University of Heidelberg

Department of Economics

Discussion Paper Series | No. 421

Division of Labour and Directed Production

Marisa Ratto and
Wendelin Schnedler

July 2005

Division of Labour and Directed Production

Marisa Ratto* Wendelin Schnedler**

July 8, 2005

Abstract

We examine a situation where efforts on different tasks positively affect production but are not separately verifiable and where the manager (principal) and the worker (agent) have different ideas about how production should be carried out: agents prefer a less efficient way of production. We show that by dividing labour (assigning tasks to different agents and verifying that agents do not carry out tasks to which they are not assigned), it is possible for the principal to implement the efficient way of production. Colluding agents can undermine this implementation. However, if agents have different abilities, collusion can be prevented by a specific assignment of agents to tasks.

Keywords: hidden action, moral hazard, specialisation, job design

JEL-Codes: L23, M52, D82, J24

---

*Centre for Market and Public Organisations (CMPO), University of Bristol. m.l.ratto@bristol.ac.uk

**Department of Economics, University of Heidelberg.
wendelin.schnedler@awi.uni-heidelberg.de

The authors wish to thank Guido Friebel, Paul Grout, Joop Hartog, Jörg Oechssler, Ulf Schiller, and Jan Zabojnik for their valuable comments. Further, this paper profited from the contributions of participants at the Public Economic Theory conference in 2003 at Duke University, at the Royal Economic Society annual meeting 2004, at the Kolloquium für Personalökonomie in Bonn 2004, at the Fifth Symposium of the German Economic Associatiation for Business Administration in Bonn 2004, and at seminars in Bristol and Erfurt.

Such schemes as these are nothing without numbers.
One cannot have too large a party.

Jane Austen, Emma Vol. III, Chap. VI

1 Introduction

In many economic situations, little is known about how an outcome, goal or production result has been achieved. For example, the success of a marketing operation may be increased by more creativity or a better-identified target audience. Exam results can be improved by augmenting the quality of teaching or by training test situations with the students. A smooth table surface can be achieved by meticulously levelling the wood or by adding an extra layer of paint. The likelihood of placing an unemployed into a job is affected by training her in a demanded profession or by brushing up her interview skills. Software may run well due to particularly careful programming or due to extensive debugging. Although, it is hard or even impossible to deduce how the observed result has been produced, it may nevertheless be important. Some ways of producing might be more efficient than others. A particular mix of targeting and creativity may bind customers longer. Students should not just learn how to cope with test situations but also receive quality teaching. The table may be more durable if the right amount of levelling and painting are applied. Too much focus on interview skills does not improve the match between labour demand and labour supply. Finally, it may be easier to maintain software when more emphasis has been put on careful programming. If the person who carries out the work (agent, worker, he) does not prefer to produce the result efficiently, there is a need for the person who organises production (principal, manager, she) to influence the activity and direct production to improve efficiency. But how can production be directed if it is not possible to verify what exactly the agent did?

The question how to organise production has a long tradition which starts with Adam Smith (1776) and encompasses Taylor's idea (1911) of "scientific management". Taylor was interested to identify the "one best method" to achieve a particular production goal. Here, we are concerned with the question how "one best method" can be implemented in a sparse contractual environment. We analyse a multiple-task setting, where tasks contribute to a production result, which is the only verifiable measure of effort. There is no additional information about the effort of individual agents or the effort exerted at specific tasks which can be contracted upon. The tasks are not conflicting with respect to the production result: effort on any task is (weakly) beneficial and (weakly) increases the production result. The principal cares about how the production result has been achieved, but this cannot be deduced from the result itself. If multiple agents are employed, she can assign them to tasks and ensure that they do not carry out tasks to which they are not assigned. In other words, it is possible to divide labour.

We show that the division of labour enables the principal to direct production: by assigning one task to each agent, the principal can influence how much effort is exerted at each task. The intuition is the following. Without division of labour, the agent has many ways to affect the production result. If labour is divided, an agent can only influence the production result by changing the effort on his task. The performance pay to this agent thus allows to directly manipulate his gains from exerting effort. By adjusting these gains, the principal can obtain any desired effort level. As this is possible for every task and agent, any effort allocation can be obtained.

This result provides a novel explanation for the division of labour. Gains of separating tasks between workers are usually attributed to economies of scale or comparative advantages or combinations of both. Our explanation complements these rationales: separating tasks may create gains even if it neither enables agents to master individual tasks in a better way nor allows to exploit differences in skills.

The key element for implementing the efficient production method is that agents are prevented from doing the tasks of other agents. Otherwise, they would obtain production results by exerting effort on their preferred task. But even if agents cannot do these tasks themselves, they can potentially undermine the directions of the principal by colluding: they may talk to their colleague on the respective task and agree to produce the result with the lowest joint costs rather than in the most efficient way. We examine the possibility of collusion and find that the principal can avoid collusion by assigning workers to tasks according to ability. The required assignment is not necessarily in line with the idea of a comparative advantage. Thus, the prediction from our model differs from that of classical ideas of specialisation.

