

Verbeck, Matthias

Conference Paper

Contracting with Researchers

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Contracts, No. A12-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Verbeck, Matthias (2015) : Contracting with Researchers, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Contracts, No. A12-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/112963>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Contracting with Researchers

Matthias Verbeck^{*†}

March 1, 2015 - Preliminary stage, please do not quote!

Abstract

We study a setting in which one or two agents conduct research on behalf of a principal. The agents' binary performance level (success or failure) depends on their invested research effort, and their choice of a research technology that is uncertain in respect of its aptness to generate a success. While in the single-agent-setting the agent has no incentive to deviate from the principal's preferred technology choice, this is not generally true for the multiagent-setting. When technologies are mutually exclusive - only one of them will be suitable for yielding a high output - we show that there exists a contract that aligns the principal's and the agents' interests. However, under the plausible assumption of scientists' free technology choice, our results suggest that there is a bias towards mainstream-research: Agents choose promising technologies more often than socially optimal.

Keywords: Moral hazard, Hidden action, Incentives for research performance, Motivation of scientists, Academic organization

^{*}Department of Economics, Philipps-Universität Marburg

[†]For very helpful comments I thank the participants of the Economics department's internal research seminar of Philipps-Universität Marburg. I am especially grateful to Elisabeth Schulte.

1 Introduction

Although some time has passed since Max Weber wrote his famous essay "Science as a vocation", one fact about the character of scientific research is as true as it was almost 100 years ago:

"Yet it is a fact that no amount of such enthusiasm, however sincere and profound it may be, can compel a problem to yield scientific results" (Weber 1946 [1917], p.137).

By its nature, scientific progress is uncertain and hardly predictable. There are numerous anecdotes of researchers who failed to attain scientific success, despite large amounts of effort and enthusiasm, whereas in other instances major discoveries were simply the result of serendipitous findings. Nevertheless, most scientists would certainly agree that scientific success is - by and large - the result of hard work that finally paid off. It is therefore reasonable to assume that a researcher's efforts are positively related to his prospects of creating meaningful outputs.

For any researcher - be it in academia or in the private sector - there are clear incentives to maximize *individual* research outputs. In academia, reward for success highly concentrates on the individual. Those researchers that manage to publish in high-ranked journals get the recognition of their peers and increase the prospects of obtaining top academic positions. Moreover, they will also often get more money, as pay for performance has arrived in the academic field. But also a researcher that is employed in the private sector will improve his prospects for personal growth by individual success. Given these individual incentives to perform well, there seems to be no room for a misalignment of the interests of principal and agents. Hence, one could uncritically accept that (academic) researchers are equipped with academic freedom, and most voices from academia will rarely question this freedom. If anything, freedom of research is seen as a precondition for academic suc-

cess. In this contribution however, we make the case for why this view might fall short and why coordination of research efforts might outperform uncoordinated research. Our analysis rests on the assumption that a researcher chooses an effort level, combined with a certain technology for teckling the actual scientific problem. By "technology" we mean the approach, the method or the "school" which a scientist employs. Two motivating examples will illuminate this issue:

Example 1: Cancer research

A national health agency provides funding for research on how to cure lung cancer. Academic researchers can now choose from several methods to attain this goal. They could for example conduct research on improved chemotherapies, better methods for early diagnosis, or a completely new genetic therapy approach. Methods differ in their prospects for success. Which technology will a rational researcher choose when he is motivated by personal success? Is this technology in the best interest of the principal?

Example 2: Oil drilling

An oil company hires one or several experts, (say geologists), to conduct exploratory drillings with the aim of finding a new oil spring. There are two possible locations for test drillings available that exhibit different probabilities of containing oil. Again, which location will a rational expert choose?

As these examples suggest, technology choice - if unobservable - might not always be aligned with the principal's interest. If we regard the principal as a social planner who wants to maximize society's benefits from research, this can be interpreted as a conflict of individual and social interests. Researchers have an incentive to use the technology that maximizes the chance of individual high output, whereas - in the case of multiple researchers - the principal wants to maximize the overall probability of a high output. Our model analyzes this conflict in detail. We will show that in the multiagent-case, a

bias towards mainstream- research can occur: More promising technologies are chosen for a larger parameter range than socially optimal, given that free technology choice is guaranteed.

2 Related Literature

Our research contributes to several branches of the economic literature. First and foremost it is related to the literature on contract theory, moral hazard and hidden actions. While in the classic papers of Holmstrom (1979) and Grossman and Hart (1983) the effort level is unobservable, leaving the principal with lower expected return, in our model we extend the principal's information restrictions and make technology choice an unobservable part of any agent's strategy. Since agents have a broader action space than usual, our work is related to contributions on moral hazard and multitasking, e.g. Dewatripont et al. (2000). Moreover, as we allow the principal to employ several agents, our work is also related to organization theory and incentives in teams, like Holmstrom (1982), Mookherjee (1984) and Itoh (1991). In contrast to much of the multiagent literature however, free-riding due to unobservable individual output is excluded by assumption in our model, as this problem is not in the center of our interest.

Our research is also connected to the literature on "centralization vs. decentralization" (e.g. Spence (1975), Radner (1992), Hurwicz (1973) and Mookherjee (2006)). Similar to earlier findings, we confirm that a centralized arrangement is superior to a decentralized one.

Moreover we build on the vast body of literature on the functioning and the shortcomings of academic research. Academic research has been consistently criticized from various sides. Among others, Frey (2003), Starbuck (2005) and Grey (2010) criticize the prevailing system of academic peer-reviewed publication as unreliable, intransparent and discouraging to innovative research.

Ioannidis (2005) points to a publication bias towards wrong results as they are often more interesting than correct ones. Doemeland and Trevino (2014) ask the valid question, how much of academic research actually diffuses into the public. However none of the contributions that we know deals with the issue of too much mainstream-research from a principal-agent-perspective. Hence, our research is new to the best of our knowledge.

