

# Delegation and Nonmonetary Incentives

Attila Ambrus      Georgy Egorov

Duke University

Northwestern  
University - Kellogg

December 4, 2015

ERID Working Paper Number 200

This paper can be downloaded without charge from the Social  
Science Research Network Electronic Paper Collection:

<http://ssrn.com/abstract=2700821>

**Economic Research Initiatives at Duke  
WORKING PAPERS SERIES**



# Delegation and Nonmonetary Incentives\*

Attila Ambrus<sup>†</sup>

Duke University, Department of Economics

Georgy Egorov<sup>‡</sup>

Northwestern University, Kellogg-MEDS

December 2015

## Abstract

In many contracting settings, actions costly to one party but with no direct benefits to the other (money-burning) may be part of the explicit or implicit contract. A leading example is bureaucratic procedures in an employer-employee relationship. We study a model of delegation with an informed agent, where the principal may impose money-burning on the agent as a function of the agent's choice of action, and show that money-burning may be part of the optimal contract. This result holds even if action-contingent monetary transfers are possible, as long as transfers from the principal to the agent are bounded from below (as in limited liability or minimal wage requirements). In fact, the optimal contract can involve a combination of both efficient monetary incentives and inefficient nonmonetary incentives through money burning. Our model delivers some results novel to the delegation literature. First, money-burning is more likely if the principal is more "sensitive" to the choice of action than the agent. This is consistent with the perception that there is more bureaucratization in large organizations. Second, money-burning is more likely if the agent's limited liability constraint is tighter relative to his participation constraint. This implies that a higher minimum wage distorts employment contracts towards using socially wasteful nonmonetary incentives, leading to a Pareto inferior outcome as the agent is still held down to his reservation value through increased money burning.

**Keywords:** delegation, organizational procedures, money burning

**JEL Classification:** D23, D82, D86.

---

\*We thank Daron Acemoglu, Ricardo Alonso, Volodymyr Baranovskyi, Peter Esó, Drew Fudenberg, Navin Kartik, Tymofiy Mylovanov, Marco Ottaviani, Satoru Takahashi, and seminar participants at Harvard, MIT, the New Economic School, and the University of Quebec for useful comments and suggestions, and Peter Landry and Ricardo Pique for excellent research assistance.

<sup>†</sup>E-mail: aa231@duke.edu

<sup>‡</sup>E-mail: g-egorov@kellogg.northwestern.edu

# 1 Introduction

In many economic and political interactions, the use of monetary incentives is ruled out or limited. For example, members of a legislation typically receive salaries that compensate them financially for their work, but their payments do not depend on how they vote, or how many amendments they make. It is even more common that in an economic relationship financial incentives are possible, but they are constrained. For example, a minimum wage requirement bounds monetary transfers from employers to employees from below. Another common source of limitations on monetary transfers is that one or both parties might be liquidity-constrained.<sup>1</sup> Despite these restrictions on transfers, nonmonetary incentives are often available, and thus may be used to align the interests of participants. In many settings these incentives take the form of imposing activities that are costly for one agent and do not directly benefit any of the rest. Following the standard terminology, we refer to such activities as money burning.

To study such situations formally, we consider a scenario where an uninformed principal delegates the task of choosing the action from a unidimensional action space to an agent, who before making the decision receives private information about a state variable. The state variable affects the well-being of both parties in a way that a higher state is associated with a higher optimal action choice for both of them. The principal delegates by offering the agent a contract which for any possible action prescribes a nonnegative amount of money burning the agent has to make if he takes that action. We investigate this contracting problem in two different contexts: first, we assume monetary transfers conditional on the action choice are completely ruled out (only an ex ante wage payment is possible); second, we add the possibility of such transfers. Throughout the paper we assume that the principal has to satisfy a participation constraint for the agent, and that the amount of monetary transfer from the principal to the agent is bounded from below.<sup>2</sup> The latter can come from a nonnegativity constraint on transfers to the agent, liquidity constraints on the part of the agent, or a minimum wage requirement.

The model we describe is applicable to many different situations. One important example is organizations, where a primary form of providing such incentives is bureaucracy. If managers of organizational units are biased towards requesting a higher budget for their units than what would be optimal for the organization, a manager applying for a higher budget might be required to fill out more paperwork, or get stamps of approval from different offices. In fact, it is a common

---

<sup>1</sup>For this reason, cash penalties are hardly ever used in employment contracts, except for employees with very large incomes relative to the damage that their actions can cause (p.249 of Milgrom and Roberts (1992)).

<sup>2</sup>In some applications it is reasonable to drop one of these constraints, which corresponds to the limiting case of our model in which the corresponding lower bound is  $-\infty$ .

perception that there is an excess of bureaucratic procedures.<sup>3</sup> Another application of our model is situations in which an activist group can launch a campaign or boycott against a company violating environmental or human rights standard, as in Baron and Diermeier (2007) and Abito et al. (2015).<sup>4</sup> Our model facilitates investigating such situations in a contract theory setting, assuming that the activist group can ex ante commit to a money-burning scheme. Lastly, academic journals might use editorial delay, as opposed to submission fees, to deter excessive submissions of low-quality manuscripts (see Azar (2006)).

Within the delegation literature, launched by the seminal work of Holmstrom (1977), the papers closest to ours are Athey et al. (2004), Amador et al. (2006) and Amador and Bagwell (2013a,b).<sup>5</sup> Athey et al. (2004), Amador et al. (2006) and Amador and Bagwell (2013a) feature models that are similar to ours, but they impose restrictions that ultimately imply that money burning is not part of the optimal contract. In contrast, our paper suggests that the use of money burning can be widespread in many different types of situations. Amador and Bagwell (2013b, from now on AB) is the only other paper in the delegation literature, as far as we know, that considers cases when the optimal contract involves money-burning. Our analysis is more general in the following dimensions: (i) the results that we provide outside the uniform-quadratic setting (the results in Section 4) allow for more general payoffs than AB restricts attention to; (ii) we allow for a lower bound constraint on ex ante transfers from the principal, such as a minimal wage; (iii) we provide an extension that allows for state-contingent monetary transfers; (iv) in the uniform-quadratic setting with no contingent transfers we provide a full characterization of the optimal contract, while AB only focuses on specifications when the optimal contract is one of two particular kinds. On the other hand, AB offers sufficient conditions for the optimality of the latter two kinds of contracts for more general specifications than the uniform-quadratic.

Our analysis reveals that money-burning is more likely to be used as an incentive device when the principal's utility, measured in monetary terms, is more sensitive to the implemented action

---

<sup>3</sup>Indeed, most of the related economic literature takes this stance, looking for explanations for excessive bureaucratization as in Strausz (2006), or connect bureaucracy and corruption as in Banerjee (1997) and Guriev (2004). In contrast, in this paper we suggest that bureaucratic procedures, as well as other costly and wasteful activities, can improve the (ex-ante) efficiency of an organization.

<sup>4</sup>Although these papers abstract away from the campaign imposing direct costs on the interest group, it is reasonable to assume that such actions are not free. This would require an extension of our model, to allow for money-burning to impose costs directly on the principal as well. This extension would be straightforward though, and as long as the cost of money-burning for the principal is small relative to the cost it imposes on the agent, the qualitative conclusions of the model would not change. For a different economic situation involving money-burning as a disciplinary device that is costly for all participants, see Padro i Miquel and Yared (2012).

<sup>5</sup>There is also a less directly related literature suggesting that signaling, even if costly for individual agents, may improve economic efficiency in allocation problems. For example, Chakravarty and Kaplan (2013) show that allocation of goods without transfers may be more efficient if agents are able to send costly signals to the mechanism designer (see also Condorelli (2012) and Yoon (2011)).

than the agent's utility. This is consistent with the common perception that bureaucracy is more widespread in large organizations, as the agent's choices typically influence the well-being of more other agents (represented by the principal).

Second, we find that money-burning is more likely to arise when there are more stringent restrictions on monetary transfers between the contracting parties, such as a higher minimum wage or other forms of wage control, and when outside option of the agent is low. Intuitively, in these cases the principal can impose money-burning without violating the participation constraint. This points to an additional distortionary effect of increasing the minimum wage, besides the ones commonly discussed in policy debates: an increase in minimal wage makes it more likely that employers use socially inefficient nonmonetary incentives, rather than efficient monetary incentives. In fact, increasing the minimum wage in our model may lead to a Pareto inferior outcome, in which the agent is held down to the same reservation utility as before, with the increase in wage being offset by more money burning imposed, whereas the principal is strictly worse off.<sup>6</sup>

Our paper proceeds as follows. We first show some basic general results, such as existence of the optimal contract, that the implemented action scheme is monotonically increasing in the state, and that the action specified in the contract is never below the optimal action of the principal. Under some regularity conditions, we also show that both money burning and the action choice are continuous functions of the state, and that the implemented action is always between the optimal points of the principal and the agent. Under the same regularity conditions we transform the principal's problem to an intuitive and tractable form. We demonstrate the usefulness of this result by explicitly solving the transformed problem in the popular uniform-quadratic specification of the delegation model, with the extra parameter capturing the relative importance of the action chosen for the principal and the sender (in monetary terms).<sup>7</sup> We show that this parameter and the agent's outside option relative to the minimum bound on transfers are the two crucial parameters determining whether there is money burning in the optimal contract.

---

<sup>6</sup>Wessels (1980a, 1980b) makes an observation that is a counterpart of our point: If a firm can provide fringe benefits to its workers in a socially efficient way (namely the benefit for the workers is larger than the cost for the firm), for example by maintaining better working conditions, then an increase in the minimum wage can induce the firm to cut down in these socially efficient fringe benefits. Both the latter papers and our work points out a substitution effect between monetary and nonmonetary compensation schemes, effecting social surplus. Of course we recognize that both Wessels (1980a, 1980b) and our paper only focus on particular effects of increasing the minimum wage, and that the issue is a complex one - as reflected by the large and diverse literature devoted on the topic.

<sup>7</sup>This extended uniform-quadratic setting is the only class of specifications for which we explicitly solve for the optimal contract. We note however, that ours is the first paper in the literature that characterizes the optimal contract for all specifications within this class.

In the version of the model that allows for monetary transfers, we show that the optimal contract can involve transfers only, or money burning only, or transfers in some states and money burning in other states, providing an explanation for why inefficient nonmonetary incentives are used even in settings where financial incentives are feasible. If the agent's outside option is high enough, there is no wasteful money burning, and the optimal contract achieves jointly efficient action choices. In this case, the principal and the agent essentially form a partnership. If the agent's outside option is very low, the agent receives exactly the minimum wage in each state, but money burning may be used. For an intermediate range of parameters, both positive and negative incentives can be used in the optimal contract, in a way that monetary transfers are used to reward the agent for actions he is unlikely to prefer, while money burning is used to punish him for actions that the principal is unlikely to prefer. Intuitively, using both the carrot and the stick (transfers and money-burning) allows the principal to provide incentives while keeping the participation constraint bind.

Lastly, we note that often the principal may choose between different costly activities to provide incentives to the agent, and some of these may provide benefits to the principal. However, the qualitative conclusions of the model would generalize to a setting in which the principal benefits from these activities, as long as this activity implies some efficiency loss.

## 2 Related literature

Our work continues the literature on constrained delegation started by Holmstrom (1977).<sup>8</sup> Holmstrom, as well as Melumad and Shibano (1991) and Alonso and Matouschek (2007, 2008), considers deterministic delegation with no monetary transfers, in which the principal can restrict the action space of the agent, but cannot make different actions differentially costly.<sup>9</sup> In our framework the principal can always achieve such delegation schemes by setting some actions free while the remaining ones prohibitively costly, hence she is at least weakly better off.

Kováč and Mylovanov (2007) and Goltsman et al. (2009) investigate stochastic delegation mechanisms in the constrained delegation context, assuming quadratic utilities.<sup>10</sup> Quadratic utilities imply that the utilities of both parties are additively separable to a term that only depends on the expectation of the induced action and a common negative term that depends

---

<sup>8</sup>Dessein (2002) considers delegation in which restricting the agent's action space is not allowed, but the principal can potentially retain a veto power. See also Aghion and Tirole (1997) and Szalay (2005) for models of delegation less related to ours. There is also a literature in political science on delegation and control: see, for example, Bendor et al. (1987), and McCubbins et al. (1987).

<sup>9</sup>For a recent more detailed description of this line of literature, see Armstrong and Vickers (2010).

<sup>10</sup>Stochastic delegation implies that the principal can commit to different probabilistic action choices after different reports by the agent. See also Strausz (2006) on stochastic mechanisms in a different setting.

on the variance of the induced action. In our model, money burning is only a direct cost for the agent, but it is ex ante costly for the principal if the participation constraint is binding, leading to similar terms in the principal’s objective function in the contracting problem.

Ottaviani (2000), Krahmer (2004), and Krishna and Morgan (2008) investigate delegation with monetary transfers, although either not characterizing the optimal contract, or not incorporating a participation constraint for the agent, making the results of the above papers difficult to compare to ours.<sup>11</sup> Koessler and Martimort (2012) consider a delegation problem in a 2-dimensional policy space, which is somewhat analogous to our 2-dimensional setting with money burning and a one-dimensional policy choice, although with completely different preferences over the policy space, leading to different results.

A major alternative of delegation is cheap talk communication between the informed and the uninformed parties, as in Crawford and Sobel (1982) and a large literature building on it. In a cheap talk game the uninformed party cannot commit to let the informed party to choose an action, therefore her action choice is required to be sequentially rational. The closest papers to our work in this literature are Austen-Smith and Banks (2000) and Kartik (2007), who consider communication with money burning by the informed party. The focus of these papers is very different from ours: they investigate how money burning can expand the set of cheap talk equilibria in the Crawford and Sobel model. The existing literature on both delegation and cheap talk only considers incentive compatibility constraints. A contribution of the current paper is incorporating a participation constraint for the agent.

The formal literature on procedural rules and organizational bureaucracy, despite its practical importance, is relatively scarce and not directly related to our paper.<sup>12</sup>

### 3 The basic model

In this Section we set up the basic model, in which the principal can set costly procedural rules for the agent, but contingent monetary transfers are not possible. For the extension of the model which allows for contingent transfers, see Section 6.

We consider the following principal-agent problem. There is an uninformed principal, and an informed agent who observes the realization of a random variable  $\theta \in \Theta = [0, 1]$ . From now on we will refer to  $\theta$  as the state. The c.d.f. of  $\theta$  is  $F(\theta)$ , and we assume it has a density function  $f$  that is strictly positive and absolutely continuous on  $[0, 1]$ . The principal in our model *delegates*

---

<sup>11</sup>See the end of Section 4 for a partial comparison with the results in Krishna and Morgan (2008). The main qualitative difference is that while monetary transfers are always used to some extent in the optimal contract, money burning might be a too costly incentive device for the principal, and hence not used at all.

<sup>12</sup>See Tirole (1986), Garicano (2000), Prendergast (2007) and Crémer et al. (2007).

decision-making, hence the agent has to choose an action  $y \in Y = [y_L, y_H]$ , after observing the state. Both the state and the action affect the well-being of both parties. We assume that both the principal and the agent are von Neumann and Morgenstern expected utility maximizers. If action  $y$  is chosen at state  $\theta$ , then the principal and the agent get utilities  $u^p(\theta, y) = -l^p(\theta, y)$ , while the corresponding utility for the agent is given by  $u^a(\theta, y) = -l^a(\theta, y)$ . We refer to  $l^p$  and  $l^a$  as the loss functions of the principal and the agent, and we assume that both functions are twice continuously differentiable and strictly convex in  $y$ . We assume that for fixed  $\theta$ ,  $u^p(\theta, y)$  reaches its maximum value 0 at  $y^p(\theta) = \theta$ , while  $u^a(\theta, y)$  reaches its maximum value 0 at  $y^a(\theta) = \theta + b(\theta)$  for some  $b(\theta) > 0$ . We refer to  $y^p(\theta)$  and  $y^a(\theta)$  as the ideal actions of the principal and the agent at state  $\theta$ , and to  $b(\theta)$  as the bias of the agent at state  $\theta$ . We assume that  $Y$  contains the interval  $[0, 1 + b(1)]$ . We also assume the single-crossing condition  $\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} < 0$ ; this implies, in particular, that  $\theta + b(\theta)$  is continuous and strictly increasing.<sup>13</sup> Finally, we assume that all parameters of the model are commonly known to the two parties involved.

So far the model is just the standard workhorse model of the delegation literature, that builds on the framework provided in Holmstrom (1977). The novel features of the model are the following:

(i) The principal can impose costs on the agent which may depend on his choice of action. Formally, the principal can specify a function  $m : Y \rightarrow \mathbb{R}^+$ . For any  $y \in Y$ ,  $m(y)$  is a non-recoverable loss for the agent, which does not directly affect the principal's utility, and we interpret it as the amount of paperwork needed to pick policy  $y$ . Following standard terminology for purely wasteful activities, we refer to  $m(y)$  as the amount of money burning required when choosing action  $y$ . Money burning enters the agent's utility as a cost, in an additively separable manner. We note that delegation with differential costs encompasses standard delegation agreements considered in the existing literature, where the principal restricts the set of available policies for the agent to  $D \subset Y$ : in our framework this could be replicated by setting  $m(y)$  to be zero if  $y \in D$ , and  $m(y)$  to be prohibitively high if  $y \in Y \setminus D$ . Hence, a principal who can set differential costs is at least weakly better off than a principal who can only choose a set of feasible actions for the agent.

(ii) The principal has to hire the agent by offering an acceptable contract. We assume that contracting happens *ex ante*, i.e., before the agent observes the state. The contract specifies the

---

<sup>13</sup>We therefore rule out the possibility that the agent's ideal point is state-independent. It is not unrealistic (the agent might always prefer to exert zero effort), but in this paper, we focus on the case where the interests of the principal and the agent are somewhat aligned. We want to thank an anonymous referee for pointing out that in this case, too, an optimal contract may involve money burning if the agent's willingness-to-pay depends on the state even if his ideal action does not.

cost function  $m$  (interpreted as the description of the paperwork requirements for all possible actions), and a constant transfer payment  $T$  (interpreted as a wage) that enters the agent's utility function in an additively separable manner. In our basic model, we assume that monetary transfers contingent on either  $\theta$  or  $y$  are not possible (in Section 6 we relax this requirement). The agent has an outside option  $\tilde{u}$ , therefore we assume that he accepts any contract that gives him at least this much expected utility, given the ex ante distribution of  $\theta$ . We assume that the unconditional transfer  $T$  must satisfy  $T \geq \tilde{w}$  (which may or may not be a binding constraint). We allow  $\tilde{w}$  to be either positive or negative:  $\tilde{w} \leq 0$  would correspond to the case where the agent has liquidity constraints, with  $\tilde{w} = 0$  meaning that no transfer from the agent is allowed, while  $\tilde{w} > 0$  would naturally model a minimal wage requirement.

## 4 Properties of the optimal contract

In this section we derive some qualitative features of the optimal contract. We first establish properties that hold for the most general specification of the model that we introduced above. Then we derive additional properties that require certain regularity conditions on the loss functions and the prior distribution of states to hold.

As standard in the literature, we rewrite the delegation problem to its direct mechanism interpretation:<sup>14</sup> the principal's task is to set a transfer  $T$  and a pair of measurable functions  $y(\theta)$ , the action that the agent takes in state  $\theta$ , and  $m(\theta)$ , the amount of paperwork or money burning in this state, that solve the following problem:

$$\max_{T, \{y(\theta), m(\theta)\}_{\theta \in \Theta}} \int_{\Theta} u^p(\theta, y(\theta)) dF(\theta) - T \quad (1)$$

$$\text{s.t. } \int_{\Theta} (u^a(\theta, y(\theta)) - m(\theta)) dF(\theta) + T \geq \tilde{u}, \quad (2)$$

$$\forall \theta, \theta' \in \Theta : u^a(\theta, y(\theta)) - m(\theta) \geq u^a(\theta, y(\theta')) - m(\theta'), \quad (3)$$

$$\forall \theta \in \Theta : m(\theta) \geq 0, \quad (4)$$

$$T \geq \tilde{w}. \quad (5)$$

In other words, the principal maximizes his payoff subject to the agent's individual rationality

---

<sup>14</sup>Standard arguments establish that for every mechanism in which the agent sends a message and the principal commits to a message-contingent action plan there is an equivalent direct mechanism, in the sense of having the same set of perfect Bayesian Nash equilibrium outcomes. See pp. 25-6 in Chapter II of Holmstrom (1977), and Krishna and Morgan (2008) for extending the arguments to settings with imperfect commitment on the part of the principal. In this paper we restrict attention to deterministic mechanisms, in which any message of the agent induces action and money-burning choices deterministically.

and incentive compatibility constraints (equations (2) and (3), respectively), as well as exogenous constraints on money burning and transfers.

