed a tax incentive on profit sharing schemes (PSS) based on the grounds of theoretical arguments. By PSS we mean any form of compensation to employees¹ linked to a measure of the company’s success, such as: direct participation to profits, bonus pay, stock options and stock grants. The idea of a tax incentive on PSS stems from the empirical observation that R&D intensive firms and industries appear to offer PSS compensations to employees who are capable to innovate, in order to steer their behavior by means of monetary incentives toward investing more effort into job tasks that can produce innovations (Balkin and Gomez-Mejia, 1984, Ittner *et al.*, 2003, Lerner and Wulf, 2007). d’Andria (2014) demonstrated that a policy mix relying not exclusively on tax incentives on firms’ profits but also including some level of tax reductions on PSS income could better foster innovation. Another appealing property of this tax incentive is that it does not require *a priori* definitions of what an R&D expenditure is (as it is for standard R&D tax credits and deductions), nor it needs the innovation to be subject to intellectual property rights (as it is the case for patent box incentives). All innovations and improvements that are expected to increase the profitability of the firm can be captured by PSS monetary incentives, and as a consequence, a tax incentive on PSS is in principle capable to affect a wider set of potential innovative behaviors.

In knowledge-intensive firms, however, PSS compensations are offered to employees for multiple reasons. In a survey done by Ittner *et al.* (2003) respondents from ‘New Economy’² US firms reported that important reasons for them to offer stock options and stock grants are to attract and retain skilled staff and as a way to provide incentives to perform. A similar conclusion follows from a larger sample of European firms (d’Andria and Übelmesser, 2014). PSS is used also to provide incentives to employees to participate in product or process improvements. This motive directly links the use of PSS to agent-led innovation.

Knowledge-intensive labor markets also present issues related to asymmetric infor-

¹For the sake of convenience, throughout the text a worker is also denoted as agent, as employee or simply as a male (‘he’); a firm is also denoted as principal, an employer or simply as a female (‘she’). These terms are used as perfect substitutes.

²We adopt, here and in the following parts, the same definition of ‘New Economy’ firms used by Ittner *et al.* (2003): ‘*organizations competing in the computer, software, internet, telecommunications, or networking fields*’.

mation. As already pointed out in Acharya *et al.* (2013) in relation to managers in the financial sector, the ability of these employees to generate higher returns often requires time to be properly identified by employers. We argue that the same problem arises more generally with knowledge workers such as R&D managers, researchers and technical workers (also defined overall as “technologists” in Cohen and Sauermann, 2007). When labor mobility is high (as it was for example during the ‘New Economy’ boom during the second half of the 1990s) or when workers’ skill are qualitatively very heterogeneous due to high specialization (as in superstar models like Rosen, 1981) the assessment of a worker’s ability becomes difficult: workers change employers often, they have lots of alternative opportunities to their current job, and the nature of fast changing technology makes it hard to properly understand and evaluate how well an employee’s talent and human capital fit into the current market for ideas.

The present research allows for pay-for-innovation and attraction motives in offering PSS. A pure knowledge-based economy is represented where firms try to attract best minds and to provide employees with the type of compensation contract that is believed to produce more profits. But, employers are assumed to be unable to monitor how skilled each (potential) employee is, and they are also unable (as in standard principal-agent literature) to observe how their workers allocate effort among competing job tasks.

This setting is used to shed light on the expected effects of different tax incentives for innovation. Two types of incentives are compared: one is a patent box incentive which reduces the tax rate levied on the profits generated by innovations on the side of employers; the other incentive is a reduced tax rate on the labor income obtained by workers as PSS. The two incentives are compared in a rich model setting, where workers can differ in terms of their skill and risk aversion, employers can have different technological capabilities (productivity) due to innovations performed in the past, and the job market can be more or less fluid. The effects of labor market mobility, presence of technological spillovers, the rate of technological change and risk attitude of agents on the efficacy of the tax incentives are of particular interest.

The kind of market we model is purely knowledge-driven, meaning that we do not include capital investments as an input factor (except for our third scenario where R&D investments constrain the ability of workers to commit to innovative job tasks). This choice naturally makes profit taxation (and consequently, tax incentives on the side of corporate taxation) less able to affect firms’ choices. What we achieve is therefore an analysis of the relative efficiency of the tax incentives in a setting where a profit tax incentive has low power (or no power at all) to foster innovation, thus making the availability of the alternative incentive (on PSS) desirable. Real-world markets sharing close traits to the one we model are those related to the New Economy software productions, media and creativity productions, and various types of high-end services.

This study is structured as follows. Section 2 presents the literature linked to this study and highlights the reasons behind some design choices for our ABM. Section 3 describes the basic mechanisms used in the model. In Section 4 we distinguish several scenarios: the baseline one where technology level is fixed, the one with evolutionary technology, and the scenario with evolutionary technology and capital constraints on firms investing in R&D. Main results are presented and discussed in that section. Section 5 provides a quantitative assessment of the tax incentives’ efficacy. Section 6 concludes.

2 Related literature

Research on variable pay has always been strongly interdisciplinary, blending together contributions from economics, business, psychology, sociology and law studies. Our interest here is purely economic, but even then the present model is related to different strands of economic literature.

In our model employers offer payment contracts that may include PSS with the double aim to attract qualified staff and to motivate them to perform, as in ‘supermanagers’ models like Baranchuk *et al.* (2011). Differently from these models, the type of motivation that employers try to instill into workers is not to ‘work more’ but to ‘innovate more’. In this respect, our model is closer in spirit to pay-for-innovation models (Manso, 2011, Hellmann and Thiele, 2011), which are based on principal-agent multitasking settings (as already discussed in the seminal contributions of Holmstrom, 1989 and Holmstrom and Milgrom, 1991), where PSS can drive agents into investing more effort in an innovative job task.

Our main interest is not in how contracts are optimally set by employers, but rather it is on how different tax designs may change such decisions and, as a result, affect innovation performance. Previous tax research has paid attention to tax incentives for innovation in the form of tax credits or deductions related to corporate income taxation (good examples are: Hall and Van Reenen, 2000, Myles, 2009, Griffith *et al.*, 2011, Gemmell *et al.*, 2011), but not as much related to labor income taxation. The latter is usually limited to analyzing the effects on human capital accumulation obtained by reducing personal income taxation and consequently increasing net returns to education. The focus on corporate tax incentives is also observed in real-world policies: countries in recent years have increasingly adopted ‘patent box’ tax incentives for R&D, while incentives to innovation on the side of labor income are adopted in few countries and are limited in extent (Evers *et al.*, 2013). Also, existing tax incentives on R&D wages are not differentiated based on the type of compensation (PSS or other). On the contrary, our ABM considers the case of a selective tax incentive on PSS.

Another strand of literature related to our model is the research on bonus taxes (Grossmann *et al.*, 2011, Radulescu, 2012, Dietl *et al.*, 2013), where the effects of taxation on compensation contracts, invested effort and risk taking are examined in a multitasking principal-agent analytical framework. Differently from those works which are related to the financial intermediation industry and its specific characteristics, however, our model focuses on innovative knowledge-intensive firms and compares taxation and incentives both at firm and at labor level, in this being closer to works like Russo (2004). Similarly to Russo (2004) we employ a simulation to explore different tax policy settings and their impact on innovation, tax revenues collected, cost-effectiveness and other variables of interest for a cost-benefit analysis. Contrary to Russo (2004), our methodology of choice is not Computable General Equilibrium (CGE), but rather Agent-Based Model (ABM) simulation.

There are several reasons for this choice. First, ABMs allow us to introduce information constraints more easily, thus adding an important element of realism given that, in real markets, employers face many limitations in their ability to monitor workers’ skills and efforts (in this we come closer to limited information models like Acharya *et al.*, 2013). Second, the nature of innovations itself is predominantly considered as a ‘disturbing term’ for an equilibrium analysis and it becomes more natural to model it in an evolutionary context (Nelson and Winter, 2009). The competition for innovation is

modeled here in a way that recalls previous works like Yildizoglu (2002), where firms learn from their past choices and outcomes, and they take decisions based on procedural rationality following a trial-and-error heuristic. This also allows us to explore heterogeneity of employers and workers along several dimensions, as well as different conditions of the market with regard to labor mobility and knowledge diffusion. Last not least, there are several local interaction effects on the side of employers (learning from each other about contracts, absorbing spillovers from leading firms in the industry) and employees (in scenario 3, indirectly competing for a job at a firm investing in R&D).

Following Cohen and Sauermann (2007) and Sauermann and Cohen (2010) we partly depart from a more traditional modeling of technological progress. Economists use to model technological change as the sole outcome of rational, profit-maximizing firms' decisions as was first proposed in seminal contributions like Arrow (1962) and in the literature surveyed in Nelson (1959). We extend this theoretical literature by focusing on individual choices within private firms. Our ABM represents a stylized pure knowledge economy, where the only input factor is individual effort supplied by highly skilled workers. We limit the analysis to a single industry in one country, thus ruling out effects due to cross-sectoral or cross-border mobility of factors.

This study is the first, to our knowledge, to analyze the effects of a tax incentive on PSS for R&D and compare it against a traditional incentive for R&D based on profit taxation in a multi-agent and multi-principal context. It is also the first to model an economy where contract design by private firms is examined in a setting which includes competition for labor mobility, information asymmetries with regard to both skill and effort, and truly dynamic technological innovation so that path dependency on previous choices and outcomes affect future states of the world.

3 The model

We first describe a baseline ABM which is meant to study how the model performs without introducing technological capabilities. In the baseline ABM, therefore, principals are heterogeneous only with respect to the contract they offer, which is endogenously determined by the ABM. This section describes how agents and principals are modeled in the baseline ABM and the main algorithms used to match jobs and select compensation contracts. The numerical values of the parameters used and a legend of the notation are reported in Appendix A.

There are two types of taxes in the ABM. The first is a corporate income tax levied on profits and is represented by τ . The second is a labor income tax, represented as t . The tax rates τ_S and t_S are the general tax rates levied on profits and incomes obtained by standard tasks. The tax rates τ_I and t_I are levied only on profits and incomes obtained by a successfully produced innovation. With $\tau_S = \tau_I$ and $t_S = t_I$ there is no tax incentive. When the tax rates are such that $\tau_S > \tau_I$, a 'patent box' incentive on profits is levied. When $t_S > t_I$, a tax incentive on PSS is levied. The following tests are run by starting from a case without any tax incentive, and then gradually introducing a reduction in τ_I or t_I until they reach zero. We are especially interested in understanding how effective each incentive is in fostering aggregate innovative outcome. Important to note is that the current ABM does not aim to present a proper cost-benefit analysis,³ but rather to

³Otherwise, one would need to build a more general and complex model including not only the knowledge economy market (to account for links with other economic sectors) and do a proper model

show which tax incentive is more attractive under given conditions and why. In addition, using the model we can compare the efficiency of the tax incentives between the different scenarios employed, thus providing guidance for future empirical investigations and, possibly, for field experiments aimed at introducing a PSS tax incentive in selected industries or regions.

3.1 The Agents

Agents can choose an employer, which job task to invest effort into, and how much effort to invest in the task chosen. The two possible tasks are either a standard job task, or an innovative task, which can be thought of as representing a tension between going for ‘business as usual’ versus new practices and strategies, improvements of products or of processes, in this following Hellmann and Thiele (2011). In comparison to previous models of innovation where the choice whether to innovate or not is taken by the entrepreneur or the CEO of a company, here such choice is made by the knowledge worker who observes an opportunity for a firm-specific innovation. Because the property rights over such innovation, if successfully developed, are assumed to be fully retained by the employer, a PSS is necessary to provide an *ex ante* monetary incentive to the worker in order to make him or her commit effort to the task.

In the *ex ante* situation (that is, before the uncertainty about whether the innovation is successful or not is lifted) the agent computes the expected utility payoff from the standard and from the innovative tasks, and then chooses the one which, in expected value, provides him with the highest utility. Utility increases with the expected monetary return, and decreases by private cost of effort represented by means of a standard quadratic cost function $\frac{e^2}{2}$, as used in many other works like Grossmann *et al.* (2011), Radulescu (2012) or Benabou and Tirole (2013).

Agents take the following variables as given: the tax rates levied on the income obtained from standard and innovative tasks (t_S and t_I , respectively); the terms of compensation contracts, where each contract is defined by the three items (w, β, γ) detailed later in the text; and the distribution of innovation values which is assumed to follow a Rectified Gaussian distribution (that is, a Gaussian distribution where negative values are substituted with zeros), $\mathcal{N}^R(Y, \sigma)$. Each agent chooses how much effort to invest in the standard task (this effort is labeled as e_S) or in the innovative task (e_I). Because of the functional form chosen for the utility function, effort is always committed in one task, but never in both tasks at the same time.