3 The Model

3.1 Assumptions and Main Setting

There is a technology space J consisting of two elements l (left) and r (right) and a corresponding state space $\Omega = \{l, r\}$ that defines the technology's quality. For example, if the state of the world is r , this means that a success can only be achieved with technology r , but not with technology l . Hence we can say that the quality of r is "good" and the quality of l is "bad". As we restrict the state space to only two states, it must necessarily be that exactly one technology is good and the other is bad, so that we can speak of mutually exclusive technologies. This assumption might seem restrictive, since a larger state space (e.g. both technologies are good, technology r is good and technology l is bad, technology l is good and technology r is bad, both technologies are bad) appears to be more realistic. We will discuss this issue in greater detail in the discussion section.

We have a risk-neutral principal (She) who gains a positive utility R from a successful research. Her expected utility is then defined as

$$V(R, q, w) = R \cdot q - E[W]. \tag{1}$$

Let $q_i \in \{0, 1\}$ denote agent i 's binary output of the research project, which can be either a success or a failure. Furthermore, $W = \sum_{i=1}^n w_{ij}$ denotes the overall compensation that is paid to n agents, where n either equals one or two, and the respective cases are analyzed separately in sections 3.2 and 3.3.

Moreover, w_{ij} indicates the compensation that agent i obtains for research with technology j . The contract may condition the payment on the agent's behavior (if observable) or the realized output.

Every agent i experiences his utility according to the following utility function, which is separable in income and effort:

$$U_i = u_i(w_{ij}) - e_{ij}. \quad (2)$$

Any agent can invest research effort $e_{ij} \in [0, \infty)$, causing individual costs of e_{ij} . Note, we make the crucial assumption that an agent can only exert effort by using *one technology* at a time. Hence, the case of an agent who simultaneously handles multiple technologies is ruled out by assumption. An agent's strategy can therefore be fully described as $(e_{ij}, j) \in \mathbb{R}_0^+ \times J$.

As standard in the literature, we assume that

$$u'(\cdot) > 0, \quad u''(\cdot) \leq 0.$$

The agent's reservation utility level is defined as $\bar{U} = 0$ for every agent. An agent who exerts research effort with technology j will have the following probability of success:

$$P(q_{ij} = 1 \mid e_{ij}) = \begin{cases} \rho(e_{ij}), & \text{if } \omega = j \\ 0, & \text{else.} \end{cases} \quad (3)$$

Here $\rho(e_{ij})$ defines the probability of a high output, given that technology j is a good technology. As usual, the probability of success is increasing and concave in the agent's effort. We add the following technical assumptions that guarantee an interior solution:

$$\rho'(0) \cdot R \cdot \pi_j > 1, \quad \rho(0) = 0, \quad \rho(\infty) = 1.$$

Technological uncertainty is introduced by $\pi_{ij} = P(\omega = j)$ which denotes the probability that technology j is good and the alternative technology

is bad. Vice versa, $1 - \pi_{ij}$ indicates the probability that the alternative technology is good and technology j is bad. We refer to the more promising technology as the mainstream-technology and the lower technology as the outsider-technology.

We assume that $\pi(\cdot)$ represents a common prior for all players, such that $\pi_{1j} = \pi_{2j} = \pi_j$. An agent's overall probability of success, provided the usage of technology j , is then given as

$$P(q_{ij} = 1 \mid e_{ij}) = \rho(e_{ij}) \cdot \pi_j. \quad (4)$$

We further assume that all of the above (number of agents, cost functions, utility functions, etc.) is common knowledge. Note that neither the principal, nor any agent has an endogenous preference regarding the usage of a certain technology and effort costs are independent of technology choice. Hence, the likelihood to bring about a success is the only dimension in which technologies differ.

The course of action is as follows:

1. Nature chooses ω according to $\pi(\cdot)$.
2. The principal offers a take-it-or-leave-it-contract to the agent(s) which is either accepted or rejected.
3. If an agent accepts the contract, he will choose optimal effort levels and optimal technology, given his characteristics and the conditions of the contract. If the contract is rejected, the game ends.
4. Output realizes and each contracting obtains remuneration according to the specified conditions.

3.2 One Single Researcher (n=1)

3.2.1 Symmetric Information

The optimal contract solves the following program

$$E(V_j(\cdot)) = \max_{j, e_j, \underline{w}_j, \bar{w}_j} = \rho(e_j) \cdot \pi_j \cdot (R - \bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot (-\underline{w}_j) \quad (\text{P I:SI})$$

subject to

$$\rho(e_j) \cdot \pi_j \cdot u(\bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot u(\underline{w}_j) - e_j \geq 0. \quad (\text{IR I:SI})$$

Here \bar{w}_j indicates the wage level that is paid to the agent in the case of success, whereas \underline{w}_j denotes the wage level in the event of fruitless research. As the IR constraint is going to be binding, and postponing the choice of j , we can define the resulting Lagrangian as

$$\mathcal{L} = \rho(e_j) \cdot \pi_j \cdot (R - \bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot (-\underline{w}_j) + \lambda \cdot [\rho(e_j) \cdot \pi_j \cdot u(\bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot u(\underline{w}_j) - e_j] \quad (5)$$

First-order conditions:

$$\frac{\partial \mathcal{L}}{\partial e_j} = \rho'(e_j) \cdot \pi_j \cdot [R - \bar{w}_j + \underline{w}_j] + \lambda \cdot [(\rho'(e_j) \cdot \pi_j \cdot [u(\bar{w}_j) - u(\underline{w}_j)] - 1)] = 0 \quad (6)$$

$$\frac{\partial \mathcal{L}}{\partial \bar{w}_j} = \rho(e_j) \cdot \pi_j \cdot (-1) + \lambda \cdot [\rho(e_j) \cdot \pi_j \cdot u'(\bar{w}_j)] = 0 \quad (7)$$

$$\frac{\partial \mathcal{L}}{\partial \underline{w}_j} = (1 - \rho(e_j) \cdot \pi_j) \cdot (-1) + \lambda \cdot [(1 - \rho(e_j) \cdot \pi_j) \cdot u'(\underline{w}_j)] = 0 \quad (8)$$

