Delegation and Procedural Rules

Attila Ambrus*
Harvard University, Department of Economics

Georgy Egorov†
Harvard University, Department of Economics

October 2008

Abstract

This paper investigates delegating decision-making to an informed agent with bureaucratic procedural rules. We assume that these rules are purely wasteful and refer to them as money burning. We show that the optimal delegation contract can involve money burning, both when contingent monetary transfers are not possible (as in most of the delegation literature), and when contingent transfers are allowed, but payments from the principal to the agent are bounded from below. We establish some general properties of the optimal contract with money burning, and show that under certain regularity conditions the contract imposes zero money burning in low states, and that the amount of money burning is continuous and increasing in the state. The regularity conditions also imply that the implemented action is always between the ideal points of the principal and the agent. We explicitly solve for the optimal contract in a class of models with quadratic loss functions, and investigate how the contract changes as the choice of action becomes more important to the principal than to the agent, in monetary terms. We also investigate features of the optimal contract in cases when the regularity conditions do not hold. In the case when both transfers and monetary transfers are allowed, we show that whether the optimal contract involves inefficiencies and money burning depends on the outside option of the agent, relative to the required minimum wage.

Keywords: delegation, organizational procedures, money burning
JEL Classification: D23, D82, D86.

*E-mail: ambrus@fas.harvard.edu
†E-mail: gegorov@fas.harvard.edu

1 Introduction

There is little doubt that bureaucracy, typified by formal processes, standardization, hierarchic procedures and written communication, has a large presence in organizations, both in the government sector and private companies. In fact, it is a common perception that there is an excess of bureaucratic procedures. According to a recent study released by MeaningfulWork.com, the number 1 complaint workers had about their job was "too much workplace bureaucracy."[1] Indeed, most of the related economic literature takes this stance, looking for explanations for excessive bureaucratization as in Martimort (1997) and Strausz (2006), or connect bureaucracy and corruption as in Banerjee (1997) or Guriev (2004).[2]

While not denying the possibility of excessive bureaucratization in certain organizations, this paper investigates a theory in which bureaucratic procedures improve the efficiency of an organization by aligning the interests of workers with the organizational objectives. In general, a principal might have various tools to induce an agent with private information to take certain actions: restricting the set of actions that can be chosen by the agent, committing to an action-contingent monetary transfer scheme, or as in the current paper, require the agent to perform a costly activity, the amount of which depends on the action chosen. This activity can take the form of going to the headquarters to get a stamp of approval, writing a proposal of a certain required length, or filling out various forms, all of which are commonly used in practice. These activities might serve some useful purpose even directly, for example by providing a more precise documentation of the agent's activities. However, for conceptual simplicity throughout the paper we assume that the bureaucratic procedures imposed on the agent are purely wasteful, and refer to them as money burning.

The mechanism we investigate can be observed in many settings. Taking clients to a more expensive restaurant from the corporate budget usually requires extra paperwork and written explanation. Universities often offer small research grants to their faculty and students, with the understanding that receiving the grant is fairly automatic, but requires turning in a few page long proposal, to screen out people who have a positive but very small benefit from these grants. The usage of electronic surveillance devices by federal agents in the United States requires acquiring warrants through procedures prescribed by the Foreign Intelligence Surveillance Act, to ensure that these devices are only used if other means to continue an investigation or operation are likely to be ineffective. Recently, the Bush administration tried to circumvent these regulations

[1] See Leonard (2000).

[2] In Banerjee (1997) red tape also serves a useful function, in that it is used by corrupt bureaucrats to allocate the good to some people who really need the good but are credit constrained.

several times, arguing that the prescribed procedures were too complicated and bureaucratic. However, this might very well be in the interest of the legislation, that plays the role of the principal in this setting, if members of the legislation are concerned about the widespread use of electronic surveillance.

Formally, we consider a scenario in which an uninformed principal delegates the task of choosing the level of a one-dimensional variable to an agent, who before making the decision receives a 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. However, we assume that at any state the agent would like to choose a higher action than what is optimal for the principal. The delegation involves the principal setting a procedural rule scheme, which for any possible action prescribes a nonnegative amount of money burning imposed on the agent for taking that action. On top of this, the principal has to offer the agent an ex ante transfer large enough to induce the agent to accept the job. The required amount depends on the procedural rule scheme, since the latter influences both what actions the agent expects to take at different states, and how much money burning she anticipates.

