Information, Decisions and Incentives

Julia Nafziger*

University of Bonn

February 10, 2007

Abstract

In this paper we endogenize the point in time where an agent (and the principal) observe the realization of an additional signal: before the agent's effort choice (ex ante information) or after (ex post information). We show that there is no difference between incentive and decision problems if the signal is uninformative about the agent's effort: ex ante information is never worse and – if there are gains from tailoring effort to the new information – strictly better than ex post information. If output is informative and there are no gains from tailoring, ex ante information does strictly worse.

Keywords: Moral Hazard, Symmetric Ex Ante Information, Limited Liability
JEL Classification: D82, D86, J33

*jnafziger@uni-bonn.de. I would like to thank seminar participants in Bonn, Brussels and Oxford, as well as Alexander Koch, Sandra Ludwig, Burkhard Schipper, Urs Schweizer and especially Ian Jewitt and Meg Meyer for valuable comments and discussions. Furthermore, I would like to thank Nuffield College and the University of Oxford for their kind hospitality.

1 Introduction

From principal agent theory it is well known that additional information about the agent's effort is (weakly) beneficial: the principal can use it to provide more cheaply incentives to the agent (Holmstrom 1979, Holmstrom 1982). This lead to the conclusion that "more information" is better. One assumption this literature makes is that the agent observes the signal after he has chosen his effort[1] and leaves open the question about the optimal timing of information. This paper addresses the issue whether "more information" in the sense that a signal is revealed before rather than after the action choice of the agent is beneficial for the principal.
While for decision problems it is well known that observing the signal before rather than after the action choice (weakly) increases profits, this is less clear for incentive problems. For incomplete contracts settings there are many examples that more information in the form of an additional signal or earlier revelation of such a signal can hurt (Cremer 1995, Meyer and Vickers 1997, Dewatripont, Jewitt, and Tirole 1999, Prat 2005). But there seems to be a consensus that when contracts are complete more information in any form is always better. In this paper we will question this – showing that "more information" in the sense of information revelation before or after the effort choice can be worse for the principal. For this we endogenize the timing of information in a simple limited liability moral hazard model with two output levels (high and low) and continuous effort. We then ask at what point the principal wants an agent (and herself) to observe the signal realization (that can also be high or low) – before the agent's effort choice (ex ante information, which is in our setting equivalent to intermediate information) or thereafter (ex post information).

Compared to the ex post information case, ex ante information allows the principal to tailor the agent's effort level to the signal realization and thus (weakly) increase her expected revenues. If effort implementation was a simple decision problem where incentives play no role, ex ante information would also enhance expected profits, as these are the difference between revenues and the costs of implementing a certain effort level. However, a change in timing of information also affects implementation costs as effort is unobservable. Our first result is that implementation costs for the same (expected) effort are strictly higher under ex ante information than under ex post information if the signal is informative about the agent's effort.
To understand this, consider first the incentive scheme for the ex post information case. Suppose that the output-signal combination high-high provides the strongest indication

[1] If the information is verifiable this is equivalent to a situation, where he does not observe it all. We subsume this under ex post information

that the agent worked hard. The principal exploits this and concentrates rewards for the agent in this state.[2]

Suppose now that the agent learns the signal realization before he chooses his effort. Then he knows that after a low signal realization his output is very uninformative about his effort. This forces the principal to increase the reward in this state. As a consequence, ex ante implementation costs increase: rewards are not longer concentrated in the high-high state, which is most informative about the agent's effort. If the signal is uninformative about the agent's effort, ex ante information does not affect incentives and hence implementation costs.

We then combine the effects of ex ante information on expected revenues from tailoring of effort with those on implementation costs. First, if the signal is uninformative, then providing ex ante information is never worse and sometimes strictly better than ex post information. The reason is the following: if the signal is not informative about effort, incentives and hence implementation costs are unaltered by the observation of it. The incentive problem behaves as if it was a simple decision problem and therefore ex ante information does strictly better if and only if there are gains from tailoring of effort. This changes if the signal is informative about the agent's effort – the case our second result deals with. Here we show that in the absence of gains from tailoring of effort ex ante information does strictly worse than ex post information. These results emphasize the role of the signal's informativeness about the agent's effort for the optimal timing of information in incentive problems and contrast them with role of effort tailoring in decision problems.

The paper is structured as follows. After discussing the related literature, we introduce the model in Section 2. In Section 3 we solve the first best information scenario. In Section 4 we first review the relevant information concepts. Then we derive the optimal wage schemes for the two informational scenarios ex post versus ex ante information and show the driving forces behind the differences in implementation costs. In Section 5 we derives conditions under which one of the two information scenarios yields higher overall profits. Section 6 concludes. All proofs are in the appendix.

Related Literature

The starting point for this paper is the literature that asks about the value of (ex post) information in principal agent moral hazard models (Holmstrom 1979, Holmstrom

[2]This well known from the literature. See Mookherjee (1984), Holmstrom (1979) or Holmstrom (1982) for risk averse agents or the first part of Che and Yoo (2001)'s model for an example with risk neutral agents, who are protected by limited liability.

1982, Gjesdal 1982, Grossman and Hart 1983, Kim 1995, Jewitt 1997). Our contribution is to endogenize the point in time at which an agent should receive this information – comparing the value of ex ante with that of ex post information. In incomplete contract settings more information may not be in the interest of the principal, both in the sense of more informative ex post signals, or in sense of observing earlier signal realizations (e.g., Cremer (1995), Meyer and Vickers (1997), Dewatripont, Jewitt, and Tirole (1999), Prat (2005), Amaya (2005)). In contrast, we employ in this paper a complete contract setting and ask whether more information – in the sense of more information before the action choice is beneficial. In this respect, our approach is related to Nafziger (2007) who keeps the timing fixed and asks, whether, more information – in the sense of a more informative *ex ante* observable signal is better.

Concerning the endogenous timing of information, the paper is most closely related to Lizzeri, Meyer, and Persico (2002), who consider this topic in a dynamic model: the agent produces in two periods with independent probabilities and may receive feedback about his first period output. If he does not, they show that the optimal incentive scheme rewards the agent iff he has a high output in both periods. Once he receives feedback, the principal has to pay a positive wage both after a low and a high output in the first period. But rewarding him after a low output reduces first period incentives.[3], which makes a feedback policy always worse in their setting.

Although, some of the results are similar to ones in our paper, the driving forces behind them are very different. As informativeness of the signal determines in our paper the structure of the incentive scheme, we can consider more cases than they can: the signal changes or does not change incentives. By this we can point out that a change is needed for ex ante information to do worse. But also the case where incentives change is different: the resulting negative effect on post-information incentives is not present in Lizzeri, Meyer, and Persico (2002): here *only* a negative effect on pre-information incentives emerges. Last, we introduce gains from tailoring effort to the state world, which are not present in their model. In this respect our model is related to Ederer (2004), who considers a model with such effects, taking the wage scheme however as exogenously given.

2 The Model

There is one principal, who employs one agent. The agent is risk neutral, has no wealth and the value of his reservation utility is zero. He produces a verifiable and observable

[3]Schmitz (2005) provides a solution for this dilemma (taking the intermediate information scenario as given): he shows that the principal does sometimes better by employing a different agent in the second period than in the first period.

output x (which equals the principal's revenue), which can be either high ($\bar{x}$) or low ($\underline{x} = 0$). There is a second – verifiable – signal $z \in \{\underline{z}, \bar{z}\}$. The probability that output x and signal z realize depends on the agent's effort $e \in \mathcal{E}$, $\mathcal{E} = [0, b]$ and is denoted by $f(xz|e)$, $f \in (0,1)$. We assume that $f \in C^3$ and:

$$\frac{\partial f(xz|e)}{\partial e} \equiv f_e(xz|e) \begin{cases} > 0 & \text{for } x = \bar{x} \\ < 0 & \text{for } x = \underline{x} \end{cases}$$

$$\frac{\partial^2 f(xz|e)}{\partial e^2} \equiv f_{ee}(xz|e) \begin{cases} < 0 & \text{for } x = \bar{x} \\ > 0 & \text{for } x = \underline{x} \end{cases}$$

The marginal distribution of output is $f(e) = \sum_z f(\bar{x}z|e)$ and $\sum_z f(\underline{x}z|e) = 1 - f(e)$; and for the signal it is $p = \sum_x f(x\bar{z}|e)$ and $1 - p$. The agent's cost of effort function is $c(e)$, $c : \mathcal{E} \to I\!R_0^+$, $c \in C^3$, $c(0) = 0$, $c_e(e) > 0\ \forall e > 0$, $c_e(0) = 0$, $c_{ee}(e) > 0\ \forall e > 0$ and $c_{ee}(0) = 0$.