From (7) and (8) we can easily derive the optimal co-insurance conditions (Borch 1962) between principal and agent:

$$\frac{1}{u'(\bar{w}_j)} = \frac{1}{u'(\underline{w}_j)} \Leftrightarrow u'(\bar{w}_j) = u'(\underline{w}_j) \Leftrightarrow \bar{w}_j = \underline{w}_j = w_j. \quad (9)$$

Plugging the uniform wage w_j into (6) and by using (7) and (8), we obtain

$$\frac{1}{u'(w_j)} = \lambda = \frac{\rho'(e_j) \cdot \pi_j \cdot [R - w_j + w_j]}{1 - \underbrace{\rho'(e_j) \cdot \pi_j \cdot [u(w_j) - u(w_j)]}_0} \quad (10)$$

For a given location j optimal effort and wage levels are implicitly defined by

$$\rho'(e_j) = \frac{1}{\pi_j \cdot R \cdot u'(w)} \quad (11)$$

and

$$w_j = u^{-1}(e_j) \quad (12)$$

It is evident that the optimal effort e_j will rise in R and π_j and so does w_j . What remains to be shown is the choice of the optimal technology j . We can prove the intuitive result, that choosing the more promising technology is optimal from the principal's perspective:

Proposition 1. *Technology j is optimal for the principal, iff $\pi_j = \max(\pi_j)$.*

Proof. Suppose it is true that $\pi_j > \pi_k$ but $\exists k \neq j$ that yields a higher payoff, so that j is not optimal. Then it must be that $\rho(e_k) \cdot \pi_k \cdot R - w(e_k) > \rho(e_j) \cdot \pi_j \cdot R - w(e_j) \geq \rho(e_k) \cdot \pi_j \cdot R - w(e_k) \Rightarrow \pi_j < \pi_k$ which was ruled out by assumption. It follows that k is not optimal and hence j must be optimal, as an optimum exists. \square

We see that the optimal allocation of a single researcher is easy. If $\tilde{\pi}_{Pj}$ denotes the principal's cutoff value that defines optimality for choosing technology j , then j should be selected whenever $\pi_j \geq \tilde{\pi}_{Pj} = 0.5$. As we assume that the principal can perfectly observe the agent's actions, choosing a technology other than the principal's optimum is never rational, since the principal could inflict a punishment on the agent for doing so.

3.2.2 Asymmetric Information

In the case of unobservable action, the agent's incentive constraint becomes a part of the principal's optimization problem:

$$E(V_j^{SB}(\cdot)) = \max_{j, e_j, \underline{w}_j, \bar{w}_j} = \rho(e_j) \cdot \pi_j \cdot (R - \bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot (-\underline{w}_j) \quad (\text{P I:AI})$$

subject to

$$\rho(e_j) \cdot \pi_j \cdot u(\bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot u(\underline{w}_j) - e_j \geq 0 \quad (\text{IR I:AI})$$

$$e \in \operatorname{argmax}_{\hat{e}} \rho(\hat{e}_j) \cdot \pi_j \cdot u(\bar{w}_j) + (1 - \rho(\hat{e}_j) \cdot \pi_j) \cdot u(\underline{w}_j) - \hat{e}_j \quad (\text{IC I:AI})$$

To solve this problem, we can use the common First-Order-Condition approach (Holmstrom 1979), given our assumptions on $\rho(\cdot)$, $u(\cdot)$ and π_j . Thus, the agent's original incentive constraint is replaced by:

$$\rho'(e_j) \cdot \pi_j \cdot [u(\bar{w}_j) - u(\underline{w}_j)] = 1. \quad (13)$$

This leads to the Lagrangian

$$\begin{aligned} \mathcal{L} = & \lambda \cdot [\rho(e_j) \cdot \pi_j \cdot u(\bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot u(\underline{w}_j) - e_j] + \\ & \mu \cdot [\rho'(e_j) \cdot \pi_j \cdot [u(\bar{w}_j) - u(\underline{w}_j)] - 1]. \end{aligned} \quad (14)$$

From the first-order conditions we derive

$$\frac{1}{u'(\bar{w}_j)} = \lambda + \mu \cdot \frac{\rho'(e_j)}{\rho(e_j)} \quad (15)$$

$$\frac{1}{u'(\underline{w}_j)} = \lambda - \mu \cdot \frac{\rho'(e_j) \cdot \pi_j}{1 - \rho(e_j) \cdot \pi_j}. \quad (16)$$

In optimum it must be that $\mu > 0$, since otherwise $\underline{w}_j = \bar{w}_j$ and the agent's IC constraint could not hold. Hence, we obtain the typical result that a high output (i.e. a successful research project, leading to $q = 1$) is rewarded, and fruitless research ($q = 0$) is punished. As usual, we see that unobservable

effort entails an efficiency loss, because any specific effort level requires an additional risk premium to be accepted by an agent. Therefore we know that the second-best contract will yield the principal a strictly lower expected payoff than the symmetric one, or more formally:

$$E(V_j^{SB}(\cdot)) < E(V_j). \quad (17)$$

We now extend the degree of asymmetric information and assume that also the agent's technology choice is not observable by the principal. This issue requires some explanation: Depending on the scenario at hand, it might or might not be reasonable to assume that the principal is able to observe (or to verify in court) the actual technology which the agent used for executing his job. If the technology can be observed by the principal, she has a finer information set, as she is able to tell apart q_r from q_l . In our oil-drilling example the technology's observability seems plausible. For each site separately, the principal can see whether or not oil is bubbling. However in other settings, observability is far from obvious, since any non-expert principal will be unable to evaluate which of the available technologies has caused the success.