First we observe that in an optimal contract either the agent's participation constraint (2) binds, or the transfer is minimal, so (5) binds: otherwise the principal could reduce the ex-ante transfer without violating the participation constraint (and not affecting the IC constraints) and achieve a higher expected payoff. Denote the total principal's loss from the contract by  $L^p$  and the agent's loss, conditional on  $\theta$ , by  $L^a(\theta)$ :

$$L^p(y(\cdot), m(\cdot), T) = \int_{\Theta} l^p(\theta, y(\theta)) dF(\theta) + T, \quad (6)$$

$$L^a(\theta) = L^a(\theta, y(\theta), m(\theta)) = l^a(\theta, y(\theta)) + m(\theta). \quad (7)$$

We find it convenient to rewrite the problem in the following way (we denote the principal's loss from contract  $(y(\cdot), m(\cdot))$  by  $L^p(y(\cdot), m(\cdot))$ ):

$$\min_{T, \{y(\theta), m(\theta)\}_{\theta \in \Theta}} L^p(y(\cdot), m(\cdot), T) = \min_{T, \{y(\theta), m(\theta)\}_{\theta \in \Theta}} \int_{\Theta} l^p(\theta, y(\theta)) dF(\theta) + T \quad (8)$$

$$\text{s.t. } \int_{\Theta} (l^a(\theta, y(\theta)) + m(\theta)) dF(\theta) \leq T - \tilde{u}, \quad (9)$$

$$\forall \theta, \theta' \in \Theta : l^a(\theta, y(\theta)) + m(\theta) \leq l^a(\theta, y(\theta')) + m(\theta'), \quad (10)$$

$$\forall \theta \in \Theta : m(\theta) \geq 0, \quad (11)$$

$$T \geq \tilde{u}. \quad (12)$$

Our model is not a standard hidden information principal-agent model with monetary transfers (Baron and Myerson, 1982, Guesnerie and Laffont, 1984), since money burning enters parties' utility functions differently than monetary transfers. Furthermore, as opposed to most principal-agent models, the agent's ideal action is a nontrivial function of the state.<sup>15</sup> For these reasons, existing results from the above literature cannot be directly used in our setting. However, some basic results can be derived in an analogous manner to the standard model. Claims 1–3 states these results. Since the proofs are straightforward and similar to proofs of analogous results in the literature, they are relegated to the Supplementary Appendix.<sup>16</sup>

**Claim 1** *If a pair of functions  $\{y(\theta), m(\theta)\}_{\theta \in \Theta}$  satisfies (10), then  $\theta_2 \geq \theta_1$  implies  $y(\theta_2) \geq y(\theta_1)$ . Moreover, if  $y(\theta_1) = y(\theta_2)$ , then  $m(\theta_1) = m(\theta_2)$ .*

<sup>15</sup>Related to the point that our model differs from standard principal-agent models, the optimal contract that we derive below does not satisfy some standard results like “no distortion at the top.” The agent never chooses his ideal action for  $\theta = 1$ , and it is possible that he does not choose his ideal action for any  $\theta$ .

<sup>16</sup>One additional difficulty that we face, relative to the most standard principal-agent problems is that the  $m(\cdot) \geq 0$  condition is difficult to translate into a condition on payoffs. For this reason we cannot follow the standard approach of solving for the optimal contract in payoffs space.

**Claim 2** *If there exists a solution to the problem (8)-(12) then there is a solution  $\{y^*(\theta), m^*(\theta)\}_{\theta \in \Theta}$  such that  $\inf_{\theta \in \Theta} m^*(\theta) = 0$ .*

**Claim 3** *If a pair of functions  $\{y(\theta), m(\theta)\}_{\theta \in \Theta}$  satisfies (10), then for agent's loss function  $L^a(\theta)$  the following is true:<sup>17</sup>*

(i)  $L^a(\theta)$  is Lipschitz continuous with parameter  $\Delta_\theta = \max_{\theta \in \Theta, y \in Y} \left| \frac{\partial l^a(\theta, y)}{\partial \theta} \right|$ .

(ii)  $L^a(\theta)$  has left derivative for each  $\theta_0 > 0$  and has right derivative for each  $\theta_0 < 1$ , given by:

$$\begin{aligned} \frac{d^l L^a(\theta_0)}{d\theta} &= \frac{\partial l^a(\theta_0, \lim_{\theta \rightarrow \theta_0^-} y(\theta))}{\partial \theta}, \\ \frac{d^r L^a(\theta_0)}{d\theta} &= \frac{\partial l^a(\theta_0, \lim_{\theta \rightarrow \theta_0^+} y(\theta))}{\partial \theta}. \end{aligned}$$

(iii)  $L^a(\theta)$  is differentiable at  $\theta_0 \in (0, 1)$  if and only if  $y(\theta)$  is continuous at  $\theta_0$ , and then<sup>18</sup>

$$\frac{dL^a(\theta_0)}{d\theta} = \frac{\partial l^a(\theta_0, y(\theta_0))}{\partial \theta}. \quad (13)$$

Claim 3 implies that the IC constraint (10) pins down the amount of money-burning in each state  $m(\theta)$  given an increasing function  $y(\theta)$ , up to a constant (which is itself pinned down by the condition  $\inf_{\theta \in \Theta} m(\theta) = 0$ ). The reformulation (15) below is similar to known results in settings different from ours (see for example Appendix 1 in Jullien (2000) or Lemma 1 in Noldeke and Samuelson (2006)).<sup>19</sup> The proof of this and subsequent results are in the Appendix. Take a pair of functions  $(y(\theta), m(\theta))_{\theta \in \Theta}$  satisfying (10) and define function  $\tilde{m}(y)$  as the amount of money required to burn when action  $y \in R$  is chosen (where  $R$  is the range of  $y(\theta)$ ) by:

$$\tilde{m}(y) = m(\theta) \text{ where } \theta \in \Theta \text{ satisfies } y(\theta) = y; \quad (14)$$

notice that  $\tilde{m}(\cdot)$  is well-defined (this follows from Claim 1). We now define  $\tilde{\theta}(\cdot)$  as the inverse of  $y(\cdot)$  if  $y(\cdot)$  is continuous and strictly increasing; otherwise, we let  $\tilde{\theta}(\cdot)$  be any monotone single-valued function  $[y(0), y(1)] \rightarrow \Theta$  such that  $y(\tilde{\theta}(y_0)) = y_0$  for any  $y_0 \in R$ .<sup>20</sup>

<sup>17</sup>We could apply Theorem 1 in Milgrom and Segal (2002) to obtain absolute continuity of  $L^a(\theta)$ . However, we need the stronger results that require a separate proof.

<sup>18</sup>Throughout,  $\frac{\partial l^a}{\partial \theta}$  and  $\frac{\partial l^a}{\partial y}$  denote the partial derivatives of  $l^a(\theta, y)$  with respect to the first argument and the second arguments, respectively.

<sup>19</sup>Our analysis here is also similar to that in Goldman et al. (1984), who examine a context in which a monopolist can offer a menu of price-quantity pairs. Quantity ( $q$ ) plays an analogous role to  $y$  in our model, while price ( $R$ ) does to  $m$  in our model. Their  $R(q)$  function is similar to  $\tilde{m}(y)$  in our setting. However, both the assumptions they impose and the type of results they derive differ from ours, hence their results are not applicable in our investigation.

<sup>20</sup>Monotonicity on  $R$  must follow from Claim 1; to achieve monotonicity on the entire  $\tilde{\theta}(y)$ , we define  $\tilde{\theta}(y)$  for  $y \in [y(0), y(1)] \setminus R$  by  $\tilde{\theta}(y) = \sup\{\theta : y(\theta) < y\} = \inf\{\theta : y(\theta) > y\}$ . This last part is not necessary if  $y(\cdot)$  is continuous, and thus  $R = [y(0), y(1)]$ .

**Claim 4** If  $(y(\theta), m(\theta))_{\theta \in \Theta}$  satisfies (10), then for any  $\hat{y} \in [y(0), y(1)]$ ,

$$\tilde{m}(\hat{y}) = \tilde{m}(y(0)) + \int_{y(0)}^{\hat{y}} \left( -\frac{\partial l^a(\tilde{\theta}(y), y)}{\partial y} \right) dy. \quad (15)$$

Conversely, if  $y(\cdot) : \Theta \rightarrow Y$  is a monotone function and  $\tilde{m}(y)$  satisfies (15), then  $(y(\theta), m(\theta))_{\theta \in \Theta} = (y(\theta), \tilde{m}(y(\theta)))_{\theta \in \Theta}$  satisfies (10).

This formula holds both for continuous actions schemes and for ones with jumps; in the latter case, for the purpose of integration, we pretend that there is a vertical segment connecting the two sides of the jump, and  $\tilde{\theta}(y)$  is constant. It is obvious that money burning is (weakly) increasing as long as the prescribed action stays below the optimal point of the agent (that is, if there is no overshooting). This result has a convenient graphical representation when the agent has a quadratic utility function. In this case,  $-\frac{\partial l^a(\tilde{\theta}(y), y)}{\partial y} = 2(\tilde{\theta}(y) + b - y)$ , therefore the change in the amount of money burning is proportional to the area between the ideal points curve of the agent and the actions scheme (with negative sign if the action scheme increases above the agent's ideal curve  $y = \theta + b$ ), as illustrated by Figure 1.

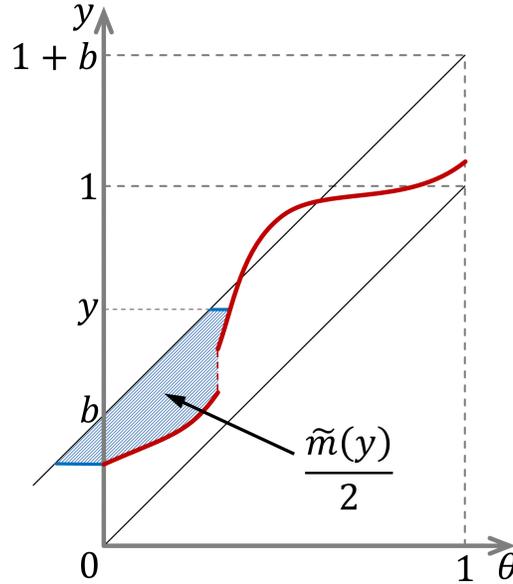


Figure 1: Representation of money burned as an integral.

The problem is thus reduced to finding a monotone action scheme  $y(\theta)$ , as then  $m(\theta)$  will be uniquely defined by (15) (also using  $\inf_{\theta \in \Theta} m^*(\theta) = 0$  implied by Claim 2). This implies the following.

**Theorem 5** *There exists a solution to problem (8)-(12).*

This result follows from Theorem 2.2 in Balder (1996).<sup>21</sup> It is not immediately applicable as  $T$  is not taken from a compact set. However, clearly, very high  $T$  is suboptimal, and by restricting  $T$  from above we can make this result applicable (see the proof in the Appendix for the formal argument).

We next establish some qualitative properties of the optimal contract. The first one, that the implemented action is never below the ideal point of the principal, holds in general. However, the other two, namely that the implemented policy is a continuous function of the state and that the implemented action is never above the ideal point of the agent, hold only under additional regularity conditions, discussed below. Most importantly for us, if there is no overshooting ( $y^*(\theta) \leq \theta + b(\theta)$ ), then money-burning has to be monotonically increasing in the state. This follows from Claim 4 and means that it is enough for us to ensure that  $m(0) \geq 0$ . (If there is overshooting, then money-burning is *not* monotonically increasing in the state, provided that  $y^*(\theta)$  is non-constant.) In the Supplementary Appendix we provide an example which shows that without the regularity conditions, the latter properties need not hold in the optimum. Intuitively, if money-burning has to be monotone, then it is very costly for the principal to impose it in low states, because he will have to increase money-burning everywhere. Overshooting in an interval of intermediate states would allow him to decrease money-burning in higher states. This sacrifices ex-post utility in the intermediate states, but decreases the level of money-burning in higher states. In the example we provide, the density function of the state takes high values in low and high states, but low values for an intermediate range of states. This violates regularity conditions and features overshooting in the optimum.

**Theorem 6** *Suppose  $(y^*(\cdot), m^*(\cdot))$  solves the problem (8)-(12) and  $y^*(\cdot)$  is continuous at  $\theta = 0$  and  $\theta = 1$ .<sup>22</sup> Then:*

1. *For every  $\theta \in \Theta$ ,  $y^*(\theta) \geq \theta$  (i.e., there is no undershooting);*
2. *If  $\frac{\partial l^p(\theta_0, y)}{\partial y} / \left( -\frac{\partial^2 l^a(\theta_0, y)}{\partial \theta \partial y} \right)$  is increasing in  $y$  for any  $\theta_0 \in \Theta$  and any  $y > \theta_0$ , then  $y^*(\cdot)$  is continuous on  $\Theta$ ;*

---

<sup>21</sup>An earlier draft of the paper contained a direct proof of existence for this model and is available from the authors upon request.

<sup>22</sup>We can always adjust  $y(\theta)$  and  $m(\theta)$  for  $\theta = 0, 1$  so that  $y(0) = \lim_{\theta \searrow 0} y(\theta)$ ,  $y(1) = \lim_{\theta \nearrow 1} y(\theta)$ ,  $m(0) = \lim_{\theta \searrow 0} m(\theta)$ ,  $m(1) = \lim_{\theta \nearrow 1} m(\theta)$ , and this will preserve (10) and other constraints and will not affect the objective function as the changes are made on the set of measure 0 only. In what follows, we restrict our attention on such contracts only. However, we cannot assume that  $y^*(\cdot)$  is continuous in general, as there are examples (see Supplementary Appendix) where this is not the case.

3. If  $f(\theta) \frac{\partial l^p(\theta, \theta + b(\theta))}{\partial y} / \left( -\frac{\partial^2 l^a(\theta, \theta + b(\theta))}{\partial \theta \partial y} \right)$  is non-decreasing in  $\theta$ , then for every  $\theta \in \Theta$ ,  $y^*(\theta) \leq \theta + b(\theta)$  (i.e., there is no overshooting).

For symmetric loss functions and constant agent bias (which is assumed in most of the literature), that is when  $l^p(\theta, y) = l(y - \theta)$  and  $l^a(\theta, y) = l(y - \theta - b)$ , the condition in part 2 simplifies to requiring that  $\frac{l''(x-b)}{l'(x)}$  is decreasing in  $x$  for  $x > 0$ . Furthermore, for any loss function of the principal that satisfies our basic assumptions (including ones with state-dependent bias), the condition is satisfied whenever the agent's loss function is quadratic, that is when  $l^a(\theta, y) = A(y - \theta - b)^2$  for some  $A > 0$ . To see this, note that in this case the denominator in the relevant expression is constant, hence the strict convexity of  $l^p$  implies that the condition holds.

A sufficient condition for the condition in part 3 to hold is that  $\frac{\frac{\partial l^p(\theta, \theta + b(\theta))}{\partial y}}{-\frac{\partial^2 l^a(\theta, \theta + b(\theta))}{\partial \theta \partial y}}$  is non-decreasing in  $\theta$  and  $f(\theta)$  is non-decreasing in  $\theta$ . For a constant bias  $b$ , the first condition holds automatically, therefore the condition is equivalent to the simple requirement that  $f(\theta)$  is non-decreasing in  $\theta$ .

We prove each part by contradiction, i.e., by assuming the contrary and finding another feasible contract which decreases principal's loss function. For Part 1, we consider increasing the action  $y(\theta)$  if  $y(\theta) < \theta$ ; this relaxes the participation constraint of the agent and decreases the loss of the principal, while simultaneously decreasing the amount of money-burning needed to keep (15) fulfilled. However, such decrease in  $m(\theta)$  may violate the nonnegativity constraint (11), and we need to adjust  $y(\theta)$  so that this does not happen.

The idea of the proof of Part 2 is to take a candidate optimal contract  $y^*(\cdot)$  that contains a discontinuity, say from  $y_1$  to  $y_2$  at  $\theta = \theta_0$ , and making it "less discontinuous" by adding an in-between action, with an amount of money-burning that attracts a small interval of types around  $\theta_0$ . This on the one hand is beneficial for the principal since types above  $\theta_0$  who choose the new action now induce an action closer to the principal's ideal point. On the other hand, types below  $\theta_0$  who choose the new action now induce an action farther away the principal's ideal point. We derive a condition for this modification of the contract to be profitable for the principal close to the limit when the interval of types attracted to the new action goes to zero (by increasing the implied money burning). It turns out that this condition always holds for strictly convex loss functions if  $y_1$  is below the agent's ideal point (if the jump involves no overshooting). Moreover, we show that the condition in Part 2 of Theorem 6 is sufficient for the inequality to hold for any kind of jump. This means that by making the jump in actions more "gradual", the principal could improve her welfare, contradicting that the optimal contract involves discontinuity.

Regarding Part 3, it is straightforward to show that it is suboptimal for the principal to specify an overshooting action at state 0: a deviation lowering the prescribed action on an

interval around 0 to the ideal curve of the agent would be in the common interest of the players and hence unambiguously increase the well-being of the principal. Let now  $\theta_0$  be the infimum of states with overshooting. We consider a deviation which keeps the implemented action on the agent's ideal curve for a small interval on the right of  $\theta_0$ . The direct effect of this would be an increase in the welfare of the principal, from the implemented action getting closer to her ideal point over the interval. However, this action would negate the decrease in money burning that the original contract would induce over the interval. We show that the condition in part 3 of Theorem 6 implies that for small enough deviations like the one specified above the deviation is beneficial for the principal, contradicting that the original contract is welfare-improving.

Note that Claim 4 and Part 3 of Theorem 6 together imply that money burning is monotonically increasing as a function of state  $\theta$  in the optimal contract (provided that the conditions in Part 3 of Theorem 6 hold). Consequently, we may assume  $m(0) = 0$ . This allows us to rewrite the maximization problem, in this case, in the following simpler way. Taking into account (13), we have

$$\begin{aligned} \int_{\Theta} L^a(\theta) f(\theta) d\theta &= \int_0^1 \left( L^a(0) + \int_0^\theta \frac{\partial l^a(\xi, y(\xi))}{\partial \theta} d\xi \right) f(\theta) d\theta \\ &= L^a(0) + \int_0^1 \int_\xi^1 \frac{\partial l^a(\xi, y(\xi))}{\partial \theta} f(\theta) d\theta d\xi = L^a(0) + \int_0^1 \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} (1 - F(\theta)) d\theta. \end{aligned}$$

Now  $m(0) = 0$  implies  $L^a(0) = l^a(0, y(0))$ , and thus the optimization problem is equivalent to the following one:

$$\begin{aligned} &\min_{T, \{y(\theta)\}} \int_0^1 l^p(\theta, y(\theta)) dF(\theta) + T && (16) \\ \text{s.t. } &l^a(0, y(0)) + \int_0^1 \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} (1 - F(\theta)) d\theta \leq T - \tilde{u} \\ &y(\cdot) \text{ is non-decreasing and continuous,} \\ &\theta \leq y(\theta) \leq \theta + b(\theta) \\ &T \geq \tilde{w}. \end{aligned}$$

The rewritten form of the optimization problem has the advantage that the incentive constraints are incorporated in a simpler integral constraint. The reformulation indicates that there can be a trade-off between decreasing the first term  $l^a(0, y(0))$ , which is minimized at  $y(0) = b(0)$ , and the second integral term, which can be minimized pointwise with the minimizing  $y(0)$  being possibly strictly below  $b(0)$ . The trade-off is caused by the requirement that  $y(\cdot)$  is non-decreasing and continuous. Intuitively, this reflects the tension between minimizing the agent's loss (from money burning and from the implemented policy being away from the agent's

ideal point), which serves the purpose of decreasing the ex-ante transfer to the agent, and the principal’s loss from the implemented policy being away from the principal’s ideal point.

## 5 The optimal contract in uniform-quadratic settings

The reformulated problem is tractable enough that we can explicitly solve for the optimal contract in a class of games that are generalizations of the canonical uniform-quadratic specification of the delegation problem. We will derive economic insights from how the optimal contract and in particular money burning is imposed on the agent depends on various parameter values of the problem.

Specifically, we assume that  $\theta$  is distributed uniformly on  $[0, 1]$  and restrict ourselves to the quadratic loss functions

$$l^p(\theta, y) = A(y - \theta)^2, \quad (17)$$

$$l^a(\theta, y) = (y - \theta - b)^2, \quad (18)$$

where  $A, b > 0$ . These loss functions imply that the agent has a constant bias  $b(\theta) = b$ . Parameter value  $A = 1$  corresponds to the uniform-quadratic example frequently used in the literature. The extra parameter  $A$  allows us to change the sensitivity of the loss function of the principal relative to the sensitivity of the loss function of the agent, independently of the size of bias. Values  $A < 1$  imply that the principal is less sensitive to policy choice than the agent (the same deviation from the ideal point means a smaller loss); values  $A > 1$  imply the opposite.<sup>23</sup> Alternatively, if we interpret the principal and the agents as firms/organizations rather than individuals, then  $A$  may be interpreted as the ratio of people in the principal organization to that in the agent organization. In other words, a large  $A$  may be thought of as a large corporation (or the public sector) employing an individual, and a small  $A$  corresponds to an individual delegating the task to a large organization. As we show below, the qualitative features of the optimal contract, including whether money burning is used in equilibrium, depend crucially on this parameter. In particular, if  $b < 1$ , which is the typical scenario in the delegation literature, money-burning will *always* be part of the optimal contract if  $A > 1$ .