The result offers a new perspective on the problem of providing incentives when there are multiple task. This problem has first been formalised in a seminal article by Holmström and Milgrom (1991). Holmström and Milgrom assume that effort at different tasks is measured with different accuracy, while we consider situations where task-wise information is not revealed but only the overall production result is contractible. Since tasks are measured with different accuracy in Holmström and Milgrom's model, there is a conflict between incentives to optimally allocate effort and insurance considerations. Under our assumption, the incentive problem arises from a different source: the production result does not reflect how effort has been allocated across tasks. Hence, it cannot be employed to obtain the desired allocation. Note, that this only hinges on the fact that the measure confounds the effort on several tasks - so the problem remains even if the production result is measured without error. Confounding performance measures have received considerable attention in recent contributions to the multi-tasking literature (Feltham and Xie 1994, Baker 2002, Schnedler 2003 and 2004). While these authors deal with the effect of the composition of performance measures on how a job is carried out, they do not consider other organisational responses to influence the agent's behaviour. In particular, they neglect the possibility of dividing tasks between agents. Like here, Dewatripont and Tirole (1999) examine the effect of dividing labour in a situation where the only verifiable outcome depends on the effort exerted at two tasks. Differently, they assume that the effort on one task increases this outcome while the other effort reduces it. This assumption makes sense if the outcome is a decision and effort is exerted to gather information for and against this decision (advocacy setting). It is less suitable for most production settings, where beneficial effort affects the outcome in the same direction, irrespective on which task it is exerted. By assuming such non-conflicting efforts, we depart from Dewatripont and Tirole (1999) at a crucial point. This departure has several consequences. First, incentives in Dewatripont and Tirole's article can be produced by introducing competition between agents. These become adversary advocates. On the contrary, when tasks are non-conflicting, agents become non-competing co-workers. Second, because the advocates compete, they also have a motive to sabotage the work of the other – a negative side-effect which is absent from our setting. Finally, collusion can never be prevented with conflicting efforts: advocates jointly gain by ending their competition and slacking on effort. In our setting collusion can sometimes be avoided. In order to enable an agent to slack, the other agent will have to work for him. If this agent has high effort costs, they cannot improve their joint situation by undermining the incentive scheme.

When the production result is the only variable which reflects effort on various tasks, dividing labour implies that a group of agents is paid according to a joint performance measure. In principle, such arrangements are vulnerable to free-riding. Holmström (1982) has shown that the free-rider problem does not occur if the principal acts as a budget breaker (as in our setting). Holmström as well as Alchian and Demsetz (1972) study the provision of incentives taking the existence of the team as given. Alchian and Demsetz justify the existence of the team by interactions in the production function. Itoh (1991) also motivates why multiple agents work on one task. In his article, effort costs are convex at each task and additively separable between tasks; hence, agents prefer splitting their effort between two tasks rather than concentrating it. Here, there are no reasons neither on the production nor on the cost side which favour multiple agents. Still, we find that having multiple agents improves efficiency. We thus provide a justification for the existence of teams defined solely by reward systems. Such teams do indeed exist. An example are the team-based bonus schemes piloted recently in three government agencies in the United Kingdom (Burgess and Ratto 2003).

While we examine the effect of division of labour in a multiple task setting with a confounding measure, such measures also appear in other contexts. One example are career concern models (Fama 1980, Holmström 1999, Gibbons and Murphy 1992). Dewatripont, Jewitt, and Tirole (1999) consider organisational responses to separate talent from effort in these models. Another example is the simultaneous decision on a project and effort. Athey and Roberts (2001) suggest that a hierarchical structure improves incentives in this setting. Intuitively, several agents may be more productive than fewer because some of them are hired as supervisors and control the activity of the others. Strausz (1997) has shown elegantly that an intermediary between principal and agent can improve incentives in a single task setting, even if this intermediary has to base judgements on the same information as the principal. In our model, incentives are also improved but not because an agent is monitoring or controlling his colleagues or acting as an intermediary.

The remainder of the paper is organised as follows. The next section sets up the model. Then, in Section 3, we derive the optimal contract with a single worker and when labour is divided. In Section 4, we discuss when collusion may occur and how it can be prevented by a specific task assignment. Finally, Section 5 concludes.

2 The model

We illustrate the idea that division of labour allows to direct production in a very simple framework. Consider a principal, who draws a benefit from two tasks one of which is easy (task E) and the other demanding (task D). Lacking time or skills to do the tasks herself, the principal employs one or two agents to do the job.

Each task involves an effort choice by agent i who carries out the task: e_i^D and e_i^E.[1] Later, we want to discuss deviations of the agents from some optimal inner solution. For agents to be able to deviate in both directions, there need to be at least three possible choices: $e_i^k \in \{0,1,2\}, k = D, E$. The costs of effort exerted by agent i at task k rise with effort $c_i^k(0) < c_i^k(1) < c_i^k(2)$ and providing effort gets increasingly difficult ("convexity"):

$$c_i^k(2) - c_i^k(1) > c_i^k(1) - c_i^k(0). \tag{1}$$

Convexity of costs will be crucial for the implementation of the optimal contract, as we shall consider below. We assume that the costs of effort at the two tasks are independent. This assumption ensures that there are no cost advantages or disadvantages of employing more or less agents.[2] Consequently, any division of labour is driven entirely by incentive effects and not by economies of scale or scope.

We want to represent that task D is more demanding, by assuming the following. When the effort level is low at the difficult task, an increase in effort is more costly than an increase in effort on the easy task starting from a medium level:

$$c_i^D(1) - c_i^D(0) > c_i^E(2) - c_i^E(1). \tag{2}$$

This assumption introduces an asymmetry in the agents' preferences about how tasks are carried out relative to the principal's interest. This asymmetry will play a crucial role to explain why efficient production cannot be achieved with a single agent. From the convexity assumption (1) and the task asymmetry assumption (2), it follows immediately that increasing effort is more costly on the demanding task whatever the initial level:

$$c_i^D(1) - c_i^D(0) > c_i^E(1) - c_i^E(0) \text{ and } c_i^D(2) - c_i^D(1) > c_i^E(2) - c_i^E(1). \tag{3}$$

[1]More generally, the agent may not exert effort but have a choice which affects himself as well as the principal. However, ordering these choices according to costs and speaking of effort greatly simplifies the exposition.

[2]We do not claim that those advantages do not occur in reality but we want to abstract from them to isolate incentive effects.

In other words, the agent has a preference for increasing effort on the easy rather than on the difficult task.