Following Proposition 1, it is still optimal to choose the mainstream technology from the principal's standpoint. It is easy to see, that a rational agent will also choose this technology.

Proposition 2. *Technology j is optimal for an agent, iff $\pi_j = \max(\pi_j)$.*

Proof. The agent's expected gain equals $\rho(e_j) \cdot \pi_j \cdot u(\bar{w}_j) + (1 - \rho(e_j) \cdot \pi_j) \cdot u(\underline{w}_j) - e_j$. If we replace π_j by some smaller π_k , the expected gains are strictly lower. \square

This means that the agent's and the principals threshold values must coincide ($\tilde{\pi}_{ij} = \tilde{\pi}_{Pj} = 0.5$), and there is no collision of interests between both parties.

3.3 A Team of Researchers (n=2)

3.3.1 Symmetric Information

We now turn to the case where the principal can employ a team of two agents. The structure is similar to the one-agent-case. Each agent is assigned to a specific technology and exerts research effort. The principal can choose to employ both agents, who either both use the same technology, or else, use different technologies. We will refer to the former option as "duplicated efforts" and the latter as "disjoined efforts". We will make the important assumption that *individual output* of agents is always observable, such that free-riding problems cannot occur in our setting.

From the principal's perspective, employing a second agent is - in expectation - always more profitable than employing only one agent. This is self-evident, since the principal can replicate the one-agent-case by offering a zero-wage to the second agent which makes her weakly better off. The interesting questions that arise in this context concern the optimal allocation of reserachers and the contract design that aims at inducing such an optimal allocation. As before, we postpone the optimal choice of technologies and take it as given.

Case 1: Disjoined Efforts:

If the reseraching agents are assigned to different technologies, optimal wage and effort levels for each agent respectively can be readily derived by equations (11) and (12). Hence there is no need to present the problem in detail here. The principal's overall expected payoff equals $E(V_{rl}) = E(V_r) + E(V_l)$.

Case 2 : Duplicated Efforts:

When both agents use the same technology, one technology of course remains deserted. When the unused technology's probability of being good is rather low (i.e. it is a clear outsider-technology), this can be rational from the principal's perspective, as bundling efforts with one technology yields a higher

expected return.

We will now derive optimal effort and wage levels when both agents use technology j . If we make use of equation (9) ($\bar{w}_{ij} = \underline{w}_{ij} = w_{ij}$) the maximization problem becomes

$$E(V_{jj}(\cdot)) = \underset{e_{i,j}, w_1, w_2}{max} = \pi_j \cdot [1 - (1 - \rho(e_{1j})) \cdot (1 - \rho(e_{2j}))] \cdot R - w_{1j} - w_{2j}. \quad (\text{P II 2:SI})$$

The overall probability of a high output equals the technology probability π_j multiplied with the counter probability of a double failure. The problem is subject to each agent i 's individual rationality constraint

$$u(w_{ij}) - e_{ij} \geq 0. \quad (\text{IR II 2: SI})$$

The resulting Lagrangian equals:

$$\mathcal{L} = \pi_j \cdot [1 - (1 - \rho(e_{1j})) \cdot (1 - \rho(e_{2j}))] \cdot R - w_{1j} - w_{2j} + \lambda_1 \cdot [u(w_{1j}) - e_{1j}] + \lambda_2 \cdot [u(w_{2j}) - e_{2j}] \quad (18)$$

We draw partial derivatives for e_{1j} and w_{1j} . Likewise we can draw the partial derivatives for agent 2:

$$\frac{\partial \mathcal{L}}{\partial e_{1j}} = \pi_j \cdot [-\rho'(e_{1j}) + \rho'(e_{1j}) \cdot \rho(e_{2j})] \cdot R + \lambda_1 \cdot [-1] = 0 \quad (19)$$

$$\frac{\partial \mathcal{L}}{\partial w_{1j}} = -1 + \lambda_1 \cdot [u'(w_{1j})] = 0 \quad (20)$$

From the last equations we can conclude that

$$\rho'(e_{1j}) = \frac{1}{\pi_{1j} \cdot R \cdot u'(w_{1j}) \cdot (1 - \rho(e_{2j}))}. \quad (21)$$

We will now show that identical effort levels for both agents are optimal and - as a corollary - a team of agents endowed with a certain technology yields higher expected returns than a single agent, using that technology:

Proposition 3. *Symmetric effort, i.e. $e_{1j} = e_{2j} = e_{ij}$, is optimal when two agents use the same technology.*

Proof. To see that this is true, we change the principal's objective function to

$$\begin{aligned} \pi_j \cdot [1 - (1 - \rho(e_{ij} + \epsilon_1)) \cdot (1 - \rho(e_{ij} - \epsilon_2))] \cdot R \\ - u^{-1}(e_{ij} + \epsilon_1) - u^{-1}(e_{ij} - \epsilon_2), \end{aligned} \quad (22)$$

where ϵ_1 and ϵ_2 are chosen such that the overall probability of finding the target remains at the optimal level. It is easy to see that this can only be done when $\epsilon_1 > \epsilon_2$. Taking into account that u^{-1} is a convex function, the sum of the wages will be higher when effort levels differ. In other words, symmetric effort levels minimize the cost for any total probability of success. Hence, different effort levels cannot be optimal. \square

Corollary 1. *A team of two agents, using both technology j , yields a strictly higher expected return than a single agent, using technology j .*

Proof. Any return that is generated by offering the optimal effort-wage combination to a single agent, can be replicated by offering the same conditions to only one of two agents while the remaining agent receives a zero-wage. However in this case, efforts would be asymmetric and therefore cannot be optimal. \square

Plugging in the result of Proposition 2 into (21), effort and wage levels can be determined by solving equation (12) and

$$\rho'(e_{ij}) = \frac{1}{\pi_j \cdot R \cdot u'(w_{ij}) \cdot (1 - \rho(e_{ij}))}. \quad (23)$$

We can easily see that in the two-agent-case a lower effort level per agent is optimal, since the solution to the equation will be smaller compared to the solution of (11).

