# CESifo Working Papers

# Private Monitoring, Collusion and the Timing of Information

Fahad Khalil
Jacques Lawarree
Troy J. Scott

**CESifo**
**Center for Economic Studies & Ifo Institute**

# Private Monitoring, Collusion and the Timing of Information

## Abstract

When a principal's monitoring information is private (non-verifiable), the agent should be concerned that the principal could misrepresent the information to reduce the agent's wage or collect a monetary penalty. Restoring credibility may lead to an extreme waste of resources—the so-called burning of money. A more realistic and efficient outcome is feasible when the private information arrives in time to rescale the agent's effort. Rescaling is more effective than pure monetary penalties because effort has different values to different parties while money is equally valuable to all parties. Furthermore, when rescaling is feasible, private monitoring is more efficient than public monitoring subject to collusion because non-monetary penalties are ineffective to deter collusion.

*Fahad Khalil*
*Department of Economics*
*University of Washington / Seattle, WA / USA*
*khalil@uw.edu*

*Jacques Lawarree*
*Department of Economics / University of*
*Washington / Seattle, WA / USA*
*lawarree@uw.edu*

*Troy J. Scott*
*RTI International*
*tjscott@uw.edu*

## 1. Introduction

Contracting parties can obtain monitoring information at various times during an agency relationship. In this article, we focus on monitoring information learned privately by the principal, i.e., private signals, and we show the importance of the timing of information. Private signals cannot be verified by a third party. For example, the signal could be the principal's subjective evaluation of the project's difficulty or of an agent's performance, a casual conversation with a customer or a co-worker about the agent's behavior, or the filtering of the agent's e-mail. Another example is the information that internal and external auditors may share with audit committees in executive sessions with details that go beyond what is documented in publicly released audit reports. In military procurement, national security concerns may prevent the Department of Defense from releasing information publicly. In these examples, the private information may arrive ex post, i.e., after the agent has performed his task, or ex ante, i.e., before or during the performance of tasks. The aim of this article is to investigate to what extent the principal may use such private signals to improve the agent's incentives and to highlight the importance of the timing of information.

When a signal is privately observed by the principal, it provides an opportunity for the principal to abuse the private information, and the principal's credibility becomes an issue. For instance, an employee should be concerned if the employer can reduce his wage or impose a penalty based on an unverifiable private signal. To resolve this concern, we consider two options. The principal can reassure the agent either by removing her own incentive to abuse private information, or by relying instead on third party monitoring.

The main part of our analysis discusses the first option, where a new constraint is needed to ensure that the principal cannot gain from a penalty based on a private signal. One way to satisfy this constraint is to force the principal to give away, for instance to a charity, any collected penalty—the so-called 'burning money' solution.[1] Although this solution is quite intriguing, it is also a little disconcerting as it involves wasting resources with respect to the contracting parties.

In this article, we argue that organizations will look for more efficient solutions than burning money to address the credibility issue with private signals. In particular, the principal may have additional alternatives depending on the timing of private information. If she obtains private information ex ante, then the agent's required effort, i.e., the scale of the project itself, can be reduced as an alternative to burning money.[2] We show that effort can be rescaled so that it has two effects: it penalizes the agent but also ensures that the principal, at that stage, does not gain from penalizing the agent. We also show that relying on non-monetary penalties like rescaling is more effective than pure monetary penalties, which require money burning. The reason is that effort has different values to different parties, whereas money is equally valuable to all parties.

To present this intuition, we use an adverse-selection model where a principal observes the output and can use an imperfect private signal of the agent's type to induce effort. The signal allows the principal to penalize the shirking agent but, because it is imperfect, complying agents are also penalized by mistake. If the penalty has to be burned, such mistakes are costly, and the signal is less valuable. In the case of an ex post private signal, we show that the signal is not valuable unless it is accurate enough. In the case of an ex ante private signal, we show that the

signal is valuable even if its accuracy is very low. The reason is that the principal can rely on rescaling rather than being limited to burning money. If rescaling allows the principal to extract all rent for low accuracy, then there is no money burning in equilibrium. Otherwise, burning money can be used to extract the remaining rent.

Because rescaling cannot take place after the project has already been completed, the timing of information becomes critical: we find that the principal always prefers to receive private information early, i.e., ex ante. In other words, the principal may be willing to incur an extra cost to receive private signals ex ante rather than ex post.

A second option for the principal to avoid her credibility problem with private monitoring is to rely instead on *public* monitoring by a third party (for instance, a public accounting firm). If third-party monitoring was free, this solution would always be preferred. As is well-known from the literature, the first best can be reached irrespective of the timing of the signal.[3]

However, in a world where information can be manipulated, we consider an endogenous cost of public monitoring due to the potential for collusion between the monitor (e.g., an auditor) and the agent. We contrast the two options faced by a principal in the presence of manipulable information: her own credibility if the signal is privately learned by her, versus the threat of collusion if she uses a public monitor instead. First, we find that rescaling effort or, more generally, non-monetary penalties, can be ineffective to deter collusion as the monitor only values monetary bribes. The reason is that the monitor does not value the output in itself, and therefore, a monetary transfer is the only instrument that can impact both the agent and the monitor. Second, we find unexpected similarities between public and private monitoring. The

standard remedy to prevent bribery – rewarding the public monitor for reporting unfavorable information about the agent[4] – has a similar effect to burning money.  The principal does not collect any penalty from the agent because it is transferred to a third party: the monitor.  Because of collusion, the signal is less effective and has to be accurate enough before it is useful.  It is not clear a priori whether the principal should prefer private monitoring suffering from a lack of credibility or public monitoring vulnerable to collusion.

We show that each manipulation issue – lack of credibility and collusion – leads to the same outcome in the ex post signal case.  In the ex ante signal case, however, the credibility issue is less severe, and private monitoring is preferred to public monitoring vulnerable to collusion.  Once again, the timing of the signal is important.  Only when it is difficult for the agent and the monitor to collude (for instance due to large collusion costs between them) that public monitoring may be preferred to private monitoring, and in that case, the timing of the signal is important: any collusion costs would render public monitoring attractive with ex post information, whereas collusion costs would have to exceed some threshold before public monitoring would become attractive with ex ante information.

This article is related to the literature on subjective evaluations, which builds upon the theory of repeated games with asymmetric information.[5]  Because this literature emphasizes moral hazard (or hidden action), the principal's private information pertains directly to the agent's action and arrives *after* that action has been taken.  MacLeod (2003) shows that it is impossible to induce effort under a single-period contract that is budget-balancing, hence the need to burn money.  Levin (2003) considers an infinitely repeated game where, instead of

burning money, the principal fires the agent when her privately observed measure of performance falls below a threshold. Also, in a dynamic environment, Fuchs (2007) characterizes equilibria in which the principal delays revealing to the agent what she has learned privately.[6]

Strausz (2006) and Dequiedt and Martimort (2014) consider adverse selection models and show the role of private information in restoring the continuity between correlated and uncorrelated environments. In these models, as in ours, the principal is able to use manipulable signals to extract rent. Strausz studies the impact of limited commitment due to private information in a multi-period game with private information arriving between periods. However, he assumes that the budget is in balance (burning money is not considered) and that collusion is not an issue. Dequiedt and Martimort (2014) study bilateral contracting in a model with multiple agents and correlated types instead of one agent and a supervisor. They do not consider the timing of information or burning money. In a sense, our result for the sub-optimality of burning money provides a rationale for their analysis. Also, in their model, if the agents can collude, multi-dimensional screening issues appear and collusion removes the value of correlated private information.