The principal is not indifferent about how effort is allocated between tasks, neither. For simplicity, we assume that the principal is satisfied as long as medium effort is chosen at each task. Thus, the principal's benefit $B(e_i^D, e_i^E)$ is defined as:

$$B(e_i^D, e_i^E) := \begin{cases} b & \text{if } e_i^E \geq 1 \text{ and } e_i^D \geq 1, \\ 0 & \text{else.} \end{cases} \tag{4}$$

Hence, there is a conflict of interest between principal and agent. The principal wants at least medium effort on both tasks while the agent prefers to concentrate his effort on the easier task.

The problem is only interesting, if the interaction between principal and agents can potentially be beneficial. Thus, we assume that the benefit exceeds the effort costs for a medium level of effort for some agent i:

$$b > c_i^D(1) + c_i^E(1). \tag{5}$$

It follows from this and the other model assumptions that it is efficient to implement a medium level of effort on both tasks. Choosing a higher effort level on any of the tasks only leads to higher costs while a lower level eliminates the benefit. Note that the agent on his own would choose an inefficient way of production. So, there is scope to increase efficiency. Hence, an intervention to direct the agent's effort may make sense.

If the input of the agent on each task could be verified by a court, the principal and the agent would write a contract to stipulate the efficient medium effort levels. This contract would then work as a device to direct the agent's effort. However, we assume that contracts about task-wise efforts cannot be written. The only information about the agent's effort which may be used in court is the production result. Anything that reveals how the final product has been generated (for example intermediate stages of production) cannot be verified. The principal may believe or even know that production has not taken place in the desired way but she cannot credibly relay this information to other parties. She may also be aware of the benefit she has received (or not) but again she is unable to "prove" this subjective quantity. The behaviour of the agent may eventually even become apparent but it may be too late because the agent cannot be held responsible anymore. So, neither the desired actions nor the consequences for the principal can be stipulated and the contractual environment we consider is that of a ”coping organisation” according to Wilson (1989). Summarising these considerations, we assume that the only verifiable quantity is a measure that is increasing in both efforts but does not allow to identify how the worker allocates effort between the two tasks. We assume that the production result “looks good” ($M = 1$) if the sum of the two efforts is above a threshold (here this threshold is two):

$$M(e_i^D, e_i^E) := \begin{cases} 0 & \text{if } e_i^E + e_i^D < 2, \\ 1 & \text{if } e_i^E + e_i^D \geq 2. \end{cases} \tag{6}$$

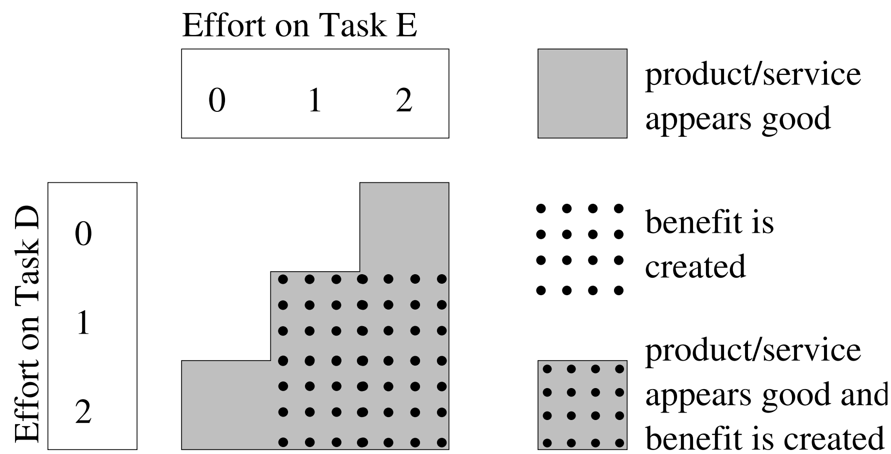
The effort allocation across tasks cannot be identified even if the observable production result M is not noisy. Of course, the allocation would still remain unidentified if the production result were noisy. Also, the fact that the production result takes two values is not crucial either. Even if the production result were a continuous function of continuous effort, it would still not be possible to identify task-wise efforts. The distinguishing and important feature of the verifiable production result used here is that it increases in both efforts and (by nature) is one-dimensional, so that effort on a specific task cannot be deduced. Because the verifiable result confounds effort on both tasks, a court of law cannot tell whether the agent allocated effort across tasks in the specific way desired by the principal. In particular, the agent can mask low effort on one task by higher effort on the other task. This is harmful to a principal who prefers a specific effort allocation. Figure 1 illustrates the situation under our simple assumptions. Essentially, the result looks good whenever the principal receives a benefit. However, there are two important exceptions: exerting a lot of effort on either task masks a low effort on the other task.

The verifiable production result gives the principal an instrument to provide incentives. She can set two transfers depending on the state: a transfer $\bar{t}_i$ which is paid whenever the production result is high ($M = 1$) and $\underline{t}_i$ which is handed to the agent i given the result is low ($M = 0$).

3 Optimal contract

We will now examine the optimal contract in the described model framework. First, we will restrict the principal to only employ one agent. Then, we will relax

Figure 1: Verifiable Production Result and Created Benefit



The same verifiable result can be obtained in beneficial or damaging ways.

this assumption.

To find the optimal contract, we assume that the principal is the mechanism designer and that the rationality of the agent is represented in the form of an incentive constraint. A participation constraint ensures that the principal does actually create surplus and is not simply exploiting the agents. For simplicity, the outside option of all agents is standardised to zero.