The standard task pays the agent a fixed wage w , which is an unconditional payment irrespective of effort by the agent, plus a variable pay linear in effort: βe_S . The value of β represents a monetary wage that is then multiplied by some observable perfect signal of effort. Both, the fixed wage and the variable pay, are reduced by the tax rate on labor income t_S . For the standard task, the analytical representation of the agent’s expected payoff is:

$$U_S = (1 - t_S)w + (1 - t_S)\beta e_S - \frac{(e_S)^2}{2} \quad (1)$$

and his optimally set effort e_S^* is obtained from first-order conditions as:

$$e_S^* = (1 - t_S)\beta. \quad (2)$$

tuning to empirical data.

For the innovative task, given an agent's level of skill k , the agent's expected payoff is

$$U_I = (1 - t_S)w + (1 - t_I)\gamma Y(n + k)e_I - \frac{r}{2}(1 - t_I)^2\gamma^2 Y^2(n + k)^2\sigma^2 - \frac{(e_I)^2}{2}, \quad (3)$$

which is a sum of:

- the net-of-tax fixed wage $(1 - t_S)w$, identically as for the case of the standard task;
- the expected value of the innovation for the agent, calculated as the product of its mean value Y , the probability of succeeding in developing it $(n + k)e_I$ (where n is a fixed scale-down parameter), and the share γ of this value which the principal promised to the agent by offering him a PSS compensation, and reduced by the tax rate t_I ;
- $-\frac{r}{2}(1 - t_I)^2\gamma^2 Y^2(n + k)^2\sigma^2$ is the risk premium demanded by the agent to be as well-off as with having a certain equivalent to the stochastically determined sum of money represented by the risky investment in the innovative task.⁴
- the last term captures the private cost of effort.

Note that agent's skill, k , positively affects the probability of success but not the mean value of the innovation Y . Also note that we assume effort costs to be the same for the two tasks, therefore ruling out the possibility that the innovative task is either more enjoyable, or more tiresome, than the standard task.

Accordingly, the worker's optimally set effort e_I^* is equal to:

$$e_I^* = \frac{(1 - t_I)\gamma Y(n + k)}{1 + r(1 - t_I)^2\gamma^2(n + k)^2\sigma^2}. \quad (4)$$

With risk neutral agents the parameter for absolute risk aversion r is equal to zero, and therefore the optimally set effort for the innovative task reduces to $e_I^* = (1 - t_I)\gamma Y(n + k)$.

Sensitivity analysis on risk aversion

In order to better explore the behavior of agents under different risk attitudes, we conduct a sensitivity analysis with respect to risk aversion and test five levels of the parameter r ranging from very large risk aversion ($r = 10$) to very large risk 'love' ($r = -10$).⁵ The simulations are calibrated in a way that a larger level of risk aversion makes agents less willing to undertake the risky tasks and, therefore, the amount of effort invested in innovative tasks is smaller than in the risk neutral case for equal monetary incentives. In contrast, risk loving agents assign higher utility to innovative tasks and are willing to undertake them even when this leads to losses (see Appendix B).

⁴This representation of non-risk neutral agents is derived by assuming a constant absolute risk aversion (CARA) utility function. For a discussion and some additional analytical results see d'Andria (2014), and for a formal derivation Bolton and Dewatripont (2005).

⁵Note that in our ABM very large values of r are needed to make the effects of risk attitudes visible. Our parametrization is chosen so that, within the interval of payoff values agents can earn, for a parameter $r = 5$ we mimic the risk attitude (in terms of constant absolute risk aversion) found empirically in Beetsma and Schotman (2001) where an absolute risk aversion of about 0.12 was estimated looking at a TV show with real monetary prizes. In cases of risk propensity where we set $r < 0$ the same scale of constant absolute risk aversion is mirrored in terms of negative risk premia demanded.

3.2 The Principals

Principals are assumed to suffer from severe information constraints: they can neither directly observe an agent's ability, nor they can monitor *ex ante* the type of task chosen or effort invested in an innovative task. This somewhat polar scenario is used to explore a setting that does not allow for the use of traditional profit maximization. As a result, principals base their decisions on a trial-and-error approach where they learn from their own past and from the observation of their competitors. By choosing the compensation scheme, principals attract agents and generate profits; and by observing the most profitable competitors they can copy some elements of the best performing contracts (see Section 3.5 for details). The long-term 'convergence' contracts⁶ are then those arising from such an evolutionary process.

Principals take taxes (τ_S, τ_I) as given. The amount of profits generated by a principal is a direct function of the agent's choices. For simplicity we do not model complementarities between multiple agents, so that profits produced by each agent are separated from profits generated by other agents.⁷ In order to attract agents and provide incentives to them, each principal offers a single contract to all agents who choose to get employed by her. A contract is made of three parts: a fixed wage w , paid regardless of agents' choices and effort; a variable pay β linked to effort invested in the standard task; and a PSS γ on the value produced by means of innovation.

The principal's expected payoff from the standard task performed by an employed agent is a function of the amount of effort e_S invested and of the skill of the agent (k), the latter being unobserved by the principal. It is equal to a gross profit $\frac{(e_S^*)^a(1+k)}{b}$, minus the cost of labor (fixed plus variable pay), $-\beta e_S^* - w$, and reduced by the tax rate τ_S :

$$\Pi_S = (1 - \tau_S) \left[\frac{(e_S^*)^a(1+k)}{b} - (\beta e_S^* + w) \right]. \quad (5)$$

Eq. (5) assumes that both the fixed and the variable pay are fully deductible from the taxable base.

The principal's expected payoff from the innovative task is a function of the probability that the innovation is successful times the expected value of the innovation, less the share which is going to be paid to the agent by means of PSS, and everything reduced by the tax rate τ_I :

$$\Pi_I = (1 - \tau_I)[(1 - \gamma)Y(n + k)e_I^* - w], \quad (6)$$

where again, as for the agents, the probability of successfully developing an innovation is $(n + k)e_I^*$. The principals possess the same information about the underlying distribution of the innovation value as the agents, so that their expected profit is computed also on the

⁶It would be obviously improper to talk about equilibrium values in such a setting. We use the term 'convergence contract' in the text to refer to the stabilized dominant contract arising after a certain number of iterations of the ABM. Being the result of an evolutionary process (more on this later on), this convergence contract has stability properties in the sense that deviations from its values happen only because of stochastic events and are only temporary.

⁷In principle, one could model multiple agents who jointly collaborate toward a common profit goal. With knowledge-intensive productions, though, this task is not trivial. Agents would face an incentive to free-ride on co-workers' effort as in standard multi-agent-principal models; but they could also benefit from positive knowledge spillovers so that co-workers' effort invested in innovative task(s) would increase one's own probabilities of succeeding, as for example in the model proposed in Ederer (2013). Despite being interesting, this extension is left for further research.

mean value Y . Because the principals are always assumed to be risk neutral,⁸ the value of σ is not relevant for their decision. As for the innovative task (Π_I), all compensations (fixed, and PSS) are fully deductible from the taxable profits.

3.3 Calculation of payoffs

In each period of time, principals set their contracts, agents choose employer and task, and profits are generated. Total profits generated by a principal are just the sum of all profits Π_S and Π_I produced by all agents employed.

For the standard tasks the payoffs are obtained deterministically following previous equations (5) and (6). From agents' (principals') point of view, the realized utility (realized profits) from a standard task is exactly equal to the expected utility (expected profits) they evaluated *ex ante*.

On the contrary, for innovation the outcome is stochastic. The probability of success in producing the innovation is equal to effort e_I^* (from eq. (4)) times $(n + k)$:

$$\text{Probability of success} = (n + k)e_I^* = \frac{(1 - t_I)\gamma Y(n + k)^2}{1 + r(1 - t_I)^2\gamma^2(n + k)^2\sigma^2}. \quad (7)$$

If the innovation is successful, the agent obtains a fraction γ of its value, reduced by taxes, plus the fixed net wage, and less the private cost of effort:

$$(1 - t_S)w + (1 - t_I)\gamma y - \frac{(e_I^*)^2}{2}, \quad (8)$$

where the exact value of innovation y is drawn from a Rectified Gaussian distribution,⁹ whose related non-rectified Gaussian distribution has mean Y and standard deviation σ .

The principal gets the remaining value of the innovation minus the fixed wage and reduced by taxes:

$$(1 - \tau_I) [(1 - \gamma)y - w]. \quad (9)$$

If the innovation fails, the agents only get a fixed wage, but still suffer the private cost of effort:

$$(1 - t_S)w - \frac{(e_I^*)^2}{2}. \quad (10)$$

The principals earn zero, and still have to pay the fixed wage: $-w$.

3.4 Labor market

For a principal it is profitable to attract additional agents doing innovative tasks as long as the expected value they produce net of PSS pay is not lower than the fixed pay they get; and it is profitable to attract agents doing standard tasks as long as their gross product exceeds the sum of fixed and variable pay.

In case when the average profitability of innovation is larger than for standard production, if principals could optimally set contracts and workers were immobile, they would

⁸This is a very standard assumption in the principal-agent literature: principals, that is shareholders, can normally diversify their investments relatively easily.

⁹In practice, because of the parameters we employ, the probability to get values below zero is irrelevant. But in principle for alternative parameter specifications negative values should be deleted from the possible outcomes of y and substituted with zeros.

offer a contract with $(w = 0, \beta = 0, \gamma = 0.5)$ (d’Andria, 2014). This is in fact the profit-maximizing contract that would drive all employees to commit effort into innovative tasks (assuming the innovative task is more profitable, in expected terms, than the standard task at the contract $\gamma = 0.5$). To put it differently: a PSS of $\gamma = 0.5$ would make the innovative task more profitable for the principal and it would maximize expected profit. The incentive the principal faces is to drive agents into committing effort into it, and this can be done by setting the opportunity cost of the other task to zero. Given the parameter values we use (see Table 4 in Appendix A), innovation is on average less profitable than standard production in absence of tax incentives. Therefore, the optimal contract with immobile agents would be set to $(w = 0, \beta = \frac{a(1+k)}{2be_s^{1-a}}, \gamma = 0)$. The fixed wage does not play any role (offering any $w > 0$ would increase costs for the principal, without adding any more effort to the task) but is included to add realism to the model and demonstrate the distribution of values it reaches under different scenarios.

Analytically in a frictionless job market with homogeneous principals, the possibility that agents switch to different employees makes principals raise γ (and β above the value maximizing standard profit under no mobility, because now the opportunity cost of the standard task depends upon the whole market and not just upon the single employer) in order to attract agents. Because the optimal fixed pay is still $w = 0$, this means that the labor market collapses in the end to a situation where all principals offer $w = 0$, $\gamma = 1$, and β at the level which produces zero profits for principals. Put in different words, all the profits are extracted by agents.

Our ABM features a *job search algorithm*. In each period agents observe a number of potential employers (between 2 and M)¹⁰ and the compensation contract these principals are offering. Agents compute expected utility for both standard and innovative tasks for all observed candidate employers, and then choose to be hired by the one having the maximum expected utility value. In doing so, agents then commit effort to the task chosen. A mobility cost mc is introduced in this model, so that agents only change employer if a competing principal offers an expected utility $(1 + mc)$ times higher than the utility expected from working at the current principal. That is, agents compare the maximum utility U_{max} they can obtain from the task generating the highest utility when employed by the current principal, with the maximum utility $U_{candidate\ employer}$ again from the task generating the highest utility by this agent when employed by an alternative employer. Formally, in order to switch principal the following inequality must be satisfied:

$$U_{candidate\ employer} \geq (1 + mc) U_{max}. \quad (11)$$

Henceforth, we refer to ‘*high mobility*’ labor market as the ABM scenarios where agents compare in each period all available employers and face zero mobility cost ($mc = 0$), while in the ‘*low mobility*’ labor market agents can compare their current employer with only one candidate employer (randomly chosen) per period, and have a mobility cost of 100% ($mc = 1$). This way to model job search has the appealing property to allow for different degrees of job mobility. By setting the mobility cost to zero and the number of observed candidate employers equal to the number of total principals, a fully mobile job market can be explored. By increasing mobility costs and reducing the number of observed employers, the job market becomes less transparent and more viscous.

¹⁰ $M = 10$ is the total number of principals in the market.