## 2. The model

We consider a principal (she) - agent (he) model with adverse selection. The agent exerts an effort $e$ which, together with a productivity parameter $\theta$, determines output $x = \theta e$. The parameter $\theta$ takes one of two positive values, with $\theta_1 < \theta_2$, where the low type $\theta_1$ obtains with probability $\xi$, and the high type $\theta_2$ obtains with probability $1 - \xi$. We normalize $\theta_2 = 1$ to simplify the exposition.[7] The agent's disutility of effort is given by the function $\psi(e) = e^2/2$.

The output belongs to the principal, who compensates the agent with a transfer $t$. The principal maximizes the output net of transfers: $\pi = x - t$. The agent maximizes his utility $U = t - \psi(e)$, and we assume that the agent's reservation utility is zero.

The first-best contract equates the marginal cost and marginal product of the agent's effort ($e_i^{FB} = \theta_i, i = 1,2$) and pays the agent just enough to compensate for his disutility of effort $\left(t_i^{FB} = \theta_i^2/2, i = 1,2\right)$. The first best cannot be implemented if $\theta$ and $e$ are private information of the agent. The principal must then offer a contract based on the observable output rather than the unobservable effort. As is well known (see, Baron-Myerson (1982) for example), the principal must solve the following problem: Max $\xi(\theta_1 e_1 - t_1) + (1 - \xi)(e_2 - t_2)$ subject to the participation constraint of the low type, $t_1 - \frac{(e_1)^2}{2} \geq 0$, and the incentive compatibility constraint of the high type, $t_2 - \frac{(e_2)^2}{2} \geq t_1 - \frac{(\theta_1 e_1)^2}{2}$.[8] The optimal contract is the standard second-best contract. We refer to it as the *no-monitoring* contract and use the superscript *NM*. It distorts the effort of the low type $\left(e_1^{NM} = \xi\theta_1/\left\{\xi + (1 - \xi)(1 - \theta_1^2)\right\} < \theta_1 = e_1^{FB}\right)$ to reduce the rent of the high type. The high type produces the first best level of effort $(e_2^{NM} = 1 = e_2^{FB})$ and the low type receives no rent.

So far, we have a standard adverse selection model based on observable output, but we now introduce a new feature in the model. During the course of the principal-agent relationship, the principal receives additional private information, which could enhance efficiency. For example, improved cost estimates about the project would allow better planning of the output. Here we focus on additional *private* information, and later in section 3, we will examine *public* monitoring by a third party who can be bribed.

This additional information is learned privately by the principal and is therefore not verifiable in a court of law.[9] The principal would certainly want to use the private information but the agent should worry about being 'held up' if the principal can alter the terms of the contract based on unverifiable information.[10] In such a case, one can ask whether a contract should be conditioned on this signal.

We use a standard way to model the principal's signal, denoted by $\sigma$: $\sigma \in \{\sigma_1, \sigma_2\}$ and $\text{Prob}(\sigma_1|\theta_1) = \text{Prob}(\sigma_2|\theta_2) = v$, and $\text{Prob}(\sigma_1|\theta_2) = \text{Prob}(\sigma_2|\theta_1) = 1 - v$, with $v > \frac{1}{2}$. That is, the signal is correct with probability $v$ and wrong with probability $(1 - v)$. The signal is obtained at zero cost. Because the agent will have an incentive to misreport his type only when he is the high type, the signal is relevant only after the agent has claimed to be $\theta_1$.[11]

Because the signal is private information of the principal, the contract can only be based on the principal's public assertion about the signal $\sigma$. We assume that the principal makes such a public announcement about the signal, and we label it $\hat{\sigma}$, where $\hat{\sigma} \in \{\hat{\sigma}_1, \hat{\sigma}_2\}$.

The timing of the signal and its announcement are important, and it matters whether the signal arrives before or after the agent makes his effort. When the signal arrives before the agent makes his effort, the principal can potentially use the contract to rescale the agent's optimal effort based on the report. We call it the *ex ante monitoring* case. When the signal arrives after the effort has already been taken, we call this the *ex post monitoring* case. With ex post *monitoring*, the only option for the principal is to use the report of the signal to affect the transfer to the agent, for instance by imposing a penalty on the agent.

Our main focus will be on the ex ante case but we will start by discussing briefly the ex post case to introduce the concept of burning money in the context of our model. We also derive a result on the accuracy of the signal that is used in the rest of the article.

*Private monitoring with an ex post signal: the optimality of burning money*

In the ex post case, the principal receives a private signal about the agent's type after the output has been produced. This case has been analyzed in the context of moral hazard by MacLeod (2003) and Fuchs (2007), and they show the need to burn money to make private signals useful. In this section, we extend their result to an adverse selection model, and we also highlight how the loss of the penalty by way of burning money implies that private monitoring can only be used if it is accurate enough. We summarize the timing as follows:

(1)     The agent observes $\theta$.

(2)     The principal offers the agent a contract; the agent rejects the contract, ending the game, or accepts the contract and reports $\hat{\theta}$.

(3)     The agent exerts effort, and output is publicly observed.

(4)     The principal observes the signal $\sigma$ and reports $\hat{\sigma}$.

(5)     The transfers are realized.

The optimal contract is based on the agent's report $\hat{\theta}$ and the principal's report $\hat{\sigma}$. It specifies efforts $e_1^p$ and $e_2^p$ based on the agent's report and transfers based on both the agent's (first subscript) and the principal's (second subscript) reports: $t_{11}^p$, $t_{12}^p$. The superscript "$p$" reminds the reader that this signal comes ex post. Later we will use the superscript "$a$" for the

ex ante case.  Because the signal is not used after the agent reports high type, the corresponding

transfer has only one subscript: $t_2^p$.

The principal would like to use the private signal to penalize the agent for under-reporting

his type and reduce his information rent.  At the same time, the agent needs to be assured that

the principal will not abuse the private signal to penalize the agent.  For example, suppose $t_{11}^p >$

$t_{12}^p$.  The principal would always report the private signal that leads him to pay the lower transfer

$t_{12}^p$.  To induce the principal to tell the truth, the contract must satisfy two "Principal's Incentive

Compatibility" (PIC) constraints, one for each signal.[12]  As known from the literature on moral

hazard, these constraints cannot be satisfied by a 'balanced budget', and it is necessary to

introduce a gap between what the principal pays and what the agent receives.  We represent this

gap by the variables $B_{11}^p$ and $B_{12}^p$.  When the principal pays $t_{11}^p$ (resp. $t_{12}^p$) to the agent, she also

burns $B_{11}^p$ (resp. $B_{12}^p$). [13]  The first constraint guarantees that the principal will not report state 2

($\hat{\sigma}_2$) when the signal is $\sigma_1$ and the second guarantees that the principal will not report state 1

($\hat{\sigma}_1$) when the signal is $\sigma_2$.  Therefore, the two (PIC) constraints for ex post monitoring are

$$-\left(t_{11}^p + B_{11}^p\right) \geq -\left(t_{12}^p + B_{12}^p\right)$$

$$-\left(t_{12}^p + B_{12}^p\right) \geq -\left(t_{11}^p + B_{11}^p\right),$$

and together they imply:

(PIC^p)        $t_{11}^p - t_{12}^p = B_{12}^p - B_{11}^p,$

which means that the principal cannot gain ex post from penalizing the agent.  Note that even

though the (PIC) requires the principal to be indifferent between reporting publicly $\hat{\sigma}_1$ or $\hat{\sigma}_2$, that

does not imply that the principal cannot benefit from private monitoring: it means that any benefit must come from its impact on the initial contract.