3.1 One-agent case

If the principal can only employ one agent, she will maximise the joint surplus subject to the following incentive and participation constraint:

$$\begin{aligned} &\max_{\bar{t}_i,\underline{t}_i} b(e_i^D,e_i^E) - M(e_i^D,e_i^E)(\bar{t}_i-\underline{t}_i)-\underline{t}_i \\ \text{such that } &(e_i^D,e_i^E) \in \operatorname{argmax}_{(\tilde{e}^D,\tilde{e}^E)}(\bar{t}_i-\underline{t}_i)\cdot M(\tilde{e}^D,\tilde{e}^E)+\underline{t}_i-c_i^D(\tilde{e}^D)-c_i^E(\tilde{e}^E) \\ \text{and } &(\bar{t}_i-\underline{t}_i)\cdot M(e_i^D,e_i^E)+\underline{t}_i-c_i^D(e_i^D)-c_i^E(e_i^E)\geq 0. \end{aligned} \tag{7}$$

Because the high production result can be observed when the agent chooses the desired middle level on both tasks, a first intuition might be to reward the agent for the high result. However, if the agent is rewarded for a good looking product, he will produce it in the cheapest and not necessarily in the best way. Indeed, production will be inadequate at the demanding task and the agent will mask this inadequacy with the easy task. Formally, the following result holds:

Proposition 1 (Window-dressing by a single agent)**.** *If there is a single agent i and transfers are set to enable efficient production* ($e_i^D = 1, e_i^E = 1$)*, the agent will engage in cheap window dressing* ($e_i^D = 0, e_i^E = 2$)*.*

Proof. To implement efficient production, it is necessary to meet the incentive constraint and pay the agent sufficiently much so that he exerts the medium rather than the low effort level: $\bar{t}_i - \underline{t}_i \geq c_i^D(1) - c_i^D(0) + c_i^E(1) - c_i^E(0)$. The agent has the choice whether to generate a high or a low result. The high result can be generated in six ways. Three of which, $(e_i^D = 1, e_i^E = 2)(e_i^D = 2, e_i^E = 1), (e_i^D = 2, e_i^E = 2)$, are strictly dominated by $(e_i^D = 1, e_i^E = 1)$. It is not profitable to reduce effort on the easy task by increasing effort on the demanding task (expensive window dressing):

$$c_i^D(2) - c_i^D(1) \overset{(3)}{>} c_i^E(2) - c_i^E(1) \overset{(1)}{>} c_i^E(1) - c_i^E(0),$$

which implies $c_i^D(2) + c_i^E(0) > c_i^D(1) + c_i^E(1)$. On the other hand, reducing effort on the demanding task and exerting more effort on the easy task is profitable: $c_i^E(2) + c_i^D(0) < c_i^E(1) + c_i^D(1)$ by equation (2). Thus, the cheapest way to produce the high result is cheap window dressing $(e_i^D = 0, e_i^E = 2)$, and yields: $\bar{t}_i - c_i^E(2) - c_i^D(0)$. For a low result, exerting low effort, $(e_i^D = 0, e_i^E = 0)$, strictly dominates the other options, $(e_i^D = 0, e_i^E = 1)$ and $(e_i^D = 1, e_i^E = 0)$. It costs $c_i^E(0) + c_i^D(0)$ and leaves the agent with $\underline{t}_i - c_i^E(0) + c_i^D(0)$. Because transfers are set to enable efficient production $\bar{t}_i - \underline{t}_i \geq c_i^D(1) - c_i^D(0) + c_i^E(1) - c_i^E(0)$ and hence $\bar{t}_i - \underline{t}_i > c_i^E(2) + c_i^D(0) - c_i^E(0) - c_i^D(0) = c_i^E(2)$, so that the agent maximises his surplus by generating the high result through cheap window dressing. □

The main message of this proposition is that any attempt to achieve the efficient production is doomed because incentives lead the agent to prefer a different allocation of effort across tasks. More specifically, he rather slacks on the difficult task and brushes up the appearance on the easy task. The crucial feature of the contractible outcome, which drives this result, is that it is not a perfect representation of the principal's gains from production and it confounds the efforts of two (non-conflicting) tasks. The production result looks better irrespective of where effort is concentrated, and this opens up the opportunity for the agent to windowdress.[3] Note that it is the problem of identifying effort on tasks at the heart of the matter here. This problem can also exist if the verifiable production result is noisy, and effort is continuous. As long as it matters on which task the agent exerts how much effort and as long as this cannot be identified by looking at the production result, window-dressing is possible.

Principal and agent foresee that the agent engages in window dressing rather than delivering the desired product. As providing incentives is costly but produces no benefit, the principal gives up on incentives.

Proposition 2 (No incentive provision)**.** *In the single agent case, it is optimal to choose the transfers such that no incentives to exert effort are provided:* $\bar{t}_i - \underline{t}_i < c_i^E(2) - c_i^E(0)$.

Proof. The agent will produce the high result by window dressing as long as $\bar{t}_i - \underline{t}_i \geq c_i^E(2) - c_i^E(0)$. If $\bar{t}_i - \underline{t}_i < c_i^E(2) - c_i^E(0)$, the agent will not produce the high result but the low result. The respective costs, benefit, and losses are zero, while window dressing leads to a loss. So, it is optimal not to provide incentives. □

The ability of the agent to cover up bad performance on the demanding task by a good performance on the easy task undermines any well-meant incentive scheme. Of course, the principal could also stop window-dressing by not allowing the agent to perform the easy task. But then, she will also fail to achieve efficient production which requires some input on each task. A possibility for the principal to get round this problem is to assign the easy task to another agent. This avenue will be explored in what follows.

3.2 Two-agent case

When there are multiple agents, the principal can assign specific tasks to each agent. A central assumption of our model is that the principal cannot enforce effort choices using contracts. However, the principal may well be able to prevent agents from working on tasks to which they were not assigned. This can be done by withholding material or instruments needed for this task, not granting access to a locality at which the task has to be performed, or not training or preparing

[3] A more in-depth discussion of the relationship between contractible performance measure and principal's gains and its consequences on the allocation of effort between tasks can be found in Schnedler 2004.

the worker to carry out the task. Also, it may be possible for the principal to provide evidence that a worker was doing something he was not supposed to do, even though she cannot prove that an agent is not doing what he should do. Whatever, the reason may be, we assume for the following that the principal can prevent agents from exerting effort on tasks which are not assigned to them. If the principal has this "power", she can ensure that one agent is working only on the easy task and the other one is only working on the difficult task.