The timing is as follows. At date 0 the principal decides when the agent should observe the realization of the signals: before he provides effort (ex ante information) or after (ex post). Furthermore, she offers a wage scheme to the agent, which specifies four wages: $\mathbf{w} = (w(\bar{x}\bar{z}), w(\bar{x}\underline{z}), w(\underline{x}\bar{z}), w(\underline{x}\underline{z}))$.[4]
Under the ex post information scenario the agent provides effort at date 1 and then the signal realizes. Effort is unobservable. Under the ex ante information scenario first the signal realizes – which is observable to both the principal and the agent – and then the agent provides unobservable effort: $e(\bar{z}) \equiv \bar{e}$ after $z = \bar{z}$ and $e(\underline{z}) \equiv \underline{e}$ after $z = \underline{z}$. After the realization of all outputs, payoffs realize: the principal receives the revenues net of wage payments, while the agent receives his wage minus his effort costs.
Note that strictly speaking ex ante information corresponds in our setting to intermediate information as the agent has already signed the contract when he receives the information. In the appendix we show however that there is no difference between the two, because the agent's reservation utility is zero. Hence, we simply refer to the case, where the agent observes the signal before the effort choice to ex ante information.

3 First Best

As a benchmark we consider first the informational scenario the principal would choose if she could observe efforts, i.e. if there is no incentive problem and she only has to

[4] For the intermediate information scenario we require that the ex ante participation constraint is satisfied. For the ex ante information case that the intermediate participation constraints are.

decide which effort to implement. Here she compensates the agent exactly for his effort costs and sets $w = c(e)$ if he provides the desired effort of e and pays him zero else.
The problem of the principal for the ex post information scenario is hence to maximize $f(e)\bar{x} - c(e)$ over e. For the ex ante information scenario – where she can make efforts state contingent – her problem is to maximize $E\left[f(\bar{x}z|e(z))\bar{x} - c(e(z))\right]$ over $(\bar{e}, \underline{e})$. We see that if we require $\underline{e} = \bar{e}$ the profit functions for the two scenarios coincide. Thus, the maximization problem of the principal is the same for both structures, except that for the ex post information scenario the restriction $\bar{e} = \underline{e}$ applies. But the possibility to let $\bar{e}$ differ from $\underline{e}$, cannot make the principal worse off – if it increases her profits she allows these two effort levels to differ from each other, otherwise she sets them equal. Hence, ex ante information does never worse than ex post information.
It follows that ex ante information has only a chance to do strictly better if the principal sets $\bar{e} \neq \underline{e}$, i.e. only if she makes second period effort state contingent. But this has two opposing effects. To consider solely the first effect assume that output and the signal are independent, which implies $f(\bar{x}|\bar{z}, e) = f(\bar{x}|\underline{z}, e) = f(e)$. Then the principal can for any effort vector of the ex ante information scenario with $\bar{e} \neq \underline{e}$ implement the same expected effort for the ex post scenario leading to strictly higher profits. This follows immediately from Jensen's Inequality as the profit function is strictly concave: the convex combination of profits (having with probability p profits from $\bar{e}$ and with $1-p$ from $\underline{e}$) is lower than the profit from the convex effort combination. Thus, to avoid this decrease the principal would ignore the information and set $\bar{e} = \underline{e}$. Hence, ex ante information cannot increase profits.
The question is hence, whether there is a second effect (for dependent outputs) that can outweigh the negative effect of setting $\bar{e} \neq \underline{e}$, such that ex ante information is strictly better:

Proposition 1 *Iff $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e}) \neq pf(\bar{e}) + (1-p)f(\underline{e})$, then ex ante information does strictly better than ex post information.*

This is the standard result that ex ante information strictly benefits in decision problems iff there are gains from tailoring effort to the state of the world, which is the case iff $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e}) \neq pf(\bar{e}) + (1-p)f(\underline{e})$.

4 The Wage Scheme

The first best analysis shows that if effort implementation is a simple decision problem, ex ante information can never harm. So what happens if effort implementation is an incentive problem? Here the principal's problem is to maximize for each informational scenario her expected profits over effort and wages, subject to the agent's incentive, limited liability and participation constraint. Call an effort that satisfies these three constraints for a given wage scheme *implementable*.

As usual in a moral hazard setting we decompose the problem into two parts. In the first step, we fix an implementable effort – e for the ex post and $(\bar{e}, \underline{e})$ for the ex ante information scenario – and minimize expected wage payments $\sum_x [f(x\bar{z}|\bar{e})w(x\bar{z}) + f(x\underline{z}|\underline{e})w(x\underline{z})]$, where $\underline{e} = \bar{e} = e$ for the ex post information scenario. Call the solution to this problem the implementation cost functions $C(\mathbf{xz}|e)$ for the ex post and $C^A(\mathbf{xz}|\bar{e}, \underline{e})$ for the ex ante information scenario, where $\mathbf{xz} = \{\bar{x}\bar{z}, \bar{x}\underline{z}, \underline{x}\bar{z}, \underline{x}\underline{z}\}$. In the second step, we maximize the principal's total profits $\Pi(e) = f(e)\bar{x} - C(\mathbf{xz}|e)$ and $\Pi^I(\bar{e}, \underline{e}) = [f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e})]\,\bar{x} - C^I(\mathbf{xz}|\bar{e}, \underline{e})$ over efforts.

In this section we consider the first part of the solution of the incentive problem described above for the two informational scenarios and compare implementation cost functions. To understand however the impact of ex ante information on them, we have to to clarify what it means for a signal to be informative. Maximizing solely revenues minus costs e.g. in the first best is a simple decision problem and hence ex ante information is beneficial iff there are gains from tailoring effort to the state of the world. For incentive problems this depends on how ex ante information changes the agent's incentives, i.e. on whether the signal is informative about the agent's effort.

4.1 Review of Information Concepts

We review here briefly the information concepts relevant to our analysis: information about efforts and gains from tailoring of effort. As we want to point out the differences between incentive and decision problems we are especially interested in examples, where the one holds but the other not. Furthermore, we will relate these concepts to informativeness about outputs to see why the latter would not suffice for our purpose. To illustrate, we often use the following examples, which refer to the distribution function in Table 1:

Definition 1 *The "times-c model" is defined by:* $k = 0$*. The "plus-k model" is defined by:* $c = 1$*.*

The term "times-c model" refers to the fact that in this model $f(x|e)p$ is multiplied by c. The "plus-k model" adds the term k to $f(x|e)p$. Note that one can generate the plus-k model by defining $f(\bar{x}|e, \sigma) = e + \sigma$ and the times-c one by defining $f(\bar{x}|e, \sigma) = e\sigma$, where σ is a shock that is correlated with the signal. This fits to our notation by setting $f(e) = Ef(\bar{x}|e, \sigma)$, $c = \frac{E\sigma z}{E\sigma Ez}$ (hence $c > 1 \leftrightarrow Cov(\sigma, z) > 0$) and $k = Cov(\sigma, z)$.