Some intuition about the impact of (PIC$^p$) on the principal's problem can be presented by looking at the relevant participation and incentive compatibility constraints:[14]

$(IR^p\text{-}1)$    $vt_{11}^p + (1-v)t_{12}^p - \frac{1}{2}(e_1^p)^2 \geq 0$

$(IR^p\text{-}2)$    $t_2^p - \frac{1}{2}(e_2^p)^2 \geq 0$

$(IC^p)$    $t_2^p - \frac{1}{2}(e_2^p)^2 \geq (1-v)t_{11}^p + vt_{12}^p - \frac{1}{2}(\theta_1 e_1^p)^2$

As usual, $(IR^p\text{-}1)$ and $(IC^p)$ will be binding. Using the binding $(IR^p\text{-}1)$ and re-arranging the terms, the high-type's rent, given by the right-hand-side of $(IC^p)$, can be re-written as:

$$u_2^p = \left[\frac{1}{2}(e_1^p)^2 - \frac{1}{2}(\theta_1 e_1^p)^2\right] - (2v-1)(t_{11}^p - t_{12}^p)$$

The first terms in the brackets represent the cost differential between a truthful low type and a shirking high type. The second term represents a penalty effect created by the difference in the agent's transfers $(t_{11}^p - t_{12}^p)$; the greater the difference in transfers, the lower the information rent.

We note first that $(PIC^p)$ puts restrictions on the penalty by determining how much money must be burned ($B_{11}^p$ and $B_{12}^p$). It is proven in appendix A that $B_{11}^p$ is zero. Intuitively, a positive $B_{11}^p$ would make it harder to satisfy $(PIC^p)$ by decreasing the penalty.[15] Thus, we can rewrite the rent expression as:

$(Rent^p)$  $\qquad\qquad u_2^p = \left[\frac{1}{2}(e_1^p)^2 - \frac{1}{2}(\theta_1 e_1^p)^2\right] - (2v-1)B_{12}^p.$

Although increasing $B_{12}^p$ reduces rent by increasing the penalty, it is also costly: in equilibrium, the low-type is penalized by mistake (unless $v=1$), and this cost is borne by the principal because the agent's participation constraint $(IR^p$-1) binds.[16] Thus the possibility of reducing the rent of a cheating high type has to be balanced against the cost of erroneous penalties on the truthful low type.

**Proposition 1**:

> *Ex post private monitoring is useful only if it is accurate enough, i.e., when $v >$*
>
> $\frac{1}{2-\xi}$ ; *money is burned and the agent earns zero rent. The no-monitoring contract is*
>
> *optimal when $v \leq \frac{1}{2-\xi}$.*

**Proof**: See appendix A.

Because the penalty $(t_{11}^p - t_{12}^p)$ is burned and therefore lost to the principal, there is a cost of increasing the penalty -- the truthful low-type is penalized by mistake. Without $(PIC^p)$ this cost would not be there, because the principal could then use the collected penalty to compensate the risk-neutral agent for the erroneous penalties. It is useful to understand that the critical level of accuracy is determined by two first-order effects. A unit increase in the penalty reduces the information rent, which is paid with probability $1-\xi$, by an amount $2v-1$ and requires a unit of money to be burnt with probability $\xi(1-v)$. Therefore it is optimal to

burn money to eliminate the rent if $(1 - \xi)(2\nu - 1) > \xi(1 - \nu)$, which can be rewritten as $\nu > 1/(2 - \xi)$.

The ex post monitoring case has similarities with a pure moral hazard model because only the transfers can be modified based on the signal. However, in the moral hazard models used in the literature on subjective evaluations (Levin (2003, part IV), Fuchs (2007) and MacLeod (2003)), the output is not publicly observed. The principal receives a signal about the output, which the agent does not see. In our model, when the agent chooses his effort, he knows what the publicly observable output will be but not the exogenous signal $\sigma_i$ about the agent's type, which is privately observed by the principal. Because output can be contracted upon, there is a feasible contract even absent any signal, and the principal will only use the signal if it is accurate enough.

Rent can be fully extracted as soon as the signal is useful, and although the distortion in effort can be reduced, effort is less than the first best unless $\nu = 1$. Therefore, the classic separation result of Baron and Besanko (1984), which says that rent is extracted before effort is restored towards the first best, can be extended to the case of ex post private monitoring.

Next, we analyze ex ante private monitoring and show that burning money is no longer optimal.

*Private monitoring with an ex ante signal*

*Preliminaries*

In this section, we consider the case where the principal receives a private signal before the agent takes his action, i.e., we flip points 3 and 4 in the timing of the ex post case above. The main

difference with respect to the ex post monitoring case is that the agent's effort may now depend on both the agent's (first subscript) and the principal's (second subscript) reports: $e_{11}^a$, $e_{12}^a$. The superscript "$a$" denotes the ex ante case. The individual rationality and incentive compatibility constraints now reflect this dependence of effort on the principal's report (the participation constraint of the high type is unchanged):

($IR^a$-1) $\qquad v\left(t_{11}^a - \frac{1}{2}(e_{11}^a)^2\right) + (1-v)\left(t_{12}^a - \frac{1}{2}(e_{12}^a)^2\right) \geq 0$

($IR^a$-2) $\qquad t_2^a - \frac{1}{2}(e_2^a)^2 \geq 0,$

($IC^a$) $\qquad t_2^a - \frac{1}{2}(e_2^a)^2 \geq (1-v)\left(t_{11}^a - \frac{1}{2}(\theta_1 e_{11}^a)^2\right) + v\left(t_{12}^a - \frac{1}{2}(\theta_1 e_{12}^a)^2\right).$

The principal's incentive compatibility constraints also now reflect the dependence of effort on the principal's public report:[17]

$$\theta_1 e_{11}^a - t_{11}^a - B_{11}^a \geq \theta_1 e_{12}^a - t_{12}^a - B_{12}^a$$

$$\theta_1 e_{12}^a - B_{12}^a - t_{12}^a \geq \theta_1 e_{11}^a - t_{11}^a - B_{11}^a.$$

These two conditions can be satisfied simultaneously only if $\theta_1 e_{11}^a - t_{11}^a - B_{11}^a = \theta_1 e_{12}^a - t_{12}^a - B_{12}^a$, which can be re-written in the more intuitive form,

($PIC^a$) $\qquad t_{11}^a - t_{12}^a = B_{12}^a - B_{11}^a + \theta_1(e_{11}^a - e_{12}^a)$

The principal wants to set $t_{11}^a > t_{12}^a$ to penalize the high-type agent who under-reports. To be credible, the principal must then either burn more money ($B_{12}^a > B_{11}^a$) or scale down effort ($e_{12}^a < e_{11}^a$) after paying the lower transfer. The choice between these two instruments to satisfy

($PIC^a$) captures the main departure from the ex post case. A key question for this section is whether the principal will rely on rescaling efforts or burning money.