Under these assumptions on the functional form, both conditions in Theorem 6 are satisfied. Let  $\mathcal{F}$  be the set of all continuous functions  $f : [0, 1] \rightarrow \mathbb{R}$ , which are nondecreasing and satisfy

---

<sup>23</sup>Martimort and Semenov (2007) introduce a similar multiplicative parameter for a policymaker in an unrelated model of lobbying.

$\theta \leq f(\theta) \leq \theta + b$ . The problem (16) may be rewritten as

$$\min_{(y(\cdot), T) \in \mathcal{F} \times \mathbb{R}} \int_0^1 A(y(\theta) - \theta)^2 d\theta + T, \quad (19)$$

$$\text{s.t. } (y(0) - b)^2 - \int_0^1 2(y(\theta) - \theta - b)(1 - \theta) d\theta - T + \tilde{u} \leq 0, \quad (20)$$

$$\tilde{w} - T \leq 0. \quad (21)$$

In what follows, we formulate and prove the following result, and then use it to study comparative statics. The following auxiliary function is critical:

$$\Omega(A, b) = \begin{cases} \frac{8}{3} \left( \frac{Ab}{A+1} \right)^3 & \text{if } A < \min\left(1, \frac{1}{2b-1}\right); \\ \frac{1}{3} A^2 b^3 & \text{if } 1 \leq A \leq \frac{1}{b}; \\ \frac{1}{12} + \left( \frac{Ab}{A+1} \right)^2 & \text{if } \frac{1}{2b-1} \leq A \leq 2b-1; \\ b^2 + b - \omega(A, b) & \text{if } A > \max\left(\frac{1}{b}, 2b-1\right), \end{cases}$$

where we define  $\omega(A, b)$  by

$$\omega(A, b) = \frac{(10 + 6A^3b - 2A - 8A^2b - 6A^2) \sqrt{\frac{2A^2b - A - 1}{A-1}} + 2A^3 + 18A^2b + 3A^2 - 8A - 10}{3A^4}.$$

For each  $b$ ,  $\Omega(A, b)$  is an increasing function of  $A$ , tending to 0 as  $A \rightarrow 0$  and to  $b^2 + b$  as  $A \rightarrow \infty$ . As  $b$  increases for any fixed  $A$ ,  $\Omega(A, b)$  increases from 0 to  $\infty$ .

**Theorem 7** *There exists a unique solution to the problem (19) s.t. (20), (21). This solution takes the following form:*

1. *If  $\tilde{w} - \tilde{u} < \Omega(A, b)$  then the agent receives a transfer larger than minimal. The exact solution is given by (we give only  $y^*(\cdot)$  and  $T^*$ , and the corresponding  $m^*(\cdot)$  may be found using Claim 4):*

(a) *If  $A \leq \min\left(1, \frac{1}{2b-1}\right)$  then the contract takes the form of “free choice with a cap”, and there is no money-burning; more precisely,*

$$\begin{aligned} y^*(\theta) &= \min\left(\theta + b, 1 + \frac{1-A}{1+A}b\right); \\ T^* &= \frac{8}{3} \left( \frac{Ab}{A+1} \right)^3 + \tilde{u}; \end{aligned}$$

(b) *If  $1 < A < \frac{1}{b}$  (this is possible for  $b < 1$  only) then the contract takes the form of “free choice, then money burning”; more precisely,*

$$\begin{aligned} y^*(\theta) &= \min\left(\theta + b, \frac{1}{A} + \left(1 - \frac{1}{A}\right)\theta\right); \\ T^* &= \frac{1}{3} A^2 b^3 + \tilde{u}; \end{aligned}$$

(c) If  $\frac{1}{2b-1} < A < 2b - 1$  (this is possible for  $b > 1$  only) then the contract is a constant and there is no money-burning; more precisely,

$$\begin{aligned} y^*(\theta) &= \frac{1}{2} + \frac{b}{A+1}; \\ T^* &= \frac{1}{12} + \left(\frac{Ab}{A+1}\right)^2 + \tilde{u}; \end{aligned}$$

(d) If  $A \geq \max\left(\frac{1}{b}, 2b - 1\right)$  then the contract takes the form “free action, then money burning”; more precisely,

$$\begin{aligned} y^*(\theta) &= \max\left(\frac{1}{A} + \left(1 - \frac{1}{A}\right)\theta, \frac{1 + \sqrt{(A-1)(2A^2b - A - 1)}}{A^2}\right); \\ T^* &= b^2 + b - \omega(A, b) + \tilde{u}; \end{aligned}$$

2. If  $\Omega(A, b) \leq \tilde{w} - \tilde{u} < b^2 + b$  then the agent receives a minimal transfer,  $T^* = \tilde{w}$ . To find  $y(\cdot)$  and  $m(\cdot)$ , consider the “effective principal’s weight”  $D = D(\tilde{w} - \tilde{u}, b)$ , which is the unique solution to equation  $\Omega(D, b) = \tilde{w} - \tilde{u}$  (this solution satisfies  $D \geq A$ ). Then  $y^*(\cdot)$  and  $m^*(\cdot)$  are the same as in the previous case, with  $A$  replaced by  $D$ ; namely (we characterize  $y(\cdot)$  only):

(a) If  $\tilde{w} - \tilde{u} \leq \min\left(\frac{1}{3}, \frac{b^3}{3}\right)$  then the contract takes the form of “free choice with a cap”, and there is no money-burning; more precisely,

$$y^*(\theta) = \min\left(\theta + b, 1 + \frac{1-D}{1+D}b\right);$$

(b) If  $\frac{b^3}{3} < \tilde{w} - \tilde{u} < \frac{b}{3}$  (this is possible for  $b < 1$  only) then the contract takes the form of “free choice, then money burning”; more precisely,

$$y^*(\theta) = \min\left(\theta + b, \frac{1}{D} + \left(1 - \frac{1}{D}\right)\theta\right);$$

(c) If  $\frac{1}{3} < \tilde{w} - \tilde{u} < b^2 - b + \frac{1}{3}$  (this is possible for  $b > 1$  only) then the contract is a constant and there is no money-burning; more precisely,

$$y^*(\theta) = \frac{1}{2} + \frac{b}{D+1};$$

(d) If  $\max(b^2 - b, 0) + \frac{\min(b, 1)}{3} \leq \tilde{w} - \tilde{u} < b^2 + b$  then the contract takes the form “free action, then money burning”; more precisely,

$$y^*(\theta) = \max\left(\frac{1}{D} + \left(1 - \frac{1}{D}\right)\theta, \frac{1 + \sqrt{(D-1)(2D^2b - D - 1)}}{D^2}\right);$$

3. If  $\tilde{w} - \tilde{u} \geq b + b^2$  then principal's ideal action is implemented:  $y^*(\theta) = \theta$ , and money-burning in state  $\theta$  is  $m^*(\theta) = \theta b$ . The agent receives a minimal transfer  $T^* = \tilde{w}$ .

Figures 2 and 3 illustrate the optimal contract for specific parameter regions.

Existence follows from Claim 5, while uniqueness follows, since (19)–(21) is a maximization problem on a convex domain with convex constraints. The simplest way to understand the characterization result is the following. The space of parameter values,  $(A, b, \tilde{u}, \tilde{w})$ , is split into three areas, depending on which of the two constraints, the participation constraint (20) and the minimum wage constraint (21), are binding. For  $\tilde{w} - \tilde{u} < \Omega(A, b)$ , only the participation constraint is binding, and the principal is paying the agent more than a minimal transfer. It is more likely to be satisfied for large  $A$  or large  $b$ , as well as for small  $\tilde{w} - \tilde{u}$  (in particular, if  $\tilde{w} \leq \tilde{u}$ , then the minimal wage is lower than the outside option and is thus non-binding). For  $\Omega(A, b) < \tilde{w} - \tilde{u} < b^2 + b$ , both the participation and the minimum wage constraints bind; the agent gets the minimum wage and a contract that makes him indifferent between working for the principal and not. Finally, if  $\tilde{w} - \tilde{u} > b^2 + b$ , the minimum wage constraint is binding, but the participation constraint is not; in this case, the principal implements his ideal action, the agent gets expected disutility  $-b - b^2$ , but is still willing to work for the minimum wage as he has too low outside option. The precise contract  $y^*(\cdot)$  is depicted on Figure 2 for  $b < 1$  and on Figure 3 for  $b > 1$ ; the dotted green lines correspond to  $y^*(\theta)$  for  $\theta$  such that  $m^*(\theta) = 0$ , and the solid red lines correspond to  $y^*(\theta)$  for  $\theta$  such that  $m^*(\theta) > 0$ .

The explicit characterization makes it easy to obtain comparative statics results.

**Theorem 8** *In the uniform-quadratic setting, the following is true in the optimal contract:*

1. *Money-burning is used if either (a)  $A > 2 \max(b, 1) - 1$  or (b)  $\tilde{w} - \tilde{u} > \frac{b^3 - (\max(b, 1) - 1)^3}{3}$ . If there is money-burning for some parameters  $(A, b, \tilde{u}, \tilde{w})$ , then there is also money-burning for higher  $A$ , higher  $\tilde{w}$ , lower  $\tilde{u}$ , or lower  $b$ .*
2. *The total amount of money-burning is (weakly) increasing in  $A$  and  $\tilde{w}$ , decreasing in  $\tilde{u}$ ; the effect of  $b$  is ambiguous.*
3. *The agent receives higher-than-minimum wage  $T^* > \tilde{w}$  if and only if  $\tilde{w} - \tilde{u} < \Omega(A, b)$ . If this is the case for  $(A, b, \tilde{u}, \tilde{w})$ , then it is also the case for higher  $A$ ,  $b$ ,  $\tilde{u}$ , or for lower  $\tilde{w}$ .*
4. *The principal implements his ideal contract if  $\tilde{w} - \tilde{u} \geq b^2 + b$ ; this condition does not depend on  $A$ .*

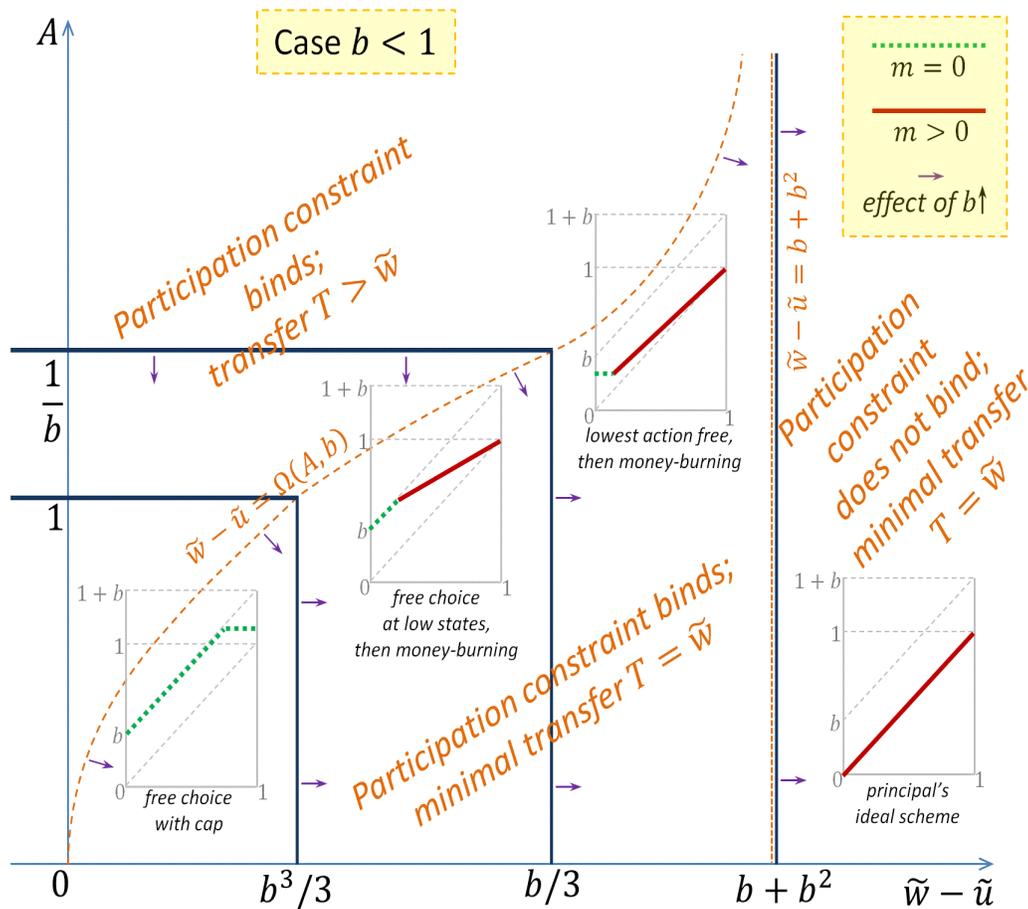


Figure 2: Optimal contract if  $b < 1$ .

This result implies that a high minimal wage increases both the incidence of money-burning and the expected amount, while a higher outside option decreases both. This is intuitive: a higher  $\tilde{w} - \tilde{u}$  relaxes the participation constraint, making the agent willing to work for the principal, which the latter compensates by both enforcing a more favorable outcome  $y(\cdot)$  (less favorable for the agent) and more money-burning. A higher  $A$  means the principal is willing to set  $y(\cdot)$  closer to his ideal, which requires more money-burning, and is willing to compensate the agent by a higher wage. As a result, a higher  $A$  only affects the contract if the agent is given higher than minimum wage (the next theorem formalizes this). The effect of  $b$  is ambiguous. On the one hand, a lower  $b$  makes the use of money-burning more likely. On the other hand, it decreases the absolute cap on expected money-burning,  $b^2 + b$  (this amount is sufficient to enforce the principal's ideal contract). Thus, a decrease in  $b$  may increase or decrease the expected money-burning.

The principal will pay more than the minimal transfer if the agent has a high outside option

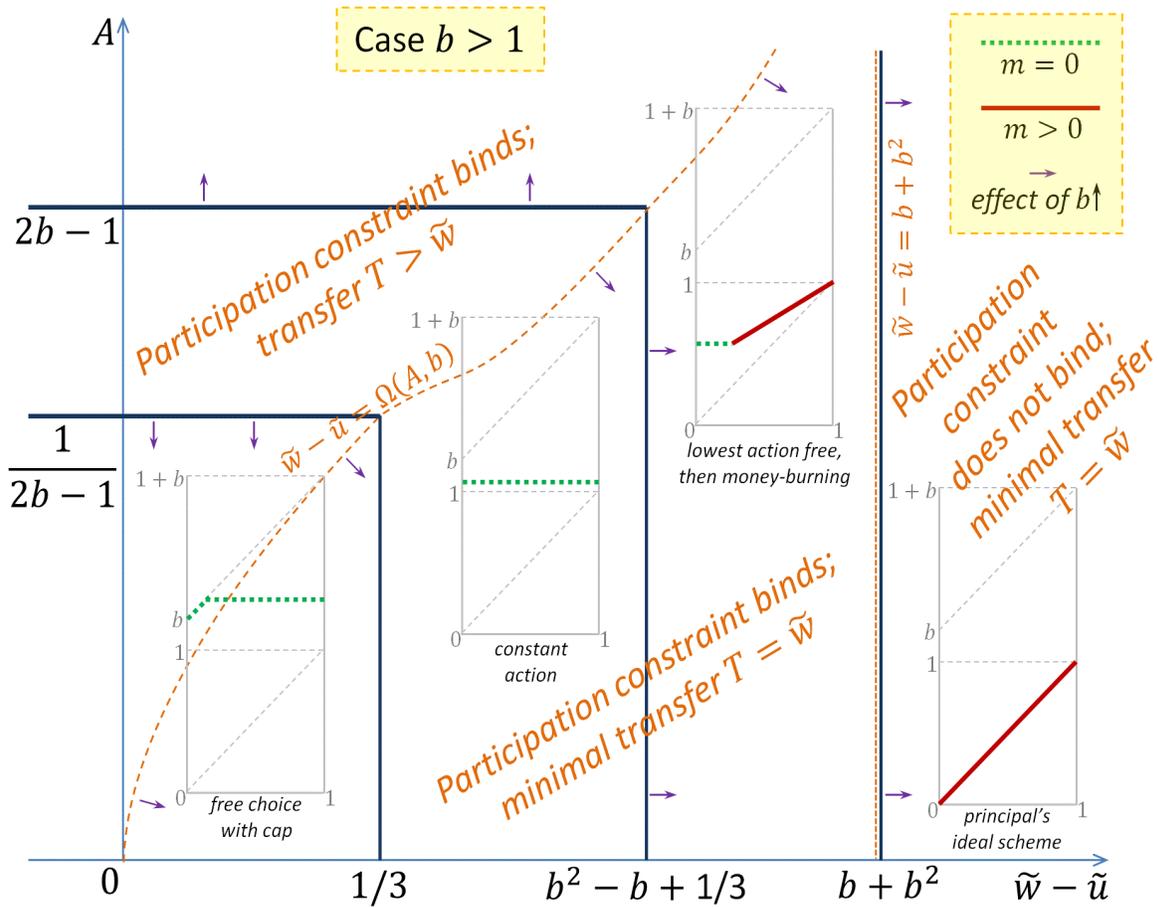


Figure 3: Optimal contract if  $b > 1$ .

(this is natural), or if the action is important for the principal (here, the principal will make the agent choose the actions that the agent dislikes and make him burn money, and he has to be compensated for that). A higher  $b$  also makes a high wage more likely, because it takes a higher amount of money-burning to keep the implemented actions closer to the principal's ideal line, which needs to be compensated through a higher wage. Not surprisingly, a higher  $\tilde{w}$  makes more likely that the minimum wage is paid. It should be emphasized, however, that paying a higher-than-minimum wage is fully compatible with money-burning. For example, if  $b < 1 < A$  and  $\tilde{w} < \tilde{u}$ , the principal's payoff is sufficiently important to provide strong incentives to the agent, but the binding constraint is his outside option rather than the minimum wage. Thus, transfers higher than minimal do not crowd out money-burning.

The next result shows that all comparative statics results depend on whether the agent receives the minimum wage or a higher wage.

**Theorem 9** *Consider the uniform-quadratic setting, and suppose that  $\tilde{w} - \tilde{u} < \Omega(A, b)$ , so in the optimal contract, the agent receives a higher-than-minimum wage. Then a marginal change in  $\tilde{w}$  or  $\tilde{u}$  does not change the action  $y^*(\cdot)$  or money-burning  $m^*(\cdot)$  (moreover, a marginal change in  $\tilde{w}$  does not affect the optimal contract, whereas an increase in  $\tilde{u}$  is matched by a higher transfer, one-to-one). An increase in  $A$  changes the optimal contract: it brings  $y(\cdot)$  closer to the principal's ideal action schedule, weakly increases the amount of money-burning, and compensates the agent by a higher wage.*

*Now suppose that  $\tilde{w} - \tilde{u} > \Omega(A, b)$ , so in the optimal contract, the agent receives the minimum wage. Then a marginal change of  $A$  has no effect on the contract. An increase in  $\tilde{w}$ , or a decrease in  $\tilde{u}$ , move  $y^*(\cdot)$  closer to the principal's ideal action schedule, and increase the total amount of money-burning.*

We see that if the agent receives the minimum wage, his contract will depend on  $\tilde{w} - \tilde{u}$ ; an increase of this difference will increase money-burning and bring  $y(\cdot)$  closer to the principal's ideal. But an increase in  $A$  will not change the contract: the principal is already getting from the agent everything he can for the minimum wage, and he will not get more until  $A$  becomes high enough so that he is ready to pay extra to the agent. However, once the principal pays a higher wage, the comparative statics becomes different. At this point, a marginal change in  $\tilde{w}$  plays no role, and a change in  $\tilde{u}$  is matched by a higher wage. An increase in  $A$  would lead to lower  $y^*(\cdot)$  (closer to the principal's ideal), more money-burning, and also a higher wage.

The following thought experiment is of interest. Suppose that an agent, who is working for the principal, got an increase of his outside option  $\tilde{u}$ , for example, by getting an outside offer. Suppose that the principal is willing to keep the agent. Then the way he will do so will be different, depending on whether the agent receives the minimum wage allowed or not. If the wage is at the minimum, the principal responds by giving the agent more freedom, so  $y^*(\cdot)$  becomes closer to the agent's ideal action scheme, and there is less money-burning. However, if the wage is higher than minimum, the principal just increases the wage. To get a further intuition, imagine that this scenario repeats multiple times (say, the agent gets experience and better outside offers over time). Then in the first few instances, the agent is retained by better working conditions, and starting at some point he just gets raises, while the action and money-burning, if any, remains the same.