3.5 Contract selection

Because principals can neither evaluate agents' skill values k nor know whether an agent will commit effort in an innovative task (they can only know this *ex post*, and only in case the innovation was successfully developed), a fully constrained-optimization problem cannot be solved to find the single compensation contract that would maximize expected profits. What principals do, instead, is to experiment with a contract setting for a certain number of periods, and then compare the resulting profits made with the profits of the competitors.

At the beginning of each simulation run all initial contracts are randomly assigned to principals. The ABM is calibrated so that the values of w , β and γ are always between zero and one.¹¹ After that, each five periods¹² a contract selection algorithm orders principals based on the amount of profits achieved per money spent (a form of return-to-investment or R.O.I., which we also refer to as *relative profits*)¹³ made in the last five periods. The principal who produces most profits does not change her contract. The principal who ranks worst copies one single element (w , β or γ randomly chosen) from the best performing principal. Those principals who rank in between (not the best and not the worst) try to change one of the three elements of their contract randomly (add/deduct a random addendum). In this way, the contract selection we employ gets close to a genetic algorithm, where the best contract over time tends to survive and dominate the market. To sum up, the algorithm is meant in an evolutionary fashion where principals respond to the selection taking place on the market (Metcalf, 1994) and learning from own and others' past experience actively looking for a good mix of contracts.

Summary

The main characteristics of the model can be summarized as follows (see Figure 1). In each period workers observe possible employers (their contracts) and taxes on income they will have to pay, evaluate the expected utility they can obtain from each observed employer and from both standard and innovative tasks, and decide where to be employed and in what task to commit effort. After the tasks are performed, each principal generates revenues and pays remuneration to her workers for their contribution. The difference between revenues and labor costs constitutes gross profits of firms they have to pay taxes from. The resulting net profit divided by total expenditures is nothing else than the average return that firms try to maximize. Hence, after observing their relative performance each principal updates her contract mix and the loop is started again.

¹¹Therefore, the initially assigned values of each element of the contracts are randomly drawn from a uniform continuous distribution between zero and one.

¹²This number is considered to be sufficiently large for principals to evaluate the relative efficiency of a contract.

¹³We use the sum of profits made divided by total money spent as a ranking criterion, because using total profits (without normalizing them in some way) would impose a preference in favor of the largest firms, so that firms with more employees would on average always be ranked highest. Our intent is instead to make the contract selection algorithm represent the struggle principals face to increase the productivity of their workforce, and to this end, a ranking based on return-to-investment seems a more sensible choice. As an alternative ranking criterion we also tried average profits generated per employee and results obtained were almost identical. In a set of small-scale tests we employed non-normalized total profits and, although some notable quantitative difference arise in comparison to the use of return-to-investment or profits per employee indicators, qualitatively the models behave very similarly.

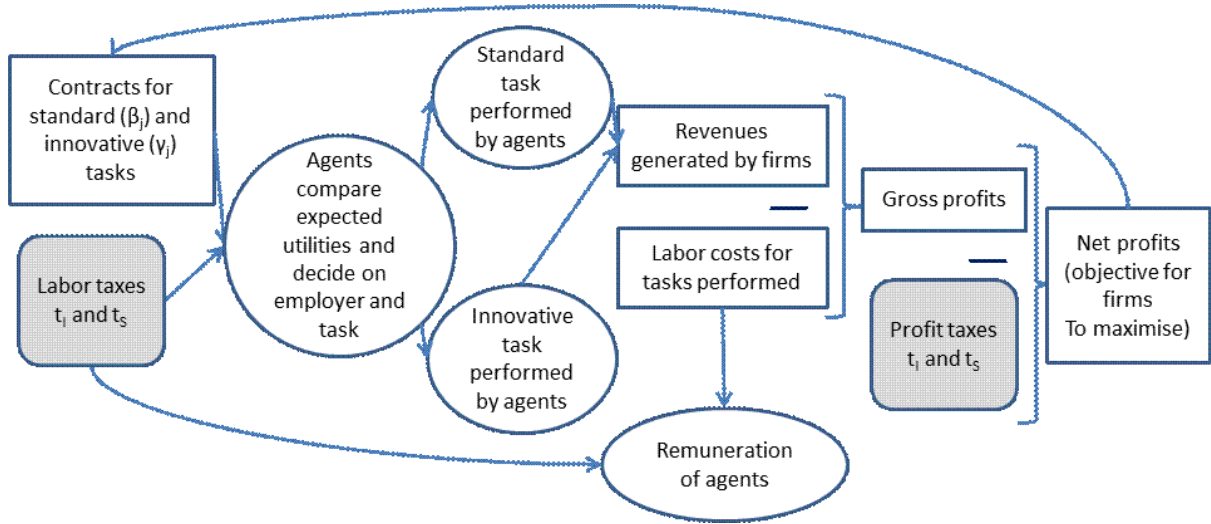


Figure 1: The model at a glance

4 Simulation results for different scenarios

We run our ABM for 2,000 periods reporting averaged values (over the last 500 periods) for the variables of interest, such as the number of innovations successfully introduced, which contracts being selected and how many tax revenues being collected.¹⁴ The parameters were chosen so that without any tax incentive innovative tasks are on average slightly less profitable for the principals. With tax incentives, innovation becomes progressively more profitable in comparison to standard tasks. This is meant to represent in a simple and stylized way the existence of some market failure preventing innovation to succeed as often when no policy intervention is present.¹⁵ In ABM scenarios 2 and 3 featuring evolutionary technology, knowledge spillovers are also included, thus providing an additional source of potential market failure (Arrow, 1962) and consequent space for policy intervention.

4.1 Baseline model

A first batch of tests is done comparing a highly mobile labor market (by setting mobility costs to zero and making agents able to observe all potential employees in every period) and a low mobility market (where an agent switches employer only if new utility is at least double the current utility, while observing only one alternative firm).

Our model can reproduce a situation observed in real-world economies (Andersson *et al.*, 2009), where knowledge-intensive firms compete in a race-to-the-top of PSS compensation offered to lure skilled workers and agents extract almost all the profit from the principals. But, the race-to-the-top of compensation only happens when the expected value from innovation is on average larger than from standard productions. As stated already, we focused instead on a case where absent tax incentives the average profitability of innovation is lower than for standard production.

The following Figure 2 shows the values of the compensation contracts to which the ABM converges in the end, for both high and low mobility scenarios. From the left, contract values for fixed pay w , standard pay β , innovation pay γ , and innovation pay in

¹⁴The current version of our ABM is written in MATLAB. A single restart of the ABM requires from 10 to 30 seconds using Intel i5 3.3 GHz (depending on the scenario under consideration).

¹⁵If no market failures were present, then there would be no need for any tax incentive.

value γY are displayed. This figure (and the ones that follow) represent boxplot graphs of the average outcome across multiple replications. The rightmost boxplots corresponding to the 30% point on the x-axis stands for a no-incentive case (so that all tax rates are set equal to 30%). As one moves leftward, one of the two taxes (either τ_I , represented by red boxplots, or t_I represented by blue boxplots) is reduced from 30% to 0%. In this way a direct comparison is provided of the effects due to equal reductions in the tax rates.

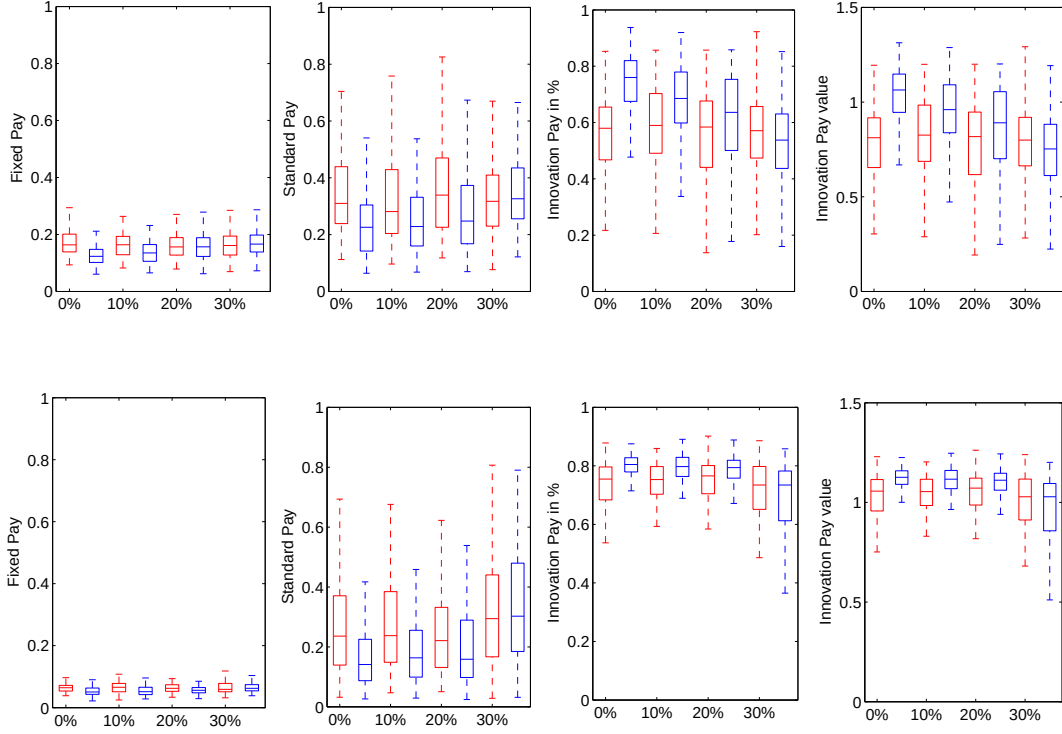


Figure 2: Convergence in contracts

Note: Contracts are represented here for the high mobility scenario (upper panel) and for the low mobility scenario (lower panel). Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained over 100 replications for each parameter combination.

Because with high mobility, market information about contract profitability is produced much faster (as agents can observe at the same time all contracts offered, and they change employer without incurring into any cost), principals focus their contract competition on the most profitable task (standard). As a consequence innovation pay has larger variance and tends to gravitate around the mean value of 0.5. In the low mobility case principals enjoy some more contractual power (but less information), and the convergence contract features a larger innovation pay γ (and with less variance). Low mobility makes principals able to retain a larger share of the profits, as they are able to offer lower wages and PSS in comparison to a high mobility setting. Accordingly, we observe larger profits per worker under low mobility (Figure 3).

Figure 4 represents boxplot results for high and low mobility, for a different set of indicators. The graphs on the far left side plot the tax-elasticity of innovation measuring how much aggregate innovation (computed as the total innovation value produced) I is increased (in %) for a 1% loss in total tax revenues T . The elasticity is calculated at the values I_1 and T_1 which are computed for the case with no tax incentives and is expressed as:

$$\frac{\Delta I}{I_1} / \frac{\Delta T}{T_1}. \quad (12)$$

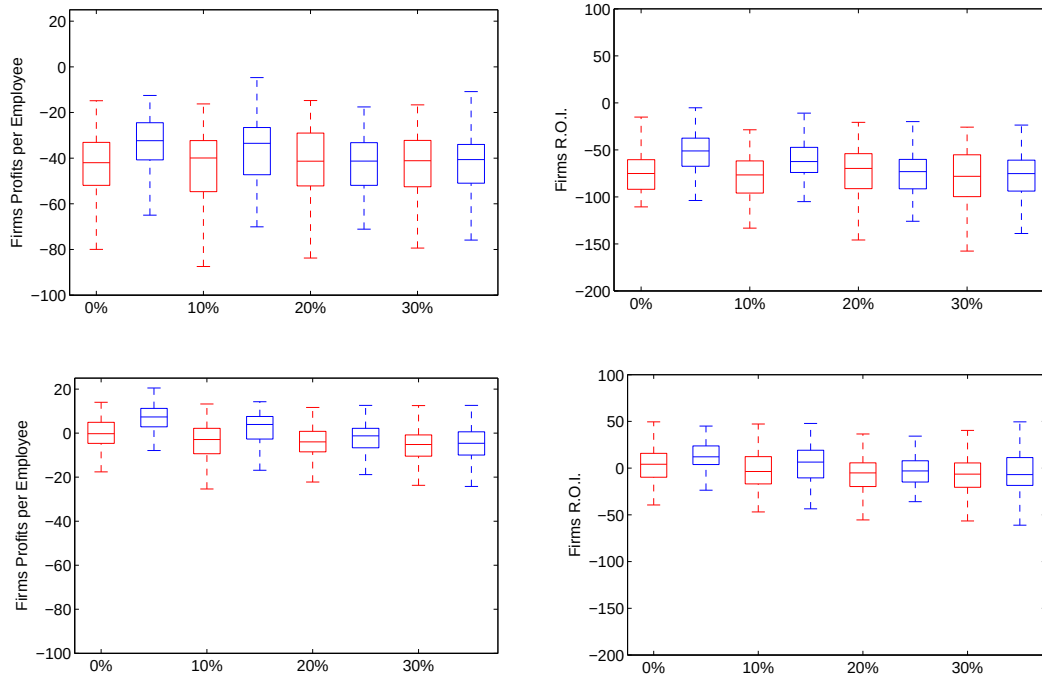


Figure 3: Relative profits

Note: Indicators are represented here both on R.O.I. and profit per employee for the low mobility scenario (bottom panel) and for the high mobility scenario (upper panel). Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained over 100 replications for each parameter combination.