	$\bar{x}$	$\underline{x}$	
$\bar{z}$	$cf(e)p - k$	$p(1 - cf(e)) + k$	p
$\underline{z}$	$f(e)(1 - cp) + k$	$1 - f(e) - p - cf(e)p) - k$	$1 - p$
	$f(e)$	$1 - f(e)$	1

Table 1: Joint Distribution for Definition 1

4.1.1 Informativeness about Effort

As mentioned above we are interested whether the signal is informative about the agent's effort. For this we compare the likelihood ratios $l(\bar{x}\bar{z}|e) \equiv \frac{f_e(\bar{x}\bar{z}|e)}{f(\bar{x}\bar{z}|e)}$ and $l(\bar{x}\underline{z}|e) = \frac{f_e(\bar{x}\underline{z}|e)}{f(\bar{x}\underline{z}|e)}$ and $l(\bar{x}z|e) = \frac{f_e(\bar{x}|e)}{f(\bar{x}|e)}$. Those tell us how likely it is that the agent provided an effort of e, given that the output-signal realization is $\bar{x}z$. Suppose $l(\bar{x}|e) \neq l(\bar{x}\bar{z}|e) \neq l(\bar{x}\underline{z}|e)$. Then we say that *the signal is informative about the agent's effort.* For $l(\bar{x}|e) = l(\bar{x}\bar{z}|e) = l(\bar{x}\underline{z}|e)$ we say it is *uninformative.*

This definition fits what Demougin and Fluet (1998) call "mechanism sufficiency". A statistic is said to be mechanism sufficient if the implementation costs for a certain effort depend only on this statistic. In our case it turns out that iff likelihood ratios in $\bar{x}\bar{z}$ and $\bar{x}\underline{z}$ are equal, then implementation costs $C(\mathbf{xz}|e)$ depend only on $x = \bar{x}$ and not on the signal. Hence, the agent's output is mechanism sufficient and for brevity we say the signal is *uninformative.* Iff the likelihood ratios are unequal implementation costs depend on the signal and we say it is informative.

Note the difference to the definitions for risk averse agents. Here x is called informative (uninformative) – in the sufficient statistic sense – if likelihood ratios are equal (unequal) also for output signal combinations $\underline{x}z$.[5] Thus, for them the sufficient statistic and the sufficient mechanism coincide.

Turning to our two examples, we see that in the plus-k model the signal is informative about the agent's effort for $k \neq 0$. For $k = 0$ output is uninformative. Hence, we can vary the informativeness of the signal by varying k . For the times-c model output is uninformative no matter which value c takes.

4.1.2 Gains from Tailoring of Effort

For the first best we already introduced the concept of gains from tailoring of effort. For our model specification with $\underline{x} = 0$ these are not simply captured by dependence. Consider for example the plus-k model: here effort and the term that captures dependence (k) are not linked. Thus, the principal cannot increase the expected probability

[5] See Holmstrom (1982). Gjesdal (1982), Grossman and Hart (1983), Kim (1995) and Jewitt (1997) use the term more informative in the sense of Blackwell (1953) and Lehmann (1988).

of obtaining output $\bar{x}$ by making effort state contingent while keeping the sum of effort across states fixed (*tailoring of effort*). To see this, note that Jensen's Inequality implies that such a policy must decrease the expected probability of state $\bar{x}$ in the absence of gains from tailoring effort:

$$f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e}) = pf(\bar{e}) + (1-p)f(\underline{e}) \leq f(p\bar{e} + (1-p\underline{e}))$$

In contrast, for the times-c model with $c \neq 1$, effort and the term that captures dependence (c) are directly linked and the principal can increase the expected probability by making effort state contingent.

4.1.3 Summary and Relation to Dependence

Why did we introduce two information concepts rather than simply consider dependence between signal and output distributions? The latter can be characterized using copulas. The copula "couples a bivariate distribution function to its one dimensional marginals" (Nelsen (1995)): define a copula $K(u,v)$, $K : [0,1]^2 \rightarrow [0,1]$, $u = f(x|e)$, $v = f(z)$. It lies between the Fréchet bounds: $\max\{u+v-1, 0\} \leq K(u,v) \leq \min\{u,v\}$, which give bounds for most negatively (positively) dependent random variables. Furthermore, $K(u,v) = uv$ for independent random variables.
Consider the plus-k model. By varying k, we shift the copula between the Fréchet bounds: for $k < 0$, the copula lies above the uv and one says that the random variables are positively quadrant dependent; for $k > 0$, they are negatively quadrant dependent, and for $k = 0$, they are independent. As explained before, by varying k we affect also the degree of informativeness of the signal about effort. Hence, we can vary informativeness about output and effort simultaneously by shifting k: the plus-k model is an example where the signal is either informative about effort *and* output or uninformative about both. Note however that dependence fails to fully capture what our two information concepts do: the two random variables are dependent (for $k \neq 0$) but there are no gains from tailoring of effort for any value of k.
For the times-c model, $c > 1$ corresponds to positive quadrant dependence, $c < 1$ to negative quadrant dependence, and $c = 1$ to independence. Again dependence does not capture fully what we are interested in: remember that here the signal is uninformative about the agent's effort, no matter which value c takes, but that there are gains from tailoring of effort for $c \neq 1$.
To summarize (see also Table 2), dependence between output and the signal does not imply that there are gains from tailoring of effort or that the signal is informative about effort, and none of these latter two concepts does implies the other. In the other direction, however gains from tailoring of effort or informativeness about effort imply informativeness about output.

	$c = 1$	$c \neq 1$
$k = 0$	NO,NE,NG	O,NE,G
$k \neq 0$	O,E,NG	O,E,G

Table 2: Relation between Parameters in Table 1 and Information Concepts. O: information about outputs. E: information about effort. G: gains from tailoring of effort. NY: no information about $Y \in \{O, E, G\}$.

4.2 Ex Post Information

In this section we consider the first part of the principal's problem: minimizing the expected wage payments for a given implementable effort level. We start with the ex post information scenario. The problem is to minimize expected wage payments:

$$\min_{\mathbf{w}} \sum_x \sum_z f(xz|e)w(xz)$$

s.t. the limited liability constraints $w(xz) \geq 0 \ \forall xz$ and the incentive constraint:

$$\hat{e} \in \arg\max_{e \in \mathcal{E}} \quad \sum_x \sum_z f(xz|\hat{e})w(xz) - c(e),$$

The participation constraint is satisfied if the other two constraints are, and we therefore omit it in the above problem. Furthermore, higher wages after a low output decrease not only the agent's incentives, but also the principal's profits. Hence, $w(\underline{x}\bar{z}) = w(\bar{x}\underline{z}) = 0$. The reduced maximization problem regarding the wages $w(\bar{x}\bar{z})$ and $w(\bar{x}\underline{z})$ is graphically illustrated in Figure 1: as lower wages increase the principal's profits her aim is to push the isowageline as far down as possible, i.e. as far down as the incentive constraint permits. As this is a linear problem, the optimal wage scheme sets one of the wages $w(\bar{x}\bar{z})$ or $w(\bar{x}\underline{z})$ larger than zero and the other one equal to zero, depending on the slope of the two lines. For example, if the isowageline is strictly steeper than the incentive constraint $(-\frac{f(\bar{x}\bar{z}|e)}{f(\bar{x}\underline{z}|e)} > -\frac{f_e(\bar{x}\bar{z}|e)}{f_e(\bar{x}\underline{z}|e)} \leftrightarrow l(\bar{x}\bar{z}|e) > l(\bar{x}\underline{z}|e))$, then $w(\bar{x}\underline{z}) = 0$ and $w(\bar{x}\bar{z})$ is chosen such that the incentive constraint is satisfied: $w(\bar{x}\bar{z}) = c_e(e)/f_e(\bar{x}\bar{z}|e)$. Iff the slopes (and hence the likelihood ratios) are equal, all wage combinations that satisfy the incentive constraint are optimal.