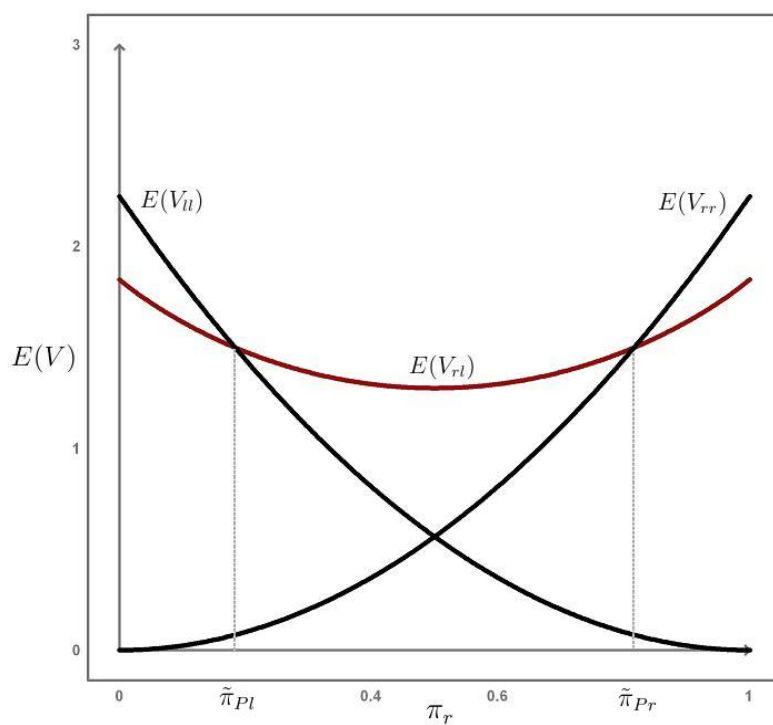


Figure 1: The principal's optimal technology choice

As we have derived optimal effort-wage-combinations for disjointed and duplicated efforts, we now turn to the question which of the two options is optimal. It is worthwhile to remember that the principal has basically three options to choose from:

- (i) Both agents are assigned to technology r (Duplicated Efforts I)
- (ii) Both agents are assigned to technology l (Duplicated Efforts II)
- (iii) Both agents are assigned to one technology each (Disjoined Efforts)

Following the reasoning of Proposition 1, assigning both agents to the outsider-technology cannot be optimal. Consequently, only the two remaining alternatives (duplicated efforts using mainstream-technology and disjointed efforts) have to be compared to determine the optimal strategy. Let e_{ij} and w_{ij} be the optimal effort and wage levels of agent i in case both agents use technology r , and let e'_{ij} and w'_{ij} denote the respective levels in case the agents split up. Assigning both agents to technology r is optimal whenever

$$\begin{aligned} E(V_{rr}(\cdot)) &= [2 \cdot \rho(e_{1r}) - \rho(e_{1r})^2] \cdot R - 2 \cdot w_{1r} \geq \\ E(V_{rl}(\cdot)) &= [\pi_r \cdot \rho(e'_{1r}) + (1 - \pi_r) \cdot \rho(e'_{2l})] \cdot R - w'_{1r} - w'_{2l}. \end{aligned} \quad (24)$$

This inequation can be rearranged with respect to π_r , to define a threshold value $\tilde{\pi}_{Pr}$ that equalizes the principal's expected payoffs of both possible allocations:

$$\tilde{\pi}_{Pr} = \frac{-\rho(e'_{2l}) \cdot R - 2 \cdot w_{1r} + w'_{2l} + w'_{1r}}{[-2 \cdot \rho(e_{1r}) + \rho(e_{1r})^2 - \rho(e'_{2l}) + \rho(e'_{1r})] \cdot R}. \quad (25)$$

Proposition 4. *For $n=2$, there exists a unique threshold value $\tilde{\pi}_{Pr}$ with $0.5 \leq \tilde{\pi}_{Pr} < 1$ that determines the optimal allocation of agents. When $\pi_r > \tilde{\pi}_r$, allocating both agents to technology r is optimal, otherwise splitting up the agents is optimal.*

Proof. We have to show that the principal's respective expected returns as a function of π_r necessarily satisfy a single-crossing condition, i.e. there

always exists one unique intersection for both payoff functions, $V_{rr}(\cdot)$ and $V_{rl}(\cdot)$. The proof consists of four parts:

- (i) Both payoff functions are strictly increasing in π_r for $\pi_r > 0.5$.
- (ii) Both payoff functions do not intersect in $0 < \pi_r \leq 0.5$, and it is optimal to allocate agents to different technologies when $\pi_r = 0.5$.
- (iii) For $\pi_r = 1$ it is optimal to allocate both agents to the same technology. Hence there must exist *at least* one intersection of the payoff functions in $0.5 < \pi_r < 1$.
- (iv) The intersection is *unique*, as for $\pi_r > \tilde{\pi}_r$ it is guaranteed that $V_{rr}(\cdot) > V_{rl}(\cdot)$.