*The optimal contract under ex ante private monitoring: the sub-optimality of burning money*

Having characterized the feasible set of contracts, we are ready to state the principal's problem in the ex ante private monitoring case. First, like in the ex post case, we can anticipate that $B_{11}^a = 0$, and we prove it formally in appendix A. The principal chooses the contract $\{e_{11}^a, e_{12}^a, e_2^a, t_{11}^a, t_{12}^a, t_2^a, B_{12}^a\}$ to solve the following problem:

(P)   maximize $\xi[v(\theta_1 e_{11}^a - t_{11}^a) + (1-v)(\theta_1 e_{12}^a - t_{12}^a - B_{12}^a)] + (1-\xi)(e_2^a - t_2^a)$

subject to (IC$^a$), (IR$^a$-1), (IR$^a$-2), and (PIC$^a$).

As usual, (IR$^a$-1) and (IC$^a$) will be binding. We can rewrite the high-type's rent, given by the right-hand-side of (IC$^a$), by using the binding (IR$^a$-1) and by replacing the 'penalty' $(t_{11}^a - t_{12}^a)$ using (PIC$^a$) as in the ex post case:

(Rent$^a$)    $u_2^a = \left[ v\frac{1}{2}(e_{11}^a)^2 + (1-v)\frac{1}{2}(e_{12}^a)^2 \right] - \left[ (1-v)\frac{1}{2}(\theta_1 e_{11}^a)^2 + v\frac{1}{2}(\theta_1 e_{12}^a)^2 \right]$

$$-(2v-1)[(e_{11}^a - e_{12}^a)\,\theta_1 + B_{12}^a]$$

The first two terms in brackets represent the cost differential between a truthful low type and a shirking high type. The last term differs from the ex post case because the (PIC$^a$) in the ex ante case also includes the difference in outputs, $(e_{11}^a - e_{12}^a)\,\theta_1$, besides $B_{12}^a$. Although the principal could rely on $B_{12}$ as in the ex post case, it is unclear how she might use the effort terms.

Requiring a lower effort when paying a lower transfer $(e_{11}^a > e_{12}^a)$ affects the cost differentials, and the net impact on the rent has to be examined.[18]

We present the main results in Proposition 2, and we explain them below.

**Proposition 2:**

    i.       *Ex ante private monitoring is useful for all $v$. The agent's rent decreases with $v$,*

          *and it becomes zero if the signal is accurate enough $(v > \tilde{v}$, where $\tilde{v} < 1/(2 - \xi))$.*

    ii.      *After a private signal contradicts the agent's claim, the project is rescaled $(e_{11}^a >$*

          $e_{12}^a)$ *and the low type agent is penalized: $t_{11}^a - \frac{1}{2}e_{11}^{a\ 2} > 0 > t_{12}^a - \frac{1}{2}e_{12}^{a\ 2}$. No money is*

          *ever burned, i.e., $B_{11}^a = B_{12}^a = 0$.*

    iii.     *The principal prefers ex ante private monitoring to ex post private monitoring.*

**Proof:** See appendix A

The first remarkable result is that private monitoring is useful for the full range of the signal, i.e., for all $v > \frac{1}{2}$, whereas in the ex post case (Proposition 1), private monitoring is useful only if it is accurate enough, i.e., $v > 1/(2 - \xi)$. The second striking departure from the ex post case is that the principal does not burn money in equilibrium. In our ex ante monitoring model, the principal has a choice between rescaling the effort and burning money, and she chooses the first option exclusively.

To provide some intuition for both results, note that, after replacing the transfers using the binding constraints, the principal's problem boils down to choosing $e_{11}^a, e_{12}^a, e_2^a, B_{12}^a$ to maximize

$$\xi\left[v\left(\theta_1 e_{11}^a - \frac{(e_{11}^a)^2}{2}\right) + (1-v)\left(\theta_1 e_{12}^a - \frac{(e_{12}^a)^2}{2} - B_{12}^a\right)\right] + (1-\xi)\left[e_2^a - \frac{(e_2^a)^2}{2} - u_2^a\right]$$

where the rent $u_2^a$ is given above. Observe first that $B_{12}^a$ appears as a cost in the objective function (term multiplied by $\xi(1-v)$) but as a benefit as a reduction in the rent (term multiplied by $(1-\xi)(2v-1)$). Only when $v > \frac{1}{2-\xi}$ is the benefit higher than the cost. Money burning has a similar trade-off as in the ex post case; it can be optimal only if the signal is accurate enough.

The trade-off is different when the principal uses effort rescaling instead of burning money. To see this, observe that the benefit of effort rescaling appears in the last term in the rent with $\theta_1(e_{11}^a - e_{12}^a)$, which acts as a penalty and reduces rent. Consider first the case where the signal is uninformative, i.e., $v = \frac{1}{2}$; the optimal efforts would be $e_{11}^a = e_{12}^a = e_1^{NM}$, and the principal would obtain the no-monitoring payoff. Now consider a small increase in accuracy. If $v < \frac{1}{2-\xi}$, burning money is not optimal, but the principal can obtain a first order gain in her payoff by creating a penalty when making $e_{11}^a$ higher than $e_{12}^a$. There is also a cost of distorting the effort away from the no-monitoring level, but it is a second-order loss as the expected effort $[ve_{11}^a + (1-v)e_{12}^a]$ stays close to the effort of the no-monitoring contract $e_1^{NM}$.

What makes effort rescaling optimal is its asymmetric effect on the output, which impacts the principal, and on the cost of effort, which impacts the agent. The impact on the principal is determined by $(PIC^a)$, $t_{11}^a - t_{12}^a = \theta_1(e_{11}^a - e_{12}^a)$, and $(PIC^a)$ limits the maximum penalty to the difference in the output levels. However, due to under-production in equilibrium, a rescaling of efforts leads to a larger impact on the agent through the penalty because the difference in outputs is larger than the difference in costs:

$$\theta_1(e_{11}^a - e_{12}^a) > \frac{1}{2}[(e_{11}^a)^2 - (e_{12}^a)^2].$$

Hence the principal can use the penalty to reduce rent, which can be seen by the fact that the lying high-type faces a penalty if the signal contradicts his claim of being a low type,

$$\left(t_{11}^a - \frac{1}{2}(\theta_1 e_{11}^a)^2\right) > \left(t_{12}^a - \frac{1}{2}(\theta_1 e_{12}^a)^2\right)$$

even when the $(PIC^a)$ is satisfied.[19] Thus, rescaling is more effective than pure monetary penalties because effort has different values to different parties, whereas money is equally valuable to all parties.

Rescaling effort allows the principal to extract the entire rent for a value of $v$ lower than the one that would make burning money profitable, namely $v = 1/(2 - \xi)$. Even when the signal is more accurate than $1/(2 - \xi)$, the principal continues to rely on the gap in efforts to satisfy (PIC) rather than burn money.

Comparing with Proposition 1, we can conclude that, if it were a choice, the principal would prefer ex ante private monitoring to ex post private monitoring. Technically, the principal's problem for the ex post monitoring case is a special case of the ex ante monitoring problem. The optimal contract in the previous ex post monitoring case satisfies all constraints in the ex ante monitoring case. Moreover, because $e_{11}^a > e_{12}^a$, which was impossible in the ex post case, we can infer that ex ante monitoring is strictly preferred. This result suggests that the principal will want to look for an alternative to burning money when relying on private monitoring. She will try to induce reporting of private signals as early as possible and rescale the project size rather than be forced to burn money.