To understand the intuition, suppose the likelihood ratio in state $\bar{x}\bar{z}$ is larger than the one in state $\bar{x}\bar{z}$. In other words, it is more likely in state $\bar{x}\bar{z}$ than in state $\bar{x}\underline{z}$ that the agent provided a certain effort. Hence, the principal optimally concentrates rewards for the agent in this output state ($w(\bar{x}\bar{z}) > w(\bar{x}\underline{z}) = 0$). If the signal contains no additional information about effort (likelihood ratios are equal), the principal cannot reduce information rents by conditioning the wage scheme on the signal. Doing so however does not harm as the agent is risk neutral for all positive wage combinations.

Thus, the principal is indifferent whether or not to condition an agent's wage on the signal.

The following lemma that we prove formally in the appendix summarizes the discussion:

Lemma 1 *Suppose the principal wants to implement e. Then the wages after a low output are zero:* $w(\underline{x}\bar{z}) = w(\underline{x}\underline{z}) = 0$. *Furthermore:*

(i)

$$If \quad l(\bar{x}\bar{z}|e) \begin{cases} > l(\bar{x}\underline{z}|e) \\ < l(\bar{x}\underline{z}|e) \end{cases} \quad then \quad w(\bar{x}\bar{z}) = \begin{cases} \frac{c_e(e)}{f_e(\bar{x}\bar{z}|e)} \\ 0 \end{cases} \quad and \quad w(\bar{x}\underline{z}) = \begin{cases} 0 \\ \frac{c_e(e)}{f_e(\bar{x}\underline{z}|e)} \end{cases}.$$

(ii) If $l(\bar{x}\bar{z}|e) = l(\bar{x}\underline{z}|e)$, *any* $w(\bar{x}\bar{z})$ - $w(\bar{x}\underline{z})$ *combination that satisfies the incentive constraint is optimal:* $\sum_z f_e(\bar{x}z|e)w(\bar{x}z) = c_e(e)$.

In the following we want to focus on either on Case (i) or (ii) and hence impose:

Assumption 1 $sign\left(l(\bar{x}\bar{z}|e) - l(\bar{x}\underline{z}|e)\right) = constant\ \forall e.$

This assumption is for example met by the times-c or the plus-k model. In the appendix we derive from Lemma 1 the expected implementation costs for e for the different wage schemes:

$$C(s|e) \equiv \frac{c_e(e)}{l(s|e)}, \tag{1}$$

where $s \in \{\bar{x}\bar{z}, \bar{x}\underline{z}\}$ depending on the ordering of the likelihood ratios, as described in Lemma 1, Case (i); for Case (ii) $s = \bar{x}$. We make the following assumption:

Assumption 2 $\frac{c_e(e)}{l(s|e)}$ *is convex in e.*

This is a sufficient condition for the principal's problem to be strictly concave in effort and for example met by the times-c or the plus-k model.

4.3 Ex Ante Information

In this section we derive the wage scheme for the agent if he can observe the signal realization before he chooses his effort. This implies that we have to consider two incentive constraints – one after a high signal and one after a low one:[6]

$$\text{after observing } \bar{z}: \ \bar{e} \in \arg\max_{e \in \mathcal{E}} \ f(\bar{x}|\bar{z}, \hat{e})w(\bar{x}\bar{z}) - c(e),$$
$$\text{after observing } \underline{z}: \ \underline{e} \in \arg\max_{e \in \mathcal{E}} \ f(\bar{x}|\underline{z}, \hat{e})w(\bar{x}\underline{z}) - c(e),$$

[6]As aforementioned it is optimal to set an agent's wage equal to zero in case this agent's output is low (i.e. $w(\underline{x}\bar{z}) = w(\underline{x}\underline{z}) = 0$) – a fact already used in the stated constraints. Note that because the agent has reservation utility of zero the participation constraint still does not bind: he receives a positive rent in each state. Using this argument, it is easy to see that there is no difference between ex ante and intermediate information in our setting.

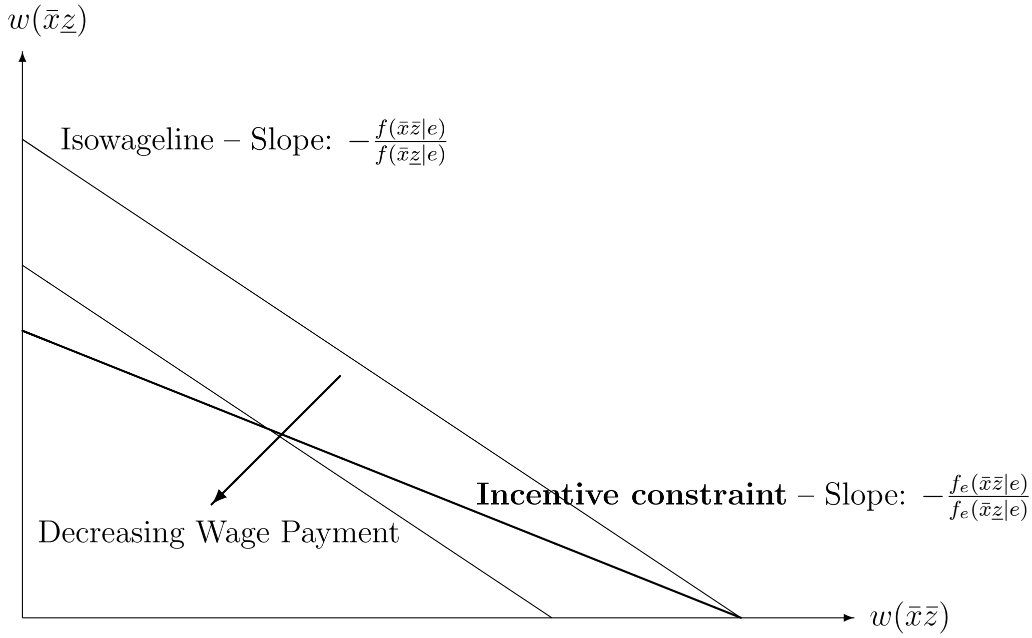

Figure 1: Derivation of the Wage Scheme

where the posterior probabilities are $f(\bar{x}|z,e) = \frac{f(\bar{x}\bar{z}|e)}{p}$ and $f(\bar{x}|z,e) = \frac{f(\bar{x}\underline{z}|e)}{1-p}$. Using these formulas we obtain the following wages that are necessary to implement effort levels $(\underline{e}, \bar{e})$: $w(\bar{x}\bar{z}) = \frac{pc_e(\bar{e})}{f_e(\bar{x}\bar{z}|\bar{e})}$ and $w(\bar{x}\underline{z}) = \frac{(1-p)c_e(\underline{e})}{f_e(\bar{x}\underline{z}|\underline{e})}$. This yields the implementation costs for $(\underline{e}, \bar{e})$ given an informative signal:

$$C^A(\bar{x}\bar{z}, \bar{x}\underline{z}|\bar{e}, \underline{e}) \equiv p\, C(\bar{x}\bar{z}|\bar{e}) + (1-p)\, C(\bar{x}\underline{z}|\underline{e}). \tag{2}$$

For an uninformative signal $C(\bar{x}\bar{z}|e) = C(\bar{x}\underline{z}|e) = C(\bar{x}|e)$ and hence reduces to:

$$C^A(\bar{x}|\bar{e}, \underline{e}) \equiv p\, C(\bar{x}|\bar{e}) + (1-p)\, C(\bar{x}|\underline{e}). \tag{3}$$

4.4 Comparison of Implementation Costs

The purpose of this section is to develop the intuition for why ex ante information may be harmful. For this we compare here only the implementation costs for the two information structures, working with the hypothetical scenario that the principal implements the same expected effort level for both structures. That is, the vector $(\underline{e}, \bar{e})$ for the ex ante information scenario and $e = p\,\bar{e} + (1-p)\,\underline{e}$ for the ex post one. This helps us later – when we compare overall profits – to construct sufficient conditions for one structure to be optimal: $(\bar{e}, \underline{e})$, say, can be *any* effort vector for the ex ante information scenario, including the optimal one. This vector then determines the effort level for the ex post scenario. As a consequence of this restriction, the optimal effort level for this scenario may be excluded though. However, if in the ex ante information scenario the implementation costs are higher for all effort vectors (one of which is the optimal one) than those in the ex post information scenario with the restricted effort level, then a fortiori they are higher than those in the ex post information with the optimal effort level. Thus, such a comparison gives us – when we later on add revenues – *sufficient* conditions for one informational scenario (in the example the ex post one) to lead to higher profits.