We show that purely wasteful bureaucratic procedures can be part of an optimal delegation contract. First we show this in the context in which contingent monetary transfers between the principal and the agent are not possible. As pointed out elsewhere, there are various reasons why in a given context monetary transfers are not possible: legal restrictions (as in between regulated firms and a regulatory body), organizational culture, or the possibility that one of the parties is liquidity/credit constrained. In fact, cash penalties are hardly ever used in employment contracts, except for employees with very large incomes relative to the damage that their misbehavior can do.[3] In this context, we first derive some basic properties of an optimal contract: (i) the implemented action is increasing in the state; (ii) the implemented action never falls below the principal's ideal point (no undershooting); (iii) low enough actions can always be chosen freely by the agent; (iv) the utility of the agent is continuous in the state. Furthermore, if some regularity conditions hold, then both the implemented action and money burning are continuous in the state, money burning is increasing in the state, and the implemented action never exceeds the agent's ideal point (no overshooting). For state-independent biases and symmetric loss functions the regularity conditions for the above results require that the density function of the state is non-decreasing, and that the principal's loss function is "convex enough" in a particular formal sense. An important special case when the regularity condition on the loss

[3]See p.249 of Milgrom and Roberts (1992). The authors also point out that "when cash penalties are limited, other details of the transaction may have to be manipulated to provide suitable incentives to parties" - which is exactly what costly procedural rules achieve in our model.

functions hold is when the agent has a quadratic utility function (while the principal can have any strictly convex loss function).

We explicitly solve for the optimal contract in the frequently used uniform-quadratic specification of the delegation model. We extend this example by introducing an extra parameter determining the relative importance of the action chosen for the principal and the sender (in monetary terms). For example if the principal represents the government, the chosen action greatly influences the well-being of citizens, and the bias of the agent results from small private benefits from the chosen action, then the action choice matters much more for the principal, reflected by a large parameter value. We show that the optimal contract depends crucially on this parameter: if the relative weight of the action choice's impact on the principal's utility measured in the numeraire good is smaller or equal to the agent's, then money burning is not used in equilibrium, and the optimal scheme allows the agent to choose any low enough action freely, while not letting her choose actions above a certain cap. The intuition behind this is that if the agent cares about the decision at least as much as the principal, then money burning is a too costly incentive device for the principal. However, if the principal cares about the decision more than the agent, then the optimal contract does involve money burning. In particular, actions that are low enough can still be freely chosen by the agent, but there is a threshold above which the amount of money burning is strictly increasing in the action chosen. As the relative importance of the principal goes to infinity, this threshold converges to the lowest possible action (that is, there is money burning in a larger and larger interval), and the implemented action choices converge to the principal's optimum.

We also investigate cases when the regularity conditions do not hold. We find that the optimal contract might prescribe discontinuities in the action choice and the amount of money burning, that the amount of money burning might be non-monotonic in the state, and that the implemented action choice might involve overshooting, that is selecting an action that is strictly higher than the agent's optimal point at the given state. Moreover, we show that these features of the optimal contract are interrelated. The intuition is that in order to keep the action choice between the optimal points of the principal and the agent, the amount of prescribed money burning needs to be monotonically increasing. This makes aligning interest through money burning in low states very costly for the principal, since this amount of money needs to be burnt in all higher states as well. The only way the principal can decrease money burning at some state in an incentive compatible way is if she prescribes an overshooting action. We show that this can indeed be optimal for the principal. For example if the density of the prior distribution is high for both low and high states but very low for an interval of states in the middle, then the optimal policy involves increasing money burning in low states, then a discontinuous jump and overshooting when reaching the interval of unlikely states (which brings the level of money burning back to zero), and increasing money burning again in high states. Intuitively, the principal sacrifices utility in the unlikely states (the implemented action is too high for both parties), in order to better align incentives in the more likely states and at the same time do not accumulate too high levels of money burning (which is costly from the ex ante point of view).