Finally, we present a result on the utilities of the two players. If the agent's participation constraint binds, he receives  $\tilde{u}$ , and any changes in parameters affect the principal's utility only. All parameters  $\tilde{u}, \tilde{w}, A, b$  decrease the principal's utility:  $\tilde{u}$  does so because the agent needs to be compensated by a higher wage or a contract which is more beneficial to him,  $\tilde{w}$  has a negative

impact whenever the minimum wage constraint binds. A higher  $A$  increases the principal's disutility from actions away from the principal's ideal scheme, so it has a negative impact; a higher  $b$  increases the conflict of interest, which also hurts both players. The only case where the agent's IC constraint does not bind is where the principal is able to implement his ideal point, in which case the agent receives  $\tilde{w} - b^2 - b$ , and the principal pays this minimum wage, thus getting  $-\tilde{w}$ . In this case, the principal only cares about the minimum wage; the agent's utility increases in the minimum wage and decreases in  $b$  (as he has to implement the policy further from his ideal and burn more money). Thus, if the participation constraint does not bind, the contract does not depend on  $A$  or  $\tilde{u}$ . This is summarized in the following theorem.

**Theorem 10** *In the uniform-quadratic setting, in the optimal contract, the agent's utility equals*

$$W_A^* = \max(\tilde{u}, \tilde{w} - b^2 - b).$$

*The agent gets utility higher than his outside option if and only if the principal's ideal action is implemented. In this case, the principal's utility is  $-\tilde{w}$ , and an increase in minimum wage results in a transfer from the principal to the agent.*

*The principal's utility is weakly decreasing in  $\tilde{u}$ ,  $\tilde{w}$ ,  $A$ , and  $b$ . It reaches its maximum,  $-\tilde{w}$ , whenever  $\tilde{w} - \tilde{u} \geq b^2 + b$ .*

It is instructive to compare these optimal contracts with the ones obtained in Krishna and Morgan (2008) – from now on KM – for the case of delegation with one-sided transfers, no participation constraint, and symmetric quadratic loss functions (corresponding to  $\tilde{u} = -\infty$  and  $A = 1$ , without the possibility of money-burning). In this environment, the optimal transfer scheme sets a positive transfer to the agent when choosing low actions, and it is monotonically decreasing. This is parallel to our results that the money burning scheme specifies zero money burning at the lowest implemented action, and that it is monotonically increasing. Furthermore, the implemented action scheme is monotonically increasing in both models, with a possible cap on the highest action that can be chosen by the agent. An important qualitative difference is that while contingent monetary transfers are always used to some extent in the optimal contract in KM, money burning might be a too costly incentive device for the principal and hence not used at all in the optimal contract. This results from the fact that as opposed to our model, there is no participation constraint in KM.<sup>24</sup>

---

<sup>24</sup>Other relevant benchmarks include Melumad and Shibano (1991) and Kovač and Mylovánov (2009). Melumad and Shibano (1991) consider a delegation problem without transfers and money-burning. In the setting with commitment (where the principal can commit to a certain action as a function of the agent's message, as in delegation but not cheap talk models), our models share a common particular case. If we assume uniform-

## 6 Delegation with both Conditional Transfers and Money Burning

So far, we have ruled out transfers other than a fixed wage to focus on money-burning. In this section, we allow for transfers conditional on the agent's report (and, by revelation principle, on the state  $\theta$ ). In other words, we will study the same problem, except that the transfer  $t(\theta)$  may depend on the state  $\theta$  rather than satisfy the restriction  $t(\theta) \equiv T$ . As before, this transfer must satisfy the minimum wage requirement, which we assume to hold in each state:  $t(\theta) \geq \tilde{w}$  for all  $\theta \in \Theta$ . The principal therefore has two means to incentivize the agent: money-burning or contingent monetary transfers. In what follows, we show that even though the former is a less efficient way to create incentives for the agent than the latter, money-burning nevertheless can be used in the optimal contract (at different states). In fact, depending on parameter values, both means of providing incentives may be used, or only one of them. The primary factors determining which case applies are once again: (i) the outside option of the agent  $\tilde{u}$ , (ii) the minimum wage  $\tilde{w}$ , and (iii) the relative importance  $A$  of the action choice for the principal versus the agent.

Formally, in this Section the contract is given by a triple  $(y(\cdot), m(\cdot), t(\cdot))$  consisting of policy  $y(\theta)$ , money burnt by the agent  $m(\theta) \geq 0$ , and transfer from the principal to the agent  $t(\theta) \geq \tilde{w}$ . The agent's loss function is now given by

$$L^a(\theta) = l^a(\theta, y(\theta)) + m(\theta) - t(\theta). \quad (22)$$

The principal solves the following problem (we immediately write it as a minimization problem,

---

quadratic setting with  $A = 1$  in our model (and also that the participation constraint is not binding), this would correspond to the case where parameter  $a = 1$  and  $k < 0$  in their model. In this case, our contract (for the interesting case  $b < 1$ ) takes the form  $y(\theta) = \min\{\theta + b, 1\}$ , coinciding with their contract.

Kovač and Mylovanov (2009) do not have money-burning, but they explicitly allow for stochastic mechanisms. They assume the agent to have a quadratic utility function, which means that if the action is stochastic, the agent loses an extra term corresponding to the variance, which could be interpreted as money burning if the principal were not risk-averse as well and did not get disutility from non-deterministic contract herself. This means that the principal could do at least as well with money-burning as with stochastic contracts; moreover, once we allow money-burning, considering deterministic contracts, as we do, is without loss of generality. As expected, our contracts coincide if the principal also has a quadratic utility function with a constant bias, and  $A = 1$ , and there is no participation constraint. More generally, if the bias in the uniform-quadratic setting is not constant (but the parameter  $A$  still equals 1), the authors show that the optimal contract allows the agent to implement his ideal action, as long as it is above some floor and below some floor. We do not get a similar result because in our model, money-burning does not directly affect the principal's payoff, and in general he can choose among a broader set of contracts. However, whenever money-burning is not optimal (for  $A < 1$ ), our optimal contract exhibits this property.

similar to (8)–(12)):

$$\min_{\{y(\theta), m(\theta), t(\theta)\}_{\theta \in \Theta}} \int_{\Theta} (l^p(\theta, y(\theta)) + t(\theta)) f(\theta) d\theta \quad (23)$$

$$\text{s.t. } \int_{\Theta} (l^a(\theta, y(\theta)) + m(\theta) - t(\theta)) f(\theta) d\theta \leq -\tilde{u}, \quad (24)$$

$$\forall \theta, \theta' \in \Theta : l^a(\theta, y(\theta)) + m(\theta) - t(\theta) \leq l^a(\theta, y(\theta')) + m(\theta') - t(\theta'), \quad (25)$$

$$\forall \theta \in \Theta : m(\theta) \geq 0, t(\theta) \geq \tilde{w}. \quad (26)$$

Many of the properties of the problem (23)–(26) are analogous to the case with fixed transfer. For example, under the same conditions as in Theorem 6,  $y^*(\cdot)$  is continuous and satisfies  $\theta \leq y^*(\theta) \leq \theta + b(\theta)$ . Moreover,  $m^*(\theta)$  is non-decreasing on  $\Theta$  and  $t^*(\theta)$  is non-increasing, and there is  $\theta_0 \in \Theta$  such that  $m^*(\theta_0) = 0$  and  $t_0 = t^*(\theta_0) = \min_{\theta \in \Theta} t(\theta)$ ; in other words, there is state  $\theta_0$  where there is no money-burning, and the transfer is minimal among all other states (it typically, but not always, satisfies  $t(\theta_0) = \tilde{w}$ ). If money burning is used, it is used at high states ( $\theta > \theta_0$ ), while conditional transfers will be made in low states ( $\theta < \theta_0$ ). We can then rewrite the problem in a tractable way (details are available in Supplementary Appendix). The following is the result of these transformations for the uniform-quadratic case, which we focus on from now on:

$$\min_{(y(\cdot), \theta_0, t_0) \in \mathcal{F} \times [0, 1] \times \mathbb{R}} \int_0^1 A(y(\theta) - \theta)^2 d\theta + \quad (27)$$

$$\int_0^{\theta_0} \left( (y(\theta) - \theta - b)^2 - 2(y(\theta) - \theta - b)\theta \right) d\theta - (y(\theta_0) - \theta_0 - b)^2 \theta_0 + t_0,$$

$$\text{s.t. } (y(\theta_0) - \theta_0 - b)^2 - \int_0^1 2(y(\theta) - \theta - b)(1 - \theta) d\theta + \int_0^{\theta_0} 2(y(\theta) - \theta - b) d\theta - t_0 + \tilde{u} \leq 0, \quad (28)$$

$$\tilde{w} - t_0 \leq 0. \quad (29)$$

The explicit characterization of the optimum is cumbersome, and it is relegated to the Supplementary Appendix. Here we only state a result that establishes some basic characteristics of the optimal contract.

**Theorem 11** *There exists a unique solution to the problem (23)–(26). More precisely:*

1. *If  $\tilde{w} - \tilde{u} < -Ab \frac{1+A-Ab}{(1+A)^2}$ , then for all  $\theta$  the agent receives a transfer larger than minimal, and there is no money-burning. The optimal contract implements*

$$y^*(\theta) = \frac{b}{A+1} + \theta,$$

which maximizes the joint utility of the agent and the principal. For an arbitrary  $\theta$ ,  $t^*(\theta) = \tilde{u} + \left(\frac{Ab}{A+1}\right)^2 - \frac{Ab}{A+1}\theta$  and  $m^*(\theta) = 0$ .

2.  $-Ab\frac{1+A-Ab}{(1+A)^2} \leq \tilde{w} - \tilde{u} < b^2 + b$ , then the agent receives a minimal transfer at  $\theta = 1$ . A precise characterization of the set of states for which  $t^*(\theta) > \tilde{w}$  or  $m^*(\theta) > 0$  is available in the Supplementary Appendix.
3. If  $\tilde{w} - \tilde{u} \geq b + b^2$ , then principal's ideal action is implemented:  $y^*(\theta) = \theta$ , and money-burning in state  $\theta$  is  $m^*(\theta) = \theta b$ . The agent receives a minimal transfer  $t^*(\theta) = \tilde{w}$  for all  $\theta$ .

While a complete characterization is in the Supplementary Appendix, Figure 4 provides an example of what the optimal scheme  $y(\cdot)$  looks like, and whether and when money-burning and contingent transfers are used, if  $b$  is sufficiently small. Several observations are in place. First, one of the incentive schemes – either contingent transfers or money-burning – is always used (this is not the case for  $b > 1$ , where a constant scheme with no incentives is possible for some parameter values). Money-burning is being used if and only if  $y^*(1) = 1$ , and if  $y^*(1) > 1$ , then there is no money-burning and the principal never gets his ideal point. Contingent transfers are being used if and only if  $y^*(0) = \frac{b}{A+1}$ , which is the action that maximizes the joint utility of the principal and the agent for  $\theta = 0$ ; otherwise,  $y^*(0) < \frac{b}{A+1}$  and contingent transfers are not being used.

Using this characterization, we can obtain the comparative statics results:

**Theorem 12** *In the uniform-quadratic setting with contingent transfers, the following is true in the optimal contract:*

1. The social optimum is implemented if and only if

$$\tilde{w} - \tilde{u} \leq -Ab\frac{1+A-Ab}{(A+1)^2}.$$

2. The principal implements his ideal contract if  $\tilde{w} - \tilde{u} \geq b^2 + b$ ; this condition does not depend on  $A$ .
3. If money-burning is used in the optimal contract for some values of  $(A, b, \tilde{u}, \tilde{w})$  then it is also used for higher  $\tilde{w}$ , lower  $\tilde{u}$ , or higher  $A$ .
4. If contingent transfer is used in the optimal contract for some values of  $(A, b, \tilde{u}, \tilde{w})$  then it is also used for lower  $\tilde{w}$ , higher  $\tilde{u}$ , or higher  $A$ .

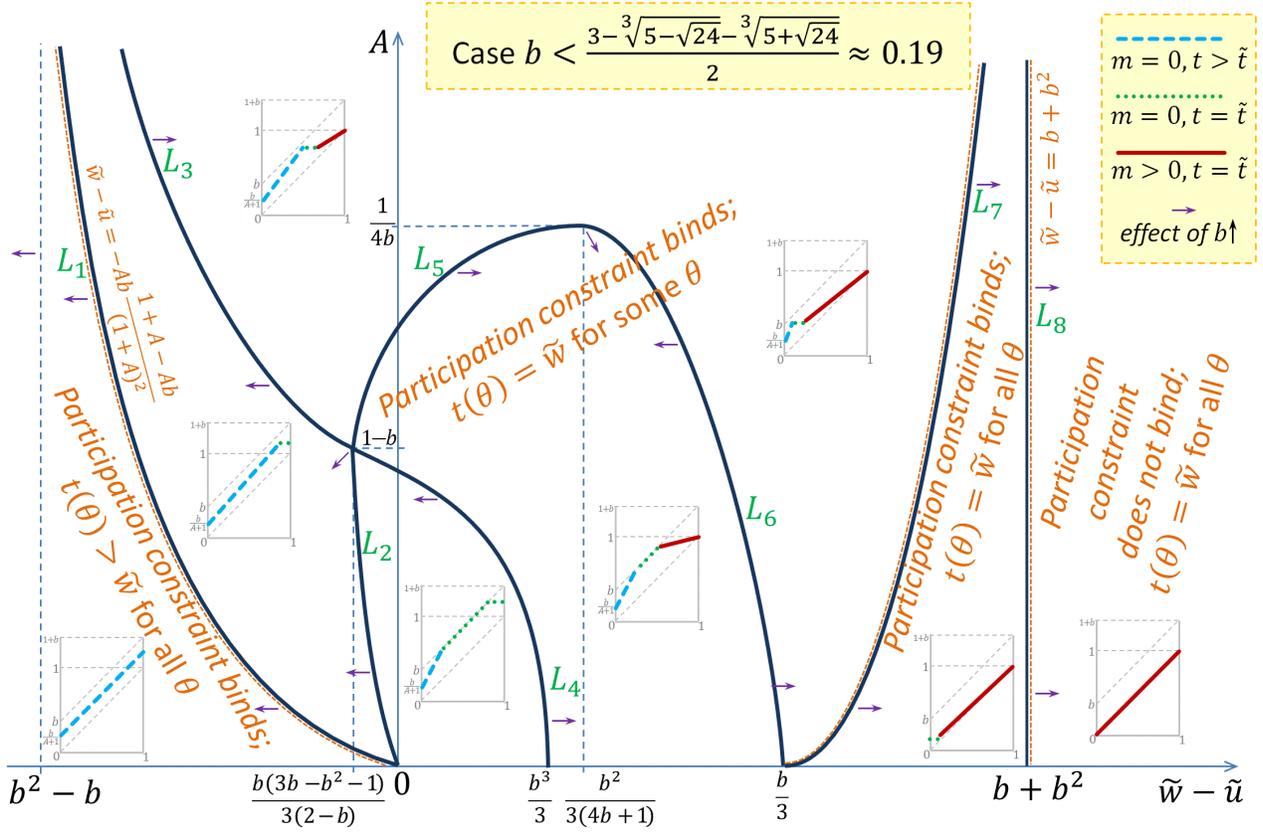


Figure 4: Optimal contract with transfers for  $b$  sufficiently low.

5. The effect of  $b$  on the use of either of the instruments (money-burning or contingent transfers) is ambiguous.

These results are intuitive. First, if the minimal wage requirement is sufficiently small or the agent's outside option is sufficiently high, then the social optimum is implemented and, moreover, only contingent transfers are used (i.e., no money-burning). Conversely, if the minimal wage is sufficiently high or the agent's outside option is sufficiently small, the principal implements his ideal action, and does so through money-burning. Indeed, for such combination of parameters, the agent is 'overpaid', and making the agent burn enough money so that the principal's ideal is implemented is costless for the principal. The use of the two instruments clearly depends on the parameters. If the minimal wage increases, the principal is more willing to adopt minimal wage, and subsequently abandon conditional transfers. If the agent's outside option increases, then he needs to be compensated more, and the principal does so by using contingent transfers, while at the same time abandoning money-burning. Interestingly, a higher  $A$  makes the use of both instruments more likely: intuitively, it just increases the importance of having  $y(\cdot)$  closer

to the principal's ideal, and the principal will use all means available to achieve this.

The welfare implications of contingent transfers are summarized in the following theorem.

**Theorem 13** *In the uniform-quadratic setting with contingent contracts, in the optimal contract, the agent's expected payoff is:*

$$W_A^* = \max(\tilde{u}, \tilde{w} - b^2 - b).$$

*The agent's expected payoff is higher than his outside option if and only if the principal's ideal action is implemented. The principal's utility is weakly decreasing in  $\tilde{u}$ ,  $\tilde{w}$ ,  $A$ , and  $b$ .*

*Furthermore, for a given tuple of parameters  $(A, b, \tilde{u}, \tilde{w})$ , allowing for contingent transfers does not affect the utility of the agent, while making the principal weakly better off (strictly if he uses contingent transfers in the optimal contract with contingent transfers).*

We conclude the section by pointing out a nonmonotonicity of the implemented action scheme in the optimal contract, as a function of the outside option of the agent. If the outside option is very low, then for any  $A$ , the implemented action scheme is equal to the principal's ideal line. As the outside option increases, the implemented action scheme shifts towards the principal's ideal points. As a result, for low values of  $A$ , there is an intermediate range of outside options in which the agent can choose his ideal action in a large set of states. However, a further increase in the outside option results in the jointly optimal action being implemented in all states, meaning that the agent cannot choose his ideal action in any of the states.

## 7 Conclusion

Our model of delegation with nonmonetary transfers may be developed in many different directions. Monetary versus nonmonetary incentives are extensively discussed in the economics of crime literature, starting from Becker (1968).<sup>25</sup> The models offered in this literature differ in many crucial aspects from ours: for example it is assumed that in the absence of any deterrents every criminal would choose the highest possible crime activity level, while our approach would assume that the optimal crime activity, from the criminal's viewpoint, is state-specific. Applying the delegation framework in this area might provide new insights on the structure of optimal monetary fines and prison sentences.

An intriguing question is that why bureaucratic procedures and paperwork seem to be the primary types of costly activity that organizations impose on their workers. A possible explanation for this, which we would like to investigate in future research, is that bureaucratic

---

<sup>25</sup>See for example Shavell (1987), Mookherjee and Png (1994), and Levitt (1997).

paperwork has the feature that the same level of activity is less costly in higher states. For example, when applying for a research grant requires turning in a long proposal, writing the proposal is less costly for an applicant who indeed has a good idea for a research project than for one who does not. Similarly, when an employee has to explain it in a report when taking a guest to an expensive restaurant from corporate budget, writing the report is less costly for employees who indeed had good reasons to select the expensive restaurant. This suggests incorporating costs of lying as in Kartik (2009) into our model.

## 8 References

Abito, J., D. Besanko and D. Diermeier (2015): “Corporate reputational dynamics and activist pressure,” *Advances in Strategic Management*, forthcoming.

Aghion, P. and J. Tirole (1997): “Formal and real authority in organizations,” *Journal of Political Economy*, 105, 1-29.

Alonso, R. and N. Matouschek (2008): “Optimal delegation,” *Review of Economic Studies*, 75, 259-293.

Alonso, R. and N. Matouschek (2007): “Relational delegation,” *Rand Journal of Economics*, 38, 1070-1089.

Amador, M. and K. Bagwell (2013a): “The theory of optimal delegation with an application to tariff caps,” *Econometrica*, 81, 1541-1599.

Amador, M. and K. Bagwell (2013b): “Money burning and the theory of delegation,” mimeo, Stanford University.

Amador, M., I. Werning and M. Angeletos (2006): “Commitment vs. flexibility,” *Econometrica*, 74, 365-396.

Armstrong, M. and J. Vickers (2010): “A model of delegated project choice,” *Econometrica*, 78, 213-244.

Athey, S., A. Atkeson and P. Kehoe (2005): “The optimal degree of discretion in monetary policy,” *Econometrica*, 73, 1431-1475.

Athey, S., K. Bagwell and C. Sanchirico (2004): “Collusion and price rigidity,” *Review of Economic Studies*, 71, 317-349.

Austen-Smith, D. and J. Banks (2000): “Cheap talk and burned money,” *Journal of Economic Theory*, 91, 1-16.

Azar, O. (2007): “The slowdown in first-response times of economics journals: Can it be beneficial?,” *Economic Inquiry*, 45, 179-187.

Balder, E. (1996): "Comments on the existence of equilibrium distributions," *Journal of Mathematical Economics*, 25, 307-323.

Banerjee, A. (1997): "A theory of misgovernance," *Quarterly Journal of Economics*, 112, 1289-1332.