Proposition 2 *Suppose Assumption 2 holds. Then for any* $(\bar{e}, \underline{e}) \neq \mathbf{0}$ *of the ex ante information scenario the principal can implement at lower costs the same expected effort* $e = p\,\bar{e} + (1-p)\,\underline{e}$ *under the ex post information scenario. The decrease in implementation costs is strict for:*

(i) $\frac{c_e(e)}{l(\bar{x}\bar{z}|e)}$ *strictly convex and* $\bar{e} \neq \underline{e}$*,*

(ii) $l(\bar{x}\bar{z}|e) \neq l(\bar{x}\underline{z}|e)$*.*

There are two forces that cause the strict inequality: convex implementation costs (Case (i) of Proposition 2) and a change in the feasible wage scheme (Case (ii)). To illustrate the first force suppose that the likelihood ratio in state $\bar{x}\bar{z}$ is equal to the one in $\bar{x}\underline{z}$. Hence, $C(\bar{x}\bar{z}|e) = C(\bar{x}\underline{z}|e) = C(\bar{x}|e)$ for $\bar{e} = \underline{e} = e$ and implementation costs for the two informational scenarios are equal: $p\,C(\bar{x}|\bar{e}) + (1-p)\,C(\bar{x}|\underline{e}) = C(\bar{x}|e)$. Intuitively, an uninformative signal does not change the agent's incentives and the principal can simply ignore it. Making effort state contingent ($\bar{e} \neq \underline{e}$) would then imply that the left hand side of the previous equation is a convex combination of the costs to implement efforts $\underline{e}$ and $\bar{e}$, respectively, while the right hand side represents the implementation costs of the convex effort combination $e = p\bar{e} + (1-p)\underline{e}$. Hence, by Jensen's inequality, implementation costs for the ex post information scenario would be (strictly) lower iff the implementation cost function is (strictly) convex given $\underline{e} \neq \bar{e}$.
If the signal is informative about the agent's effort this changes: we now have $C(\bar{x}\bar{z}|e) < C(\bar{x}\underline{z}|e)$, which implies that:

$$p\,C(\bar{x}\bar{z}|e) + (1-p)\,\underbrace{C(\bar{x}\underline{z}|e)}_{>C(\bar{x}\bar{z}|e)} > C(\bar{x}\bar{z}|e).$$

Hence, the principal cannot simply ignore the information by setting $\bar{e} = \underline{e} = e$.[7] Making effort state contingent does again only increase the value of the implementation costs

[7]Note however that for linear implementation costs functions the strict inequality holds only for $\underline{e} > 0$: if this effort level was zero, not only the convexity effect disappears, but also the disadvantage of high implementation costs in the uninformative state. However, under our standard assumptions we $\underline{e} > 0$ is optimal.

further. At first glance, it seems surprising that the principal cannot reduce implementation costs by conditioning effort on the informative state of the world. However, because the agent observes the signal realization she has to pay a positive wage in both $\bar{x}\bar{z}$ and $\bar{x}\underline{z}$ and not only in state $\bar{x}\bar{z}$. Paying a positive wage in the state $\bar{x}\underline{z}$ – in which it is less likely that the agent provided effort than in the other state – offsets the informational advantage that the principal can achieve by tailoring of effort. Thus, if the signal is informative, the principal can reduce implementation costs by conditioning an agent's *incentive scheme* on this output, but not by conditioning his *effort* on it.

5 Comparing the Structures – Total Profits

In this section we compare overall profits for both information scenarios to make statements about the optimal timing of information. To get necessary and sufficient conditions for optimality, we need to proceed to the second step of the principal's problem and maximize her profits over effort, using the implementation costs derived in the first step. While we do so for the case where the signal is uninformative about the agent's effort, we will use the sufficient conditions to compare structures for the case where output is informative: solving the principal's problem generally is not tractable here and the sufficient conditions allow us to show that ex ante information can do strictly worse in incentive problems what it can never do in decision problems.

5.1 Uninformative Signal

If the signal is uninformative about the agent's effort we know from the previous section that setting $\underline{e} = \bar{e}$ makes the implementation cost functions equal for both structures (Proposition 2, Case (i)). This hinges on the likelihood ratios being equal: the information does not change the agent's incentives as it does not help to estimate his effort.
Furthermore, expected revenues for the ex ante information scenario are equal to the ones for the ex post information scenario if $\bar{e} = \underline{e}$ as then $pf(\bar{x}|\bar{z}, \bar{e}) + (1-p)f(\bar{x}|\underline{z}, \underline{e}) = f(e)$.Therefore, if we require $\underline{e} = \bar{e}$ the ex ante information scenario's profit function coincides with the ex post information scenario's one. Hence – as for the first best – the principal cannot do worse when maximizing without the restriction $\underline{e} = \bar{e}$ and can do strictly better iff there are gains from tailoring of effort:

Proposition 3 *Suppose $l(\bar{x}\bar{z}|e) = l(\bar{x}\underline{z}|e)$. Then the ex ante information scenario yields at least as high profits as the ex post one. Iff $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e}) \neq pf(\bar{e}) + (1-p)f(\underline{e})$, the ex ante information scenario yields strictly higher profits.*

Note the similarity between Proposition 1 and 3. Key for the understanding is that the second best maximization problems parallels the one for the first best: the wage scheme does not change when moving from the ex post to the ex ante information scenario if the signal is uninformative. This implies that the principal can ignore the information if there are no gains from tailoring of effort and exploit these if there are some. The joint presence of gains from tailoring of effort and an uninformative signal is possible as we showed with the times-c model.

Discussion

Proposition 3 also shows that even for concave profit functions ex ante information is optimal because of the gains from tailoring of effort. This result parallels the one for the "multiplicative model" (which is equivalent to our times-c model) of Ederer (2004). However, he assumes that the wage scheme is the same for both structures. Our analysis showed that the argument for optimality relies on information not altering the wage scheme. As the latter is driven by the uninformativeness of the signal about effort, we had to point out that there exist cases where this is possible, although there are gains from tailoring effort.

In the endogenous wage model with one agent of Lizzeri, Meyer, and Persico (2002) $\bar{e} \neq \underline{e}$ is optimal so as to provide first period incentives. This drives up convex implementation costs and makes ex ante information suboptimal because there are no gains from tailoring of effort. In comparison, in our model with an uninformative signal the principal chooses $\bar{e} \neq \underline{e}$ to exploit the gains from tailoring of effort and not to influence incentives: ex ante information does not change the wage scheme , which is impossible in their dynamic model. Hence, in our setting providing ex ante information can be strictly better, although wages are endogenous.

5.2 Informative Signal

When the signal is informative the implementation cost functions do not coincide for $\underline{e} = \bar{e}$. Hence, we cannot apply the argument we used to show Proposition 3. To get necessary and sufficient conditions for the present case we would have to fully determine optimal efforts and compare the value functions. Instead tackling directly this intractable problem, we build on our previous results to derive sufficient conditions that enable us to show that ex ante information does strictly worse than ex post information in the absence of gains from tailoring of effort:

Proposition 4 *Suppose* $l(\bar{x}\bar{z}|e) \neq l(\bar{x}\underline{z}|e)$ *and* $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e}) = pf(\bar{e}) + (1-p)f(\underline{e})$. *Then the ex ante information scenario leads to strictly lower profits than the ex post information scenario.*

As there are no gains from tailoring of effort, expected revenue is lower under ex ante than under ex post information: $pf(\bar{e}) + (1-p)f(\underline{e}) \leq f(p\bar{e} + (1-p)\underline{e})$ by Jensen's inequality. Regarding implementation costs remember from Proposition 2 that for any $(\bar{e}, \underline{e})$ in the ex ante information scenario, the principal can implement the same expected effort for the ex post information scenario, but with strictly lower implementation costs. Taking both together we see that for any $(\bar{e}, \underline{e}) \neq 0$ (which is optimal for $f_e(\bar{x}z|e) > 0$) we can implement the same expected effort under the ex post scenario leading to strictly higher profits. This profit is lower than the value function: the expected effort from the ex ante information case need not be optimal for the ex post scenario. But the principal cannot do worse with the optimal effort.