(i): For $V_j(\cdot)$, monotonicity indirectly follows from equations (11) and (12) and our assumptions on $u(\cdot)$. First it is evident that e_{ij} and w_{ij} both increase in π_j . The expected gain of a single agent $V_j(\cdot)$ is a convex function of π_j since $u^{-1}(e_{ij}) = w_{ij}$ is also a convex function and in optimum it must be that $\Delta\rho(e_{ij}) \cdot R > -\Delta w_{ij}$. Here $\Delta\rho(\cdot)$ and Δw_{ij} indicate the adjustments of probabilities and wages, given a changed value of π_r . To see that $V_{rl}(\cdot) = V_r(\cdot) + V_l(\cdot)$ as a whole increases in π_r for $\pi_r > 0.5$, consider the case of $\pi_r = \pi_l = 0.5$ where effort and wage levels are identical for both agents. When π_r increases to $\pi_r + \epsilon$, it must be that π_l decreases to $\pi_r - \epsilon$. Since $V_r(\cdot)$ and $V_l(\cdot)$ are both convex functions, it is clear that the increase of $V_r(\cdot)$ is larger than the decrease of $V_l(\cdot)$. Hence the overall payoff function strictly increases.

Likewise the monotonicity of $V_{rr}(\cdot)$ follows from equations (23) and (12). A higher level of π_j will always result in a higher levels of efforts and wages, where in optimum the increase in expected gain has to outweigh the increase in wages, since otherwise the old wage level would be preferred by the principal. Therefore $V_{rr}(\cdot)$ is strictly increasing in π_r .

(ii): For $0 < \pi_r \leq 0.5$ it must be that $E(V_{rr}(\cdot)) < E(V_{rl}(\cdot))$. To see that this is true, one can plug in the optimal effort-wage-combination for $V_{rr}(\cdot)$ into $V_{rl}(\cdot)$. We obtain

$$\begin{aligned}
E(V_{rr}(\cdot)) &= [\pi_r \cdot (2 \cdot \rho(e) - \rho(e)^2)] \cdot R - 2 \cdot w < \\
E(V_{rl}(\cdot)) &= [\pi_r \cdot \rho(e) + (1 - \pi_r) \cdot \rho(e_{1r})] \cdot R - 2 \cdot w \\
&\Leftrightarrow \pi_r \cdot (2 \cdot \rho(e) - \rho(e)^2) > \rho(e) \\
&\Leftrightarrow \pi_r < \frac{1}{2 - \rho(e)}
\end{aligned} \tag{26}$$

(iii) For $\pi_r = 1$ it must be that $E(V_{rr}(\cdot)) > E(V_{rl}(\cdot))$. To see that this is true, we once more plug in the optimal effort-wage-combination for $V_{rr}(\cdot)$ into $V_{rl}(\cdot)$. We have

$$\begin{aligned}
E(V_{rr}(\cdot)) &= [\pi_r \cdot (2 \cdot \rho(e) - \rho(e)^2)] \cdot R - 2 \cdot w > \\
E(V_{rl}(\cdot)) &= [\pi_r \cdot \rho(e) + (1 - \pi_r) \cdot \rho(e)] \cdot R - 2 \cdot w \\
&\Leftrightarrow \pi_r \cdot (2 \cdot \rho(e) - \rho(e)^2) > \pi_r \cdot \rho(e) \\
&\Leftrightarrow \rho(e) < 1
\end{aligned} \tag{27}$$

(iv): The proof is completed by showing that for any value of $\pi_r \geq \tilde{\pi}_r$, it is guaranteed that $V_{rr}(\cdot) > V_{rl}(\cdot) = V_r(\cdot) + V_l(\cdot)$. For $\pi_r = \tilde{\pi}_r$, by definition, we have $V_{rr}(\cdot) = V_{rl}(\cdot) = V_r(\cdot) + V_l(\cdot)$. For a level of $\pi_r > \tilde{\pi}_r$ we claim that $V_{rr}(\cdot) + \Delta V_{rr} > V_r(\cdot) + \Delta V_r(\cdot) + V_l(\cdot) + \Delta V_l(\cdot)$, where $\Delta V_j(\cdot)$ denotes the increase of expected payoff, given a higher value of π_r . The claim must be true, since first $V_l(\cdot)$ decreases in π_r so that $V_l(\cdot) + \Delta V_l(\cdot) < V_l(\cdot)$ and second $V_{rr}(\cdot) + \Delta V_{rr}(\cdot) > V_r(\cdot) + \Delta V_r(\cdot)$ is guaranteed by Proposition 2 and its corresponding corollary. \square

Since $\pi_l = 1 - \pi_r$, we can define a threshold value for π_l as $\tilde{\pi}_{Pl} = \tilde{\pi}_{Pr}$. Therefore the optimal allocation of agents to technologies is tripartite. For $0 \leq \pi_r \leq 1 - \tilde{\pi}_{Pr} = \tilde{\pi}_{Pl}$, both agents use technology l , for $1 - \tilde{\pi}_{Pr} < \pi_r < \tilde{\pi}_{Pr}$ agents are separated across technologies, and for $\tilde{\pi}_{Pr} \leq \pi_r \leq 1$ both agents use technology r .

3.3.2 Asymmetric Information

Like in the symmetric information setting, we focus on the case where agents conduct research with identical technologies, because again, effort and wage levels for disjoined efforts agents can be easily deduced from section 3.2.2.

The principal's maximization problem then becomes

$$\begin{aligned}
E(V_{jj}^{SB}(\cdot)) = \max_{e_{i,j}, \bar{\bar{w}}_{ij}, \bar{w}_{ij}, \underline{w}_{ij}, \underline{\underline{w}}_{ij}} &= \pi_j \cdot [\rho(e_{1j}) \cdot \rho(e_{2j}) \cdot \\
& (R - \bar{\bar{w}}_{1j} - \bar{\bar{w}}_{2j}) + \\
& \rho(e_{1j}) \cdot (1 - \rho(e_{2j})) \cdot \\
& (R - \bar{w}_{1j} - \underline{w}_{2j}) + \\
& (1 - \rho(e_{1j})) \cdot \rho(e_{2j}) \cdot \quad (\text{P II 2:AI}) \\
& (R - \underline{w}_{1j} - \bar{w}_{2j}) + \\
& (1 - \rho(e_{1j})) \cdot (1 - \rho(e_{2j})) \cdot \\
& (-\underline{\underline{w}}_{1j} - \underline{\underline{w}}_{2j})] + \\
& (1 - \pi_j) \cdot (-\underline{\underline{w}}_{1j} - \underline{\underline{w}}_{2j}).
\end{aligned}$$