Baron, D. and D. Diermeier (2007): "Strategic activism and nonmarket strategy," *Journal of Economics and Management Strategy*, 16, 599-634.

Baron, D. and R. Myerson (1982): "Regulating a monopoly with unknown costs," *Econometrica*, 50, 911-930.

Becker, G. (1968): "Crime and punishment: an economic approach," *The Journal of Political Economy*, 76, 169-217.

Bendor, J., S. Taylor, and R. Van Gaalen (1987): "Politicians, bureaucrats, and asymmetric information," *American Journal of Political Science*, 31, 796-828.

Chakravarty, S. and T. Kaplan (2013): "Optimal allocation without transfer payments," *Games and Economic Behavior*, 77(1), 1-20.

Condorelli, D. (2012): "What money can't buy: Efficient mechanism design with costly signals," *Games and Economic Behavior*, 75 (2), 613-624.

Crawford, V. and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50, 1431-1452.

Crémer, J., L. Garicano and A. Prat (2007): "Language and the theory of the firm," *Quarterly Journal of Economics*, 122, 373-407.

Dessein, W. (2002): "Authority and communication in organizations," *Review of Economic Studies*, 69, 811-838.

Garicano, L. (2000): "Hierarchies and the organization of knowledge in production," *Journal of Political Economy*, 108, 874-904.

Goldman, M., H. Leland and D. Sibley (1984): "Optimal nonuniform pricing," *Review of Economic Studies*, 51, 305-320.

Goltsman, M., J Hörner, G. Pavlov, and F. Squintani (2009): "Mediated cheap talk," *Journal of Economic Theory*, 144, 1397-1420.

Guesnerie, R. and J. Laffont (1984): "A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm," *Journal of Public Economics*, 25, 329-369.

Guriev, S. (2004): "Red tape and corruption," *Journal of Development Economics*, 73, 489-504.

Holmstrom, B. (1977): "On incentives and control in organizations" (PhD dissertation,

Stanford University).

Jullien, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory* 93, 1-47.

Kartik, N. (2007): "A note on cheap talk and burned money," *Journal of Economic Theory*, 136, 749-758.

Kartik, N. (2009): "Strategic communication with lying costs," *Review of Economic Studies*, 76, 1359-1395.

Koessler, F. and D. Martimort (2012): "Optimal delegation with multi-dimensional decisions," *Journal of Economic Theory*, 147, 1850-1881.

Kováč, E. and T. Mylovanov (2009): "Stochastic mechanisms in settings without monetary transfers," *Journal of Economic Theory*, 144, 1373-1395.

Krähmer, D. (2006): "Message-contingent delegation," *Journal of Economic behavior & Organization*, 60, 490-506.

Krishna, V. and J. Morgan (2008): "Contracting for information under imperfect commitment," *RAND Journal of Economics*, 39, 905-925.

Levitt, S. D. (1997): "Incentive compatibility constraints as an explanation for the use of prison sentences instead of fines," *International Review of Law and Economics*, 19, 179-192.

Luenberger, D. G. (1969) *Optimization by Vector Space Methods*. John Wiley & Sons: New York.

Martimort, D. and A. Semenov (2007): "The pluralistic view of politics: Asymmetric lobbyists, ideological uncertainty and political entry," *Economics Letters*, 97, 155-161.

McCubbins, M., R. Noll, and B. Weingast (1987): "Administrative procedures as instruments of political control," *Journal of Law, Economics and Organization*, 3, 243-277.

Melumad, N. and T. Shibano (1991): "Communication in settings with no transfers," *Rand Journal of Economics* 22, 173-198.

Milgrom, P. and J. Roberts (1992): "Economics, organization and management," Prentice-Hall, Englewood Cliffs, NJ.

Mookherjee, D. and I. Png (1994): "Marginal deterrence in enforcement of law," *Journal of Political Economy*, 102, 1039-1066.

Noldeke, G., and L. Samuelson (2007): "Optimal bunching without optimal control," *Journal of Economic Theory*, 134(1), 405-420.

Ottaviani, M. (2000): "The economics of advice," Mimeo London Business School.

Padro i Miquel, G., and P. Yared (2012) "The Political Economy of Indirect Control," *Quarterly Journal of Economics*, 127, 947-1015.

Prendergast, C. (2007): “The motivation and bias of bureaucrats,” *American Economic Review*, 97, 180-196.

Shavell, S. (1987): “The optimal use of nonmonetary sanctions as a deterrent,” *American Economic Review*, 77, 584-592.

Strausz, R. (2006): “Deterministic versus stochastic mechanisms in principal-agent models,” *Journal of Economic Theory*, 128, 306-314.

Szalay, D. (2005): “The economics of extreme options and clear advice,” *Review of Economic Studies*, 72, 1173-1198.

Tirole, J. (1986): “Hierarchies and bureaucracies: On the role of collusion in organizations,” *Journal of Law, Economics, and Organization*, 2, 181-214.

Wessels, W. (1980a): “The effect of minimum wages in the presence of fringe benefits: An expanded model,” *Economic Inquiry*, 18, 293-313.

Wessels, W. (1980b): Minimum wages, fringe benefits, and working conditions, American Enterprise Institute, Washington D.C.

Yoon, K. (2011): “Optimal mechanism design when both allocative inefficiency and expenditure inefficiency matter,” *Journal of Mathematical Economics*, 47 (6), 670-676.

## Appendix

**Proof of Claim 4.** Recall that  $R$  is the range of  $y(\theta)$ . Consider  $Z = [y(0), y(1)] \setminus R$  and let us extend the definition of function  $\tilde{\theta}(\cdot)$  on the set  $Z$ . Take any  $y \in Z$  (then  $y(0) < y < y(1)$ ) and let  $\tilde{\theta}(y) = \sup_{y^*(\theta) \leq y} \theta = \inf_{y^*(\theta) \geq y} \theta$  (these coincide because  $y^*$  is monotone). Let us now define, for  $y \in Z$ ,  $\tilde{m}(y) = \tilde{m}(y^*(\tilde{\theta}(y))) + l^a(\tilde{\theta}(y), y^*(\tilde{\theta}(y))) - l^a(\tilde{\theta}(y), y)$ ; in other words, we define  $\tilde{m}(y)$  as an amount of money burning that would leave the type  $\tilde{\theta}(y)$  indifferent between  $(y, \tilde{m}(y))$  and  $(y^*(\tilde{\theta}(y)), \tilde{m}(y^*(\tilde{\theta}(y))))$ . We will have  $\tilde{m}(y) \geq 0$  because  $l^a$  is convex in its second argument. Let us now prove that  $\tilde{m}(\cdot)$  is continuous at any  $y \in [y_1, y_2]$ ; its left derivative exists at  $y$  if  $y > y_1$ , and its right derivative exists at  $y$  if  $y < y_2$ , and they are equal to:

$$\begin{aligned} \frac{d^l \tilde{m}(y)}{dy} &= -\frac{\partial l^a(\tilde{\theta}_{min}(y), y)}{\partial y}, \\ \frac{d^r \tilde{m}(y)}{dy} &= -\frac{\partial l^a(\tilde{\theta}_{max}(y), y)}{\partial y}, \end{aligned}$$

where  $\tilde{\theta}_{min}(y) = \inf_{y^*(\theta)=y} \theta$  and  $\tilde{\theta}_{max}(y) = \sup_{y^*(\theta)=y} \theta$ .

Indeed, if  $y > y_1$ , take sufficiently small  $\varepsilon > 0$ . Let us prove that

$$\begin{aligned} l^a \left( \tilde{\theta}(y - \varepsilon), y - \varepsilon \right) + \tilde{m}(y - \varepsilon) &\leq l^a \left( \tilde{\theta}(y - \varepsilon), y \right) + \tilde{m}(y); \\ l^a \left( \tilde{\theta}(y), y \right) + \tilde{m}(y) &\leq l^a \left( \tilde{\theta}(y), y - \varepsilon \right) + \tilde{m}(y - \varepsilon). \end{aligned} \quad (30)$$

We prove the first inequality (the second one can be proved analogously). Applying (10) to type  $\tilde{\theta}(y - \varepsilon)$ , we get

$$\begin{aligned} l^a \left( \tilde{\theta}(y - \varepsilon), y^* \left( \tilde{\theta}(y - \varepsilon) \right) \right) + \tilde{m} \left( y^* \left( \tilde{\theta}(y - \varepsilon) \right) \right) \\ \leq l^a \left( \tilde{\theta}(y - \varepsilon), \lim_{\delta \rightarrow +0} y^* \left( \tilde{\theta}(y) - \delta \right) \right) + \lim_{\delta \rightarrow +0} \tilde{m} \left( y^* \left( \tilde{\theta}(y) - \delta \right) \right); \end{aligned}$$

indeed, the agent does not want to pretend to be any type arbitrarily close to  $\tilde{\theta}(y)$ . By construction, we have

$$l^a \left( \tilde{\theta}(y - \varepsilon), y^* \left( \tilde{\theta}(y - \varepsilon) \right) \right) + \tilde{m} \left( y^* \left( \tilde{\theta}(y - \varepsilon) \right) \right) = l^a \left( \tilde{\theta}(y - \varepsilon), y - \varepsilon \right) + \tilde{m}(y - \varepsilon)$$

(this inequality is trivial if  $y - \varepsilon \in R(y^*)$ ). We also have that if  $y \in R(y^*)$ , then

$$l^a \left( \tilde{\theta}(y - \varepsilon), y^* \left( \tilde{\theta}(y) \right) \right) + \tilde{m} \left( y^* \left( \tilde{\theta}(y) \right) \right) = l^a \left( \tilde{\theta}(y - \varepsilon), y \right) + \tilde{m}(y),$$

and it remains to prove that for  $y \in Z$ , we must have

$$l^a \left( \tilde{\theta}(y - \varepsilon), y \right) + \tilde{m}(y) \geq l^a \left( \tilde{\theta}(y - \varepsilon), \lim_{\delta \rightarrow +0} y^* \left( \tilde{\theta}(y) - \delta \right) \right) + \lim_{\delta \rightarrow +0} \tilde{m} \left( y^* \left( \tilde{\theta}(y) - \delta \right) \right).$$

But the agent of type  $\tilde{\theta}(y)$  must be indifferent between the contracts  $(y, \tilde{m}(y))$  and  $\left( \lim_{\delta \rightarrow +0} y^* \left( \tilde{\theta}(y) - \delta \right), \lim_{\delta \rightarrow +0} \tilde{m} \left( y^* \left( \tilde{\theta}(y) - \delta \right) \right) \right)$ . As  $y \geq \lim_{\delta \rightarrow +0} y^* \left( \tilde{\theta}(y) - \delta \right)$  by monotonicity, and  $\tilde{\theta}(y - \varepsilon) \leq \tilde{\theta}(y)$  the type  $\tilde{\theta}(y - \varepsilon)$  must strictly prefer the latter. This establishes the first inequality of (30).

The inequalities in (30) imply

$$l^a \left( \tilde{\theta}(y - \varepsilon), y - \varepsilon \right) - l^a \left( \tilde{\theta}(y - \varepsilon), y \right) \leq \tilde{m}(y) - \tilde{m}(y - \varepsilon) \leq l^a \left( \tilde{\theta}(y), y - \varepsilon \right) - l^a \left( \tilde{\theta}(y), y \right).$$

Since this holds for any function  $\tilde{\theta}(\cdot)$  that satisfies  $y^* \left( \tilde{\theta}(y) \right) = y$ , by continuity, we have

$$l^a \left( \tilde{\theta}_{\min}(y - \varepsilon), y - \varepsilon \right) - l^a \left( \tilde{\theta}_{\min}(y - \varepsilon), y \right) \leq \tilde{m}(y) - \tilde{m}(y - \varepsilon) \leq l^a \left( \tilde{\theta}_{\min}(y), y - \varepsilon \right) - l^a \left( \tilde{\theta}_{\min}(y), y \right).$$

Dividing all parts by  $\varepsilon$ , we notice that the leftmost and the rightmost parts tend to  $-\frac{\partial l^a(\tilde{\theta}_{\min}(y), y)}{\partial y}$ , because  $\lim_{\varepsilon \rightarrow +0} \tilde{\theta}_{\min}(y - \varepsilon) = \tilde{\theta}_{\min}(y)$ . This shows that  $\frac{d\tilde{m}(y)}{dy}$  exists and it is given by the formula. The same argument works for the right derivative; in either case,  $\tilde{m}(y)$  is

continuous at  $y$ . This also implies that  $\tilde{m}(\cdot)$  is differentiable at  $y$  if and only if  $(y^*)^{-1}(y)$  is a singleton (in particular,  $\tilde{m}(\cdot)$  is differentiable in any point in  $Z$ ).

Given that  $\tilde{m}(\cdot)$  is almost everywhere differentiable with bounded derivative, it satisfies the Lipschitz conditions and is thus absolutely continuous. It can then be reconstructed from its derivative. ■

**Proof of Theorem 5.** To be able to apply Theorem 2.2 in Balder (1996), we need to make the space of feasible contracts compact. First note that we can assume that  $T \leq \bar{T}$  for some  $\bar{T}$  sufficiently high, so  $T$  is taken from a compact set  $[\tilde{w}, \bar{T}]$ : indeed, the contract  $y(\theta) = \theta + b(\theta)$ ,  $m(\theta) = 0$ ,  $T = \max\{\tilde{w}, \bar{T}\}$  is feasible and yields  $L^p = Ab^2 + \max\{\bar{u}, \bar{w}\}$ , and thus any contract featuring  $T > \bar{T} = \int_0^1 l^p(\theta, \theta + b(\theta))dF(\theta) + \max\{\bar{u}, \bar{w}\}$  is suboptimal. The set of monotonic mappings from  $\Theta$  to  $Y$  is compact in sup-metrics, so the set of feasible  $y(\theta)$  is compact. Finally, by Claim 4 we have

$$m(\theta) = \tilde{m}(y(0)) + \int_{y(0)}^{\hat{y}} \left( -\frac{\partial l^a(\tilde{\theta}(y), y)}{\partial y} \right) dy, \quad (31)$$

which means that the entire function  $m(\theta)$  may be derived from  $y(\theta)$  and  $m(y(0))$ . Since  $m(y(0))$  may be assumed to be taken from some compact set  $[0, \bar{m}]$ , the set of feasible contracts  $(T, y(\theta), m(\theta))$  may be assumed to be compact. With this modification of the space the conditions of Theorem 2.2 from Balder (1996) trivially hold, therefore there exists a solution to the problem. ■

**Proof of Theorem 6. Part 1.** Assume  $y^*(\theta) < \theta$  for some  $\theta \in [0, 1]$ . Below we show that the principal can offer a contract that improves his expected payoff.

Since we are focusing on  $y(\cdot)$  that are (right- and left-)continuous at  $\theta = 0$  and  $\theta = 1$ , there must exist  $(\underline{\theta}, \bar{\theta})$  such that  $y^*(\theta) < \theta$  for every  $\theta \in (\underline{\theta}, \bar{\theta})$ . Take  $0 \leq \underline{\theta} < \bar{\theta} \leq 1$  such that (i)  $y^*(\theta) < \theta$  for every  $\theta \in (\underline{\theta}, \bar{\theta})$ ; (ii) either  $y^*(\underline{\theta}) \geq \underline{\theta}$  or  $\underline{\theta} = 0$ ; (iii) either  $y^*(\bar{\theta}) \geq \bar{\theta}$  or  $\bar{\theta} = 1$ . We consider two situations.

Case 1. Suppose that for all  $\theta > \bar{\theta}$ ,  $y^*(\theta) \leq y^a(\theta)$ . Then define action function  $y'(\cdot)$  such that  $y'(\theta) = \theta$  for  $\theta \in [\underline{\theta}, \bar{\theta}]$  and  $y'(\theta) = y^*(\theta)$  otherwise, and define  $m'(\cdot)$  such that  $m'(\theta) = m^*(\theta)$  for  $\theta \leq \underline{\theta}$  and  $m'(\cdot)$  and  $y'(\cdot)$  together satisfy (10). In this case,  $y'(\theta) \leq y^a(\theta)$  for every  $\theta \geq \underline{\theta}$  and therefore  $m'(\cdot)$  is increasing over this range. This implies that  $m'(\theta) \geq 0$  for every  $\theta \in [0, 1]$ . Furthermore, note that every type  $\theta$  is weakly better off given  $(y'(\cdot), m'(\cdot))$  than given  $(y^*(\cdot), m^*(\cdot))$ ; the reason for this is that every action between  $y^*(\underline{\theta})$  and  $\max(y^*(\bar{\theta}), 1)$  that was used by some type in  $y^*(\cdot)$  now involves equal or less money-burning, plus perhaps some

actions that were not available (as  $y^*(\cdot)$  is not assumed to be continuous) are now available. Consequently, all actions now involve weakly less money-burning, and this makes every type  $\theta$  of agent weakly better off. Therefore the participation constraint of the agent holds with transfer  $T^*$ . The above imply that  $(y'(\cdot), m'(\cdot), T^*)$  is a feasible contract. For  $\theta \in (\underline{\theta}, \bar{\theta})$  the implemented action is strictly better for the principal (while at other states the implemented action remains the same), hence the principal is strictly better off ex ante, contradicting the optimality of  $(y^*(\cdot), m^*(\cdot), T^*)$ .

Case 2. Suppose that for some  $\theta > \bar{\theta}$ ,  $y^*(\theta) > y^a(\theta)$ . Then there are  $\theta', \theta''$  such that  $\bar{\theta} < \theta' < \theta''$  and for all  $\theta \in (\theta', \theta'')$ ,  $y^*(\theta) > y^a(\theta)$ ; we can pick  $\theta'$  such that the left limit  $\lim_{\theta \nearrow \theta'} y^*(\theta) \leq y^a(\theta')$ . Then there exist  $\alpha, \beta \in (0, 1)$  such that  $y'(\theta)$  and  $m'(\theta)$  defined as below is a feasible contract:  $y'(\theta) = y^*(\theta)$  for  $\theta \notin [\underline{\theta}, \bar{\theta}] \cup [\theta', \theta'']$ ,  $y'(\theta) = \alpha y^*(\theta) + (1 - \alpha)\theta$  for  $\theta \in (\underline{\theta}, \bar{\theta})$ ,  $y'(\theta) = \beta y^*(\theta) + (1 - \beta)y^a(\theta)$  for  $\theta \in (\theta', \theta'')$ , and  $m'(\theta) = m^*(\theta)$  for  $m < \underline{\theta}$  and  $m > \theta''$  (whereas for  $\theta \in [\underline{\theta}, \theta'']$ ,  $m'(\theta)$  is defined so that  $m'(\cdot)$  and  $y'(\cdot)$  satisfy (10)). To see this, note that if  $\alpha$  is close enough to 1 then since the amount of money-burning in any state is continuous in  $\alpha$  and  $\beta$ , there exists  $\beta \in (0, 1)$  such that the increase of money-burning on  $(\theta', \theta'')$  is exactly offset by the decrease of money-burning on  $(\underline{\theta}, \bar{\theta})$ , and therefore  $m'(\theta) \geq 0$  for every  $\theta \in [0, 1]$ . Moreover, every type  $\theta$  is weakly better off given  $(y'(\cdot), m'(\cdot))$  than given  $(y^*(\cdot), m^*(\cdot))$ . To see this, notice that we can augment the plan  $(y'(\cdot), m'(\cdot))$  by allowing agents to choose any action  $y \in [y'(0), y'(1)]$  and choosing  $\tilde{m}(y)$  such that agent with type  $\tilde{\theta} = \sup\{\theta : y'(\theta) \leq y\}$  is exactly indifferent between  $(y'(\tilde{\theta}), m'(\tilde{\theta}))$  and  $(y, \tilde{m}(y))$ , in this case no agent will be willing to switch, so no agent is strictly better off from having these additional options. Compared to the plan  $(y^*(\cdot), m^*(\cdot))$ , every action that was available under  $(y^*(\cdot), m^*(\cdot))$  now involves weakly less money-burning, and perhaps additional actions were made available. Thus, similarly to Case 1, the participation constraint of the agent holds under  $(y'(\cdot), m'(\cdot))$  with the same transfer  $T^*$ . Notice, however, that for  $\theta \in (\underline{\theta}, \bar{\theta}) \cup (\theta', \theta'')$  the implemented action  $y'(\cdot)$  is strictly better for the principal than  $y^*(\cdot)$ , while at other states the implemented action remains the same, hence the principal is strictly better off ex ante, contradicting the optimality of  $(y^*(\cdot), m^*(\cdot), T^*)$ .

**Part 2.** We start by proving the following result. Suppose that  $\frac{\frac{\partial l^\alpha(\theta, y_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta, y_0)}{\partial \theta}}{l^p(\theta, y_0) - l^p(\theta, y_1)} > \frac{\frac{\partial l^\alpha(\theta, y_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta, y_2)}{\partial \theta}}{l^p(\theta, y_2) - l^p(\theta, y_1)}$ , for every  $\theta \in (0, 1)$ , and  $\lim_{\theta' \searrow \theta} y^*(\theta') \geq y_2 > y_0 > y_1 \geq \theta$ . Then  $y^*(\theta)$  and  $m^*(\theta)$  are continuous on  $(0, 1)$ .