The possibility of discontinuities in the optimal action scheme points out that optimal control methods like in Krishna and Morgan (2008) cannot be used to derive the optimal contract in our setting. Indeed, like Kováč and Mylovanov (2007), throughout the paper we only use techniques that do not impose continuity on the solution.

We conclude the paper with an extension of the model in which we allow the principal to be able to specify contingent monetary transfers in the contract, besides money burning. We assume that there is a lower bound on the transfer that the principal has to pay to the agent. This can correspond to a minimum wage requirement, or if negative then to the maximal amount of fee that the principal can impose on the agent. We show that qualitative features of the optimal contract remain the same if the same regularity conditions hold as before. The optimal contract can imply only transfers, or only money burning, or transfers in some states and money burning in other states. In general the optimal contract prescribes decreasing positive monetary transfers in low states and increasing money burning in high states. The principal might use money burning as an incentive tool, despite being more inefficient than monetary transfers, because the optimal contract when only using transfers that are required to be above the lower bound might result in the agent's participation constraint being slack. Therefore, the principal can improve her expected utility by introducing money burning in high states and further align the interests of the agent with those of the principal. We show that if the outside option of the agent is high enough, then there is no wasteful money burning, and the optimal contract achieves jointly efficient action choices. On the other hand, if the agent's outside option is low enough (relative to the required lower bound on transfers) then the contract prescribes money burning, possibly in almost every state.

Our work continues the literature on delegation started by Holmstrom (1977). Holmstrom, as well as Melumad and Shibano (1991) and Alonso and Matouschek (2007, 2008) consider deterministic delegation with no monetary transfers, in which the principal can only restrict the action space of the agent, instead of making different actions differentially costly.[4] In our

[4] 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

framework the principal can always achieve such delegation schemes by setting some actions free while the remaining ones prohibitively costly. This means that in our framework the principal has a larger set of feasible contracts and hence she is at least weakly better off. We show that in fact in a lot of cases the principal is strictly better off in this extended framework. Kováč and Mylovanov (2007) and Goltsman et al. (2007) investigate stochastic delegation mechanisms in the same context, assuming that that the principal and the agent have quadratic utility functions.[5] There is a mathematical connection between these works and our paper, for the following reason. 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 another term that only depends on the variance of the induced action. In particular the latter variance term enters negatively in both parties' utility functions. In our model, money burning is only a direct cost for the agent. However, through the participation constrained money burning is also costly for the principal, by increasing the amount of ex ante transfer required for the agent accepting the job. Ottaviani (2000), Kahmer (2004) and Krishna and Morgan (2008) investigate delegation with monetary transfers. A major alternative of delegation is cheap talk communication between the informed and the uninformed party, as in Crawford and Sobel (1982) and a large literature building on it. The closest papers to our work in this literature are Austen-Smith (2000) and Kartik (2007), who consider communication with money burning by the informed party. In essence, these papers can be viewed as the "signaling" versions of our "screening" model.

The existing literature on both delegation and cheap talk only considers incentive compatibility constraints. A conceptual contribution of the current paper is incorporating a participation constraint for the agent. Furthermore, to our best knowledge ours is the first paper investigating the effects of changing the relative importance of the action choice (in monetary terms) for the informed and the uninformed party, which is relevant for comparing different types of organizations in terms of the contracts prevailing there.

The formal literature on procedural rules and organizational bureaucracy, despite its practical importance, is scarce. Garicano (2000) investigates optimal knowledge hierarchies in an organization, which set rules for communicating and solving tasks by agents of different skills within the organization.[6] Crémer et al. (2007) investigate communication protocols within organizations. At an informal level, Walsh and Devar (1987) claim that procedural rules might have a positive effect on administrative efficiency and organizational effectiveness because they

example Bendor et al. (1987), and Mc Cubbins et al. (1987).

[5]Stochastic delagation implies that the principal can commit to different probabilistic action choices after different reports by the agent. That is, the agent cannot determine the action any more, but she can choose among probability distributions of actions after observing the state.

[6]Se also Garicano and Rossi-Hansberg (2006).

provide a set of role expectations and reduce uncertainty, while Wilson (1989) argues that complex rules and regulations are imposed on bureaucracy to reduce favoritism and discretion in order to contain corruption.