The example in Figure 1 illustrates how the optimal efforts change with the accuracy of the signal. When $v = \frac{1}{2}$, the no-monitoring contract is optimal as the signal is of no value. We have $e_{11}^a = e_{12}^a = e_1^{NM}$, and the high-type receives the no-monitoring information rent. For $v > \frac{1}{2}$, the principal uses the signal to impose a penalty $(t_{11}^a - t_{12}^a)$. Due to $(PIC^a)$, there must also be a rescaling of efforts, but no money is burned. As long as the agent's rent is not fully extracted, we show that the gap between efforts is formed by increasing $e_{11}^a$ and decreasing $e_{12}^a$, as this is mainly done to reinforce the penalty effect and reduce the rent.[21] Note that, despite the upward rescaling, $e_{11}^a$ is always less than the first best level of effort.[22]

When private monitoring is sufficiently accurate $(v > \tilde{v})$, the entire rent is extracted. Any further increase in the accuracy of monitoring (higher $v$) allows the principal to restore the effort levels toward the first best while keeping the rent at zero.[23] Both efforts can now be increased toward the first best, but to maintain the penalty, a gap between the two efforts is necessary. Thus, as $v$ approaches one, $e_{11}^a$ approaches the first-best level while $e_{12}^a$ does not. Even when relying on very precise private monitoring, the principal must rescale the project when his private information contradicts the agent's claim about his type, although the inefficient project vanishes in expectation as $v$ approaches 1.[24]

*Extensions*

Can burning money be optimal? As rescaling effort is sufficient to extract all rent before burning money becomes profitable, i.e., before $v = 1/(2 - \xi)$, no money is burned at all. This suggests that if we modify the model such that the agent retains rent even when $v = 1/(2 - \xi)$, we may

find money burning to be optimal. We show next that money burning can be optimal under ex ante monitoring only if rescaling effort is not efficient enough to extract all the rent. That is, the instruments are ranked.

We modify the model to reduce the efficacy of rescaling effort as a rent-extraction device. Consider a situation where the principal is limited in her ability to rescale the project due to a minimum scale constraint. We assume that $e_{12}^a$ cannot fall below some level $e_0$. For example, this would be the case for a construction project that requires a minimum safety or quality standard. In the online appendix, we derive the minimum value of $e_0$ such that burning money occurs in equilibrium. Basically, $e_0$ must be such that the rent is not fully extracted before $v = 1/(2 - \xi)$.[25] Note that burning money only occurs for $v \geq 1/(2 - \xi)$, and it is used in conjunction with rescaling effort. For smaller values of $v$, rescaling effort is the only instrument used.[26]

Another interesting extension would be to explore the case where the agent does not necessarily know when the principal receives the signal.[27] The principal may have an incentive to claim that she received the signal ex ante whether it is true or not. The principal's credibility problem is even more severe than before because constraints have to be added to prevent the principal from lying about *whether* and *when* she received the signal, which further limit her ability to rescale effort based on ex ante monitoring. In the online appendix we suggest a model to show the principal still benefits from the possibility of obtaining the signal ex ante.

In the next section we study third-party public monitoring as an alternative to imposing (PIC).

### 3. Private monitoring versus public monitoring by a third party monitor

*Public monitoring*

In the previous section, we showed how the principal can rely on effort adjustment to address her credibility problem (PIC) under private monitoring. In this section, the principal does not receive any private signal, and we want to contrast our analysis of section 2 with the case where the principal relies instead on third party monitoring. Because we assume that the information provided by the third party is publicly available and verifiable in court, we refer to it as public monitoring. For instance, the principal could hire an auditor from a reputable public accounting firm. However, as is known from well-publicized scandals (e.g., Arthur Andersen and Enron), even such firms are vulnerable to collusion, and manipulation of information can occur.[28] Even though the credibility problem is due to potential manipulation by the principal and the collusion problem is due to potential manipulation by the agent, there are similarities between the two problems. The principal again has two instruments to deter manipulation of information: (i) she can reward the monitor to deter bribery,[29] which has a similar effect to money burning; (ii) she can use non-monetary tools such as effort adjustment. Our main result is that, if the public monitor is potentially collusive, non-monetary penalties become ineffective. This implies that, for ex ante signals, public monitoring is worse than private monitoring if the monitor and agent can easily collude. For ex post signals, where only monetary tools are available, the principal is indifferent and may as well rely on private monitoring.

Before developing these results for public monitoring, we briefly discuss the well-known benchmark in the absence of collusion: if public signals are free, we obtain the first best as the

agent is risk neutral (e.g., Crémer and McLean (1988), Riordan and Sappington (1988)). Because the difference in the transfers is not restricted by (PIC), it is easy to find transfers satisfying $t_{11} > t_{12}$ such that the relevant (IR) and (IC) constraints are satisfied for the first-best contract, and the agent earns no rent. So, unless there are costs in obtaining third party information, the principal should always prefer public monitoring provided by a third party, and the timing of signals would be irrelevant. Note that the contract under ex ante private monitoring has more realistic features as the principal adjusts both transfers *and* effort levels with additional (private) information. In actual contracts, various features of the project itself (not just transfers) are often adjusted as new information arrives. Our results suggest that one motivation for doing so might be to satisfy the principal's incentive constraints under private monitoring.

*Collusion versus credibility*

We now clarify the setting under which collusion takes place. In order to focus on the cost due to collusion, we assume that public monitoring is otherwise free. As is typical in the monitoring literature, we assume that the monitor has the ability to present verifiable evidence to substantiate a report about his information. The contract is based on a report by the monitor, and this report is based on hard (i.e., verifiable) information. This presumes varying communication costs among the contracting parties (see Dewatripont-Tirole (2005)), but we abstract from such discussions and focus instead on collusion. We first note that the information provided by the monitor's signal cannot be completely "soft," or the monitor would be useless.[30] The information is soft if the monitor can manipulate the information by himself. Therefore, we make an assumption similar to Kofman-Lawarrée (1993) and assume that the monitor can only

manipulate the signal with the help of the agent.[31]  This restriction on manipulability makes

information hard.  Our modeling is also similar to Faure-Grimaud, Laffont and Martimort (2000,

section 4), who interpret it as the monitor being able to suppress 'good news' but not able to

create 'bad news'.  In Tirole (1986), the restriction on manipulability takes the form of only

suppressing/ignoring information, but not being able to create evidence.  As in our model, this

rules out the threat of extortion or blackmail by the monitor.[32]

We also assume that collusion between the agent and monitor occurs after the signal has

been observed by the monitor but before he makes his report to the principal.[33]  In addition, we

assume that both the agent and the monitor observe the signal, and therefore the two parties

negotiate the bribe under symmetric information.[34]  As is standard in this type of model, the side-

contract is assumed to be enforceable (Tirole (1986, 1992)).

We begin by first considering the effect of collusion on ex post public monitoring.

Referring to the earlier timing of section 2 with an ex post signal, step 4 is now modified: the

monitor observes the signal, negotiates a side-contract with the agent, and then makes a report

to the principal.  If the agent reports $\theta_1$ and the monitor reports $\sigma_2$, the agent stands to lose

$\left(t_{11}^p - t_{12}^p\right)$, which is the stake of collusion.  To deter collusion, the optimal contract must satisfy

a coalition incentive compatibility (CIC) constraint, which says that the principal has to pay a

reward to the monitor, denoted by $W^p$, that is at least as high as the stake of collusion.  Thus,

the (CIC[p]) constraint can be written as:

(CIC[p])        $W^p \geq t_{11}^p - t_{12}^p.$

The reward is paid to the monitor whenever $t_{12}^p$ is paid to the agent. The principal's problem is given in appendix B. Because the (CIC$^p$) is binding in equilibrium, imposing this constraint effectively transfers the penalty $\left(t_{11}^p - t_{12}^p\right)$ to the monitor who acts as a bounty-hunter.