To prove this, note that Claim 4 implies that  $m^*$  is continuous at  $\theta$  if  $y^*$  is continuous at  $\theta$ , hence it is enough to prove continuity of the latter. The proof below is by contradiction.

Suppose that for some  $\theta_0 \in (0, 1)$ ,  $y^*$  is discontinuous at  $\theta_0$ . Denote

$$\begin{aligned}\hat{y}_1 &= \sup_{\theta \in [0, \theta_0)} y^*(\theta), \\ \hat{y}_2 &= \inf_{\theta \in (\theta_0, 1]} y^*(\theta).\end{aligned}$$

Note that, since  $y^*(\theta)$  is monotonic, it is true that  $\hat{y}_1 = \lim_{\theta \rightarrow \theta_0^-} y^*(\theta)$ ,  $\hat{y}_2 = \lim_{\theta \rightarrow \theta_0^+} y^*(\theta)$ . Define  $\hat{m}_1 = \lim_{\theta \rightarrow \theta_0^-} m^*(\theta) \geq 0$  and  $\hat{m}_2 = \lim_{\theta \rightarrow \theta_0^+} m^*(\theta) \geq 0$ ; these limits exist by the continuity of loss function  $L^a(\theta)$ :  $\hat{m}_1 = \lim_{\theta \rightarrow \theta_0^-} L^a(\theta) - l^a(\theta_0, \hat{y}_1)$ , and similarly  $\hat{m}_2 = \lim_{\theta \rightarrow \theta_0^+} L^a(\theta) - l^a(\theta_0, \hat{y}_2)$ . It is evident that an agent of type  $\theta_0$  is indifferent between contracts  $(y^*(\theta_0), m^*(\theta_0))$ ,  $(\hat{y}_1, \hat{m}_1)$  and  $(\hat{y}_2, \hat{m}_2)$ : otherwise, if, for instance, we had  $l^a(\theta_0, \hat{y}_1) + \hat{m}_1 > l^a(\theta_0, \hat{y}_2) + \hat{m}_2$  instead, then an agent of type  $\theta_0 + \varepsilon$  would strictly prefer contract  $(y^*(\theta_0 - \varepsilon), m^*(\theta_0 - \varepsilon))$  to  $(y^*(\theta_0 + \varepsilon), m^*(\theta_0 + \varepsilon))$  by continuity, which would violate (10).

The idea of the proof is to perturb the optimal contract  $(y^*(\theta), m^*(\theta))_{\theta \in \Theta}$  around the point of discontinuity  $\theta_0$  and obtain a higher value of  $V^P$ , which would contradict the optimality of the initial contract. Take some  $a \in (0, 1)$  and define  $\hat{y}_0$  by

$$\frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} = a \frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial \theta} + (1-a) \frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial \theta}; \quad (32)$$

clearly, such  $\hat{y}_0 \in (\hat{y}_1, \hat{y}_2)$  exists (and is unique) for any  $a \in (0, 1)$ , since  $\frac{\partial l^a(\theta_0, y)}{\partial y}$  is continuous and monotonic (increasing) in  $y$ . Trivially, (32) is equivalent to

$$\frac{\frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial \theta}}{\frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta}} = \frac{a}{1-a}.$$

We now pick  $\hat{m}_0$  to be such that

$$l^a(\theta_0, \hat{y}_0) + \hat{m}_0 = l^a(\theta_0, \hat{y}_1) + \hat{m}_1 = l^a(\theta_0, \hat{y}_2) + \hat{m}_2.$$

Since  $l^a(\theta_0, y)$  is strictly convex in  $y$ , we have  $l^a(\theta_0, \hat{y}_0) < \max(l^a(\theta_0, \hat{y}_1), l^a(\theta_0, \hat{y}_2))$ , and therefore  $\hat{m}_0 > \min(\hat{m}_1, \hat{m}_2) \geq 0$ .

By construction, agent of type  $\theta_0$  is indifferent between  $(\hat{y}_0, \hat{m}_0)$ ,  $(\hat{y}_1, \hat{m}_1)$ , and  $(\hat{y}_2, \hat{m}_2)$ . In contrast, agents with  $\theta < \theta_0$  strictly prefer  $(\hat{y}_1, \hat{m}_1)$  to  $(\hat{y}_0, \hat{m}_0)$  (this immediately follows from the single-crossing condition), and prefer  $(y^*(\theta), m^*(\theta))$  to  $(\hat{y}_1, \hat{m}_1)$  (from (10), as  $(\hat{y}_1, \hat{m}_1)$  is a limit of feasible contracts), while agents with  $\theta > \theta_0$  weakly prefer  $(y^*(\theta), m^*(\theta))$  to  $(\hat{y}_2, \hat{m}_2)$ , which they strictly prefer to  $(\hat{y}_0, \hat{m}_0)$ . Consider the function

$$z(\theta) = l^a(\theta, \hat{y}_0) + \hat{m}_0 - L^a(\theta),$$

which is naturally interpreted as the ‘‘gap’’ in utility from choosing  $(y^*(\theta), m^*(\theta))$ , which agent  $\theta$  does, and choosing  $(\hat{y}_0, \hat{m}_0)$  if he had such an option. From Claim 3 it follows that function

$z(\theta)$  is continuous for  $\theta \in \Theta$ , it is positive and strictly decreasing for  $\theta < \theta_0$ , it is positive and strictly increasing for  $\theta > \theta_0$ , and it equals zero at  $\theta = \theta_0$ .

Let us take a sufficiently small  $\varepsilon > 0$  and augment the set of available choices  $(y^*(\theta), m^*(\theta))_{\theta \in \Theta}$  by adding  $(\hat{y}_0, \hat{m}_0 - \varepsilon)$  to it. From the properties of function  $z(\theta)$  it follows that players with  $\theta \in (\theta_1(\varepsilon), \theta_2(\varepsilon))$  will switch to  $(\hat{y}_0, \hat{m}_0 - \varepsilon)$  while the rest will not (and those with types  $\theta_1(\varepsilon)$  and  $\theta_2(\varepsilon)$  will be indifferent); here,  $\theta_1(\varepsilon)$  and  $\theta_2(\varepsilon)$  are continuous functions of  $\theta$  such that  $\theta_1(\varepsilon)$  is decreasing and  $\theta_2(\varepsilon)$  is increasing in  $\varepsilon$ . As  $\varepsilon \rightarrow 0$ ,  $\theta_1(\varepsilon) \rightarrow \theta_0$  and  $\theta_2(\varepsilon) \rightarrow \theta_0$ . Let us find the limit of  $\frac{\theta_0 - \theta_1(\varepsilon)}{\theta_2(\varepsilon) - \theta_0}$  (and simultaneously show that it exists and is finite). To do that, it is convenient to consider the inverse functions,  $\varepsilon_1(\theta_1)$ , defined for  $\theta_1 \leq \theta_0$ , and  $\varepsilon_2(\theta_2)$ , defined for  $\theta_2 \geq \theta_0$ .

By construction,  $\varepsilon_1(\theta_1)$  satisfies

$$L^a(\theta_1) = l^a(\theta_1, \hat{y}_0) + \hat{m}_0 - \varepsilon_1(\theta_1).$$

Hence,

$$\begin{aligned} \varepsilon_1(\theta_1) &= l^a(\theta_1, \hat{y}_0) + \hat{m}_0 - L^a(\theta_1) \\ &= l^a(\theta_1, \hat{y}_0) - l^a(\theta_0, \hat{y}_0) + L^a(\theta_0) - L^a(\theta_1). \end{aligned}$$

Therefore Claim 3 implies that  $\varepsilon_1(\theta_1)$  has a left derivative at  $\theta_1 = \theta_0$ :

$$\frac{d^l \varepsilon_1(\theta_1)}{d\theta_1} = \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial \theta}.$$

Similarly,

$$\frac{d^r \varepsilon_2(\theta_1)}{d\theta_1} = \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial \theta}.$$

We then have

$$\begin{aligned} \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial \theta} &= \lim_{\theta_1 \rightarrow \theta_0^-} \frac{\varepsilon_1(\theta_0) - \varepsilon_1(\theta_1)}{\theta_0 - \theta_1} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{-\varepsilon}{\theta_0 - \theta_1(\varepsilon)}, \\ \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial \theta} &= \lim_{\theta_2 \rightarrow \theta_0^+} \frac{\varepsilon_2(\theta_2) - \varepsilon_2(\theta_0)}{\theta_2 - \theta_0} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{\theta_2(\varepsilon) - \theta_0}. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \frac{\theta_0 - \theta_1(\varepsilon)}{\theta_2(\varepsilon) - \theta_0} &= \frac{\lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{\theta_2(\varepsilon) - \theta_0}}{-\lim_{\varepsilon \rightarrow 0^+} \frac{-\varepsilon}{\theta_0 - \theta_1(\varepsilon)}} \tag{33} \\ &= \frac{\frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial \theta}}{\frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta}} = \frac{a}{1-a}. \end{aligned}$$

We are now ready to estimate the welfare effect of this perturbation. The agent of any type is weakly better off, and for some types the agent is strictly better off: for  $\theta \in (\theta_1(\varepsilon), \theta_2(\varepsilon))$  switched to  $(\hat{y}_0, \hat{m}_0 - \varepsilon)$  which he strictly prefers to  $(y^*(\theta), m^*(\theta))$  which he was choosing before, and the rest have not changed their contract. We therefore only need to compute the change in the principal's payoff. This change equals

$$\begin{aligned} & \int_{\theta_1(\varepsilon)}^{\theta_2(\varepsilon)} (l^p(\theta, y(\theta)) - l^p(\theta, \hat{y}_0)) f(\theta) d\theta = \int_{\theta_1(\varepsilon)}^{\theta_2(\varepsilon)} \int_{\hat{y}_0}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} dy f(\theta) d\theta \quad (34) \\ & = \int_{\theta_0}^{\theta_2(\varepsilon)} (l^p(\theta, y(\theta)) - l^p(\theta, \hat{y}_0)) f(\theta) d\theta - \int_{\theta_1(\varepsilon)}^{\theta_0} (l^p(\theta, \hat{y}_0) - l^p(\theta, y(\theta))) f(\theta) d\theta. \end{aligned}$$

To check that this expression is positive, it is sufficient, given the continuity of  $f(\theta)$  at  $\theta_0$  and existence of limits  $\lim_{\theta \rightarrow \theta_0^-} \frac{\partial l^p(\theta, y)}{\partial y} = \frac{\partial l^p(\theta_0, y)}{\partial y}$  and  $\lim_{\theta \rightarrow \theta_0^+} \frac{\partial l^p(\theta, y)}{\partial y} = \frac{\partial l^p(\theta_0, y)}{\partial y}$ , to prove that

$$\lim_{\varepsilon \rightarrow 0} ((\theta_2(\varepsilon) - \theta_0) (l^p(\theta, \hat{y}_2) - l^p(\theta, \hat{y}_0)) - (\theta_0 - \theta_1(\varepsilon)) (l^p(\theta, \hat{y}_0) - l^p(\theta, \hat{y}_1))) > 0.$$

In light of (33), it suffices to prove that

$$(1 - a) (l^p(\theta, \hat{y}_2) - l^p(\theta, \hat{y}_0)) > a (l^p(\theta, \hat{y}_0) - l^p(\theta, \hat{y}_1)). \quad (35)$$

By Part 1,  $l^p(\theta, \hat{y}_2) > l^p(\theta, \hat{y}_0)$  and  $l^p(\theta, \hat{y}_0) > l^p(\theta, \hat{y}_1)$ , and (35) is equivalent to

$$\frac{l^p(\theta_0, \hat{y}_2) - l^p(\theta_0, \hat{y}_0)}{l^p(\theta_0, \hat{y}_0) - l^p(\theta_0, \hat{y}_1)} > \frac{\frac{\partial l^\alpha(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, \hat{y}_2)}{\partial \theta}}{\frac{\partial l^\alpha(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, \hat{y}_0)}{\partial \theta}}.$$

By adding 1 to both sides, we find this is equivalent to

$$\frac{l^p(\theta_0, \hat{y}_2) - l^p(\theta_0, \hat{y}_1)}{l^p(\theta_0, \hat{y}_0) - l^p(\theta_0, \hat{y}_1)} > \frac{\frac{\partial l^\alpha(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, \hat{y}_2)}{\partial \theta}}{\frac{\partial l^\alpha(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, \hat{y}_0)}{\partial \theta}}.$$

Now, rearranging (note that the denominators are positive) and changing the sign, we get

$$\frac{\frac{\partial l^\alpha(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, \hat{y}_0)}{\partial \theta}}{l^p(\theta_0, \hat{y}_0) - l^p(\theta_0, \hat{y}_1)} > \frac{\frac{\partial l^\alpha(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, \hat{y}_2)}{\partial \theta}}{l^p(\theta_0, \hat{y}_2) - l^p(\theta_0, \hat{y}_1)}. \quad (36)$$

Claim 1 and part 1 of the current claim imply that  $\theta \leq \hat{y}_1 < \hat{y}_0 < \hat{y}_2$ , hence the assumption of the lemma implies that (36) holds. This implies that the proposed deviation is profitable, contradicting that  $y^*$  is discontinuous at  $\theta_0$ .

To finish the proof of this result, notice that for any  $\theta_0 \leq y_1 < y_0 < y_2$  the following holds:

$$\frac{\frac{\frac{\partial l^\alpha(\theta_0, y_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, y_0)}{\partial \theta}}{y_0 - y_1}}{\frac{l^p(\theta_0, y_0) - l^p(\theta_0, y_1)}{y_0 - y_1}} > \frac{\frac{\frac{\partial l^\alpha(\theta_0, y_0)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, y_2)}{\partial \theta}}{y_2 - y_0}}{\frac{l^p(\theta_0, y_2) - l^p(\theta_0, y_0)}{y_2 - y_0}}.$$

This is equivalent to  $\frac{\frac{\partial l^\alpha(\theta_0, y_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, y_0)}{\partial \theta}}{l^p(\theta_0, y_0) - l^p(\theta_0, y_1)} > \frac{\frac{\partial l^\alpha(\theta_0, y_1)}{\partial \theta} - \frac{\partial l^\alpha(\theta_0, y_2)}{\partial \theta}}{l^p(\theta_0, y_2) - l^p(\theta_0, y_1)}$ , therefore the argument above implies the claim in the theorem.

**Part 3.** We first prove the following auxiliary result. Suppose  $(y(\theta), m(\theta))$  satisfies (10) and  $y(\theta)$  is continuous on  $\Theta$ . Then for any  $\theta_1, \theta_2 \in \Theta$ , we have

$$m(\theta_2) - m(\theta_1) = l^\alpha(\theta_1, y(\theta_1)) - l^\alpha(\theta_2, y(\theta_2)) + \int_{\theta_1}^{\theta_2} \left( \frac{\partial l^\alpha(\theta, y(\theta))}{\partial \theta} \right) d\theta. \quad (37)$$

Indeed, from (13), we have

$$\begin{aligned} \int_{\theta_1}^{\theta_2} \left( \frac{\partial l^\alpha(\theta, y(\theta))}{\partial \theta} \right) d\theta &= L^\alpha(\theta_2) - L^\alpha(\theta_1) \\ &= l^\alpha(\theta_2, y(\theta_2)) + m(\theta_2) - l^\alpha(\theta_1, y(\theta_1)) - m(\theta_1). \end{aligned}$$

Rearranging, we obtain (37).

Now, by Part 2,  $y^*(\theta)$  is a continuous function. Suppose, to obtain a contradiction, that there exists  $\theta_0 \in \Theta$  such that  $y(\theta_0) > \theta_0 + b(\theta_0)$ . Because  $y^*(\cdot)$  is continuous, without loss of generality we may assume that  $0 < \theta_0 < 1$ . There are two possibilities: either for all  $\theta < \theta_0$ ,  $y^*(\theta) \geq \theta + b(\theta)$ , or there exists  $\theta' < \theta_0$  such that  $y^*(\theta') < \theta' + b(\theta')$ . We start with the first possibility.

Suppose  $y^*(\theta) \geq \theta + b(\theta)$  for all  $\theta < \theta_0$ . Let  $\bar{\theta} = \inf \{\theta : y^*(\theta) < \theta + b(\theta)\}$  if such  $\theta$  exists; otherwise, let  $\bar{\theta} = 1$ . Define function  $y(\theta)$  by

$$y(\theta) = \begin{cases} y^*(\theta) & \text{if } \theta > \bar{\theta}, \\ \theta + b(\theta) & \text{if } \theta \leq \bar{\theta}. \end{cases}$$

Note that by continuity,  $y^*(\bar{\theta}) = \bar{\theta} + b(\bar{\theta})$ , hence the above function is continuous.

Suppose  $\bar{\theta} < 1$ , then let

$$m(\theta) = \begin{cases} m^*(\theta) & \text{if } \theta > \bar{\theta}, \\ m^*(\bar{\theta}) & \text{if } \theta \leq \bar{\theta}; \end{cases}$$

given that scheme  $y^*(\cdot), m^*(\cdot)$  satisfies (10) and (11), it is straightforward to verify that scheme  $(y(\cdot), m(\cdot))$  also satisfies (10) and (11). In the modified scheme, the utility of the agent at  $\theta \geq \bar{\theta}$  is unchanged. If  $\theta < \bar{\theta}$ , then, by (13)

$$\begin{aligned} L^\alpha(\theta, y(\theta), m(\theta)) &= L^\alpha(\bar{\theta}, y(\bar{\theta}), m(\bar{\theta})) - \int_{\theta}^{\bar{\theta}} \frac{\partial l^\alpha(\xi, y(\xi))}{\partial \theta} d\xi \\ &< L^\alpha(\bar{\theta}, y^*(\bar{\theta}), m^*(\bar{\theta})) - \int_{\theta}^{\bar{\theta}} \frac{\partial l^\alpha(\xi, y^*(\xi))}{\partial \theta} d\xi = L^\alpha(\theta, y^*(\theta), m^*(\theta)); \end{aligned}$$

this holds because at  $\bar{\theta}$  the contract is unchanged, and

$$\int_{\theta}^{\bar{\theta}} \frac{\partial l^\alpha(\xi, y^*(\xi))}{\partial \theta} d\xi - \int_{\theta}^{\bar{\theta}} \frac{\partial l^\alpha(\xi, y(\xi))}{\partial \theta} d\xi = \int_{\theta}^{\bar{\theta}} \int_{y(\xi)}^{y^*(\xi)} \frac{\partial^2 l^\alpha(\xi, y)}{\partial \theta \partial y} d\xi < 0,$$

since  $y(\xi) < y^*(\xi)$  whenever  $\xi < \bar{\theta}$ , at least for  $\xi$  close to  $\theta_0$ . Consequently, all types of agent are at least weakly better off. The principal, is obviously better off, since for some  $\theta$ ,  $y(\theta)$  became closer to  $\theta$  than  $y^*(\theta)$ . This contradicts that contract  $(y^*(\theta), m^*(\theta))$  solves the problem (8).

Now suppose that  $\bar{\theta} = 1$ . In this case,  $y(\theta) = \theta + b(\theta)$ , so let us take  $m(\theta) = 0$ . This makes all agents at least weakly better off (their loss becomes zero), and the principal is strictly better off. This again contradicts that contract  $(y^*(\theta), m^*(\theta))$  solves the problem (8).

Consider the second case, where there exists  $\theta' < \theta_0$  such that  $y^*(\theta') < \theta' + b(\theta')$ . Let  $\theta_1 = \min\{\theta \in [\theta', \theta_0] : y^*(\theta) = \theta + b(\theta)\}$ ,  $\theta_2 = \inf\{\theta \in [\theta_1, \theta_0] : y^*(\theta) > \theta + b(\theta)\}$ ; by continuity,  $\theta_1$  and  $\theta_2$  are well-defined and they may or may not coincide. By construction, if  $\theta \in [\theta_1, \theta_2]$ , then  $y^*(\theta) = \theta + b(\theta)$ ; moreover, for sufficiently small  $\varepsilon > 0$  we have  $y^*(\theta_1 - \varepsilon) < \theta_1 - \varepsilon + b(\theta_1 - \varepsilon)$  and  $y^*(\theta_2 + \varepsilon) > \theta_2 + \varepsilon + b(\theta_2 + \varepsilon)$ . This implies, in particular, that  $m^*(\theta)$  is bounded away from 0 on  $[\theta_1, \theta_2]$  (from Claim 4 it follows that

$$m^*(\theta) - m^*(\theta_1 - \varepsilon) = \int_{y^*(\theta_1 - \varepsilon)}^{y^*(\theta)} \left( -\frac{\partial l^a(\tilde{\theta}(y), y)}{\partial y} \right) dy > 0,$$

where the inequality is true, because  $y^*(\theta_1) = \theta_1 + b(\theta_1) > \theta_1 - \varepsilon + b(\theta_1 - \varepsilon) > y^*(\theta_1 - \varepsilon)$ , so integration is conducted over a nondegenerate interval; moreover, for almost all  $y$ ,  $\tilde{\theta}(y)$  (the inverse) is uniquely defined, and for  $\tilde{\theta} \in (\theta_1 - \varepsilon, \theta_1)$ ,  $y^*(\tilde{\theta}) < \tilde{\theta} + b(\tilde{\theta})$ , so the partial derivative under the integral is positive, and for  $\tilde{\theta} \in (\theta_1, \theta_2)$  it is at least nonnegative).