2 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 one-sided contingent transfers, see Section 5.

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 θ as the state. The c.d.f. of θ 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* 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 θ, 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 θ, $u^p(\theta, y)$ 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 points of the principal and the agent at state θ. We refer to $b(\theta)$ as the bias of the agent at state θ. 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. 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 Crawford and Sobel (1982). 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 $m : Y \to \mathbb{R}^+$. 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 offered specifies the cost function m (interpreted as the description of the paperwork requirements), and a constant transfer payment T (interpreted as a wage) that enters the agent's utility function in an additively separable manner. We assume that monetary transfers contingent on either θ or y are are not possible.[7] The agent has an outside option U_0, therefore we assume that he accepts any contract that gives him at least this much expected utility, given the ex ante distribution of θ.

3 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.

We start the analysis by writing the delegation problem in the direct mechanism interpretation. Trivially, the revelation principle applies, and the principal's task is therefore to define a pair of measurable functions $y(\theta)$ and $m(\theta)$ (where $y(\theta)$ is the action that the agent is supposed to choose in state θ) that solve the following problem:

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

$$\text{s.t.} \int_{\Theta} (u^a(\theta, y(\theta)) - m(\theta))\, dF(\theta) + T \geq U_0 \tag{1b}$$

$$\forall \theta, \theta' \in \Theta : u^a(\theta, y(\theta)) - u(\theta) \geq u^a(\theta, y(\theta')) - m(\theta') \tag{1c}$$

$$\forall \theta \in \Theta : m(\theta) \geq 0. \tag{1d}$$

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

[7] See Section 6 where we drop this assumption.

and incentive compatibility constraints (equations 1b and 1c, respectively), where we assume that agent's reservation utility is U_0.

First we observe that in an optimal contract the agent's participation constraint ($1b$) has to bind: 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. Substituting this into the principal's problem yields (we denote the principal's loss from contract $(y(\cdot), m(\cdot))$ by $V^p(y(\cdot), m(\cdot))$):

$$\min_{\{y(\theta),m(\theta)\}_{\theta\in\Theta}} V^p(y(\cdot), m(\cdot)) = \min_{\{y(\theta),m(\theta)\}_{\theta\in\Theta}} \int_\Theta (l^p(\theta, y(\theta)) + l^a(\theta, y(\theta)) + m(\theta))\, d\theta \tag{2}$$

$$\text{s.t. } \forall\, \theta, \theta' \in \Theta : l^a(\theta, y(\theta)) + m(\theta) \leq l^a(\theta, y(\theta')) + m(\theta') \tag{3}$$

$$\forall\, \theta \in \Theta : m(\theta) \geq 0 \tag{4}$$

In what follows, we solve problem $(2) - (4)$.

Next, we derive a series of lemmas which establish properties of the optimal contract that hold in general. The first one, which can be derived from the incentive compatibility constraints using the single-crossing property of the agent's loss function, is that the implemented action is a weakly increasing function of the state. All proofs are in the Appendix.

Claim 1 *Suppose that pair of functions* $\{y(\theta), m(\theta)\}_{\theta\in\Theta}$ *satisfies* (3). *Then* $\theta_2 \geq \theta_1$ *implies* $y(\theta_2) \geq y(\theta_1)$.

The next two lemmas establish that the money burning scheme has bounded variation, and that the amount of money burnt at different states cannot be bounded away from zero. The two results together imply that the amount of money burnt when using an optimal contract is bounded (for any distribution of states). The result on bounded variation follows from the incentive compatibility constraints and the fact that the agent's loss function is continuously differentiable in the state, since these imply that the amount of money burnt in one state cannot be arbitrarily higher than in another one. The result that money burning is not bounded away from zero follows from the simple observation that otherwise the amount of money burnt could be decreased in all states without affecting the IC constraints, easing the IR constraint.