Although this result is standard in the collusion literature, it is remarkable that it is identical to the result in the case of ex post *private* monitoring where the principal had to burn the penalty (or give it away to a charity) to address her credibility problem (satisfy the (PIC$^p$) constraint). With collusion, the penalty is instead given away to the monitor. Thus, for an ex post signal, *private* monitoring is equivalent to *public* monitoring that is vulnerable to collusion for any given level of accuracy. In either case, the signal is only useful if it is accurate enough, i.e., $v > 1/(2 - \xi)$.

Now consider the effect of collusion on *ex ante* public monitoring. Because the efforts can now be based on the signal, we can study the effectiveness of rescaling effort as a tool to fight collusion. The coalition incentive compatibility constraint for this case (CIC$^a$) has to account for the possibility of rescaling the effort levels. If the high-type agent under-reports and the monitor detects it, the stake of collusion is:

$$\left[t_{11}^a - \frac{1}{2}(\theta_1 e_{11}^a)^2\right] - \left[t_{12}^a - \frac{1}{2}(\theta_1 e_{12}^a)^2\right].$$

Therefore, the reward to the monitor must be at least as high as the stake of collusion and the (CIC$^a$) constraint can be written as: [35]

(CIC$^a$) $\qquad W^a \geq t_{11}^a - t_{12}^a - \frac{1}{2}\theta_1{}^2((e_{11}^a)^2 - (e_{12}^a)^2).$

We characterize in appendix B the main features of the optimal contract for ex ante public monitoring with collusion. We show that the principal does not benefit from a gap in efforts, and the optimal contract involves $e_{11}^a = e_{12}^a = e_1^a$. Therefore, the (CIC$^a$) is identical to (CIC$^p$). This implies that, for any given level of accuracy, the optimal contracts are identical whether public monitoring takes place ex ante or ex post.

This confirms our earlier intuition that non-monetary tools are effective only when the parties value them differently. Despite some similarity in form, (CIC$^a$) captures very different incentives than (PIC$^a$) under ex ante private monitoring in section 2. Namely, (CIC$^a$) removes the agent's incentive to manipulate (under-report) the signal while (PIC$^a$) eliminates the principal's incentive to manipulate (over-report) the signal. In the private monitoring case, the principal evaluated effort in terms of output, while the agent evaluated it in terms of cost, and a gap in efforts could be used to address both (PIC) and (IC). In the public monitoring case with collusion, both agent and monitor evaluate effort similarly because the agent's rent, which depends on the difference in cost of effort, is also the source of the stake of collusion. A gap in efforts is ineffective in relaxing the (IC) and (CIC). The only tool that can affect both the (IC) and the (CIC) is monetary transfers. In other words, because the monitor does not care about the output per se (and therefore the effort), a monetary transfer is the only instrument that can impact both the agent and the monitor.[36]

Hence, we have shown that ex ante private monitoring is strictly better than public monitoring that is vulnerable to collusion. We have already shown in proposition 2 that ex ante private monitoring is better than ex post private monitoring, and in this section we have argued

that ex post private monitoring is equivalent to public monitoring (ex ante or ex post) vulnerable to collusion. In short, ex ante private $>$ ex post private $\sim$ ex post public vulnerable to collusion $\sim$ ex ante public vulnerable to collusion.

We summarize our findings in Proposition 3:

**Proposition 3:**

i.   *With public monitoring vulnerable to collusion, non-monetary tools, such as rescaling, are ineffective and the timing of information does not matter.*

ii.   *Ex post private monitoring is equivalent to public (ex ante or ex post) monitoring vulnerable to collusion; monitoring is only useful if it is accurate enough.*

iii.   *Ex ante private monitoring is preferable to public (ex ante or ex post) monitoring vulnerable to collusion.*

*Collusion costs*

If collusion costs are present (due, for instance, to the psychological burden of participating in an illegal activity (see Tirole (1992)), transaction costs (Faure-Grimaud, Laffont and Martimort (2003)) or the fear of being detected in the future (Khalil and Lawarree (2006)), public monitoring becomes more profitable for the principal. Technically, the (CIC) constraint is easier to satisfy.

Denote by $k$, the collusion costs, i.e., one unit of bribe is only worth $k$ dollars to the monitor, where $k \in [0,1]$. So far we have assumed that $k = 1$. With $k < 1$, the agent can offer at

most a fraction $k$ of the benefit of misreporting the agent's type (the stake of collusion). If $k = 0$, the public signal is collusion-free, and the first best can be reached. For the ex post signal, the (CIC) is now $W^p \geq k(t_{11}^p - t_{12}^p)$, and for an ex ante signal, the (CIC) becomes $W^a \geq$

$$k\left(\left(t_{11}^a - \frac{1}{2}(\theta_1 e_{11}^a)^2\right) - \left(t_{12}^a - \frac{1}{2}(\theta_1 e_{12}^a)^2\right)\right).$$

Let us now re-examine Proposition 3 when collusion costs are present ($k < 1$).[37] Part (i) of Proposition 3 remains intact. With $k < 1$, the reward to deter collusion is less but the lower reward applies both ex ante and ex post. Therefore, the timing of information for public monitoring still does not matter. In Part (ii) of Proposition 3, we showed that ex post private monitoring was equivalent to public monitoring vulnerable to collusion. This is no longer true. With collusion costs, the public monitoring is more profitable and, therefore, public monitoring dominates private monitoring.

Finally, the principal still prefers ex ante private monitoring (Part (iii) of Proposition 3 is still true) unless collusion costs are large. With no collusion costs ($k = 1$), we showed in Proposition 3 that ex ante private monitoring was better than public monitoring. With large enough collusion cost (i.e., small enough $k$), public monitoring could become more profitable than private monitoring. Indeed, if $k = 0$, the first best can be reached with public monitoring because the signal is collusion-free. Therefore, there is a critical $k$ – say $k^p$ – such that ex ante private monitoring dominates public monitoring for all $k \geq k^p$ and is dominated by public monitoring otherwise.

## 4. Conclusion

In this article, we show that the timing of information can be critical to solve the credibility issue related to private information received by a principal. If the principal can collect her private information before the project is completed, she prefers to use non-monetary tools like rescaling the project instead of monetary penalties which induce burning money. The intuition relies on the fact that money is equally valuable to all parties, whereas a non-monetary tool, like rescaling effort, exploits the fact that parties evaluate effort differently. An immediate consequence is that the principal always prefers to obtain private information before the project is over because rescaling the project is more efficient than burning money.

By comparing private signals and third party public signals vulnerable to collusion, we have learned that relying on non-monetary penalties, like rescaling, is not effective to deter collusion; instead, the principal must use monetary penalties. Intuitively, it is because the monitor does not care about the output per se, and only a monetary transfer can affect both the agent and the monitor. For ex post signals, only monetary penalties are available, and private signals are equivalent to public signals vulnerable to collusion. However, if the signal arrives ex ante, a private signal is preferred because non-monetary penalties are available.