If different agents are working at the two tasks, this induces a game between agent i and agent j and the maximisation programme is:

$$\begin{aligned} &\max_{\bar{t}_i, \underline{t}_i, \bar{t}_j, \underline{t}_j} b(e_i^D, e_j^E) - (\bar{t}_i + \bar{t}_j - \underline{t}_i - \underline{t}_j) M(e_i^D, e_j^E) - \underline{t}_i - \underline{t}_j \\ &\text{such that } (e_i^D, e_j^E) \text{ is a Nash equilibrium of the game depicted in Table 1} \\ &\quad \text{and } (\bar{t}_i - \underline{t}_i) \cdot M(e_i^D, e_j^E) + \underline{t}_i - c^D(e_i^D) \geq 0 \\ &\text{as well as } (\bar{t}_j - \underline{t}_j) \cdot M(e_i^D, e_j^E) + \underline{t}_j - c^E(e_j^E) \geq 0. \end{aligned} \tag{8}$$

In case that the same agent works at both tasks ($i = j$), the game degenerates to the decision problem described in the incentive constraint in the programme (7).

Table 1: Payoff matrix of agent i (row) and j (column)

		Agent j (working on easy task)		
		0	1	2
Agent i (demanding task)	0	$(\underline{t}_i - c_i^D(0), \underline{t}_j - c_j^E(0))$	$(\underline{t}_i - c_i^D(0), \underline{t}_j - c_j^E(1))$	$(\bar{t}_i - c_i^D(0), \bar{t}_j - c_j^E(2))$
	1	$(\underline{t}_i - c_i^D(1), \underline{t}_j - c_j^E(0))$	$(\bar{t}_i - c_i^D(1), \bar{t}_j - c_j^E(1))$	$(\bar{t}_i - c_i^D(1), \bar{t}_j - c_j^E(2))$
	2	$(\bar{t}_i - c_i^D(2), \bar{t}_j - c_j^E(0))$	$(\bar{t}_i - c_i^D(2), \bar{t}_j - c_j^E(1))$	$(\bar{t}_i - c_i^D(2), \bar{t}_j - c_j^E(2))$

The first side condition indicates that action choices are determined interactively between agents. This disciplines the agent on the demanding task in a way which is not possible in the single agent case. Moreover, the principal has substantial influence on the structure of the game. In particular, she can disentangle incentives for the two tasks by setting the transfers to agent i, $\underline{t}_i$ and $\bar{t}_i$, different from those to agent j, $\underline{t}_j$ and $\bar{t}_j$. Exploiting this opportunity, it is possible to implement efficient production.

Proposition 3 (Implementation efficient production)**.** *If agents specialise ($i \neq j$), transfers can be chosen such that the efficient production ($e_i^D = 1, e_j^E = 1$) can be implemented as a Nash-equilibrium. The respective transfers need to fulfil the conditions:*

$$\bar{t}_i - \underline{t}_i \geq c^D(1) - c_i^D(0) \text{ and } \bar{t}_j - \underline{t}_j \geq c_j^E(1) - c_j^E(0). \tag{9}$$

Proof. Set transfers such that $\bar{t}_i - c^D(1) \geq \underline{t}_i - c_i^D(0)$ and $\bar{t}_j - \underline{t}_j \geq c_j^E(1) - c_j^E(0)$. Assume division of labour ($i \neq j$). Next, verify that under these transfers and with division of labour, ($e_i^D = 1, e_j^E = 1$) is a Nash-equilibrium. If agent i moves to a higher level, it just increases costs, if he moves to a lower level a low result is produced, and by $\bar{t}_i - c^D(1) \geq \underline{t}_i - c_i^D(0)$, agent i gets a lower utility. Thus, i has no incentive to deviate. For agent j, moving to a higher level only increases costs and moving to a lower level leads to a lower utility because of: $\bar{t}_j - \underline{t}_j \geq c_j^E(1) - c_j^E(0)$. Thus, ($e_i^D = 1, e_j^E = 1$) is a Nash-equilibrium. □

The conditions on transfers in this proposition mean that the bonus for a good looking product needs to be sufficiently large to entice agents to carry out the more costly medium effort. By setting separate transfers the principal gains an extra degree of freedom in designing an optimal performance contract. In particular she can set the gains from effort separately for each task and obtain the desired effort at each of the two tasks. As there is no other motive to separate tasks, the benefit of division of labour is entirely due to improved incentives.

The proposition also holds for continuous efforts. Then, as before, the division of labour enables the principal to set incentives separately for tasks and she can precisely determine the effort level at each task – now by altering the marginal gains. If the production result is noisy and agents are risk averse, it will not be possible to obtain efficient production anymore. By dividing labour, however, the principal can still influence the allocation of effort across tasks and improve efficiency.

When labour is divided, an agent's decision on how much effort to exert depends on what the other agent chooses. This interaction creates a positive externality between the two agents: if an agent exerts high effort instead of medium effort, he allows the other agent to slack on his task. By slacking the other agent can increase his payoff. So, the first agent creates a positive externality for the second agent. However, this externality is not internalised by the first agent: he would get the same transfer as by exerting medium effort (the result does not change), but would incur higher costs. He would not benefit from the cost savings of the second agent. Hence efficient production can be implemented because agents do not consider all payoff consequences of their actions.

The idea of inducing an externality to solve an incentive problem is also present in Dewatripont's and Tirole's paper on advocates (1999). But in their context an advocate who increases effort imposes a *negative* externality on his colleague, who, as a consequence, is less likely to win the case. Unlike in our setting, task separation generates competition in the case considered by Dewatripont and Tirole. This also means that sabotage becomes an option between advocates whereas there is no role for sabotage in our model: an agent who successfully sabotages his colleague jeopardises the production result and hence reduces his own payoff.