Introducing a tax incentive on innovation with fixed technology we expect innovations to rise while tax revenues fall, which makes the elasticity index negative, and the larger its absolute value, the larger is the innovation-per-dollar increase for the tax incentive.¹⁶

The graphs in the central-left column of Figure 4 measure the number of innovations which were successfully performed per period. The graphs in the central-right column represent the loss in tax revenues, as a percentage of the revenues collected without any tax incentive (that is, when tax rates are all set equal to 30%). We also build an aggregate indicator (shown in the rightmost graphs in Figure 4) by taking the total value of innovations produced, and subtracting from it the loss in tax revenues that tax incentives produce in comparison to the no-incentive case. The interpretation of the indicator is that as long as one monetary unit of tax revenues is evaluated by the policymaker equally to one unit of value produced by innovation, the net increase of the difference between the two provides a synthetic measure of the attractiveness of a policy. Thus, *ceteris paribus* larger values of the indicator are associated with higher efficiency of a tax incentive.¹⁷

¹⁶An important drawback of the indicator (despite the useful information it provides with regard to the effectiveness of tax incentives per unit of tax revenues lost) is the fact that its values are very unstable over restarts. Thus, having a very small tax reduction resulting from the tax incentive scheme drives the denominator close to zero while the numerator remaining large. In later scenarios as technology may improve over time, innovation can compensate the tax loss so that both, difference in taxes and difference in innovations, may be positive.

¹⁷An advantage of that indicator is its ‘numerical’ stability compared to the elasticity of innovation. It is certainly disputable how much weight a policy maker may put on each monetary unit of taxes lost and on each monetary unit of innovation value gained, as this choice would be determined by the specific target social welfare function that the policymaker is assumed to maximize. But we want to emphasize

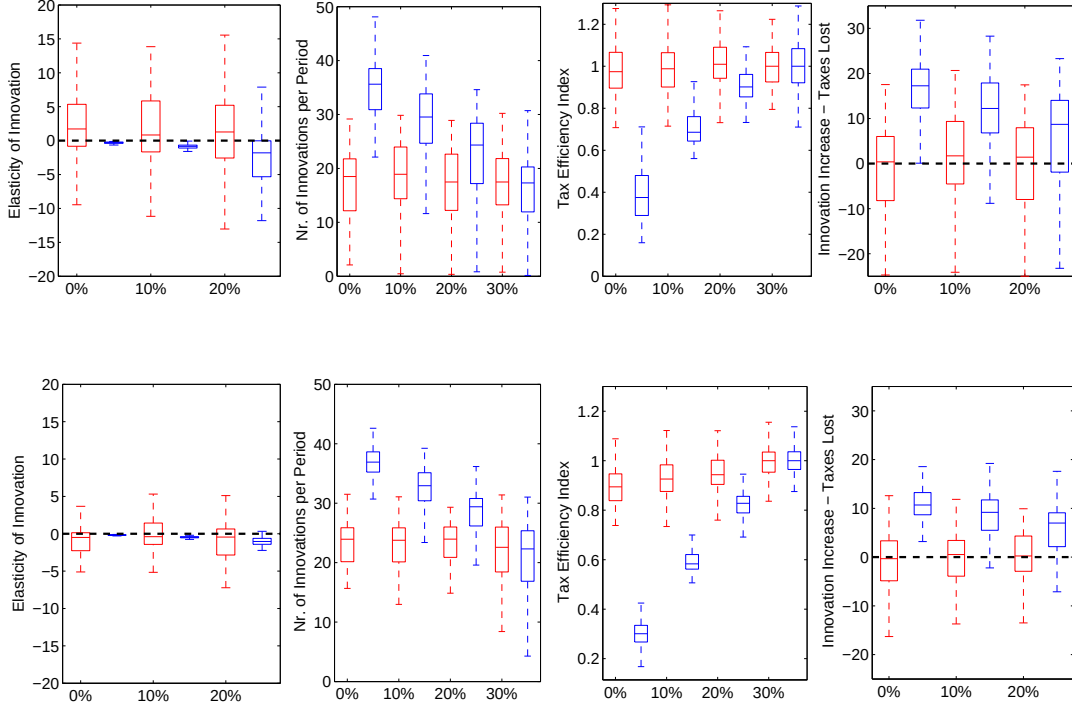


Figure 4: Efficiency indicators

Note: Indicators are represented here for the high mobility scenario (upper graphs) and for the low mobility scenario (lower graphs). Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained over 100 replications for each parameter combination.

From Figure 4 it is immediate to see how the tax incentive on PSS (a reduction in t_I) outperforms the patent box incentive (a reduction in τ_I). The effects of the latter are detrimental: one loses taxes (see graphs for the tax efficiency index) without gaining any considerable effect in terms of more innovations being introduced (the values of the elasticity index are quite evenly distributed around zero). In contrast, a reduction in labor taxes t_I by 10% is already capable to generate substantial increases in aggregate innovation, and results in an elasticity value close to -2% (2% more innovations successfully realized for 1% less taxes collected). Easy to notice from the elasticity boxplots is the decreasing marginal rate of return from the labor tax incentives introduced: the larger the incentive, the smaller the percentage gain in innovations. For the low mobility scenario it is observable how the tax incentive on profits is still inefficient. The tax incentive on PSS is capable of steering more effort into innovation although less efficiently than under high mobility (this can be seen from the bottom left graph in Figure 4, where innovation is still increasing with a reduction in t_I). Here the elasticity index reaches visibly lower values (e.g., 1% in case of moderate tax reduction).

When modifying the degree of risk aversion of agents, the relative utility gained from the two tasks is affected under a *ceteris paribus* condition with regard to compensation contracts. Figure 11 in Appendices reports boxplots for expected utility from innovative tasks, the share of job tasks done that are innovative, and the convergence contracts for innovative pay (boxplots are for the case with low mobility). A larger risk aversion makes expected utility from the risky task (innovation) less appealing for agents than

that in principle a set of weights could be very easily implemented into the model to represent such social preferences.

risk neutrality. This affects the power of incentive pay to drive them into innovation and, consequently, innovation contracts converge to a lower level in comparison to a scenario with lower risk aversion as innovative productions become less profitable for principals (less effort is invested by agents for the same amount of pay) than standard productions. Because monetary incentives are less effective in driving agents into innovation for a given compensation contract, tax incentives on PSS are expected to be less effective as well.

4.2 Scenario 2: Principals with evolutionary technology

This section presents an extension of the ABM (summarized in Figure 1) modeling technology dynamics as dependent on past innovations performed.¹⁸ All principals have the same level of technological capability at the beginning of the simulations, but now they can become heterogeneous based i) on their past performance and ii) on knowledge spillovers produced by other firms in the market. As a result, firms having higher technological capabilities become less sensitive to competition on contracts as they can offer a higher remuneration to agents thanks to the larger revenues generated. This is a fundamental feature of the following model extensions. In particular, in situations where a firm builds up a technological advantage over other firms, she may not need to improve her contract structure anymore and as a consequence will produce less innovations in the long run.

We model technology as an accumulative process. Principals who have experienced successful innovation in recent periods enjoy an increase in the expected value for both innovative and standard tasks. This represents together the gains obtained in productivity by means of process innovation and firm-specific know-how. Therefore, if a principal discovered a good compensation contract before competitors, and innovation succeeded often enough, this principal may be able to offer compensation contracts made of a lower γ in comparison to competitors and still be able to attract agents.

The accumulation of technological capabilities is modeled as follows. The sum of the innovations I_t that have been successfully produced in the last ten periods by a principal is calculated and divided by the number N of agents in the population to normalize it, then it is multiplied by a factor introducing marginally decreasing returns. The latter is meant to represent the fact that firms who are closer to the maximum achievable level of technology ($tech_{max}$) will only get progressively smaller increments of technology for equal amounts of knowledge generated in the past as it becomes more difficult to advance an already mature technology in comparison to a new one. Thus, one can think of a process innovation which allows principals to use their production structure more efficiently (see, e.g., Yildizoglu, 2002).

Formally, the increase of the current stock of technological capabilities of a principal from her own accumulation of knowledge is:

$$Internal\ knowledge = \frac{1}{200} * \frac{\sum_{t-10}^t I_t}{N} * \frac{-1}{\log \left(tech_{firm} * \frac{tech_{max} - tech_{firm}}{tech_{max}} \right)}, \quad (13)$$

where $tech_{firm}$ is the current technological level reached by this principal, $tech_{max}$ is the (exogenously set) maximum technology level achievable by this economy. The last element in the product, $\frac{-1}{\log(\cdot)}$, introduces the decreasing marginal returns so that the

¹⁸Unless stated otherwise in the text, when introducing evolutionary technology we employ the same parameters used in the baseline ABM to allow for comparability.

closer is $tech_{firm}$ to $tech_{max}$, the smaller the knowledge gain from past innovations. Internal knowledge is thus always between 0 and 1.

In some scenarios we also include knowledge spillovers between principals, so that firms can benefit from a larger knowledge stock available in the market. We follow Cantner and Pyka (1998) and Egbetokun and Savin (2014) assuming that not all firms gain the same benefits from a given stock of spillovers, because the ability to gain from them is mediated by their absorptive capacity and by how distant the firm is from the current technological frontier. We proxy absorptive capacity in a simplistic way and assume that having a larger number of highly skilled agents in recent periods makes a principal more able to absorb external knowledge. Absorptive capacity is modeled first by calculating for each principal the current relative distance from the highest current technological level:

$$tech_{gap} = \frac{tech_{frontier} - tech_{firm}}{tech_{max} - tech_{start}}, \quad (14)$$

where $tech_{frontier}$ is the highest value for $tech_{firm}$ reached across all principals in a given period, and $tech_{start}$ is just a parameter setting the initial level of technology all principals have at time zero (when the simulation is started). Therefore $tech_{frontier}$ is the best technology available in a period and is always between $tech_{start}$ and $tech_{max}$. Consequently $tech_{gap}$ is always between 0 and 1.

Each principal increases her own technological level by the following:

$$Gain\ from\ spillovers = \psi * \frac{\sum_{t=10}^t \sum_i k_{it}}{N} * (tech_{gap} - tech_{gap}^2), \quad (15)$$

where k_{it} is the level of skill of the i -th employee at time t , ψ is a parameter defining the speed of accumulation of knowledge spillovers, and the last multiplier ($tech_{gap} - tech_{gap}^2$) introduces a simple parabolic representation of the idea that being very far, or very close, to the current technological frontier makes a firm less able to benefit from knowledge spillovers. In our parametrization, spillovers are included by setting a parameter $\psi = 0.05$ (or alternatively $\psi = 0$ if no spillovers).

The increase in a principal's technological level is thus given by the current level, multiplied by the sum of one plus the gain from own accumulation plus the gain from spillovers:

$$Technological\ capability = tech_{firm}(1 + Internal\ knowledge + Gain\ from\ spillovers). \quad (16)$$

In order to make the simulations more stable, we also introduce a hard cap so that principals can never increase their level of technology more than a given percentage per single period of time. In the simulations this hard cap (ϕ) is set to 2.5% for a 'slow' technology, and to 5% for a 'fast' technology. Also, similarly to what was done with contract selection, we made principals able to change their technology levels only at a selected interval of time, set to 20 in the simulations reported.

We look at the way the ABM evolves with technology by setting different scenarios. These scenarios are differentiated based on:

1. the degree of agents' mobility, as done in the baseline model specifications;
2. whether knowledge spillovers are present or not;
3. the rate of technological accumulation ('slow' or 'fast' technology).