Discussion

In this section we showed that more information in the sense of obtaining a signal realization earlier can make the principal strictly worse off. To arrive at this result we abstracted from gains of tailoring effort in the revenue function – by choosing the joint probability function appropriately and by assuming $\underline{x} = 0$. This allowed us to show that ex ante information does *strictly worse* than ex post information when effort is unobservable in precisely those circumstances where both information structures would perform equally well if the principal faced a simple decision problem with observable effort. Adding gains from tailoring of effort (by e.g. choosing $\underline{x} \neq 0$) would create a trade-off: ex ante information is good because the principal can increase revenues by tailoring effort to the state of the world, but bad because it changes incentives. Depending on the strength of these effects ex ante information can do strictly better or worse. Analyzing this however is not possible with general functions, as one has to determine optimal effort levels and would not lead to a fundamentally new insight: Proposition 4 already shows that ex ante information can do strictly worse than ex post information, which it can never do in decision problems.

Note that the suboptimality of ex ante information in Ederer (2004) for his plus-k model has very different causes. The exogenously imposed wages mean that the principal cannot implement the same effort levels across states, only $\bar{e} \neq \underline{e}$ can arise – even though there are no gains from tailoring effort. This strictly decreases the value of the concave profit function. In contrast, the suboptimality of ex ante information in our model is driven by a change in the wage scheme, and therefore would even arise if the principal set $\underline{e} = \bar{e}$ or if the profit function was linear.

In Lizzeri, Meyer, and Persico (2002) the suboptimality of ex ante information is also caused by a change in the wage scheme. However, it is not informativeness about efforts that drives it but the fact that the agent anticipates that he will receive a positive wage also after a failure in the first period. Thus, the negative effect in their model arises *only* because pre-information incentives change, while in ours the effect comes from changes in post-information incentives.

6 Conclusion

In this chapter we considered the impact of ex ante information on the incentive and decision problems. We have seen that the optimality of ex ante information relative to ex post information depends on two forces: gains from tailoring of effort and whether the signal is informative about the agent's effort. While the relevance of the first for the optimality of ex ante information is well known from decision problems, we showed the importance of the second one for incentive problems and the interplay between the two forces. One of our surprising results is that ex ante information which provides information about the agent's effort can harm once we consider incentives.

Appendix

Proof Proposition 1.

Let (e) be any effort vector of the ex post information scenario. In the following we show first that we can find $(\underline{e}, \underline{e})$, satisfying $e = p\bar{e} + (1-p)\underline{e}$ that yield strictly higher profits if the condition in the proposition is satisfied. Note that this triple $(\underline{e}, \bar{e})$ need not to be optimal for the ex ante information scenario. However, – as profits are already higher with these non-optimal ones, profits have to be higher with the optimal ones and hence we get with this method a sufficient condition.

Consider first the cost function. By a Taylor Approximation we have (assume that $\bar{e} \geq \underline{e}$, the case $\bar{e} \leq \underline{e}$ is analogous):

$$c(\bar{e}) = c(\underline{e}) + c_e(\underline{e})(\bar{e} - \underline{e}) + \sum_k \frac{c^k(\underline{e})(\bar{e} - \underline{e})^k}{k!}$$

and

$$c(e) = c(\underline{e}) + pc_e(\underline{e})(\bar{e} - \underline{e}) + \sum_k \frac{c^k(\underline{e})p^k(\bar{e} - \underline{e})^k}{k!}.$$

Hence, we can we write the implementation costs for the ex ante information scenario as:

$$pc(\bar{e}) + (1-p)c(\underline{e}) = c(\underline{e}) + pc_e(\underline{e})(\bar{e} - \underline{e}) + p\sum_k \frac{c^k(\underline{e})(\bar{e} - \underline{e})^k}{k!}.$$

Subtracting implementation costs for the ex post information scenario from the ones for the ex ante information scenario we get the difference:

$$g(\bar{e} - \underline{e}) \equiv \sum_k \frac{c^k(\underline{e})(p - p^k)(\bar{e} - \underline{e})^k}{k!}.$$

Note that $g(0) = 0$ and $\frac{\partial g(\bar{e} - \underline{e})}{\partial(\bar{e} - \underline{e})}|_{\bar{e} = \underline{e}} = 0$. Thus, costs do not increase by introducing a small spread under the ex ante information scenario.

The difference in expected revenues is given by:

$$f(\bar{x}\bar{z}|\bar{e}) - f(\bar{x}\bar{z}|e) + f(\bar{x}\underline{z}|\underline{e}) - f(\bar{x}\underline{z}|e).$$

Again by a Taylor Approximation:

$$f(\bar{x}\bar{z}|e, \cdot) = f(\bar{x}\bar{z}|\bar{e}) - (1-p)f_e(\bar{x}\bar{z}|\bar{e})(\bar{e} - \underline{e}) - \sum_k \frac{(1-p)^k(f(\bar{x}\bar{z}|\bar{e}))^k(\bar{e} - \underline{e})}{k!}$$

and

$$f(\bar{x}\underline{z}|e) = f(\bar{x}\underline{z}|\underline{e}) + pf_e(\bar{x}\underline{z}|\underline{e})(\bar{e} - \underline{e}) + \sum_k \frac{p^k(f(\bar{x}\underline{z}|e))^k(\bar{e} - \underline{e})}{k!}.$$

Hence, we can write the difference in revenues as:

$$\begin{aligned} h(\bar{e} - \underline{e}) \quad &:= \quad [(1-p)f_e(\bar{x}\bar{z}|\bar{e}) - pf_e(\bar{x}\underline{z}|\underline{e})](\bar{e} - \underline{e}) \\ &+ \quad \sum_k \frac{(1-p)^k(f(\bar{x}\bar{z}|\bar{e}))^k(\bar{e} - \underline{e})}{k!} - \sum_k \frac{p^k(pf(\bar{x}\underline{z}|\underline{e}))^k(\bar{e} - \underline{e})}{k!}. \end{aligned}$$

Note that $h(0) = 0$ and $\frac{\partial h(\bar{e}-\underline{e})}{\partial(\bar{e}-\underline{e})}|_{\bar{e}=\underline{e}} = (1-p)f_e(\bar{x}\bar{z}|\bar{e}) - pf_e(\bar{x}\underline{z}|\underline{e})$. Dividing by p and $1-p$ shows that the sign of this derivative is given by sign$[f_e(\bar{x}|\bar{z},\bar{e}) - f_e(\bar{x}|\underline{z},\underline{e})]$. Note then that $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\bar{z}|\underline{e}) = pf(\bar{e}) + (1-p)f(\underline{e}) \Rightarrow f_e(\bar{x}|\bar{z},\bar{e}) \neq f_e(\bar{x}|\underline{z},\underline{e})$ Hence, if we can find $(\bar{e},\underline{e})$, satisfying $e = p\bar{e} + (1-p)\underline{e}$, such that $f_e(\bar{x}|\bar{z},\bar{e}) > f_e(\bar{x}|\underline{z},\underline{e})$, then the ex ante information scenario does strictly better.