Even if we have chosen transfers to implement efficient production, it is not obvious that window dressing is not an equilibrium at the same time. The next result addresses this problem.

Theorem 1 (Prevention of window dressing)**.** *If agents specialise ($i \neq j$), transfers can be chosen such that "cheap" window dressing ($e_i^D = 0, e_j^E = 2$) as well as "expensive" window dressing ($e_i^D = 2, e_j^E = 0$) can be prevented while efficient production ($e_i^D = 1, e_j^E = 1$) is a Nash-equilibrium. The respective transfers need to fulfil the conditions:*

$$\begin{aligned} c_i^D(2) - c_i^D(1) \geq \bar{t}_i - \underline{t}_i \geq c_i^D(1) - c_i^D(0) \\ \textit{and } c_j^E(2) - c_j^E(1) \geq \bar{t}_j - \underline{t}_j \geq c_j^E(1) - c_j^E(0). \end{aligned} \tag{10}$$

Proof. Set transfers such that $c_i^D(2) - c_i^D(1) \geq \bar{t}_i - \underline{t}_i \geq c_i^D(1) - c_i^D(0)$. and $c_j^E(2) - c_j^E(1) \geq \bar{t}_j - \underline{t}_j \geq c_j^E(1) - c_j^E(0)$. Note, that this is possible because of the "convexity" in costs expressed in inequality (1). Then, neither cheap nor expensive window dressing can be a Nash-equilibrium because agent j can profitably deviate under cheap window dressing to a lower effort while agent i has an incentive to choose a lower effort under expensive window dressing. At the same time, transfers fulfil the condition from Proposition 3 and thus ensure that efficient production is a Nash equilibrium. □

Window dressing can be prevented because the bonus for a good looking product is too low for high effort to be worthwhile. Setting the bonus in this way is possible because of the convexity of costs in effort.

Theorem 1 allows us to rule out window dressing as an equilibrium. Does this mean that efficient production can be implemented as the unique equilibrium? Unfortunately not: the low effort production ($e_i^D = 0, e_j^E = 0$) is also an equilibrium and cannot be eliminated at the same time as window dressing. However, this result can be regarded as an artifact of the very limited nature of the measure. Suppose the measure is less abrupt and takes on a third value, say 0.5, for ($e_i^D = 0, e_j^E = 1$). Then, we could design a contract which destabilises the low effort equilibrium by making a switch to the middle effort for agent j sufficiently attractive. At the same time, agent i would be unaffected and efficient production would remain an equilibrium.

We have shown that it is possible to implement efficient production ($e_i^D = 1, e_j^E = 1$) by separating tasks. But will this lead to the efficient surplus? The answer depends on the availability of agents. If there is an abundant supply of equally able agents, the principal will hire a second agent and ask him to take over one task. Division of labour will occur merely for incentive reasons because agents are equally able and costs are additive. If employing the second agent involves no hiring costs, the efficient surplus can be obtained. When there is no abundant supply of equally able agents or hiring additional agents leads to some other form of costs, the principal is willing to bear these costs as long as the benefit b is sufficiently large. Thus, one might observe that the principal hires a less productive or otherwise costly agent to increase efficiency – although the same visible result could be produced more cheaply by a single agent.

Efficient production is achieved while the production technology remained the same. Also, the only verifiable result of production has not changed. What is crucial is that the principal employs two agents instead of one. Thus efficient production may no longer be possible if the two agents act as if they were one agent. In other words, the two agents might collude and thereby undermine the incentive scheme. This problem is explored in the following section.

4 Collusion

Before dealing with collusion, we want to be more specific about the ability of different agents. Inequality (3) implied that the demanding task has always higher "marginal" costs. We take this to be a general property of the production technology that cannot be altered by hiring different agents. We thus assume the following generalisation of inequality (3) for the case where differently able agents work at the two tasks:

$$c_i^D(1) - c_i^D(0) > c_j^E(1) - c_j^E(0) \text{ and } c_i^D(2) - c_i^D(1) > c_j^E(2) - c_j^E(1). \quad (11)$$

In other words, there is no agent i who finds it easier to increase effort on the demanding task than another agent j who wants to increase effort on the easy task by a similar amount. We also maintain that any agent working alone prefers window dressing to selecting efficient levels, that is inequality (2) holds.

Within these restrictions imposed by the production technology, there is still scope for agents to differ. We call the population of agents *heterogeneous*, if there is a pair of agents i and j such that the increase of effort at the demanding task by agent i at a low level is less costly than the increase in effort at the easy task by agent j at a *high* level: $c_i^D(1) - c_i^D(0) < c_j^H(2) - c_j^D(1)$. This definition allows us to make a general claim about the consequences of collusion.

Collusion is only possible if the agents have a larger contract space than the principal: they need to be able to formally or informally contract on the easy task – for example the team might engage in repeated interactions while the principal is replaced.[4] We follow Tirole (1986) and model collusion as an enforceable side contract between agents. We assume that agents learn each others productivity when being paired. This facilitates collusion because agents are not required to extract information from each other.

Theorem 2 (Preventing collusion)**.** *If agents can collude, efficient production can be implemented if and only if the population is heterogeneous.*

Proof. To simulate collusion, assume agents could secretly contract on e_i^E and e_i^D. By Theorem 1, a transfer scheme with $c_i^D(2) - c_i^D(1) \geq \bar{t}_i - \underline{t}_i \geq c_i^D(1) - c_i^D(0)$ and $c_j^E(2) - c_j^E(1) \geq \bar{t}_j - \underline{t}_j \geq c_j^E(1) - c_j^E(0)$ yields efficient production in absence of collusion. The most profitable deviation which agents may achieve jointly is either the cheap or the expensive window-dressing. As the superficial appearance does not change ($M = 1$), the sum of transfers received by the agents remains the same. Deviating to the expensive window dressing thus yields $c_i^D(1) + c_j^E(1) - c_i^D(2) - c_j^E(0)$. But by assumption (11) and convexity, we know that

$$c_i^D(2) - c_i^D(1) \overset{(11)}{>} c_j^E(2) - c_j^E(1) \overset{(1)}{>} c_j^E(1) - c_j^E(0),$$