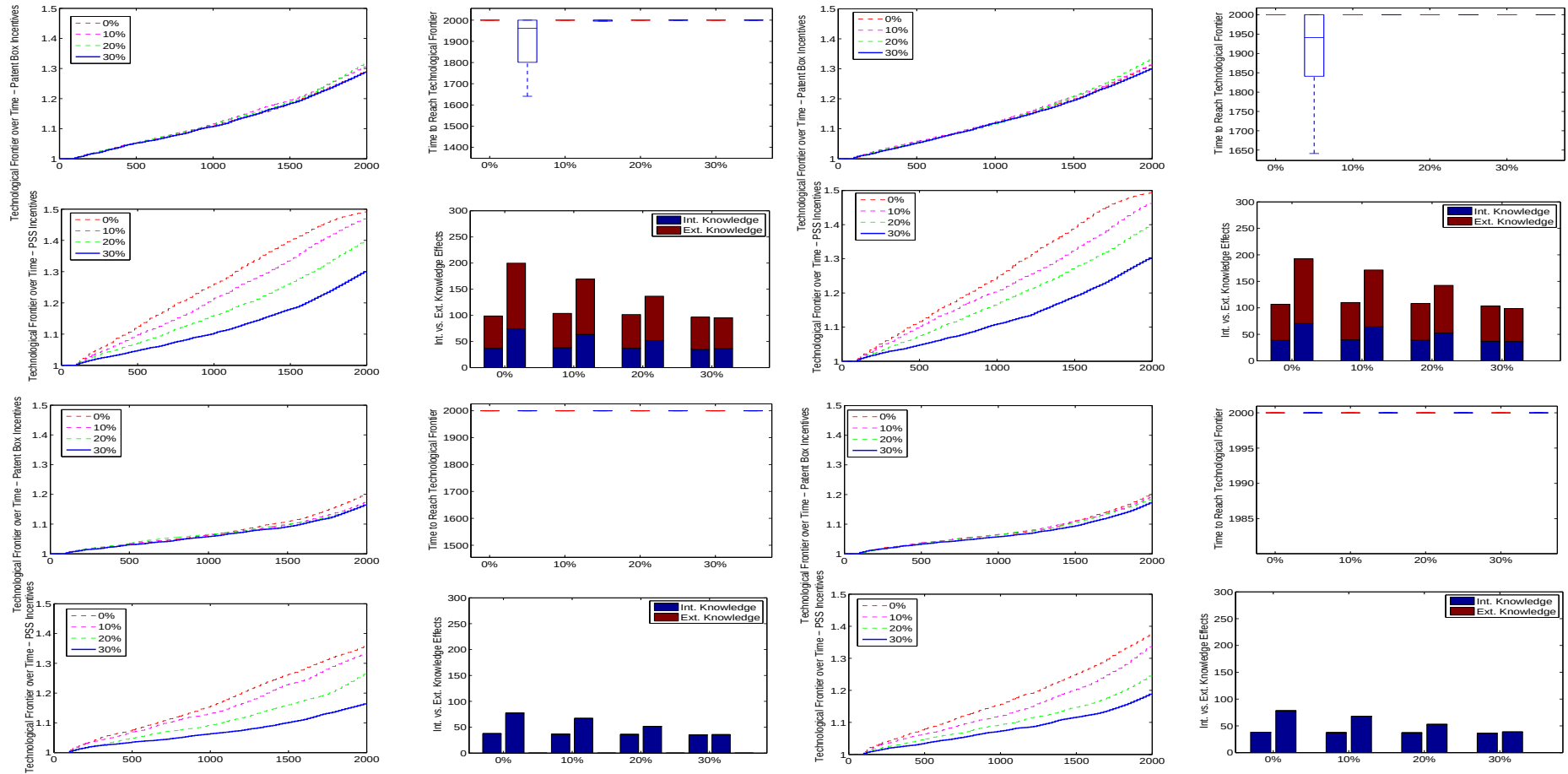


Figure 5: Technology indicators with evolutionary technology

Note: Indicators are represented here for the scenario with fast technology and spillovers (upper leftmost four graphs) and for the scenario with spillovers but slow technology (upper rightmost four graphs), all with low mobility. In addition, the two technology speeds with no spillovers are presented in the lower panels: fast technology on the left and slow technology on the right. As before, blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained with 25 replications for each parameter combination.

We calibrated the parameters ψ and ϕ so that the maximum technology $tech_{max}$ is rarely reached by our simulated economy and, if reached, this happens close to the last period of the ABM. In this way the process of knowledge accumulation remains dynamic and we can infer some interesting properties of the tax incentives. These are the results that are displayed in Figures 5, 6 and 8.

When setting different ‘speeds’ to technological improvements by changing the parameters ψ and ϕ , we obtain quantitative changes only. A ‘fast technology’ makes the model sometimes reach the maximum technology level (this parameter, $tech_{max}$, is made equal to 1.5), while a ‘slow technology’ is meant to make this economy never reaching such upper limit. The exclusion of spillover effects was also found to affect results mainly quantitatively. Generally, knowledge spillovers help principals’ to catch-up on technological capabilities and be more competitive in attracting agents by offering attractive contract schemes, which in the long run leads to a higher level of technology reached at the end of the simulation.

In Figure 5 we report information about how the evolution of technology is affected by the tax incentives. The graphs in the leftmost area of each quadrant (separately for patent box and PSS tax incentives) compare the average $tech_{frontier}$ reached in each period of the simulation (computed as the mean across all replications of the ABM) with different levels of the tax incentives. Blue lines represent $tech_{frontier}$ reached without tax incentives, while dotted lines represent gradually stronger incentives. The figure also contains information in the boxplots about the time needed for maximum technology level ($tech_{max}$) to be reached (if it is reached at all). Bar charts graphically represent how much of the increase in the technology frontier from $tech_{start} = 1$ to $tech_{frontier}$ is caused by internal knowledge generation (blue bars), and how much is attributed to external knowledge spillovers (red bars), in cases when a PSS incentive (bars on the right) or a profit tax incentive (bars on the left) are introduced. These bar graphs represent the relative cumulated contribution of internal and external (through spillovers) knowledge in increasing the technological frontier, along the whole time span of 2,000 periods.

The degree of agents’ mobility affects the endogenous variables in the ABM. In an economy with innovation more profitable than standard production, high labor mobility firms would tend to reach the technological frontier faster while the opposite happens in our parametrization, and tax incentives only marginally change this. But still, PSS incentives perform differently based on the mobility regime (refer to Section 5 for an in-depth analysis). Also we see how mobility changes the composition of internal and external sources of knowledge in shaping technological capabilities of firms. This result stems from the assumption we made that spillovers are mediated by absorptive capacity, which is proxied by the number of employees multiplied by their individual skills. In a low mobility scenario, internal knowledge is relatively more important because many firms in the market are able to retain workers for some time, thus building up some internal knowledge over time. In a high mobility scenario it is more likely that some (or just one) firms grow much faster while other firms lag far behind. This widens the technological gap for laggards and increases (up to a certain level of the gap) the effect of spillovers they benefit from.

Figure 6 again reports convergence contracts for both the high mobility and the low mobility scenarios. Contrasting it with values reported in Figure 2, the main difference lies in the rightmost charts, where the monetary value of PSS compensation (*InnValue*) is reported. The larger values obtained in the ABM with evolutionary technology are due to the fact that technology improvement through innovation affected the mean value Y

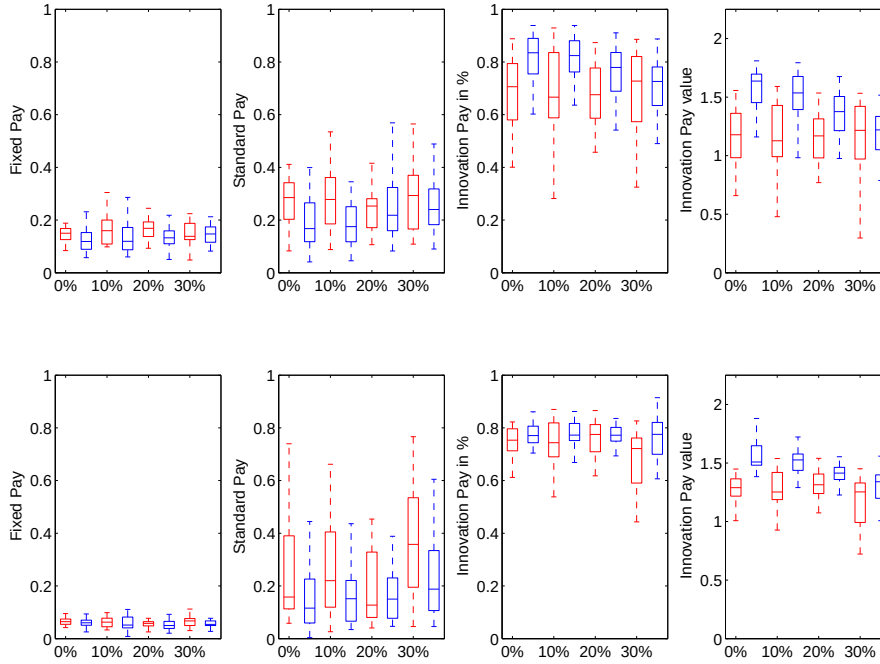


Figure 6: Convergence in contracts with evolutionary technology (scenario 2)

Note: Contracts are represented here under fast technology with spillovers, for the high mobility scenario (upper panel) and for the low mobility scenario (lower panel). Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained with 25 replications for each parameter combination.

of further innovations. In both scenarios innovation contracts react more to the PSS tax incentive under high mobility (when information about relative profitability of contracts is generated faster by the market).

Figure 7 reports profits per employee and R.O.I., and compares them in scenarios with high mobility and low mobility. As per the baseline scenario with fixed technology, under high mobility relative profits can often become negative as agents extract almost all the surplus, while they tend to be positive in the low mobility scenario. Relative profits are on average larger than in the baseline model, thanks to productivity gains due to technological improvement.

Figure 8 reports efficiency indices, again to allow for a comparison with the baseline ABM. The incentive on PSS is still performing much better than the patent box incentive.

4.3 Scenario 3: Principals with evolutionary technology and R&D investments

The economy represented in our ABM is characterized by a production function where only knowledge and individual effort enter as input factors. This is meant to represent a polar scenario to be contrasted with many existing models where only capital investments in R&D are considered. However, even not including capital investments as an additional formal input, the capacity to employ workers in innovative tasks can be a function of a minimal investment made by firms to provide employees with infrastructures that are required to conduct specific job tasks. Without the minimal endowment of capital investments (i.e. to buy specialized software and machinery) employees might be unable to

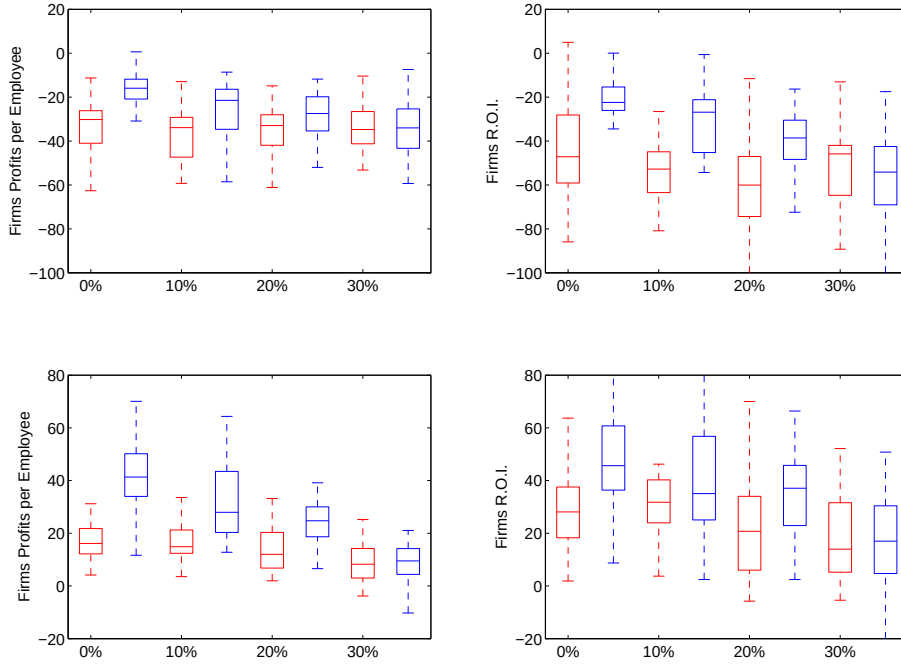


Figure 7: Relative profits with evolutionary technology (scenario 2)

Note: Profits per employed agent and R.O.I. are represented here for the high mobility (upper panel) and for the low mobility (lower panel) scenario, under fast technology with spillovers. Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained with 25 replications for each parameter combination.

commit to innovative job tasks. Given that agents observe candidate employers in each time period, lack of adequate investment capacity might induce agents to prefer, *ceteris paribus*, other employers who enable them to commit to innovation.