# References

Angelucci, C. and A. Russo (2012): "Moral Hazard and Collusion in Hierarchies," *mimeo*, Toulouse School of Economics.

Baron, D. and D. Besanko (1984): "Regulation, Asymmetric Information and Auditing," *Rand Journal of Economics*, 15, 447–470.

Baron, D. and R. Myerson (1982): "Regulating a monopolist with Unknown Costs," *Econometrica*, 50, 911–930.

Celik, G. (2009): "Mechanism Design with Collusive Supervision," *Journal of Economic Theory*, 144, 69–95.

Che, Y.-K. and J. Kim (2006): "Robustly Collusion-Proof Implementation," *Econometrica*, 74, 1063–1107

Crémer, J. and R. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1258.

Demski, J. S., and D. E. M. Sappington (1993): "Sourcing with Unverifiable Performance Information," *Journal of Accounting Research*, 31, 1–20.

Dequiedt, V. and D. Martimort (2014): "Vertical Contracting with Informational Opportunism," *mimeo,* Paris School of Economics.

Farhi, E, J. Lerner and J. Tirole (2012) "Fear of Rejection? Tiered Certification and Transparency," *mimeo*, Toulouse School of Economics.

Faure-Grimaud, A., J.-J. Laffont, and D. Martimort (2003): "Collusion, Delegation and Supervision with Soft Information," *Review of Economic Studies*, 70, 253–279.

Faure-Grimaud, A., J.-J. Laffont, and D. Martimort (2002): "Risk Averse Supervisors and the Efficiency of Collusion," *Contributions to Theoretical Economics*, 2(1), 1534-5971.

Faure-Grimaud, A., and D. Martimort (2003): "Regulatory Inertia," *RAND Journal of Economics*, 34(3), 413-437.

Fuchs, W. (2007): "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97, 1432–1448.

Holmstrom, B., and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 303–328.

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66, 627–652.

Kandori, M. (2002): "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102, 1–15.

Khalil, F. and J. Lawarree (2006): "Incentives for Corruptible Auditors in the Absence of Commitment," *Journal of Industrial Economics*, 54, 269–291.

Khalil, F., J. Lawarree and S. Yun (2010) "Bribery versus extortion: allowing the lesser of two evils" *Rand Journal of Economics*, 41(1), 179-198.

Kofman, F. and J. Lawarree (1993): "Collusion in Hierarchical Agency," *Econometrica*, 61, 629–656.

Laffont, J.-J. and D. Martimort (1997): "Collusion under Asymmetric Information," *Econometrica*, 65, 875–911.

_____ (2000): "Mechanism Design with Collusion and Correlation," *Econometrica*, 68, 309–342.

Levin, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93, 835–857.

MacLeod, W. B. (2003): "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 93, 216–240.

Mookherjee, D. "Incentives in Hierarchy." in R. Gibbons and J. Roberts, eds., *Handbook of Organizational Economics*, Princeton University Press, 2012.

Rahman, D. (2012): "But who will monitor the monitor", *American Economic Review*, 102, 2767–2797.

Riordan, M. and D. Sappington (1986): "Optimal Contracts with Public and Private Ex Post Information," *mimeo*, Stanford University.

_____ (1988): "Optimal Contracts with Ex Post Information," *Journal of Economic Theory*, 45, 189–199.

Strausz, R. (1997): "Delegation and Monitoring in a Principal-Agent Relationship," *The Review of Economic Studies*, 64, 337–357.

Strausz, R. (2006): "Interim Information in Long-Term Contracts," *Journal of Economics and Management Strategy*, 15, 1041–1067.

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

_____ (1992): "Collusion and the theory of organizations," in Jean-Jacques Laffont, ed., *Advances in economic theory: Proceedings of the Sixth World Congress of the Econometric Society*, Vol. 2. Cambridge: Cambridge University Press, pp. 151–206.

[1] See the recent literature on subjective evaluation (Levin (2003), MacLeod (2003), Fuchs (2007)). Riordan and Sappington (1986) propose an analogous remedy: the principal may commit not to accept some portion of the output *ex post*, allowing it to go to waste.

[2] Holmstrom and Milgrom (1987) also challenge the assumption that the agent is required to make a once-for-all choice of his effort during the relationship without regard to the arrival of performance information. Their focus is on the agent, instead of the principal, receiving information and adjusting his effort continuously over time.

[3] See Cremer and McLean (1988) for ex ante signals, and Baron and Besanko (1984) and Riordan and Sappington (1988) for the case of ex post signals.

[4] See, e.g., Tirole (1986) and Kofman and Lawarrée (1993).

[5] See, e.g., Kandori and Matsushima (1998). Kandori (2002) provides a survey.

[6] Demski and Sappington (1993) consider a double moral hazard problem with a private signal but rely on a second public signal in order to satisfy the principal's incentive constraint with a balanced budget.

[7] A high-type can mimic the low type by putting in a fraction $\theta_1/\theta_2$ of the low-type's effort. So, a condition saying $\theta_1 < 1$, is really the condition $\theta_1/\theta_2 < 1$. If we used $\theta_2$ instead of 1, we would need to replace $\theta_1$ by $\theta_1/\theta_2$ in the right hand side of the IC constraint below but our main results would be unaffected.

[8] The high type can mimic the observable output of the low type by exerting an effort $\tilde{e}$ such that $\theta_2 \tilde{e} = \theta_1 e_1$. With $\theta_2$ normalized at 1, the cost of mimicking is $\frac{(\theta_1 e_1)^2}{2}$.

[9] Fuchs (2007) distinguishes between private signals and observable but unverifiable signals.

[10] For instance, suppose that the agent claims that the productivity parameter is low and the private signal confirms the claim. Because the information is private, nothing stops the principal from reporting that the signal contradicts the agent's claim and penalizing him. In other words, the private information is soft.

[11] Signals received after the agent reports that he is a high type can be ignored.

[12] The Revelation Principle can be extended to the case of private signals because the signal is free for the principal. An alternative is to use a mediated contract as in Strausz (2006). On the use of the revelation principle with private information, see also Dequiedt and Martimort (2014). If obtaining the private signal is costly, our results could be extended if the principal can commit to acquire the private signal. Otherwise, Rahman (2012) shows that mediated contracts relying on correlated equilibrium can improve efficiency.

[13] MacLeod (2003) suggests that an unbalanced budget is merely a tractable metaphor for equilibrium conflict, such as "work to rule" behavior by union members. What is important is that the principal does not collect $B_{11}^p$ or $B_{12}^p$; giving them to a third party (like a charity) would also satisfy (PIC). Fuchs (2007) cites an example from Jon Levin where fines collected from baseball players are given to charity. The same is true for football players in the NFL.

[14] The (IC) constraint for the low type is redundant and suppressed here and in the sections that follow.

[15] If $B_{11}^p$ were strictly positive, the principal could reduce her cost without altering any constraint by lowering both $B_{11}^p$ and $B_{12}^p$ by equal amounts.

[16] The IR can be written as $t_{12} + \nu B_{12} - \frac{e_1^2}{2} = 0$.

[17] Note that the principal infers that the agent is of low type when computing her payoffs in (PIC). This is due to the (IC) constraints and the fact that the signal is only relevant after the agent claims to be a low-type.

[18] The agent is paid less when the signal contradicts his report, but he is also given a less arduous assignment.