Note that we have a slight abuse of notation here, because $\bar{\bar{w}}_{ij}$ ($\underline{\underline{w}}_{ij}$) now denotes the wage level that agent i obtains when he is the *only* agent that has success (fails) and the other agent fails (succeeds). Accordingly \bar{w}_{ij} and \underline{w}_{ij} denote the wage levels in case of a double success and a double failure respectively. The problem is subject to each agent i 's individual rationality constraint

$$\begin{aligned}
& \pi_j \cdot [\rho(e_{ij}) \cdot \rho(e_{ij}) \cdot u(\bar{\bar{w}}_{ij}) \\
& + \rho(e_{ij}) \cdot (1 - \rho(e_{ij})) \cdot u(\bar{w}_{ij}) \\
& + (1 - \rho(e_{ij})) \cdot \rho(e_{ij}) \cdot u(\underline{w}_{ij}) + \quad (\text{IR II 2:AI}) \\
& (1 - \rho(e_{ij})) \cdot (1 - \rho(e_{ij})) \cdot u(\underline{\underline{w}}_{ij})] + \\
& (1 - \pi_j) \cdot u(\underline{\underline{w}}_{ij}) - e_{ij} \geq 0
\end{aligned}$$

where we set $e_{1j} = e_{2j} = e_{ij}$ according to Proposition 3.

The incentive constraint becomes

$$\begin{aligned}
e_{ij} \in \operatorname{argmax}_{\hat{e}_{ij}} \pi_j \cdot [\rho(\hat{e}_{ij}) \cdot \rho(\hat{e}_{ij}) \cdot u(\bar{w}_{ij}) \\
+ \rho(\hat{e}_{ij}) \cdot (1 - \rho(\hat{e}_{ij})) \cdot u(\bar{w}_{ij}) \\
+ (1 - \rho(\hat{e}_{ij})) \cdot \rho(\hat{e}_{ij}) \cdot u(\underline{w}_{ij}) + \\
(1 - \rho(\hat{e}_{ij})) \cdot (1 - \rho(\hat{e}_{ij})) \cdot u(\underline{\underline{w}}_{ij})] + \\
(1 - \pi_j) \cdot u(\underline{\underline{w}}_{ij}) - \hat{e}_{ij}.
\end{aligned} \tag{IC II:AI}$$

Replacing the original incentive constraint with the first order condition, and solving the Lagrangian, we yield

$$\frac{1}{u'(\bar{w}_{ij})} = \frac{1}{u'(\bar{\bar{w}}_{ij})} = \lambda_i + \mu_i \cdot \frac{\rho'(e_{ij})}{\rho(e_{ij})} \tag{28}$$

and

$$\frac{1}{u'(\underline{w}_{ij})} = \frac{1}{u'(\underline{\underline{w}}_{ij})} = \lambda_i - \mu_i \cdot \frac{\rho'(e_{ij}) \cdot \pi_{ij}}{1 - \rho(e_{ij}) \cdot \pi_{ij}}. \tag{29}$$

Optimal wage levels depend on individual performance, but do not depend on the other agent's performance. As usual, a success gets rewarded and a failure gets punished. Due to the agents' risk aversion, any effort level becomes costlier for the principal when effort is unobservable, such that the principal's payoff in the second-best solution is lower compared to the first-best solution.

Analogous to equation (24) in the symmetric setting, we can determine a threshold value $\tilde{\pi}_{Pr}^{SB}$ with $0.5 < \tilde{\pi}_{Pr}^{SB} < 1$ for the present asymmetric setting. Whether $\tilde{\pi}_{Pr}^{SB}$ is smaller or lower than $\tilde{\pi}_{Pr}$ depends on the particular configuration of $\rho(\cdot)$, $u(\cdot)$ and R and both cases can occur in principle. However this question is not scrutinized here in detail, as it is of no further relevance for the analysis.

The interesting question is, whether the agent can be incentivized to select

the principal's preferred technology, although the principal cannot observe technology choice. Somewhat surprisingly in the light of the informativeness principle, it turns out that this is possible.

Proposition 5. *With $n=2$ and unobservable technology choice, there is a forcing contract that allows for obtaining the original second-best outcome.*

Proof. When duplicated efforts are optimal, agents will choose the principal's favored technology by pure self-interest according to Proposition 2 and $\tilde{\pi}_{Pj} = \tilde{\pi}_{ij}$. When disjointed efforts are optimal, and agent 1 is assigned to technology j whereas agent 2 is assigned to technology k and $\pi_k > \pi_j$, agent 1 prefers to choose technology k and $\tilde{\pi}_{Pk} \neq \tilde{\pi}_{ik}$. The principal can deter this behavior however by fining agent 1 with a penalty \overline{w}_1^F , whenever *both* agents yield a positive output. Since agent 2 will always have a high output with some positive probability, such a fine that deters agent 1 from choosing technology k must always exist. \square

We can state that the possibility of observing individual output serves as a *substitute* for observing the technology. In other words, technology choice is redundant information, given the incentives of the optimal contract and does not need to be observed. Evidently the fine is only needed in the mentioned case when it is optimal to separate agents across technologies. Note that for Proposition 4 to be true, it is important that side-contracting between agents must not be possible, because otherwise agent 1 would exert effort with technology k , whereas agent 2 would not work at all and both agents would have to argue how to share agent 1's expected surplus.