To address the existence of R&D investment and – indirectly – risk sharing between principal and agent, on top of the extension of knowledge spillovers and technological accumulation introduced in Section 4.2, we additionally model R&D investments as a capacity constraint.¹⁹ Many theorists (see for example the review in Reinganum, 1989) argue that the choice about R&D investments is complex as firms must also predict the behavior of competitors, thus participating in a strategic competitive dynamic game. This requires high-level rationality and that firms are able to access a large amount of information. Empirical evidence and models as discussed in Nelson and Winter (2009) suggest that instead many firms follow a simple rule of thumb by allocating a fixed percentage of revenues or profits to internal R&D investments. We take an in-between stance and assume firms try to improve their R&D strategy according to some observable signal from the market.

R&D investments are defined as a percentage of the net profits and are set as follows. R&D expenditures are more likely to be increased if firms observe an average return from innovative activities larger than standard production, or on the contrary – they are more likely to be reduced in the opposite case. Firms in the ABM are made able to compute the average return from innovative and standard activities across the whole market, and

¹⁹This can be also thought of as a Leontief-like production function of the form $y = f(l) * \text{ISTRUE}(k \geq \bar{k})$, where $\text{ISTRUE}(k \geq \bar{k})$ is equal to 1 *iff* per worker capital is at least $\bar{k}$ (and equal to 0 otherwise).

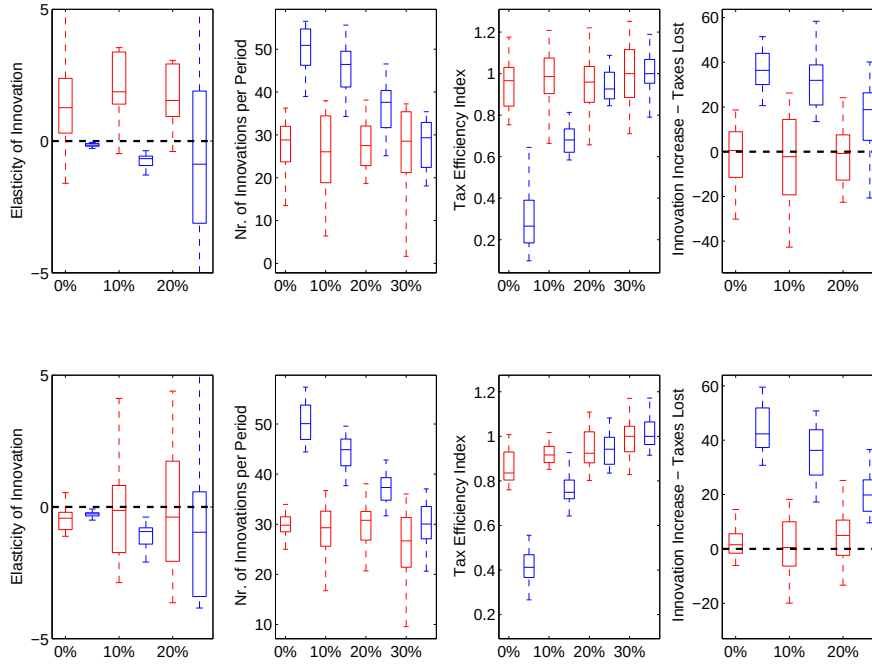


Figure 8: Efficiency indicators with evolutionary technology (scenario 2)

Note: Indicators are represented here for the high mobility (upper panel) and for the low mobility (lower panel) scenario, under fast technology with spillovers. Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained with 25 replications for each parameter combination.

they modify their current R&D strategy by summing up an addendum to the current R&D investment. The addendum is drawn as a uniformly distributed random number between -0.2 and 0.4 if a larger average return from innovation was observed, or between -0.4 and 0.2 in the opposite case.

Agents are assumed to perfectly observe R&D investments and the number of employees doing innovation in the previous period in each (observed) firm. When current R&D investments are insufficient for the number of employees committing effort to innovation, agents get an imperfect signal: the larger the gap between the employed workforce in innovation and R&D investments, the larger the probability that an agent observing this candidate employer will assign an expected utility of zero to the chance to obtain value from an innovation in that firm. This is meant to represent the expectation by agents that a too low R&D capacity generates congestion and drastically reduces the probability of successful innovation. The resulting technological evolution is displayed in Figure 9, again for the low mobility scenario.

Generally speaking the introduction of R&D costs limits the competition over contracts (Figure 10), introducing an additional dimension along which firms compete for agents, namely, R&D investments. This has two important consequences. First, there is an additional source of path dependency: the more successful firms have been in the past, the more likely they will be successful in the future as they can invest more in R&D (i.e. well-known success-breeds-success hypothesis). Second, similar to the effect of accumulated technology introduced earlier, this extension increases the probability of technological divergence between the winning firm and the competitors, where the

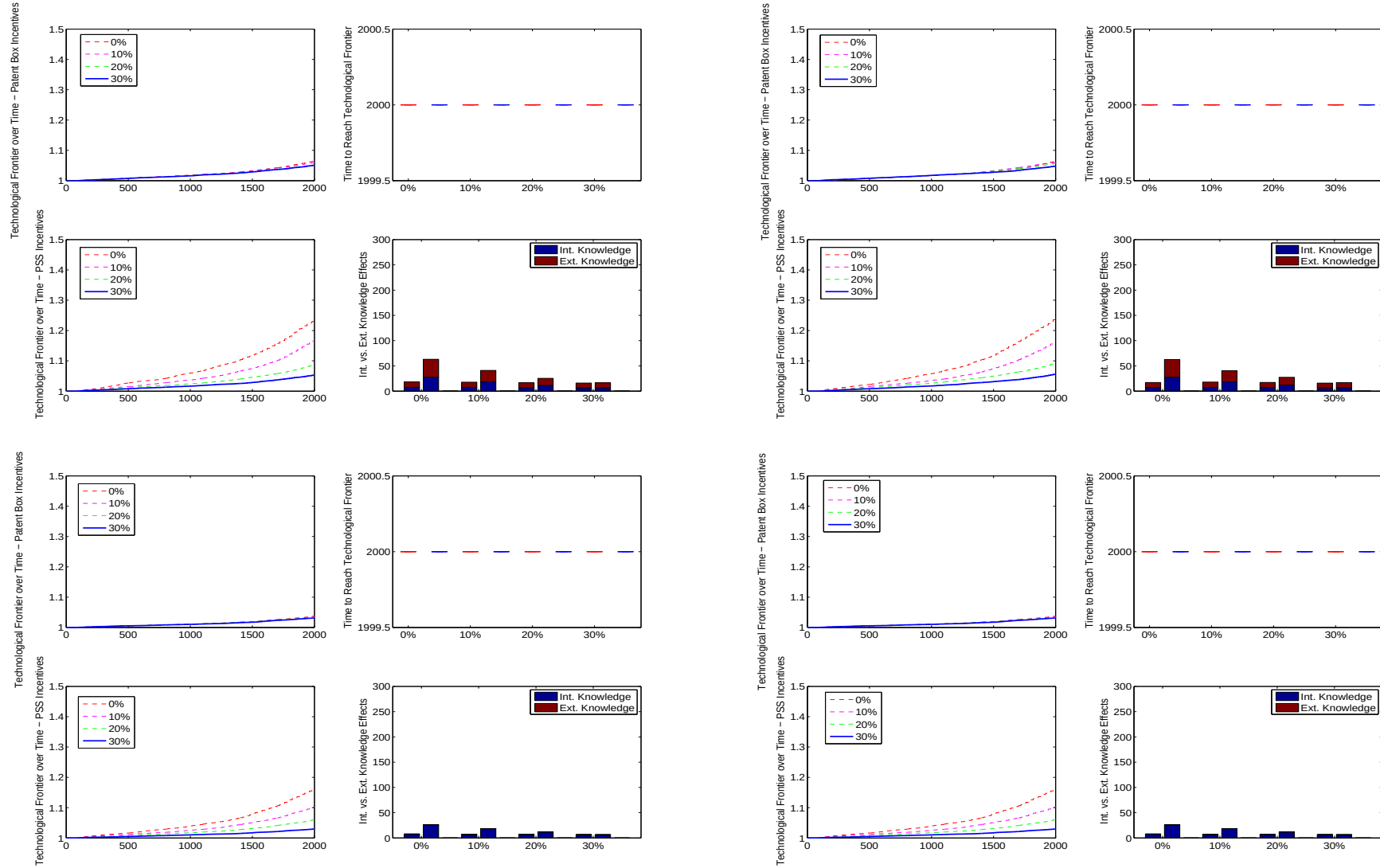


Figure 9: Technology indicators with evolutionary technology and R&D investments

Note: Indicators are represented here for the scenario with fast technology and spillovers (upper leftmost four graphs) and for the scenario with spillovers but slow technology (upper rightmost four graphs), under low mobility. In addition, the two technology speeds with no spillovers present are presented in the lower panels: fast technology on the left and slow technology on the right. As before, blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained with 25 replications for each parameter combination.

dominant firm faces no incentives to improve the compensation contract, therefore (and regardless of taxation) the market is stuck in a slower technological progress (see Figure 9 and compare with Figure 5) with agents getting, on average, larger shares of the profits (compare the upper graphs in Figure 10 with the upper graphs in Figure 6).

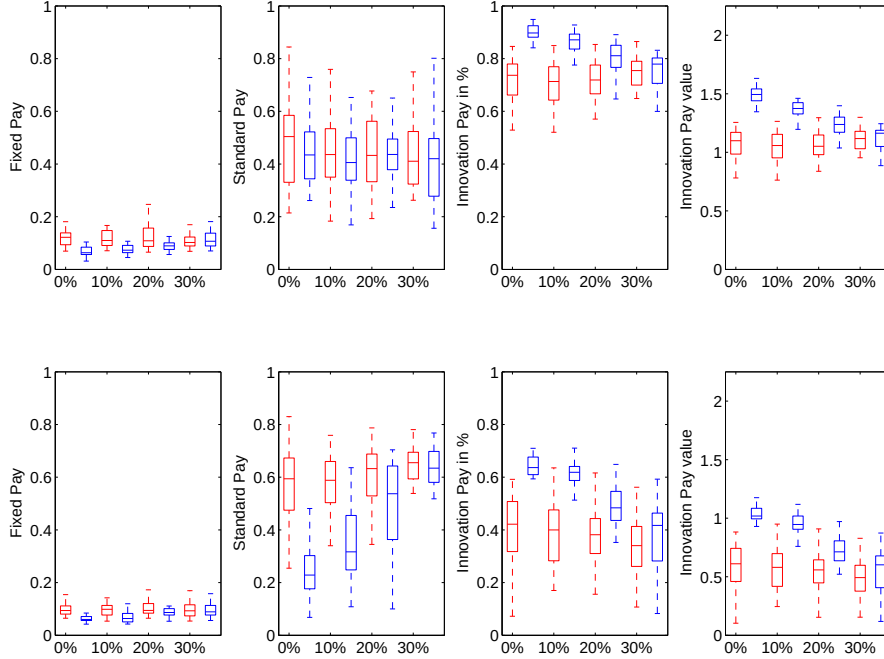


Figure 10: Convergence in contracts with evolutionary technology (scenario 3)

Note: Contracts are represented here under fast technology with spillovers, for the high mobility scenario (upper panel) and for the low mobility scenario (lower panel). Blue boxplots represent PSS tax incentives; red boxplots represent patent box incentives. Results are obtained with 25 replications for each parameter combination.