[4] Roy (1952) reports on the workers' behaviour in a production line where workers were able to enforce effort levels in order to restrict joint output, while the employer was unable observe it.

which implies that there is a loss when switching to the expensive window dressing: $c_i^D(1)+c_j^E(1)-c_i^D(2)-c_j^E(0)<0$. The deviation to cheap window dressing brings the agents $c_i^D(1)+c_j^E(1)-c_i^D(0)-c_j^E(2)$. This number is larger than zero when the population is not heterogeneous. If, however, an agent pair (i,j) exists for which $c_i^D(0)+c_j^E(2)-c_i^D(1)-c_j^E(1)>0$, the principal can assign i to the demanding task and j to the less demanding task and rule out any profitable deviations from efficient production. □

The intuition behind this result is that by assigning agents with different productivities to tasks, the principal is able to affect the costs of collusion. If the agent on the difficult task is particularly productive and his colleague on the easy task sufficiently unproductive, collusion simply does not pay for them; the gains from slacking on the difficult task by the productive agent are more than outweighed by the loss of more effort by the unproductive agent on the easy task. Again, this intuition also works for continuous effort (where being productive means having lower marginal costs) and is independent from the specific nature of the observable production result.

Theorem 2 shows that collusion leading to window dressing can be prevented by a particular form of division of labour if agents have different (marginal) production costs.[5] This requires that a marginally less productive agent j is hired and assigned to the easy task. Unlike in the previous section, where we did not consider collusion, there is now a clear optimal assignment of agents to tasks. Consequently, it might not be possible to assign agents to the task where they have a comparative advantage. Accordingly, a seemingly inefficient assignment may be the only way to avoid collusion.

While before, the principal preferred to hire two equally able agents, the possibility of collusion changes this. Even if there is abundant supply of productive agents, the principal now has to employ a less able agent because it is the only way to prevent collusion. It is worth noting that if the incentive scheme induces a negative externality, like in Dewatripont and Tirole (1999), it is always vulnerable to collusion – even if agents differ in productivity. In our case, this is not true. We

[5]Here is an example of a simple cost function which fulfils all assumptions. Effort takes the values $e=0,\frac{1}{2},1$. Costs are αce^2 where $c=1$ for the demanding task and $c=\frac{1}{4}$ for the easy task. Set $\alpha=1$ for an agent j. Suppose there is another agent i with $\alpha\in(\frac{1}{4},\frac{3}{4})$. Then, collusion can be prevented. If agent i is too productive $\alpha<\frac{1}{4}$ expensive window dressing becomes a problem; if $\alpha>\frac{3}{4}$, the gains of cheap window dressing become too attractive.

have just seen that internalising the externality by collusion is not enough when agents productivities differ in a particular way.

If the principal wants to assign productive agents to the difficult task and less productive agents to the easy task, she has to know their productivity. What if this information is not available? Will agents sort themselves to the adequate task? The principal can stop productive agents from applying for the easy task if there is a premium for the difficult task which outweighs the gains of collusion from choosing the easy task. Will unproductive agents then apply for the difficult task to obtain the premium? The principal could ask the agent on the easy task whether his colleague on the difficult task is productive or not, take his report to be true and re-assign agents appropriately. This shifts all bargaining power to the agent on the easy task, who can hold the agent on the difficult task down to the outside option and appropriate any rents. If an unproductive agent went for the difficult task in the hope to capture the premium for this task, he will be held up and loose this premium to the agent on the easy task who threatens to denounce him. Thus, an unproductive agent has no reason to select the "wrong" task. Of course, the agent on the easy task can also hold a productive agent down to the outside option. However, this outside option is not lower than the current gains of the productive agent, so the payoff to him is not altered by the denunciation, there is no rent to be appropriated, and the agent on the easy task has no reason to denounce. So, even in an environment where the principal knows little about the costs of the workers, the division of labour prevents collusion and helps to overcome incentive problems.

5 Conclusion

We have shown that splitting a production process and assigning tasks to different agents can help to improve incentives. More specifically, we examined a situation where the only verifiable result confounds the input of different tasks. Then, having many "specialised" agents allows to better influence inputs on each task and hence ensures a more efficient production than with few and unspecialised agents. Efficient production is achieved merely by increasing the number of agents, assigning them tasks and preventing them from doing the tasks of other agents; the production technology and the nature of the performance measure are not affected. In other words, a larger number of agents is now remunerated on the basis of the same joint outcome.

Our analysis presents two main insights. First, it supplies an argument for division of labour in settings where a single worker can use different aspects of the work to enhance appearance without improving the substance. We show that by employing additional workers and separating tasks amongst them, this opportunity is eliminated and incentives are improved. We thus provide an explanation for specialisation which complements the classical idea of comparative advantage. This has consequences for the literature on job enrichment and enlargement. In this literature a wider job description is thought to be advantageous because it reduces coordination and communication problems (see e.g. Becker and Murphy 1992 or Bolton and Dewatripont 1994), because there are task complementarities (Lindbeck and Snower 2000), or because workers like variation (costs are convex and additive separable between tasks, see e.g. Itoh 1992). These advantages of job enrichment are usually traded off against the advantage of specialisation put forward by Adam Smith. Lindbeck and Snower (2000) argue that new versatile technologies which make workplaces more flexible (such as computers) reduce the benefit of specific investments to improve performance on a task and thus diminish the gains from specialisation. Accordingly, the advantages of job enrichment become relatively more important and modern production processes should be less divided. Here, we have shown that there are other benefits of dividing labour. These benefits do not necessarily decrease when a new more versatile technology is introduced. In fact, if the new technology makes it more difficult to identify task-wise contributions, it may actually be advantageous to partition the production process more finely.