Although we have seen that optimal allocation of agents is possible in principle, an important result directly follows from the reasoning above:

Corollary 2. *With free technology choice and $n=2$, there is an efficiency loss whenever disjointed efforts are optimal, since the agent that is assigned to the outsider-technology chooses to employ the mainstream-technology.*

The reasoning directly follows from the proof of Proposition 4, given that no punishment \overline{w}_1^F is applied. The result has quite severe implications. It suggests that freedom of research might not be in the best interest of scientific progress in total. It therefore challenges the conventional wisdom that mostly treats freedom of research as a guarantor for academic success. We will elaborate on this *bias towards mainstream-research* in greater detail in the following discussion section.

4 Discussion

The results of the previous section give rise to a great number of questions concerning the validity of our model. From our perspective, the following issues are most controversial and therefore seem worthwhile to be discussed:

1. The limited state space and the assumption of mutually exclusive technologies
2. Binary output levels
3. The observability of individual output and free-riding
4. Homogeneity of agents and identical cost functions for both technologies
5. Economies of scale of coworking researchers

1) As already mentioned, our current model analyzes the rather special case of mutually exclusive technologies. When one technology is good, the remaining one is necessarily bad. The existence of a contract that deters agents from choosing a technology that is not wanted by the principal clearly hinges on this critical assumption. This is so, because in the case of independent technologies, both agent could have a success, although they have chosen

different technologies. Then the occurrence of two successful agents cannot tell the principal that one of the agents has cheated. It is therefore natural to look for more general mechanisms that provide optimal incentives for optimal technology choice. These mechanisms are matter of our future research and revised versions of this article. Note however that the existence of a bias towards mainstream-research is not affected by the restrictive technology assumption. For any two technologies (whether independent or not), it is always rational to choose the more promising one from any agent's perspective.

2) It is clearly a simplification to assume that only binary output levels are possible, whereas in reality a researcher's output could be described as a continuum. Nevertheless we think that this simplification is justified, since a continuum of outcomes would complicate the model and most likely would not provide any major insights.

3) Whenever several agents work as a team the problem of free-riding arises. However in our analysis we ignored this issue for two reasons: First, individual observable output does not seem to be an unrealistic assumption in the academic sector. Whenever an agent has produced a success, this success usually remains private knowledge until the knowledge is shared with the scientific community. Therefore a certain success is normally attributable to single individuals. Second, there is a vast body of literature on free-riding in teams and there exist sophisticated mechanisms to deter free-riding among agents (e.g. Legros and Matsushima (1991)). However the issue is beyond the scope of this paper which is why we neglect it.

4) A further critical assumption concerns the existence of completely homogeneous agents and identical cost functions for both technologies. One could also question the assumption that only one research technology at a time is possible. While the latter assumption seems rather uncontroversial, the former one is certainly not valid, since scientist are often experts in

methods rather than in solving a special problem. However then the question arises whether there is enough specialization in different methods, or if only those methods are learned that maximize individual prospects for success.

5) Due to our functional form of modelling individual success, the marginal return of a second agent that uses the same technology must be decreasing. One could therefore question whether it is not more realistic that overall output more than doubles, when the number of researchers that use a certain technology is multiplied by two, at least for low numbers of researchers. In any case benefits from coworking seem more plausible to occur when researchers use identical technologies, since completely different approaches are often incommensurable.

5 Conclusion

We have derived optimal contracts for researching agents whose action space encompasses an effort level and the choice of a research technology. Both players - principal and agent - follow easy to understand and intuitively plausible technology decisions which can collide when technology choice is hidden to the principal. While the agent strives for personal success, the principal wishes to maximize overall scientific success. This leads to a mismatch of interests of principal and agent, whenever there is more than one agent and separating agents across technologies is optimal. The conflicting interest can be deterred by an optimal contract without any further cost for the principal. However the existence of such a costless contract hinges on the extreme assumption of mutually exclusive technologies, and is unlikely to occur for independent technologies. Given academic freedom, i.e. free choice of research technology, there is a clear bias towards mainstream-research: No single researcher wants to sacrifice his own career for the benefit of science as a whole. Thus, a social planner's optimal research agenda is not implemented and overall scientific progress remains at a suboptimal level.

References

- Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica* 30, 424–444.
- Dewatripont, M., I. Jewitt, and J. Tirole (2000). Multitask agency problems: Focus and task clustering. *European Economic Review* 44(4-6), 869–877.
- Frey, B. (2003). Publishing as prostitution - choosing between one's one ideas and academic success. *Public Choice* 116, 205–223.
- Grey, C. (2010). Organizing studies:: Publications, politics and polemic. *Organization Studies* 31, 677–694.
- Grossman, S. and O. Hart (1983). An analysis of the principal-agent problem. *Econometrica* 51(1), 7–45.
- Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics* 10(2), 324–340.
- Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics* 13(1), 74–91.
- Hurwicz, L. (1972). The design of mechanisms for resource allocation. *American Economic Review* 63, 1–30.
- Ioannidis, J. (2005). Why most published research findings are false. *PLoS Med* 2(8), e124.
- Itoh, H. (1991). Incentives to help in multi-agent situations. *Econometrica* 59(3), 611–636.
- Legros, P. and H. Matsushima (1991). Efficiency in partnerships. *Journal of Economic Theory* 55, 296–322.
- Mookherjee, D. (1984). Optimal incentive schemes with many agents. *Review of Economic Studies* 51(3), 433–446.

- Mookherjee, D. (2006). Decentralization, hierarchies, and incentives: A mechanism design perspective. *Journal of Economic Literature* 44(2), 367–390.
- Radner, R. (1992). Hierarchy: The economics of managing. *Journal of Economic Literature* 30, 1382–1415.
- Spence, M. (1975). The economics of internal organization: An introduction. *Bell Journal of Economics Spring 1975*, 163–172.
- Starbuck, W. (2005). How much better are the most prestigious journals? The statistics of academic publication. *Organization Science* 16, 180–200.
- Weber, M. (1946). *From Max Weber: Essays in sociology - Science as a Vocation*, pp. 129–156. New York: Oxford University Press.