5 Regression analysis

In order to draw general conclusions based on the extensive number of parameter combinations tested in our model, we use regression analysis of the data generated by our ABM. This is particularly helpful as one can distinguish between the different effects of each specific parameter on our output measures and test their significance. We focus on three variables that are interesting policy-wise: the aggregate value of innovations generated per period (*InnValue*, this is the value reported in the rightmost graphs in Figures 2, 6 and 10), the tax revenues expressed as a percentage of the average revenues that were collected without any tax incentive (*TaxIndex*, this is the same tax efficiency index reported in Figures 4 and 8), and the technology level reached in the end (*TechL*, this last variable is only for scenarios with evolutionary technology and is the value of $tech_{frontier}$ the ABM produces in the last period).²⁰

²⁰We do not employ the elasticity index in the following analysis for two reasons. First, as already stated before, it is a very unstable and volatile index. The second reason is that the elasticity index is not informative alone, but only when interpreted in conjunction with measures of innovations done and tax revenues lost, therefore it is not suitable as a dependent variable in a regression analysis. As for

As all our dependent variables are continuous, we do not need to employ any specific estimation methods meant for count data. However, due to complexity of our ABM and no *a priori* knowledge on the distribution of observations and error terms, we need to use robust regression methods which are less sensitive to outliers and the assumption of normality.²¹ To better select the most fitting estimation method, we evaluate them based on their out-of-sample prediction. In short, one takes the total number of observations in each scenario (or pools them together), randomly splits the sample in two parts: 80% (to estimate coefficients) and 20% (to evaluate out-of-sample fit), and repeats that 100 times averaging the results. The method minimizing the Average Mean Absolute Error (AMAE) is considered to perform best.²²

Table 1: Out-of-sample model evaluation

Estimation method	Scenario 1	Scenario 2	Scenario 3	All Scenarios
Dependent variable is <i>InnValue</i>				
OLS	6.89 (3)	10.32 (3)	5.01 (3)	8.18 (2)
Huber	6.83 (1)	10.29 (2)	4.96 (2)	8.15 (1)
Bisquare	6.83 (2)	10.28 (1)	4.95 (1)	8.18 (3)
Dependent variable is <i>TaxIndex</i>				
OLS	6.69 (3)	10.39 (3)	5.04 (3)	8.06 (2)
Huber	6.62 (1)	10.35 (2)	5.00 (2)	8.05 (1)
Bisquare	6.63 (2)	10.34 (1)	4.99 (1)	8.08 (3)
Dependent variable is <i>TechL</i>				
OLS		10.39 (3)	5.01 (3)	7.84 (1)
Huber		10.35 (2)	4.96 (2)	7.86 (2)
Bisquare		10.34 (1)	4.95 (1)	7.90 (3)

Results are reported in AMAE with model rank in parentheses.

Results of the out-of-sample evaluation exercise are presented in Table 1 for OLS and for two robust M-estimators, the Huber estimator (Huber, 1973) and the Bisquare estimator (Beaton and Tukey, 1974).²³ As one can see, results from the three methods are very similar with Huber and Bisquare consistently outperforming OLS in terms of AMAE ranking when considering any of the scenarios separately. When all scenarios are pooled together, OLS and Huber are found as better alternatives, and the best performing between these two depends on the particular dependent variable under consideration. In the regression analysis that follows, we employ the best AMAE-ranked method for each scenario, that is, the estimation method with the lowest AMAE value which is the method marked with a (1) in Table 1 for the scenario and dependent variable under scrutiny.

Regression results with the preferred estimation methods are presented in Table 2. Note that when using *TaxIndex* as dependent variable we drop observations produced with no tax incentives (and therefore we have $N = 6,000$ and not 8,000), as *TaxIndex* is always equal to 1 without tax incentives. Some effects are worth noting. First, in all scenarios a decrease in the tax rate levied on PSS increases the production of innovation (both in number and in value). This effect is very stable across all specifications of the ABM. On the contrary tax incentives on profits are only sometimes statistically

our other measure of innovation output, *InnValueTL*, we do not employ it here as results obtained with *InnValue* and *TaxIndex* already provide a clearer and more comprehensive picture.

²¹Technically, those estimators have a higher breakdown point and Gaussian efficiency (Maronna *et al.*, 1992).

²²We consider absolute and not squared errors on purpose in order not to put much weight on single large outliers. Output indicators, even after excluding elasticity of innovation, are still quite volatile and we put emphasis on robustness of the results obtained.

²³We employ Huber and Bisquare algorithms implemented by the Stata *robreg* package (Jann, 2010), using default settings.

significant and their coefficients are always much smaller than for PSS incentives when looking at *InnValue*. Patent box incentives become somewhat effective in scenario 3, but the magnitude of the effects on *InnValue* is much smaller compared to PSS incentives.

Table 2: Robust regression results

	Baseline scenario		Scenario 2			Scenario 3		
	InnValue	TaxIndex	InnValue	TaxIndex	TechL	InnValue	TaxIndex	TechL
Tax on PSS	−.746***	.023***	−1.484***	.023***	−.003***	−1.228***	.023***	−.002***
Tax on Profits	−.037***	.005***	−.021	.007***	−.001	−.088***	.005***	−.001***
Risk aversion	−.454***	−.001	−.814***	.001	−.001***	−.611***	.001***	−.001***
Low Mobility (dummy)	6.190***	−.069***	8.544***	−.016***	.032***	−13.85***	−.021***	−.014***
Speed of technology			19.700*	−.216*	.113***	−3.098	−.104	−.018*
Knowledge spillovers			325.700***	.074	1.520***	64.020***	.048	.354
Constant	48.540***	.264***	73.230***	.204***	.643***	60.830***	.237***	.595***
N	8,000	6,000	8,000	6,000	8,000	8,000	6,000	8,000
R ²	.60	.72	.77	.59	.83	.87	.63	.85

***,**, * Statistically significant, respectively, at the 1, 5 and 10% level

Second, we see that in the scenarios 2 and 3 a larger technological level *TechL* is reached in the end when PSS incentives are introduced. Patent box incentives have smaller effects (statistically significant only in scenario 3) on technological level. These results bring powerful policy implications: as long as accelerating the speed at which the economy develops a new technology is a desirable policy outcome, the PSS incentive – through providing incentives to agents to commit to innovative tasks even under not so attractive contracts from the point of view of employers – may allow for this effect as firm-specific technological capabilities develop faster thanks to the larger amount of knowledge generated through own accumulation and spillovers from competitors. Patent box incentives can still modify the way technology evolves in scenario 3 (as firms are left with proportionally larger net profit, a portion of which they invest into R&D), but in a milder way and, as already discussed, with lower power to foster aggregate innovation.

Third, the effects of mobility differ across scenarios. In scenarios 1 and 2 low mobility (which is represented by a Low Mobility dummy equal to 1) affects convergence contracts so that PSS pay γ is on average larger (refer again to Figures 2 and 6). But in scenario 3 mobility has the opposite effect. This happens because of the following mechanisms. The general case observed in scenarios 1 and 2 (and also what intuition suggests) is that when mobility is large information about the relative profitability of different contracts is generated faster. But, principals then tend to focus contract competition on the most profitable type of job task, which in our setup for initial parameters is (on average) the standard task, and therefore more workers are pushed to commit to standard tasks when mobility is high.²⁴ On the contrary, in scenario 3 with low mobility, the combined effect of low observability of competing employers (again because of worse generation of aggregate information) together with the constraints to the availability of innovative job

²⁴We also run a series of simulations by setting the value of innovative task on average being more profitable ($Y = 2$). As a result, high mobility always fosters innovations created and technology reached at the end of the simulations in all three scenarios. An in-depth analysis of the endogenous variables produced by the ABM confirmed that high mobility generates more information about the relative profitability of the job tasks, thus making principals steer agents' task choice toward the innovative task.

tasks (because of R&D investments, which only few firms can sustain) practically shuts down innovation and leads principals to compete mostly on standard contracts anyway. Introducing high mobility in scenario 3 makes innovation visible to agents again, and in comparison to low mobility, innovation contracts become less random and more driven by the contract algorithm as per scenarios 1 and 2.

It is still information that drives the effect of mobility on contracts and innovation in scenario 3, only in a different way. The consequence of the effects of mobility on contracts is that under scenarios 1 and 2, low mobility is able to foster innovation and (in scenario 2) technological level, thanks to worse production of aggregate information about contracts and job tasks profitability. In scenario 3, high mobility is associated with lower aggregate innovation and technological level (in comparison to low mobility) as the better generation of information has the dominant effect of making firms somewhat compete over innovation contracts.

To further inquire the relative efficiency of our two tax incentives, we compare the coefficients obtained in different scenarios across a number of specifications. We focus again on estimates obtained using the previous three variables (*InnValue*, *TaxIndex*, and *TechL*) as dependent variable. Coefficients are compared by means of a z-test following Clogg *et al.* (1995) and Paternoster *et al.* (1998). Therefore, we compute:

$$Z = \frac{|b_1 - b_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad (17)$$

where b_1 and b_2 are the coefficients obtained by two distinct regressions where only one condition differs (namely: high or low mobility in all scenarios 1-2-3; scenario 2 (no R&d investments) or 3 (with R&d investments); fast or slow technology in scenarios 2-3; presence of spillovers in scenarios 2-3),²⁵ and σ_1 and σ_2 are the coefficients' standard errors. If $Z > 1.96$, one can conclude that the difference in coefficients is statistically significant at the 5% level. We run the z-tests for the coefficients obtained for both tax incentives. Results are summarized in Table 3.

From the findings in Table 3 we can infer some additional properties of the tax incentives. The mobility of the job market significantly affects the efficiency of the PSS tax incentives in terms of innovation produced, tax revenues collected and technology reached (the z-test is always larger than 1.96). Generally low mobility is associated with larger effects of the PSS tax incentive, but as discussed before the way mobility affects outcomes differs based on whether we introduce R&D investments or not in the model. In scenario 3, we obtain that high mobility makes the profit tax incentive more efficient (more innovation is produced while less tax revenues are lost). The presence of R&D investments also makes a difference for the PSS incentives, and as we see in Table 2 this difference goes in the direction of making the PSS incentive generally less effective.

In scenarios 2 and 3, the intrinsic speed at which technology progresses is never a significant driver of the relative power of the two incentives. The presence of spillovers appears to matter only for the PSS incentive as they amplify its positive effects on the innovation produced and on the final technology reached by the economy.

²⁵We run an extensive set of estimates with different degrees of risk aversion and found that PSS incentives are quantitatively less effective with larger risk aversion, while no visible effect was detected for the profit tax incentive. These observed behaviors of the ABM motivated the decision not to test for risk aversion when comparing relative efficiency of tax incentives under different scenarios (even though the data set used to build the comparisons summarized in Table 3 was produced also by changing risk aversion with the parameters explained in previous sections).

Table 3: Comparison of coefficients under different scenarios: z-test values and corresponding regression coefficients (in parentheses)

Type of comparison	Tax on PSS	Tax on Profits
	Dep. variable is <i>InnValue</i>	
Low vs. High mobility	2.27 (1.107548 - 1.053101)	1.31 (.073850 - .047105)
ABM 2 vs. ABM 3	13.93 (1.483516 - 1.228208)	4.40 (.020962 - .088701)
Faster vs. Slower technology	0.16 (1.376538 - 1.370723)	0.81 (.057090 - .080807)
Spillovers vs. No spillovers	11.34 (1.579780 - 1.195288)	0.07 (.069434 - .067166)
	Dep. variable is <i>TaxIndex</i>	
Low vs. High mobility	7.01 (.022535 - .024229)	2.03 (.006128 - .005654)
ABM 2 vs. ABM 3	(1.17) (.023601 - .023236)	6.30 (.007303 - .005348)
Faster vs. Slower technology	1.53 (.023233 - .023707)	0.64 (.006248 - .006448)
Spillovers vs. No spillovers	0.90 (.023877 - .023090)	1.27 (.006789 - .005921)
	Dep. variable is <i>TechL</i>	
Low vs. High mobility	5.96 (.003308 - .002721)	1.21 (.000174 - .000055)
ABM 2 vs. ABM 3	20.33 (.003100 - .002128)	4.14 (.000004 - .000133)
Faster vs. Slower technology	0.83 (.002776 - .002689)	0.39 (.000078 - .000110)
Spillovers vs. No spillovers	8.79 (.003206 - .002296)	0.31 (.000082 - .000108)

Results in bold highlight values larger than the threshold value ($Z > 1.96$).

6 Conclusions

We compared two different tax incentives meant to foster innovation in a knowledge-intensive economy by employing an Agent-Based Model (ABM) under different specifications with regard to labor mobility, risk aversion, and evolution of technology.