Division of labour is helpful in our setting because workers do not consider all payoff consequences of their actions. They fail to take into account the externality which they create for coworkers. A similar argument is also used by Dewatripont and Tirole (1999), who show that division of labour can help in delivering incentives because it allows to create competition between agents. In their setting, separation of tasks creates a negative externality: greater effort on one task creates a disservice for the other, competing agent. In our context involved parties are not working against each other. The division of labour creates a positive externality, in that greater effort on one task gives the opportunity to slack on the other task. While the negative externality in the tournament-type setting of Dewatripont and Tirole (1999) invites sabotage activities, there is no reason why workers would engage in such activities in our setting.

Both settings are vulnerable to collusion: if workers are able to collude, they internalise the externality. However, and this is our second main insight, while it is not possible to prevent workers from colluding and shirking when the externality is negative, in our context collusion can be prevented by hiring agents with different productivities. We show that by employing a less productive worker on the easy task, the employer can ensure efficient production. Deviating from the desired way of producing is simply too expensive for the workers. This implies that the principal may hire less productive workers even if more productive workers are available.

Our first main finding explains the incidence of "over"-specialisation: production is broken down to very small tasks which are carried out by *different* workers while some of them could easily be carried out by the same worker. Considering hiring costs, coordination costs, etc., this seems inefficient. However, it might be advisable in order to ensure that the desired amount of effort is exerted at each task. The second main finding relates to a seemingly wrong job assignment: to prevent collusion some tasks are deliberately assigned to workers who are not particularly good at them. Both aspects, "over"-specialisation and "wrong" job assignment, are often associated with bureaucracies. Prendergast (2003) pointed out that bureaucracies may be seen as an optimal incentive solution to a specific contractual environment. Here, we show that Prendergast's observation can be extended to hiring decisions and work assignment in bureaucracies for the very same reason: the optimal provision of incentives.

References

ALCHIAN, A. A., AND H. DEMSETZ (1972): "Production, Information Costs, and Economic Organization," *The American Economic Review*, 62(5), 777–795.

ATHEY, S., AND J. ROBERTS (2001): "Organizational Design: Decision Rights and Incentive Contracts," *American Economic Review*, 91(2), 200–205.

BAKER, G. (2002): "Distortion and Risk in Optimal Incentive Contracts," *Journal of Human Resources*, 37(4), 728–751, Special Issue on Designing Incentives to Promote Human Capital.

BECKER, G. S., AND K. M. MURPHY (1992): "The Division of Labor, Coordination Costs, and Knowledge," *Quarterly Journal of Economics*, 107, 1137–1160.

BOLTON, P., AND M. DEWATRIPONT (1994): "The Firm as a Communication Device," *Quarterly Journal of Economics*, CIX(4), 809–839.

BURGESS, S., AND M. RATTO (2003): "The Role of Incentives in the Public Sector: Issues and Evidence," *Oxford Review of Economic Policy*, 19(2), 285–300.

DEWATRIPONT, M., I. JEWITT, AND J. TIROLE (1999): "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies," *Review of Economic Studies*, 66, 199–217.

DEWATRIPONT, M., AND J. TIROLE (1999): "Advocates," *Journal of Political Economy*, 107(1), 1–39.

FAMA, E. F. (1980): "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88(2), 288–307.

FELTHAM, G. A., AND J. XIE (1994): "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations," *The Accounting Review*, 69(3), 429–453.

GIBBONS, R., AND K. J. MURPHY (1992): "Optimal Incentive Contracts in the Presence of Career Concerns Theory and Evidence," *Journal of Political Economy*, 100(3), 469–505.

HOLMSTRÖM, B. (1982): “Moral Hazard in Teams,” *Bell Journal of Economics*, 13, 324–340.

——— (1999): “Manegerial Incentive Problems: A Dynamic Perspective,” *Review of Economic Studies*, 66, 169–182.

HOLMSTRÖM, B., AND P. MILGROM (1991): “Multitask Principal-Agent-Analysis: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics, and Organization*, 7, 24–52.

ITOH, H. (1991): “Incentives to Help in Multi-Agent Situations,” *Econometrica*, 59(3), 611–636.

ITOH, H. (1992): “Cooperation in Hierarchical Organizations: An Incentive Perspective,” *Journal of Law, Economics and Organization*, 8(2), 321–345.

LINDBECK, A., AND D. SNOWER (2000): “Multitask Learning and the Reorganization of Work: From Tayloristic to Holistic Organization,” *Journal of Labor Economics*, 18(3), 353–376.

PRENDERGAST, C. (2003): “The Limits of Bureaucratic Efficiency,” *Journal of Political Economy*, 111(5), 929–958.

RICARDO, D. (1817): *On the Principles of Political Economy and Taxation*. John Murray, London.

ROY, D. (1952): “Goldbricking in a Machine Shop,” *American Journal of Sociology*, 57(5), 427–442.

SCHNEDLER, W. (2003): “The Prudence of Rewarding for A While Hoping for B,” Discussion Paper 03/067, CMPO.

——— (2004): “Performance Measure Congruence and Efficient Allocation Across Tasks,” mimeo, University of Heidelberg.

SMITH, A. (1999): *An Inquiry Into the Nature and Causes of the Wealth of Nations*. Penguin, London, (first published in 1776).

STRAUSZ, R. (1997): “Delegation of Monitoring in a Principal-Agent Relationship,” *Review-of-Economic-Studies*, 64(3), 337–357.

TAYLOR, F. W. (1911): *The Principles of Scientific Management*. Harper, New York.

TIROLE, J. (1986): "Hierarchies and Bureaucracies: On the Role of Collusion in Large Organisations," *Journal of Law, Economics, and Organization*, 2, 181–214.

WILSON, J. Q. (1989): *Bureaucracy: What Governmental Agency Do and Why They Do It*. Basic Books, New York.