To see the necessary part suppose to the contrary that not, i.e. profits are strictly higher under the ex ante information scenario, but $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\bar{z}|\underline{e}) = pf(\bar{e}) + (1-p)f(\underline{e})$. But then by Jensen's Inequality, we have that the profit function for the ex ante information scenario $pf(\bar{e}) + (1-p)f(\underline{e}) - pc(\bar{e}) + (1-p)c(\underline{e})$ lies below the one for the ex post information scenario $f(e) - c(e)$. Hence, profits cannot be strictly higher. ■

Proof Lemma 1.

Before we can solve the principal's problem we have to check that we can indeed omit the participation constraint and that the incentive constraint yields a global maximum of the agent's problem for the equilibrium wage scheme.

Omit the Participation Constraint:

This constraint is given by:

$$f(\bar{x}\bar{z}|\hat{e})w(\bar{x}\bar{z}) + f(\bar{x}\underline{z}|\hat{e})w(\bar{x}\underline{z}) + f(\underline{x}\bar{z}|\hat{e})w(\underline{x}\bar{z}) + f(\underline{x}\underline{z}|\hat{e})w(\underline{x}\underline{z}) - c(\hat{e}) \geq 0.$$

Since the agent's reservation utility is zero, the participation constraint is automatically satisfied if the limited liability and the incentive constraint are: by choosing an effort level of zero, the agent can achieve at least an expected utility of zero as wages are larger than zero by the limited liability constraint. Hence, by maximizing his expected utility over effort (the incentive constraint) he can never do worse then zero.

Incentive Constraint Yields Unique Global Maximum:

Concerning the incentive constraint note that a global maximum of the agent's problem exists, since the second order conditions are satisfied for all wage schemes, satisfying $w(\bar{x}z) \geq w(\underline{x}z)$ (which we show below holds in equilibrium): the cost function is strictly convex and the probability function concave for all output combinations. For out of equilibrium wages $w(\bar{x}z) < w(\underline{x}z)$ the agent provides an effort of zero.

The Principal's Problem:

We can now solve the principal's problem (as stated in the text). This is a linear optimization problem, with a convex feasible set and a continuous and concave objective function. Since for such problems any local maximum is a global maximum, the Kuhn-Tucker first order conditions are necessary and sufficient for an optimum.

Note that $w(\underline{x}\bar{z}) > 0$ (analogue $w(\underline{x}\underline{z}) > 0$) decreases incentives, compared to setting $w(\underline{x}\bar{z}) = 0$ (analogue $w(\underline{x}\underline{z}) = 0$), since $\frac{\partial f(\underline{x}z|e)}{\partial e} < 0$. This cannot be profit maximizing for the principal

since effort is smaller and the principal pays the agent more by setting $w(\underline{x}\bar{z}) > 0$ (analogue $w(\underline{x}\underline{z}) > 0$), which reduces her profit. Therefore, $w(\underline{x}\bar{z}) = 0$ (analogue $w(\underline{x}\underline{z}) = 0$).
Denoting by λ the Lagrange multiplier and by $\mathcal{L}$ the Lagrange function for the principal's problem, the first order conditions with respect to $w(\bar{x}\bar{z})$ and $w(\bar{x}\underline{z})$ are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w(\bar{x}\bar{z})} &= -f(\bar{x}\bar{z}|\hat{e}) + \lambda f_e(\bar{x}\bar{z}|\hat{e}) \leq 0, \quad w(\bar{x}\bar{z})\frac{\partial \mathcal{L}}{\partial w(\bar{x}\bar{z})} = 0, \\ \frac{\partial \mathcal{L}}{\partial w(\bar{x}\underline{z})} &= -f(\bar{x}\underline{z}|\hat{e}) + \lambda f_e(\bar{x}\underline{z}|\hat{e}) \leq 0, \quad w(\bar{x}\underline{z})\frac{\partial \mathcal{L}}{\partial w(\bar{x}\underline{z})} = 0. \end{aligned}$$

There exists a λ such that both equations hold with equality iff $l(\bar{x}\bar{z}|\hat{e}) = l(\bar{x}\underline{z}|\hat{e})$. In this case wages are determined by the incentive constraint: $[f_e(\bar{x}\bar{z}|\hat{e})w(\bar{x}\bar{z}) + f_e(\bar{x}\underline{z}|\hat{e})w(\bar{x}\underline{z})] = c_e(e)$.
So let these ratios are unequal and assume first $l(\bar{x}\bar{z}|\hat{e}) > l(\bar{x}\underline{z}|\hat{e})$. Suppose that $\frac{\partial \mathcal{L}}{\partial w(\bar{x}\bar{z})} = 0$ and hence $w(\bar{x}\bar{z}) \geq 0$. Rearranging gives us $\lambda = 1/l(\bar{x}\bar{z}|\hat{e})$. Plugging this in $\frac{\partial \mathcal{L}}{\partial w(\bar{x}\underline{z})}$ and rearranging again shows that this holds with $<$ given the assumption that $l(\bar{x}\bar{z}|\hat{e}) > l(\bar{x}\underline{z}|\hat{e})$. Hence, $w(\bar{x}\underline{z}) = 0$. From the incentive constraint it then follows that the principal sets $w(\bar{x}\bar{z}) > 0$ to implement a strictly positive effort. The case $l(\bar{x}\bar{z}|\hat{e}) < l(\bar{x}\underline{z}|\hat{e})$ follows analogue.
Summarizing, we write the incentive constraint as (assuming likelihood ratios are unequal):

$$f_e(xz|\hat{e})w(xz) = c_e(\hat{e}),$$

where $xz \in \{\bar{x}\bar{z}, \bar{x}\underline{z}\}$, depending on the likelihood ratios as described above. Hence, to implement effort levels $\hat{e}$ the wage has to satisfy:

$$w(xz) = \frac{c_e(\hat{e})}{f_e(xz|\hat{e})}. \tag{4}$$

■

Proof Equation (1).

As described in the Proof of Lemma 1 if likelihood ratios are equal, wages for a given effort vector e are determined by:

$$f_e(\bar{x}\bar{z}|e)w(\bar{x}\bar{z}) + f_e(\bar{x}\underline{z}|e)w(\bar{x}\underline{z}) = c_e(e).$$

Substituting for $f_e(\bar{x}\bar{z}|e)$ from the likelihood ratios yields:

$$f_e(\bar{x}\underline{z}|e)\left[\frac{f(\bar{x}\bar{z}|e)}{f(\bar{x}\underline{z}|e)}w(\bar{x}\bar{z}) + w(\bar{x}\underline{z})\right] = c_e(e).$$

Rearranging gives us the expected wage payment:

$$C(\bar{x}|e) = f(\bar{x}\bar{z}|e)w(\bar{x}\bar{z}) + f(\bar{x}\underline{z}|e)w(\bar{x}\underline{z}) = \frac{f(\bar{x}\underline{z}|e)}{f_e(\bar{x}\underline{z}|e)}c_e(e) = \frac{f(\bar{x}|e)}{f_e(\bar{x}|e)}c_e(e).$$

The last equality follows as $l(\bar{x}\bar{z}|\hat{e}) = l(\bar{x}\underline{z}|\hat{e}) \Leftrightarrow f_e(\bar{x}|e)f(xz|e) = f(e)f_e(xz|e)$. Also $l(\bar{x}\bar{z}|\hat{e}) = l(\bar{x}|\hat{e}) \leftrightarrow f_e(\bar{x}|e)f(xz|e) = f(e)f_e(xz|e)$.
For $l(\bar{x}\bar{z}|\hat{e}) \neq l(\bar{x}\underline{z}|\hat{e})$ to calculate the expected wage, $C(xz|e)$, multiply $w(xz)$ (see Equation

(4)) by $f(xz|e)$. ■

Proof Equation (2).