[19] Substitute $t_{11}^a - t_{12}^a$ from (PICᵃ) in the above inequality to obtain $\left(1 - \frac{\theta_1}{2}(e_{11}^a + e_{12}^a)\right) > 0$, which is true as there is under-production in equilibrium ($e_{12}^a < e_{11}^a < \theta_1$).

[20] Figure 1 is based on the example, where $\xi = 0.5$ and $\theta_1 = 0.36$. Note that both $e_1^{SB}$ and $\tilde{\nu}$ are increasing with respect to $\xi$.

[21] Increasing $e_{11}^a$ also has a negative effect on the cost differential (it increases the rent) while decreasing $e_{12}^a$ has an ambiguous effect on the cost-differential. Overall, we show that the penalty effect is stronger than the cost differential effect and the principal succeeds in decreasing the rent by increasing $e_{11}^a$ and decreasing $e_{12}^a$.

[22] In contrast to the ex post case (Proposition 1), the traditional separation result (Baron and Besanko (1984)) breaks down because the low type's efforts, $e_{11}^a$ and $e_{12}^a$, are adjusted as soon as $\nu > \frac{1}{2}$. However, because $e_{12}^a$ is only

restored once rent is fully extracted, the intuition of the traditional separation result remains largely valid: restoring efficiency in effort is not a priority for the principal as long as the rent is positive.

[23] In other words, the private nature of information restores continuity between correlated and uncorrelated environments as noted by Dequiedt and Martimort (2014).

[24] Rescaling $e_{12}^a$ is costly for the principal because, in equilibrium, the agent is always truthful and the signal is simply incorrect.

[25] The key idea is to restrict downward adjustment in $e_{12}$. For instance, a model with shut-down of effort would be another example. This could be achieved through a non-standard marginal cost function that violates Inada conditions, e.g., $\psi(e) = \frac{ae + e^2}{2}$, with a high enough $a > 0$.

[26] Other modifications of the model that limits the principal's ability to extract rent can also make money burning optimal. For example, suppose there is a lower bound on the transfer such that $t \geq \underline{t}$. Then, for $\underline{t}$ high enough, we can show that rent will remain positive at $\nu = 1/(2 - \xi)$ and money burning may occur.

[27] Recall that when arguing that ex ante monitoring is better than ex post monitoring, we assumed that the agent knew when the principal obtained her signal.

[28] In the collusion literature, even though third party monitoring is public and verifiable, the information is still manipulable (Tirole (1986), (1992)). Because the monitor provides a signal that is public and verifiable, his identity can be known by the agent who could attempt to bribe him to induce him to report a more favorable signal.

[29] This is the well-known bounty-hunter scheme (see, e.g., Tirole (1992) or Kofman-Lawarree (1993)).

[30] Note that soft information in the hands of the principal is still useful in our model as shown in section 2. The reason lies in the commitment power of the principal. Soft information for the monitor can be useful in specific settings (for instance see Baliga (1999), Faure-Grimaud, Laffont and Martimort (2003), Khalil, Lawarree and Yun (2010), or Angelucci and Russo (2012)). Rahman (2012) shows how mediated contracts can also address monitoring with soft information when there is no collusion between the agent and the monitor. His focus is on the monitor's moral hazard, whereas ours is on inducing truth-telling by the principal. In Rahman, the principal instructs the agent to shirk occasionally and thereby creates shared observations between herself and the monitor. The shared observations allow the principal to verify the monitor's report.

[31] When the signal is σ₂, the agent will try to bribe the monitor to report σ₁. However, when the signal is σ₁, the agent will not bribe the monitor who is then unable to report σ₂ on his own.

[32] For a technology where extortion is possible, see Khalil, Lawarree and Yun (2010).

[33] For a model where the timing of collusion has strategic value to the principal, see Faure-Grimaud, Laffont and Martimort l (2002), and Faure-Grimaud and Martimort (2003).

[34] See Laffont and Martimort (1997, 2000) or Che and Kim (2006) or Celik (2009) for collusion under asymmetric information. See Mookherjee (2012) for a survey.

[35] In 1appendix B, we show that the analogous constraint for a truthful low type is also satisfied.

[36] This result suggests that the principal could benefit from making the monitor a claimant for output.

[37] For the rest of the section, we just outline the arguments as the techniques are similar to earlier steps.
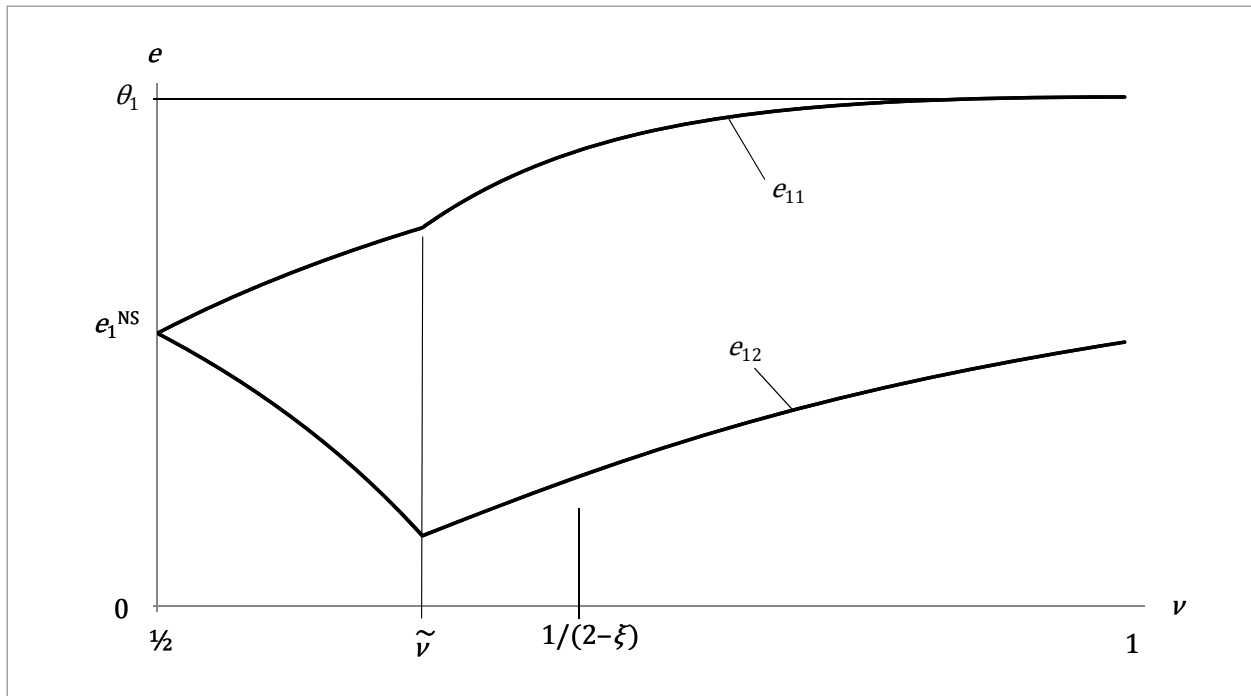
Figure 1—Optimal effort levels ($e_{11}^a > e_{12}^a$) with respect to $\nu$.[1]

[1] Figure 1 is based on the example, where $\xi = 0.5$ and $\theta_1 = 0.36$. Note that both $e_1^{SB}$ and $\tilde{\nu}$ are increasing with respect to $\xi$.