Let us construct an alternative  $y(\theta)$  as follows. We take  $\varepsilon_1$  and  $\varepsilon_2$  to be such small positive numbers such that

$$\int_{\theta_1 - \varepsilon_1}^{\theta_1} \int_{y^*(\theta)}^{\theta + b(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta = \int_{\theta_2}^{\theta_2 + \varepsilon_2} \int_{\theta + b(\theta)}^{y^*(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta, \quad (38)$$

and pick a small  $\varepsilon_0 > 0$ . We require that

$$y(\theta) = \begin{cases} y^*(\theta) & \text{if } \theta \leq \theta_1 - \varepsilon_1 - \varepsilon_0, \\ \in (y^*(\theta), \theta + b(\theta)) & \text{if } \theta \in (\theta_1 - \varepsilon_1 - \varepsilon_0, \theta_1 - \varepsilon_1), \\ \theta + b(\theta) & \text{if } \theta \in [\theta_1 - \varepsilon_1, \theta_2 + \varepsilon_2], \\ \in (\theta + b(\theta), y^*(\theta)) & \text{if } \theta \in (\theta_2 + \varepsilon_2, \theta_2 + \varepsilon_2 + \varepsilon_0), \\ y^*(\theta) & \text{if } \theta \geq \theta_2 + \varepsilon_2 + \varepsilon_0, \end{cases}$$

and that

$$\int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta = \int_{\theta_2}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \int_{y(\theta)}^{y^*(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta. \quad (39)$$

Now, if we define  $m(\theta)$  to be such that the agent's loss function  $L^a(\theta, y(\theta), m(\theta))$  satisfies (13) and coincides with  $L^a(\theta, y^*(\theta), m^*(\theta))$  for  $\theta \notin (\theta_1 - \varepsilon_1 - \varepsilon_0, \theta_2 + \varepsilon_2 + \varepsilon_0)$ , we would get

a contract  $(y(\theta), m(\theta))$  that satisfies (10) and (11), as well as (12). (To see that such  $m(\theta)$  exists, take  $m(\theta) = m^*(\theta)$  for  $\theta \leq \theta_1 - \varepsilon_1 - \varepsilon_0$ , and for  $\theta \leq \theta_1 - \varepsilon_1 - \varepsilon_0$ , define  $m(\theta)$  by

$$m(\theta) = m(\theta_1 - \varepsilon_1 - \varepsilon_0) + l^a(\theta_1 - \varepsilon_1 - \varepsilon_0, y(\theta_1 - \varepsilon_1 - \varepsilon_0)) - l^a(\theta, y(\theta)) + \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta} \left( \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} \right) d\theta.$$

In this case,  $m(\theta) = m^*(\theta)$  would be true for  $\theta \geq \theta_2 + \varepsilon_2 + \varepsilon_0$ , which immediately implies  $L^a(\theta, y^*(\theta), m^*(\theta)) = L^a(\theta, y(\theta), m(\theta))$  for  $\theta \notin (\theta_1 - \varepsilon_1 - \varepsilon_0, \theta_2 + \varepsilon_2 + \varepsilon_0)$ . Indeed, it suffices to show that  $m(\theta_2 + \varepsilon_2 + \varepsilon_0) = m^*(\theta_2 + \varepsilon_2 + \varepsilon_0)$ , because for larger  $\theta$  the schemes  $y(\cdot)$  and  $y^*(\cdot)$  coincide, and indeed (using (37)) we have

$$\begin{aligned} & m(\theta_2 + \varepsilon_2 + \varepsilon_0) - m^*(\theta_2 + \varepsilon_2 + \varepsilon_0) \\ &= (m(\theta_2 + \varepsilon_2 + \varepsilon_0) - m(\theta_1 - \varepsilon_1 - \varepsilon_0)) - (m^*(\theta_2 + \varepsilon_2 + \varepsilon_0) - m^*(\theta_1 - \varepsilon_1 - \varepsilon_0)) \\ &= \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \left( \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} \right) d\theta - \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \left( \frac{\partial l^a(\theta, y^*(\theta))}{\partial \theta} \right) d\theta \\ &= \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_1} \left( \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} - \frac{\partial l^a(\theta, y^*(\theta))}{\partial \theta} \right) d\theta + \int_{\theta_2}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \left( \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} - \frac{\partial l^a(\theta, y^*(\theta))}{\partial \theta} \right) d\theta \\ &\quad \int_{\theta_1}^{\theta_2} \left( \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} - \frac{\partial l^a(\theta, y^*(\theta))}{\partial \theta} \right) d\theta \\ &= \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta + \int_{\theta_2}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta \\ &\quad + \int_{\theta_1}^{\theta_2} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta \\ &= \int_{\theta_1}^{\theta_2} \int_{\theta + b}^{\theta + b} \left( -\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y} \right) dy d\theta = 0. \end{aligned}$$

The last thing to check is that  $m(\theta) \geq 0$  is satisfied, but this is true because  $m^*(\theta)$  is bounded away from zero, and  $\varepsilon_1$  and  $\varepsilon_2$  are small.

Under the new contract  $(y(\theta), m(\theta))$ , all agents with type  $\theta \in (\theta_1 - \varepsilon_1 - \varepsilon_0, \theta_2 + \varepsilon_2 + \varepsilon_0)$  are better off; moreover, the agents with types  $\theta \in [\theta_1, \theta_2]$  are better off by at least (38). The change in the principal's utility is given by

$$\begin{aligned} & \int_{\theta_2}^{\theta_2 + \varepsilon_2 + \varepsilon_0} (l^p(\theta, y^*(\theta)) - l^p(\theta, y(\theta))) f(\theta) d\theta - \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_1} (l^p(\theta, y(\theta)) - l^p(\theta, y^*(\theta))) f(\theta) d\theta \\ &= \int_{\theta_2}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta - \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta. \end{aligned}$$

It suffices to show that

$$\int_{\theta_2}^{\theta_2 + \varepsilon_2 + \varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta > \int_{\theta_1 - \varepsilon_1 - \varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta.$$

Dividing this by (39), we are to prove

$$\frac{\int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta}{\int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^\alpha(\theta, y)}{\partial \theta \partial y} \right) dy d\theta} > \frac{\int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta}{\int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^\alpha(\theta, y)}{\partial \theta \partial y} \right) dy d\theta}.$$

This would be true<sup>26</sup> if we prove that for any  $(\theta_L, y_L)$  and  $(\theta_H, y_H)$  such that  $\theta_1 - \varepsilon_1 - \varepsilon_0 < \theta_L < \theta_1$ ,  $y^*(\theta_L) < y_L < y(\theta_L)$ ,  $\theta_2 < \theta_H < \theta_2 + \varepsilon_2 + \varepsilon_0$ ,  $y(\theta_H) < y_H < y^*(\theta_H)$ ,

$$\frac{\frac{\partial l^p(\theta_H, y_H)}{\partial y} f(\theta_H)}{-\frac{\partial^2 l^\alpha(\theta_H, y_H)}{\partial \theta \partial y}} > \frac{\frac{\partial l^p(\theta_L, y_L)}{\partial y} f(\theta_L)}{-\frac{\partial^2 l^\alpha(\theta_L, y_L)}{\partial \theta \partial y}}.$$

Since  $\frac{\frac{\partial l^p(\theta, y)}{\partial y} f(\theta)}{-\frac{\partial^2 l^\alpha(\theta, y)}{\partial \theta \partial y}}$  is strictly increasing in  $y$  for any fixed  $\theta$ , and  $y_L < \theta_L + b(\theta_L)$ ,  $y_H > \theta_H + b(\theta_H)$ , it suffices to prove that

$$\frac{\frac{\partial l^p(\theta_H, \theta_H + b(\theta_H))}{\partial y} f(\theta_H)}{-\frac{\partial^2 l^\alpha(\theta_H, \theta_H + b(\theta_H))}{\partial \theta \partial y}} \geq \frac{\frac{\partial l^p(\theta_L, \theta_L + b(\theta_L))}{\partial y} f(\theta_L)}{-\frac{\partial^2 l^\alpha(\theta_L, \theta_L + b(\theta_L))}{\partial \theta \partial y}}.$$

However, this follows from the assumption. This completes the proof. ■

**Lemma 1** Suppose  $y^*(\cdot)$  solves

$$\min_{(y(\cdot), T) \in Z} \int_0^1 \left( A(y(\theta) - \theta)^2 - 2(y(\theta) - \theta - b)(1 - \theta) \right) d\theta + (y(0) - b)^2.$$

Then if  $\theta$  satisfies  $y^*(0) < y^*(\theta) < y^*(1)$ , then  $y^*(\theta) = \min\{z(\theta), \theta + b\}$ , where  $z(\theta) = 1 - \frac{A-1}{A}(1 - \theta)$ .

<sup>26</sup>Since both denominators are positive, let us subtract the right-hand side from the left-hand side and find the sign of the numerator. We have

$$\begin{aligned} & \int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta \times \int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^\alpha(\theta, y)}{\partial \theta \partial y} \right) dy d\theta \\ & - \int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta)}^{y(\theta)} \left( -\frac{\partial^2 l^\alpha(\theta, y)}{\partial \theta \partial y} \right) dy d\theta \times \int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta)}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} f(\theta) dy d\theta \\ = & \int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta_H)}^{y(\theta_H)} \frac{\partial l^p(\theta_H, y_H)}{\partial y} f(\theta_H) dy_H d\theta_H \times \int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta_L)}^{y(\theta_L)} \left( -\frac{\partial^2 l^\alpha(\theta_L, y_L)}{\partial \theta \partial y} \right) dy_L d\theta_L \\ & - \int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta_H)}^{y(\theta_H)} \left( -\frac{\partial^2 l^\alpha(\theta_H, y_H)}{\partial \theta \partial y} \right) dy_H d\theta_H \times \int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta_L)}^{y(\theta_L)} \frac{\partial l^p(\theta_L, y_L)}{\partial y} f(\theta_L) dy_L d\theta_L \\ = & \int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0} \int_{y^*(\theta_H)}^{y(\theta_H)} \int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1} \int_{y^*(\theta_L)}^{y(\theta_L)} \left[ \frac{\frac{\partial l^p(\theta_H, y_H)}{\partial y} f(\theta_H) \left( -\frac{\partial^2 l^\alpha(\theta_L, y_L)}{\partial \theta \partial y} \right)}{-\left( -\frac{\partial^2 l^\alpha(\theta_H, y_H)}{\partial \theta \partial y} \right) \frac{\partial l^p(\theta_L, y_L)}{\partial y} f(\theta_L)} \right] dy_H d\theta_H dy_L d\theta_L \\ > & 0; \end{aligned}$$

here, the first equality is just renaming variables of integration, and the last inequality follows from  $\frac{\frac{\partial l^p(\theta_H, y_H)}{\partial y} f(\theta_H)}{-\frac{\partial^2 l^\alpha(\theta_H, y_H)}{\partial \theta \partial y}} > \frac{\frac{\partial l^p(\theta_L, y_L)}{\partial y} f(\theta_L)}{-\frac{\partial^2 l^\alpha(\theta_L, y_L)}{\partial \theta \partial y}}$ , which holds for all  $(\theta_L, y_L, \theta_H, y_H)$  in the domain of integration. This proves that the top expression is positive, and thus the inequality with integrals holds. Consequently, this condition is indeed sufficient.

**Proof of Lemma 1.** We start by noticing that for a fixed  $\theta$ , the expression under the integral is convex in  $y$  and is minimized at  $z(\theta)$ . Furthermore,  $z(\theta) > \theta$  for all  $\theta < 1$ .

Suppose, to obtain a contradiction, that the statement does not hold. Then there is  $\theta_0$  such that  $y^*(0) < y^*(\theta_0) < y^*(1)$  (which means, in particular, that  $0 < \theta_0 < 1$ ) and  $y^*(\theta_0) \neq \min\{z(\theta_0), \theta_0 + b\}$ . First, consider the case where  $z(\theta)$  is nondecreasing (note that it is a linear function of  $\theta$ ). Suppose first that  $y^*(\theta_0) < \min\{z(\theta_0), \theta_0 + b\}$ . Then, by continuity of  $y^*(\theta)$  and the assumption that  $y^*(\theta_0) < y^*(1)$ , there exists  $\theta' > \theta_0$  such that  $y^*(\theta) < \min\{z(\theta), \theta + b\}$  for all  $\theta \in [\theta_0, \theta']$  and also  $y^*(\theta_0) < y^*(\theta')$ . But then slightly increasing  $y^*(\theta)$  for  $\theta \in (\theta_0, \theta')$  while preserving  $y^*(\theta_0)$  and  $y^*(\theta')$  would decrease the minimand, because of convexity in  $y$ . Now suppose  $y^*(\theta_0) > \min\{z(\theta_0), \theta_0 + b\}$ ; since  $y^*(\theta_0) \leq \theta_0 + b$  for functions in  $\mathcal{F}$ , we must have  $z(\theta_0) < y^*(\theta_0) \leq \theta_0 + b$ . Since  $z(\theta)$  is nondecreasing and  $y^*(\cdot)$  is continuous, we can choose  $\theta' < \theta_0$  such that  $z(\theta') < y^*(\theta') < y^*(\theta_0)$ . Then if we slightly decrease  $y^*(\theta)$  for  $\theta \in (\theta', \theta_0)$  while preserving  $y^*(\theta')$  and  $y^*(\theta_0)$ , this would again decrease the minimand. In either case, if  $z(\theta)$  is nondecreasing, we get a contradiction.

Now suppose that  $z(\theta)$  is strictly decreasing. Let us first suppose that  $y^*(\theta_0) > \min\{z(\theta_0), \theta_0 + b\}$ , which means  $z(\theta_0) < y^*(\theta_0) \leq \theta_0 + b$ . Then  $z(1) < y^*(1) \leq 1 + b$ , so we could slightly decrease  $y^*(\theta)$  for  $\theta \in (\theta_0, 1]$  while preserving  $y^*(\theta_0)$  and thereby make  $y^*(\theta)$  closer to  $z(\theta)$  on  $(\theta_0, 1]$ ; this would decrease the minimand. This means, in particular, that in this case, if for some  $\theta'$ ,  $y^*(\theta') = z(\theta')$ , then  $y^*(\theta') = y^*(1)$ : indeed, this is trivially true if  $\theta' = 1$ , while if  $\theta' < 1$  and  $y^*(\theta') \neq y^*(1)$  then there exists  $\theta > \theta'$  such that  $z(\theta) < y^*(\theta) < y^*(1)$ , which is, as we just proved, impossible. Now consider the remaining case,  $y^*(\theta_0) < \min\{z(\theta_0), \theta_0 + b\}$ . We have  $y^*(1) \geq 1 = z(1)$ , and thus there is some  $\theta' \in (\theta_0, 1]$  for which  $y^*(\theta') = z(\theta')$  (because  $y^*(\theta_0) < z(\theta_0)$  and  $y^*(1) \geq z(1)$ ), and then for  $\theta \in (\theta_0, \theta')$  we have  $y^*(\theta) < z(\theta)$ . Hence, if we slightly increase  $y^*(\theta)$  for  $\theta \in (\theta_0, \theta')$  while preserving  $y^*(\theta_0)$  and  $y^*(\theta')$ , we would decrease the minimand. In all cases, we get a contradiction to that  $y^*(\cdot)$  is optimal. This contradiction completes the proof of Lemma 1. ■

**Proof of Theorem 7.** Existence follows from Theorem 5. Uniqueness follows, since the constraints define a convex subset of a linear space, and the objective function is strictly convex in  $y(\cdot)$ , and linear in  $T$ , so multiplicity solely due to different  $T$  is not possible.

To proceed, we analyze the cases separately.

**Part 1.** Let us find the optimal contract for  $\tilde{w} - \tilde{u} \leq \Omega(A, b)$ , i.e., including the boundary  $\tilde{w} - \tilde{u} = \Omega(A, b)$ ; this will be helpful for analyzing Part 2. We now solve the problem (19) s.t. (20) while ignoring (21), and then show that (21) is satisfied. We use Theorem 8.4.1 from Luenberger (1969, p. 220; henceforth Luenberger's theorem) in the following way. We take the

Lagrange multiplier  $\lambda = 1$  and  $y^*(\cdot)$  and  $T^*$  given in the statement of the Theorem. To show that  $(y^*(\cdot), T^*)$  solve the problem (19) s.t. (20), we need to verify that  $(y^*(\cdot), T^*)$  minimizes

$$\min_{(y(\cdot), T) \in \mathcal{F} \times \mathbb{R}} \int_0^1 \left( A(y(\theta) - \theta)^2 - 2\lambda(y(\theta) - \theta - b)(1 - \theta) \right) d\theta + \lambda(y(0) - b)^2 + \lambda\tilde{u} + (1 - \lambda)T$$

for  $\lambda = 1$ . For such  $\lambda$ , this becomes

$$\min_{y(\cdot) \in \mathcal{F}} \int_0^1 \left( A(y(\theta) - \theta)^2 - 2(y(\theta) - \theta - b)(1 - \theta) \right) d\theta + (y(0) - b)^2 + \tilde{u}. \quad (40)$$

As follows from Claim 1, the solution takes the form  $\min(z(\theta), \theta + b)$ , with  $z(\cdot)$  given by  $z(\theta) = \theta + \frac{1}{A}(1 - \theta)$ , potentially with a floor and a cap. It thus remains to do the following: (1) optimize over the floor and the cap and show that the optimal contract is indeed given by  $y^*(\cdot)$ , (2) verify that the constraint (20) holds as equality; this would imply (by Luenberger's theorem) that  $(y^*(\cdot), T^*)$  indeed solves (19) s.t. (20) and, finally, (2) verify that  $T^* \geq \tilde{w}$  holds.

Before we proceed, the following calculation is useful. If  $y(\cdot)$  is a constant, then the value of the minimand equals  $(A + 1)y^2 - (A + 2b + 1)y + b^2 + b + \frac{A+1}{3} + \tilde{u}$ , and its minimal value is achieved at  $y(\theta) = \frac{1}{2} + \frac{b}{A+1}$  and is equal to  $\frac{A+1}{12} + \frac{Ab^2}{A+1} + \tilde{u}$ . In what follows, we find optimal nonconstant contracts (in which case Lemma 1 will be applicable), and then compare the minimum with  $\frac{A+1}{12} + \frac{Ab^2}{A+1} + \tilde{u}$ .

Case a:  $A \leq \min\left(1, \frac{1}{2b-1}\right)$ . Consider the case  $A \notin \left\{1, \frac{1}{2b-1}\right\}$ . In this case,  $z(\theta)$  is decreasing. If the optimal  $y(\cdot)$  is nonconstant, then it equals  $\min\{z(\theta), \theta + b\}$  on a segment of positive measure of  $\theta$ , and since  $z(\theta)$  is decreasing, we must have  $y(\theta) = \theta + b$  whenever  $y(\theta) \notin \{y(0), y(1)\}$ . If so, it must be that  $y(0) = b$  ( $y(0) > b$  is impossible for  $y(\cdot) \in \mathcal{F}$ , and  $y(0) < b$  would imply a discontinuity, which is also impossible for  $y(\cdot) \in \mathcal{F}$ ). Thus, if a nonconstant  $y(\cdot)$  is optimal, it must take the form  $y(\theta) = \theta + b$  for  $\theta < q$ ,  $y(\theta) = q + b$  for  $\theta \geq q$ , for some  $q > 0$ . To optimize with respect to  $q$ , we need to solve

$$\min_q \int_0^q (Ab^2) d\theta + \int_q^1 \left( A(q + b - \theta)^2 - 2(q - \theta)(1 - \theta) \right) d\theta + \tilde{u}. \quad (41)$$

This is a cubic polynomial in  $q$ , with a local minimum at  $q = 1 - \frac{2Ab}{A+1}$  and a local maximum at  $q = 1$ ; furthermore,  $1 - \frac{2Ab}{A+1} \in (0, 1)$ . This implies that the minimum is achieved at  $q = 1 - \frac{2Ab}{A+1}$ . The corresponding value  $y(q) = q + b = 1 + \frac{1-A}{1+A}b > 1$  (so this  $y(\cdot) \in \mathcal{F}$ ). Plugging this value of  $q$  into the minimand, we find that it equals  $Ab^2 - \frac{4}{3} \frac{A^3b^3}{(A+1)^2} + \tilde{u}$ , which is less than the optimal value under a constant contract  $\frac{A+1}{12} + \frac{Ab^2}{A+1} + \tilde{u}$  (the difference between the former and the latter equals  $Ab^2 - \frac{4}{3} \frac{A^3b^3}{(A+1)^2} - \left(\frac{A+1}{12} + \frac{Ab^2}{A+1}\right) = -\frac{1}{12}(A + 4Ab + 1) \frac{(1-2Ab+A)^2}{(A+1)^2} < 0$ ). Consequently,  $y^*(\cdot)$  indeed minimizes (40) if  $A < \min\left(1, \frac{1}{2b-1}\right)$ . If, however,  $A \in \left\{1, \frac{1}{2b-1}\right\}$ , then it follows by

continuity that  $y^*(\cdot)$  minimizes (40) in these limit cases as well. Plugging the  $y(\cdot)$  and  $T$  into (20), one immediately verifies that it holds as equality.