To obtain Equation (2) multiply $w(\bar{x}\bar{z})$ by $f(\bar{x}\bar{z}|e)$, which gives $f(e)\frac{f(\bar{x}\bar{z}|e)}{p}$ and $w(\bar{x}\underline{z})$ by $f(\bar{x}\bar{z}|e)$, which gives $f(e)\frac{f(\bar{x}\underline{z}|e)}{1-p}$. Adding up yields:

$$C^I(\bar{x}\bar{z}, \bar{x}\underline{z}|\underline{e}, \bar{e}) = p\,\frac{c_e(\bar{e})}{l(\bar{x}\bar{z}|\bar{e})} + (1-p)\,\frac{c_e(\underline{e})}{l(\bar{x}\underline{z}|\underline{e})}$$

Then use the definition of $C(s|e) = \frac{c_e(e)}{f_e(s|e)}c_e(e)$. ■

Ex Ante versus Ex Ante Information.

What differs under these two informational scenarios is the agent's participation constraint(s):

$$\sum_x \sum_z f(xz|e(z))w(xz) - p(z)c(e(z)) \geq 0,$$

versus

$$\sum_x f(x\bar{z}|e(\bar{z}))w(x\bar{z}) - p\,c(e(\bar{z})) \geq 0,$$

$$\sum_x f(xz|e(\underline{z}))w(x\underline{z}) - (1-p)\,c(e(\underline{z})) \geq 0.$$

Using the two incentive constraint we see that the agent receives in every of these two states a strictly positive rent. Hence, they are satisfied. The ex ante participation constraint (for the ex ante information scenario) then also holds. ■

Proof Proposition 2.

Suppose that likelihood ratios are equal, i.e. $l(\bar{x}\bar{z}|e) = l(\bar{x}\underline{z}|e) = l(\bar{x}|e)$. Using this in the previous equation in Equation (2), we obtain for the expected wage of the ex ante information scenario (where $f \in \{f(\bar{x}\bar{z}|\cdot), f(\bar{x}\underline{z}|\cdot)\}$):

$$E_z\left[\frac{c_e(e(z))}{l(\bar{x}|e(z))}\right].$$

For the ex post information scenario the wage is:

$$\left[\frac{c_e(E_z e(z))}{l(\bar{x}|E_z e(z))})\right].$$

Defining $g(e) = \frac{c_e(e)}{l(\bar{x}|e)}$ and applying Jensen's Inequality shows that the wage under the ex ante information scenario is larger iff $E_z g(e(z)) \geq g(E_z e(z))$, i.e. iff $g(e)$ is convex.

Suppose next that likelihood ratios are unequal. By the same argument as used before:

$$p\,\frac{c_e(\bar{e})}{l(\bar{x}\bar{z}|\bar{e})} + (1-p)\,\frac{c_e(\underline{e})}{l(\bar{x}\bar{z}|\underline{e})} \geq \frac{f(\bar{x}\bar{z}|e)}{f_e(\bar{x}\bar{z}|e)}c_e(e),$$

but then for $l(\bar{x}\bar{z}|e) > l(\bar{x}\underline{z}|e)$ (the other case is analogue), it follows that:

$$\begin{aligned} C^I(\bar{x}\bar{z}, \bar{x}\underline{z}|\underline{e}, \bar{e}) &= p\,\frac{c_e(\bar{e})}{l(\bar{x}\bar{z}|\bar{e})} + (1-p)\,\frac{c_e(\underline{e})}{l(\bar{x}\underline{z}|\underline{e})} \\ &> p\,\frac{c_e(\bar{e})}{l(\bar{x}\bar{z}|\bar{e})} + (1-p)\,\frac{c_e(\underline{e})}{l(\bar{x}\bar{z}|\underline{e})} \\ &\geq \frac{c_e(e)}{l(\bar{x}\bar{z}|e)}. \end{aligned}$$

■

Proof Proposition 3.

The principal's expected profit if the agent receives information ex post:

$$\begin{aligned}\Pi(e) &= f(e)\bar{x} - C(\bar{x}|e) \\ &= \{p\, f(\bar{x}|\bar{z},\bar{e}) + (1-p)\, f(\bar{x}|\underline{z},\underline{e})\}\,\bar{x} - \{p\, C(\bar{x}|\bar{e}) + (1-p)\, C(\bar{x}|\underline{e})\}\,,\end{aligned}$$

with $\bar{e} = \underline{e}$. This is just the profit function of the ex ante information scenario, where we, however, do not require $\bar{e} = \underline{e}$. Hence, the maximization problem of the principal is identical for both structures up to this restriction. Maximizing without it cannot make the principal worse off.

The proof that ex ante information can do strictly better iff there are gains from tailoring of effort follows then exactly the same line as the Proof of Proposition 1 ■

Proof Proposition 4.

We first argue that for any $(\bar{e},\underline{e})$ of the ex ante information scenario we can implement the same expected effort for the ex post information scenario yielding to strictly higher profits.

For the implementation cost part remember from Proposition 2 that those are strictly higher for the same expected effort under an informative signal. Furthermore, expected revenues are lower for the ex ante information scenario, since: 1) we excluded gains from tailoring of effort, i.e. $f(\bar{x}\bar{z}|\bar{e}) + f(\bar{x}\underline{z}|\underline{e}) = p\, f(\bar{e}) + (1-p) f(\underline{e})$ and 2) $f(e)$ is concave in e.

Taking implementation costs and expected revenues together we see thus that for any $(\bar{e},\underline{e}) \neq 0$ (which is optimal for $f_e(\bar{x}z|e) > 0$) for the ex ante information scenario we can implement the same expected profit for the ex post scenario leading to strictly higher profits. This effort need not to be optimal. But the principal cannot do worse with the optimal effort. Hence, if the signal is informative and there are no gains from tailoring effort, ex ante information does strictly worse than ex post information. ■

References

AMAYA, K. (2005): "Rational Ignorance in Moral Hazard Problems," mimeo, Kobe University.

BLACKWELL, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

CHE, Y.-K., AND S.-W. YOO (2001): "Optimal Incentives for Teams," *American Economic Review*, 91(3), 525–541.

CREMER, J. (1995): "Arm's Length Relationships," *The Quarterly Journal of Economics*, 110(2), 275–95.

DEMOUGIN, D., AND C. FLUET (1998): "Mechanism Sufficient Statistic in the Risk-Neutral Agency Problem," *Journal of Institutional and Theoretical Economics*, 127(4), 622–639.

DEWATRIPONT, M., I. JEWITT, AND J. TIROLE (1999): "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, 66(1), 183–98.

EDERER, F. P. (2004): "Feedback and Motivation in Dynamic Tournaments," mimeo, MIT.

GJESDAL, F. (1982): "Information and Incentives: The Agency Information Problem," *Review of Economic Studies*, 49(3), 373–90.

GROSSMAN, S. J., AND O. D. 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(1), 74–91.

——— (1982): "Moral Hazard in Teams," *Bell Journal of Economics*, 13(2), 324–340.

JEWITT, I. (1997): "Information and Principal-Agent Problems," mimeo, University of Bristol.

KIM, S. K. (1995): "Efficiency of an Information System in an Agency Model," *Econometrica*, 63(1), 89–102.

LEHMANN, E. L. (1988): "Comparing Location Experiments," *The Annals of Statistics*, 16, 521–523.

Lizzeri, A., M. A. Meyer, and N. Persico (2002): "The Incentive Effects of Interim Performance Evaluations," Penn CARESS Working Paper, University of Pennsylvania.

Meyer, M. A., and J. Vickers (1997): "Performance Comparisons and Dynamic Incentives," *Journal of Political Economy*, 105(3), 547–81.

Mookherjee, D. (1984): "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies*, 51(3), 433–46.

Nafziger, J. (2007): "Moral Hazard and Intermediate Observability," mimeo, University of Bonn.

Nelsen, R. B. (1995): "Copulas, Characterization, Correlation, and Counterexamples," *Mathematics Magazine*, 68(3), 193–198.

Prat, A. (2005): "The Wrong Kind of Transparency," *American Economic Review*, 95(3), 862–877.

Schmitz, P. (2005): "Allocating Control in Agency Problems with Limited Liability and Sequential Hidden Actions," *RAND Journal of Economics*, 36(2), 318–336.