Our results strongly support the view that a tax incentive on profit sharing schemes might be better than, or a complementary tool to, the use of tax incentives on corporate income taxation, in a context where knowledge and the willingness of agents to engage in innovative activities are more important than the capital investments made by principals. A PSS incentive appears to be a powerful policy instrument to foster effort commitment by innovative employees, and in a framework where technology is modeled as a continuous evolutionary process it was also shown to give impulse to technological progress. Simulation results suggest that, starting from a condition where innovation value is expected to be on average less profitable than standard production, a PSS tax incentive would produce the maximum increase in aggregate innovation and technological progress in a pure knowledge economy without R&D investments, with knowledge spillovers, low mobility of workers and moderate risk aversion. When R&D investments constrain the ability of workers to commit effort into innovative job tasks, a PSS tax incentive was found to have the strongest effects under high mobility of workers.

Moreover we identify an additional channel through which labor mobility might affect innovation. Previous studies (Saxenian, 1996, Almeida and Kogut, 1999, Scarpetta and Tressel, 2004, Miguélez and Moreno, 2013) have documented that a larger inter-firm mobility of technical workers is associated with larger local production of innovation. The channel identified by researchers through which mobility can be innovation-enhancing is the transfer of knowledge caused by highly skilled workers moving between companies. This

could help explain why a region where mobility is particularly high (like Silicon Valley, Fallick *et al.*, 2006) also features more intense production of innovations. Our model provides a different (albeit compatible) view: mobility also affects the generation of information about the effectiveness of new organizational practices, like new compensation policies. With large expected innovation value (as it was probably the case in Silicon Valley during the 1990s), or with low innovation value coupled with R&D investment constraints, mobility boosts innovation by letting firms more easily discover the best compensation contract to offer to their technical workforce. And as long as workplace innovation is relevant in affecting the generation of innovation (as argued by Black and Lynch, 2004), this information-related channel might amplify the innovation-enhancing properties of inter-firm labor mobility.

From a policy perspective, all tax incentives have to be evaluated under the lens of cost-benefit analysis. From our study it emerges that the benefits in terms of an increase in the aggregate production of innovation are expected to be significant. But we also see that tax incentives on PSS are costly in terms of loss of tax revenues collected, as the rapid decrease of our indices of tax efficiency has shown. Therefore, we cannot make any claim about the optimal level of such incentives, and only point to the fact that, in a knowledge-intensive economy which shares the traits of the economy represented in our model, PSS incentives are likely to be more powerful in fostering innovation than tax incentives for R&D based on corporate taxation. In a scenario with evolutionary technology, the tax incentive on PSS is found to be strongest in enhancing the aggregate production of innovation for a given loss in tax revenues. When R&D expenditures on the side of the principals are included in an evolutionary technology setting, the two tax incentives appear complementary as both could be employed at the same time thus minimizing the loss in tax revenues while maximizing the induced increase in innovation.

In real markets, though, principals could bargain with employees and switch part of the workers' compensation to PSS in order to benefit from a reduced tax rate, rather than to attract and motivate qualified workers. Although our model does not feature an explicit choice on tax avoidance, it nevertheless allows to appreciate the potential loss in tax revenues. As long as offering more PSS drives agents into innovation (even though only as a by-product of tax dodging strategies), our model and its main conclusions hold notwithstanding possible tax avoidance behavior. On the other hand it might be that tax avoidance makes principals offer a different type of PSS to agents, which is not as capable to drive them into innovative job tasks nor to make them commit more effort. Previous studies (Lerner and Wulf, 2007) highlighted how the specific design of PSS (in terms of long-term versus short-term incentives) is relevant in defining how effective these payment schemes are in fostering innovation. But future research should clarify whether (and how) PSS compensation differs in case it is the outcome of tax avoidance from PSS introduced as a motivational device.

An important question becomes – which of the constraints, financial or organizational (e.g., lack of skilled personal as it is specified in the Mannheim Innovation Survey) is more critical for firms willing to innovate. Depending on which constraint is more critical for a specific industry, the one or the other incentive becomes more important.

The present study provides results that are in principle testable by means of empirical analysis. While the kind of PSS tax incentives discussed are not really offered in any country, tax reductions exist in a number of countries for some types of PSS, for instance, stock options. Our prediction would be that PSS incentives are able to induce a larger increase in innovation in firms for which intangibles are more important (like for most of

New Economy companies), than comparable incentives on corporate income.

References

- Acharya, V.I V., M. Pagano and P. Volpin (2013). Seeking alpha: Excess risk taking and competition for managerial talent. Working Paper 18891. National Bureau of Economic Research.
- Almeida, P. and B. Kogut (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management science* **45**(7), 905–917.
- Andersson, F., M. Freedman, J. Haltiwanger, J. Lane and K. Shaw (2009). Reaching for the stars: Who pays for talent in innovative industries?. *The Economic Journal* **119**(538), 308–332.
- Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In: *The rate and direction of inventive activity: Economic and social factors*. pp. 609–626. Nber.
- Balkin, David B. and Luis R. Gomez-Mejia (1984). Determinants of R&D compensation strategies in the high tech industry. *Personnel Psychology* **37**(4), 635–650.
- Baranchuk, N., G. MacDonald and J. Yang (2011). The economics of super managers. *Review of Financial Studies* **24**(10), 3321–3368.
- Beaton, A. E. and J. W. Tukey (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics* **16**(2), 147–185.
- Beetsma, R.M.W.J. and P. C. Schotman (2001). Measuring risk attitudes in a natural experiment: data from the television game show lingo. *The Economic Journal* **111**(474), 821–848.
- Benabou, R. and J. Tirole (2013). Bonus culture: Competitive pay, screening, and multitasking.
- Black, S. E. and L. M. Lynch (2004). What’s driving the new economy?: the benefits of workplace innovation. *The Economic Journal* **114**(493), F97–F116.
- Boldrin, M. and D. K. Levine (2008). Against intellectual monopoly. Cambridge books. Cambridge University Press.
- Bolton, P. and M. Dewatripont (2005). *Contract Theor.* The MIT Press. Cambridge and London.
- Cantner, U. and A. Pyka (1998). Absorbing technological spillovers: Simulations in an evolutionary framework. **7**(2), 369–397.
- Clogg, C. C., E. Petkova and A. Haritou (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology* pp. 1261–1293.
- Cohen, W. and H. Sauermann (2007). Schumpeter’s prophecy and individual incentives as a driver of innovation. *Perspectives on Innovation, F. Malerba, S. Brusoni, eds* pp. 73–104.

- d'Andria, D. (2014). Taxation and incentives to innovate: a principal-agent approach. *Jena Economic Research Papers* 28.
- d'Andria, D. and S. Übelmesser (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, H. M., M. Grossmann, M. Lang and S. Wey (2013). Incentive effects of bonus taxes in a principal-agent model. *Journal of Economic Behavior & Organization* **89**, 93–104.
- Ederer, F. P. (2013). Incentives for parallel innovation. SSRN Scholarly Paper ID 2309664. Social Science Research Network.
- Egbetokun, A. and I. Savin (2014). Absorptive capacity and innovation: When is it better to cooperate?. *Journal of Evolutionary Economics* **24**(2), 399–420.
- Evers, L., H. Miller and C. Spengel (2013). Intellectual property box regimes: Effective tax rates and tax policy considerations. Technical Report 13-070. ZEW Discussion Papers.
- Fallick, B., C. A. Fleischman and J. B. Rebitzer (2006). Job-hopping in silicon valley: some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics* **88**(3), 472–481.
- Gemmell, N., R. Kneller and I. Sanz (2011). The growth effects of corporate and personal tax rates in the OECD. *Globalisation and Development Centre*.
- Griffith, R., H. Miller and M. O'Connell (2011). Corporate taxes and the location of intellectual property. SSRN Scholarly Paper ID 1865574. Social Science Research Network. Rochester, NY.
- Grossmann, M., M. Lang and H. M. Dietl (2011). Why taxing executives' bonuses can foster risk-taking behavior. SSRN Scholarly Paper ID 1952729. Social Science Research Network. Rochester, NY.
- Hall, B. and J. Van Reenen (2000). How effective are fiscal incentives for R&D? a review of the evidence. *Research Policy* **29**(4), 449–469.
- Hellmann, T. and V. Thiele (2011). Incentives and innovation: A multitasking approach. *American Economic Journal: Microeconomics* **3**(1), 78–128.
- Holmstrom, B. (1989). Agency costs and innovation. *Journal of Economic Behavior & Organization* **12**(3), 305–327.
- Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization* **7**, 24.
- Huber, P. J. (1973). Robust regression: asymptotics, conjectures and monte carlo. *The Annals of Statistics* pp. 799–821.
- Ittner, C. D., R. A. Lambert and D. F. Larcker (2003). The structure and performance consequences of equity grants to employees of new economy firms. *Journal of Accounting and Economics* **34**(1), 89–127.

- Jann, B. (2010). robreg: Stata module providing robust regression estimators. Available from <http://ideas.repec.org/c/boc/bocode/s457114.html>.
- Lerner, J. and J. Wulf (2007). Innovation and incentives: Evidence from corporate R&D. *Review of Economics and Statistics* **89**(4), 634–644.
- Manso, G. (2011). Motivating innovation. *The Journal of Finance* **66**(5), 1823–1860.
- Maronna, R.A., D.R. Martin and V.J. Yohai (1992). *Robust Statistics. Theory and Methods*. Wiley. Chichester.
- Metcalfe, J.S. (1994). Competing, fisher’s principle and increasing returns in the selection process. *Journal of Evolutionary Economics* **4**, 327–346.
- Miguélez, E. and R. Moreno (2013). Research networks and inventors’ mobility as drivers of innovation: evidence from europe. *Regional Studies* **47**(10), 1668–1685.
- Myles, G. D. (2009). Economic growth and the role of taxation theory. OECD Economics Department Working Paper 713. OECD Publishing.
- Nelson, R. R. (1959). The economics of invention: A survey of the literature. *The Journal of Business* **32**(2), 101–127.
- Nelson, R. R. and S. G. Winter (2009). *An Evolutionary Theory of Economic Change*. Harvard University Press.
- Paternoster, R., R. Brame, P. Mazerolle and A. Piquero (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology* **36**(4), 859–866.
- Radulescu, D. M. (2012). The effects of a bonus tax on manager compensation and welfare. *FinanzArchiv: Public Finance Analysis* **68**(1), 1–16.
- Reinganum, J. F. (1989). The timing of innovation: Research, development, and diffusion. *Handbook of industrial organization* **1**, 849–908.
- Rosen, S. (1981). The economics of superstars. *The American economic review* pp. 845–858.
- Russo, B. (2004). A cost-benefit analysis of R&D tax incentives. *Canadian Journal of Economics* **37**(2), 313–335.
- Sauermann, H. and W. M. Cohen (2010). What makes them tick? Employee motives and firm innovation. *Management Science* **56**(12), 2134–2153.
- Saxenian, A. (1996). *Regional advantage*. Harvard University Press.
- Scarpetta, S. and T. Tressel (2004). Boosting productivity via innovation and adoption of new technologies: any role for labor market institutions?. *World Bank Working Paper*.
- Yildizoglu, M. (2002). Competing R&D strategies in an evolutionary industry model. *Computational Economics* **19**(1), 51–65.

Appendices

A Details on model calibration

Table 4: Parameters used

Description	Symbol	Value
Number of agents	N	100
Number of principals	M	10
Tax on labor income from standard tasks and fixed wage	t_S	30%
Tax on labor income from innovative tasks	t_I	between 0 and 30%
Tax on profits from standard tasks	τ_S	30%
Tax on profits from innovative tasks	τ_I	between 0 and 30%
Scale-down parameter	n	0.55
Skill level of agents	k	drawn from $N(0, 0.3)$
Mean of the baseline innovation value	Y	1.4
Standard deviation of the innovation value	σ	0.2
Absolute risk aversion	r	between -10 and +10 (0 for risk neutral agents)
Parameter for standard task productivity (exponent)	a	0.8
Parameter for standard task productivity (denominator)	b	1.2
Mobility cost	mc	0 or 100%
Technological level at start	$tech_{start}$	set at 1
Maximum technological level firms can reach	$tech_{max}$	set at 1.5
Speed of accumulation of knowledge spillovers	ψ	either 0 or 0.05
Cap set to total one-time increase in technology	ϕ	either 0.025 or 0.05 (for slow and fast technology)
R&D costs for each innovation task slot	RD	0.025
Effort invested in standard task	e_S	} determined endogenously
Effort invested in innovative task	e_I	
Fixed wage	w	
Pay for the standard task (variable pay wage)	β	
Pay for the innovative task (PSS) in the form of share of innovation value	γ	

B Risk sensitivity analysis

Figure 11: Risk aversion sensitivity analysis