Case b:  $1 < A < \frac{1}{b}$ . In this case,  $z(\theta)$  is increasing, and, furthermore, satisfies  $z(0) = \frac{1}{A} > b$ . Moreover, in this case  $b < 1$ , and therefore  $\mathcal{F}$  does not contain constant functions  $y(\cdot)$ . Since the optimal  $y(\cdot)$  is nonconstant, it equals  $\min\{z(\theta), \theta + b\}$  whenever  $y(\theta) \notin \{y(0), y(1)\}$ , and this is true for a segment of positive measure. Furthermore, we must have  $y(0) \leq b = \min\{z(0), 0 + b\}$  and  $y(1) \geq 1 = \min\{z(1), 1 + b\}$ , which implies that there is neither a ‘floor’ (where  $y(\theta) = y(0)$  for a set of positive measure) nor a ‘cap’ (where  $y(\theta) = y(1)$  for a set of positive measure). Thus,  $y^*(\theta) = \min\{z(\theta), \theta + b\}$  is a function in  $\mathcal{F}$  that minimizes (40). Plugging  $y^*(\cdot)$  and  $T^*$  into (20), it is immediate to verify that it holds as equality.

Case c:  $\frac{1}{2b-1} < A < 2b - 1$ . Here, we need to prove that the optimal  $y(\cdot)$  is a constant, in which case, as we showed above, it indeed equals  $\frac{1}{2} + \frac{b}{A+1}$  (this constant is in  $\mathcal{F}$  in this case). Suppose that the optimal  $y(\cdot)$  is not a constant. First,  $\frac{1}{2b-1} < 2b - 1$  implies  $b > 1$ . Suppose first that  $Ab \leq 1$ ; this implies, in particular, that  $A < 1$ . In this case,  $z(\theta)$  is decreasing, and thus (similar to Case a)  $y(\theta)$  would have to equal  $\theta + b$  for  $\theta \leq q$  and  $q + b$  for  $\theta > q$ , for some  $q > 0$ . Optimizing with respect to  $q$  is equivalent to minimizing (41). Notice that its derivative is  $(1 - q)(2Ab + (A + 1)(q - 1))$ ; since the first term is nonnegative and the second is increasing in  $q$ , it is positive for all  $q \in (0, 1)$  if and only if it is positive at  $q = 0$ , which is true, as there it equals  $2Ab - A - 1 = A(2b - 1) - 1 > 0$ . Thus, minimum is attained at  $q = 0$ , where the function is a constant. This contradiction shows that the optimal  $y(\theta)$  is a constant in this case.

Second, suppose that  $Ab > 1$ , but  $A \leq 1$ . In this case,  $z(\theta)$  is nonincreasing, but  $z(\theta) < \theta + b$  for all  $\theta$ , so  $\min\{z(\theta), \theta + b\} = z(\theta)$ . But  $y(\theta) = z(\theta)$  cannot hold on an interval of  $\theta$  where  $y(\theta)$  is increasing, so  $y(0) < y(1)$  is impossible, and  $y(\cdot)$  must be a constant, again a contradiction.

Finally, suppose  $A > 1$ . Then  $z(\theta)$  is an increasing function; since  $z(1) = 1$ , then if  $y(\theta)$  is not a constant, it must equal  $z(\theta)$  for  $\theta \geq q$  and  $z(q)$  for  $\theta < q$  for some  $q < 1$ . Let us optimize over  $q$ ; we need to solve

$$\begin{aligned} \min_q \int_0^q \left( A \left( 1 - \frac{A-1}{A} (1-q) - \theta \right)^2 - 2 \left( 1 - \frac{A-1}{A} (1-q) - \theta - b \right) (1-\theta) \right) d\theta + \quad (42) \\ \int_q^1 \left( A \left( 1 - \frac{A-1}{A} (1-\theta) - \theta \right)^2 - 2 \left( 1 - \frac{A-1}{A} (1-\theta) - \theta - b \right) (1-\theta) \right) d\theta \\ + \left( 1 - \frac{A-1}{A} (1-q) - b \right)^2 + \tilde{u}. \end{aligned}$$

The derivative of the minimand with respect to  $q$  equals  $(A - 1) \frac{A(A-1)q^2 + 2(A-1)q + 2(1-Ab)}{A^2}$ , which is negative for all  $q$  (indeed, this is increasing in  $q$ , so it suffices to check that the sign is negative

for  $q = 1$ , and in this case it equals  $\frac{A-1}{A}(A - (2b - 1)) < 0$ . Thus, minimum is reached at  $q = 1$ , and for this  $q$  the function is actually a constant ( $y(\theta) = 1$ ), again a contradiction.

Consequently, our assertion that  $y(\theta)$  is not constant was shown to be wrong, therefore, in this case  $y(\theta)$  is a constant, and as shown above,  $y(\theta) = \frac{1}{2} + \frac{b}{A+1}$ ; this function lies in  $\mathcal{F}$  and therefore it solves (40). Now it is straightforward to plug  $y^*(\theta)$  and  $T^*$  into (20) and verify that it holds as equality.

Case d:  $A \geq \max(\frac{1}{b}, 2b - 1)$ . As in case a, start with assuming  $A \notin \{\frac{1}{b}, 2b - 1\}$ . In this case,  $A > 1$  (if  $b \leq 1$  this follows from  $A > \frac{1}{b}$  and if  $b \geq 1$  this follows from  $A > 2b - 1$ ), and therefore  $z(\theta)$  is increasing. Furthermore,  $Ab \geq 1$  implies that  $z(\theta) \leq \theta + b$  for all  $\theta$  and thus, as in the similar subcase of Case c, if  $y(\theta)$  is not a constant, then it must equal  $z(\theta)$  for  $\theta \geq q$  and  $z(q)$  for  $\theta < q$  for some  $q < 1$ . Let us optimize over  $q$ ; we need to solve (42), which equals

$$\frac{(A-1)^2}{3A}q^3 + \frac{(A-1)^2}{A^2}q^2 + 2\frac{(A-1)(1-Ab)}{A^2}q + \frac{1}{3A^2}(3A^2b^2 + 3A^2b - 6Ab - A + 3) + \tilde{u}.$$

Its derivative with respect to  $q$  equals  $(A-1)\frac{A(A-1)q^2 + 2(A-1)q + 2(1-Ab)}{A^2}$ . In this case, however, this derivative is monotonically increasing in  $q$ ; it is negative (equal to  $2\frac{A-1}{A^2}(1-Ab)$ ) if  $q = 0$  and positive (equal to  $\frac{A-1}{A}(A - (2b - 1))$ ) for  $q = 1$ . Thus, (42) has a unique minimum on  $[0, 1]$ , attained at  $q = \frac{1}{A}\left(\sqrt{1 + 2A\frac{Ab-1}{A-1}} - 1\right)$ . This function  $y(\theta)$  is in  $\mathcal{F}$ , is nonconstant (since we assumed  $A \neq \frac{1}{b}$ ), and the corresponding value of the minimand is  $\frac{2}{3}(A-1)\frac{1+A-2A^2b}{A^4}\sqrt{1 + 2A\frac{Ab-1}{A-1}} + \frac{3A^2b(A^2b+A^2-2)-(A+1)(A^2-2)}{3A^4} + \tilde{u}$ .

Now consider the possibility that  $y(\theta)$  is a constant. This cannot be the solution for  $b < 1$ , since in this case there are no constant functions in  $\mathcal{F}$ . If  $b \geq 1$ , notice that if  $y(\cdot)$  is a constant, then the minimand (40) is strictly convex in this value. In the case under consideration,  $A \geq 2b + 1$  implies  $\frac{1}{2} + \frac{b}{A+1} \leq 1$ , and since (40) is minimized (among constant  $y(\cdot)$ ) at  $y(\theta) = \frac{1}{2} + \frac{b}{A+1}$ , the constant function in  $\mathcal{F}$  that minimizes (40) must be the one closest to  $\frac{1}{2} + \frac{b}{A+1}$ , i.e.,  $y(\theta) = 1$ . Notice, however, that this function belong in the class above for  $q = 0$ . Thus, if  $A > \max(\frac{1}{b}, 2b - 1)$ ,  $y(\theta) = 1$  does not achieve minimum and thus  $y^*(\theta)$  minimizes (40), and if  $A \in \{\frac{1}{b}, 2b - 1\}$ , then  $y^*(\cdot)$  minimizes (40) by continuity. It remains to plug  $y^*(\theta)$  and  $T^*$  into (20) and verify that it holds as equality, which is straightforward.

Luenberger's theorem now implies that  $(y^*(\cdot), T^*)$  indeed solves (19) s.t. (20). It remains to check that  $T^* \geq \tilde{w}$ , but this is equivalent to  $\tilde{w} - \tilde{u} \leq \Omega(A, b)$ , which holds in the case under consideration. Furthermore, if  $\tilde{w} - \tilde{u} < \Omega(A, b)$ , so the inequality is strict, then  $T^* > \tilde{w}$ , so the transfer is higher than minimal. This completes the proof of Part 1.

**Part 2.** In these cases,  $\Omega(A, b) \leq \tilde{w} - \tilde{u} < b^2 + b$ , we again use Luenberger's theorem, this time applying it to the problem (19) s.t. (20) and (21) directly. We take Lagrange multipliers

$\lambda = \frac{A}{D}$  (where  $D = D(\tilde{w} - \tilde{u}, b)$  is defined in the statement of the Theorem; as it is easy to check,  $\Omega(A, b)$  is strictly increasing from 0 to  $b + b^2$  as  $A$  increases from 0 to  $\infty$ , and therefore  $\Omega(D, b) = \tilde{w} - \tilde{u}$  has a unique solution for  $0 < \tilde{w} - \tilde{u} < b + b^2$ , which, if  $\Omega(A, b) \leq \tilde{w} - \tilde{u}$ , satisfies  $D \geq A$ ) and  $\mu = 1 - \frac{A}{D}$ , and verify that  $(y^*(\cdot), T^* = \tilde{w})$  minimize

$$\min_{(y(\cdot), T) \in Z} \int_0^1 \left( A(y(\theta) - \theta)^2 - 2\lambda(y(\theta) - \theta - b)(1 - \theta) \right) d\theta + \lambda(y(0) - b)^2 + \lambda\tilde{u} + \mu\tilde{w} + (1 - \lambda - \mu)T. \quad (43)$$

Notice, however, that the last term  $(1 - \lambda - \mu)T$  vanishes (as  $\lambda + \mu = 1$ ) and  $\mu\tilde{w}$  is a constant, so the problem is equivalent to

$$\min_{y(\cdot) \in \mathcal{F}} \int_0^1 \left( A(y(\theta) - \theta)^2 - 2\lambda(y(\theta) - \theta - b)(1 - \theta) \right) d\theta + \lambda(y(0) - b)^2 + \lambda\tilde{u}.$$

Since  $\lambda$  is a positive scalar, we can divide by it, getting

$$\min_{y(\cdot) \in \mathcal{F}} \int_0^1 \left( D(y(\theta) - \theta)^2 - 2(y(\theta) - \theta - b)(1 - \theta) \right) d\theta + (y(0) - b)^2 + \tilde{u},$$

but this is exactly the same problem as (40), except that  $A$  is substituted for  $D$ . This immediately implies that  $y^*(\cdot)$  indeed minimizes the (43).

Let us show that (20) and (21) hold as equalities. This is trivially true for (21). For (20), notice that  $y^*(\cdot), T^*$  is also the solution for the set of parameters  $(A', b', \tilde{u}', \tilde{w}') = (D, b, \tilde{u}, \tilde{w})$  (since  $\Omega(D, b) = \tilde{w} - \tilde{u}$ , Part 1 prescribes  $T^* = \tilde{w}$ ). This means that (20) holds as equality for  $(A', b', \tilde{u}', \tilde{w}')$ . Notice, however, that it does not explicitly depend on parameter  $A$ , which implies that it holds as equality for  $(A, b, \tilde{u}, \tilde{w})$  as well. Thus, for  $y^*(\cdot)$  and  $T^*$ , both constraints (20) and (21) hold as equality. Now, Luenberger's theorem implies that  $(y^*(\cdot), T^* = \tilde{w})$  is indeed the solution to (19) s.t. (20) and (21).

**Part 3.** Notice that the value of (19) cannot be less than  $\tilde{w}$ , if the constraints are to be satisfied, and moreover it is achieved under  $y^*(\theta) = \theta, T^* = \tilde{w}$ . Let us verify that if  $\tilde{w} - \tilde{u} \geq b + b^2$ , then (20) is satisfied as well. Indeed, the left-hand side of (20) becomes

$$(0 - b)^2 - \int_0^1 2(\theta - \theta - b)(1 - \theta) d\theta - \tilde{w} + \tilde{u} = b + b^2 - (\tilde{w} - \tilde{u}) \leq 0.$$

Thus, this contract is indeed optimal. ■

**Proof of Theorem 8. Part 1.** According to Theorem 7, if  $b \leq 1$ , then money burning is used either if  $A > 1$  or if  $\tilde{w} - \tilde{u} \geq \frac{b^3}{3}$ , and if  $b > 1$ , then it is used if either if  $A > 2b - 1$  or  $\tilde{w} - \tilde{u} > b^2 - b + \frac{1}{3}$ . Together, this amounts to the condition  $A > 2\max(b, 1) - 1$  or  $\tilde{w} - \tilde{u} > \frac{b^3 - (\max(b, 1) - 1)^3}{3}$ . The comparative statics results here are straightforward.

**Part 2.** Consider first comparative statics with respect to  $A$ . It suffices to study the effect of a small increase in  $A$ . From Theorem 7 it follows that if  $\tilde{w} - \tilde{u} > \Omega(A, b)$ , a small increase in  $A$  does not change the optimal  $y(\cdot)$  and thus does not affect money-burning, thus, consider the case  $\tilde{w} - \tilde{u} \leq \Omega(A, b)$ .

If  $A < \min\left(1, \frac{1}{2b-1}\right)$  or  $\frac{1}{2b-1} < A < 2b-1$ , then there is no money-burning, and a small change in  $A$  does not change this. Consider the case  $1 < A < \frac{1}{b}$ . Taking Claim 4 into account, it suffices to show that optimal  $y^*(\theta)$  is weakly decreasing in  $A$  for all  $\theta$ . But this immediately follows from the formula  $y^*(\theta) = \min\left(\theta + b, \frac{1}{A} + \left(1 - \frac{1}{A}\right)\theta\right)$ .

Lastly, consider the case  $A > \max\left(\frac{1}{b}, 2b-1\right)$ , where  $y^*(\theta) = \max\left(\frac{1}{A} + \left(1 - \frac{1}{A}\right)\theta, \frac{1 + \sqrt{(A-1)(2A^2b-A-1)}}{A^2}\right)$ . The first expression  $\frac{1}{A} + \left(1 - \frac{1}{A}\right)\theta$  is decreasing in  $A$ , so it suffices to show that  $h(A, b) = \frac{1 + \sqrt{(A-1)(2A^2b-A-1)}}{A^2}$  also is. We have

$$\frac{\partial h(A, b)}{\partial A} = \frac{2A^2b - A^3b + A^2 - 2 - 2\sqrt{(A-1)(2A^2b-A-1)}}{A^3\sqrt{(A-1)(2A^2b-A-1)}}.$$

Notice also that

$$\frac{\partial h(A, b)}{\partial b} = \frac{1}{\sqrt{1 + \frac{2A(Ab-1)}{A-1}}},$$

and thus the sign of  $\frac{\partial^2 h(A, b)}{\partial A \partial b}$  is the same as the sign of  $\frac{d}{dA} \left(-\frac{2A(Ab-1)}{A-1}\right) = 2\frac{2Ab-A^2b-1}{(A-1)^2}$ . If  $A \geq 2$ , the numerator is negative, which means that for a fixed  $A$ ,  $\frac{\partial h(A, b)}{\partial A}$  is decreasing in  $b$ . Since  $A > \frac{1}{b}$ , it suffices to check that  $\frac{\partial h(A, b)}{\partial A} \leq 0$  for  $b = \frac{1}{A}$ , and it turns out that it equals 0. Thus, for such  $A$  and  $b > \frac{1}{A}$ ,  $\frac{\partial h(A, b)}{\partial A} < 0$ . If, however,  $A < 2$ , then  $2Ab - A^2b - 1$  is negative for  $b < \frac{1}{2A-A^2}$  and positive for  $b > \frac{1}{2A-A^2}$ . Thus, for a fixed  $A < 1$ ,  $\frac{\partial h(A, b)}{\partial A}$  is decreasing in  $b$  on  $\left(\frac{1}{A}, \frac{1}{2A-A^2}\right)$  and is increasing in  $b$  thereafter. We already know that for  $b = \frac{1}{A}$ ,  $\frac{\partial h(A, b)}{\partial A} = 0$ ; since  $A > 2b-1$  in this case, it now suffices to verify that for  $b = \frac{A+1}{2}$ ,  $\frac{\partial h(A, b)}{\partial A} \leq 0$ . But this is true, since it in fact equals  $-\frac{1}{2(A+1)} < 0$ . This proves that for all parameter values where  $A > \max\left(\frac{1}{b}, 2b-1\right)$ ,  $\frac{\partial h(A, b)}{\partial A}$  is negative. Thus,  $y^*(\theta)$  is weakly increasing in  $A$  for all  $\theta$ , which completes this part of the proof.

The reasoning in the case of  $\tilde{w}$  and  $\tilde{u}$  is similar and is omitted. To see that the amount of money-burning may be nonmonotone in  $b$ , take  $A = 2$  and  $\tilde{w} - \tilde{u} = \frac{5}{16}$ . Then if  $b = 1$ , there is no money burning, if  $b = \frac{1}{2}$ , there is a positive amount of money-burning (totally,  $\frac{1}{3}Ab^3(A-1) = \frac{1}{12}$ ), but for  $b < \frac{1}{4}$ , the total amount of money-burning equals  $b$  (and the principal implements his ideal contract) and therefore decreases as  $b$  decreases further. This completes the proof of Part 2.

**Part 3.** The first part of the statement immediately follows from Theorem 7, and the second from the fact that  $\Omega(A, b)$  is increasing in  $A$  and  $b$ .

**Part 4.** Immediately follows from Theorem 7. ■

**Proof of Theorem 9.** These statements immediately follow from Theorem 7. ■

**Proof of Theorem 10.** From Theorem 7, it follows that the agent's participation constraint (20) binds if and only if  $\tilde{w} - \tilde{u} \leq b + b^2$ , in which case agent's utility is  $\tilde{u}$ . If  $\tilde{w} - \tilde{u} \leq b + b^2$ , principal's ideal  $y(\theta) = \theta$  is implemented, and the agent's utility is  $T - \left( (y(0) - b)^2 - \int_0^1 2(y(\theta) - \theta - b)(1 - \theta) d\theta \right) = \tilde{w} - b - b^2$ , which proves the first part of the statement.

To show that the principal's utility is decreasing in  $\tilde{u}$  and  $\tilde{w}$ , notice that lower  $\tilde{u}$  or  $\tilde{w}$  relax both constraints (20) and (21), thus the principal cannot be worse off, and it is trivial to find examples where he is better off. A higher  $A$  decreases the principal's payoff from any fixed scheme  $y(\cdot)$  (except if  $y(\theta) = \theta$  where an increase in  $A$  has no effect), and thus he is worse off. Finally, suppose that  $y(\cdot)$  solved the problem for  $(A, b, \tilde{u}, \tilde{w})$  and let  $b' < b$ . Take  $y'(\theta) = \frac{b'}{b}(y(\theta) - \theta) + \theta$  and take the same transfer  $T' = \tilde{w}$ . This new contract satisfies the constraints (20) and (21) because the previous one did (it makes agent's disutility from actions less and weakly reduces money-burning), and it also makes the principal better off. Thus, there is a contract that makes the principal better off, which means that if the principal were to choose the optimal contract under new  $b'$ , he would be better off from a lower  $b'$ . This completes the proof. ■