Define

$$\begin{aligned}
\Delta_\theta &= \max_{\theta\in\Theta, y\in Y} \left| \frac{\partial l^a(\theta, y)}{\partial \theta} \right| \\
\Delta_y &= \max_{\theta\in\Theta, y\in Y} \left| \frac{\partial l^a(\theta, y)}{\partial y} \right|.
\end{aligned}$$

Claim 2 *Suppose that pair of functions* $\{y(\theta), m(\theta)\}_{\theta\in\Theta}$ *satisfies* (3). *Then* $\forall\ \theta_1, \theta_2 \in \Theta$*:*

$$|m(\theta_2) - m(\theta_1)| \leq |y(\theta_2) - y(\theta_1)|\,\Delta_y, \tag{5}$$

In particular, function $|m(\theta_2) - m(\theta_1)|$ *has bounded variation, and its variation on* Θ *does not exceed* $(y(1) - y(0))\,\Delta_y$

The next claim is almost trivial, and shows that the amount of money burnt in different states cannot be bounded away from zero at the optimum.

Claim 3 *If* (y^*,m^*) *is a solution to the problem (2)-(4) then* $\inf_{\theta\in\Theta} m^*(\theta) = 0$.

Next, let us define the total utility loss of the agent conditional on θ as follows:

$$L^a(\theta) = L^a(\theta, y(\theta), m(\theta)) = l^a(\theta, y(\theta)) + m(\theta)\,. \tag{6}$$

The following lemma states that the agent's total loss conditional on the state (and hence her ex post utility) is a Lipschitz-continuous function of the state, with Lipschitz parameter depending on the loss function of the agent. Moreover, the agent's ex post utility function has left and right derivatives, and these derivatives can be expressed directly from the agent's loss function and the actions scheme. The latter will be useful below for deriving a convenient integral representation of the amount of money burnt at different states.

Claim 4 *Suppose that the pair of functions* $\{y(\theta), m(\theta)\}_{\theta\in\Theta}$ *satisfies* (3). *Then for agent's loss function* $L^a(\theta)$ *the following is true:*
(i) $L^a(\theta)$ *is Lipschitz continuous with parameter* Δ_θ*.*
(ii) $L^a(\theta)$ *has left derivative for each* $\theta_0 > 0$ *and has right derivative for each* $\theta_0 < 1$*, given by:*

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

Part (ii) of this result implies that $L^a(\theta)$ is differentiable at $\theta_0 \in (0,1)$ if and only if $y(\theta)$ is continuous at θ_0.[8] In that case, we have

$$\frac{dL^a(\theta_0)}{d\theta} = \frac{\partial l^a(\theta_0, y(\theta_0))}{\partial\theta}. \tag{7}$$

We proceed by deriving a convenient integral representation of the money required to be burned at different states, in order for a given action scheme to be incentive compatible. In

[8] By Rademacher's theorem, L^a is differentiable almost everywhere, since it is Lipschitz-continuous.

particular, suppose that pair of functions $(y(\theta), m(\theta))_{\theta\in\Theta}$ satisfies (3). Denote the range of $y^*(\theta)$ by $R(y^*)$. Note that whenever $y(\theta_1) = y(\theta_2)$, we have $m(\theta_1) = m(\theta_2)$ (otherwise (3) is violated). Hence, we can define function $\tilde{m}(y)$, the amount of money required to burn when action $y \in R(y^*)$ is chosen by:

$$\tilde{m}(y) = m(\theta) \text{ where } \theta \in \Theta \text{ satisfies } y^*(\theta) = y. \tag{8}$$

Let $J(y_1, y_2) = \{\theta \in \Theta | y_1 \leq \sup_{\theta'<\theta} y(\theta') \neq \inf_{\theta'>\theta} y(\theta')\}$. In words, $J(y_1, y_2)$ denotes the set of states at which the action scheme has a jump, in the segment of the scheme that is between y_1 and y_2.

Claim 5 *Let $\tilde{\theta}(\cdot)$ be any single-valued function satisfying $y^*\left(\tilde{\theta}(y)\right) = y$ for any $y \in R(y^*)$; then for any $y_1, y_2 \in [y^*(\theta_1), y^*(\theta_2)]$ such that $y_1 > y_2$:*

$$\tilde{m}(y_2) - \tilde{m}(y_1) = \int\limits_{y\in[y_1,y_2]\cap R(y^*)} \left(-\frac{\partial l^a\left(\tilde{\theta}(y), y\right)}{\partial y}\right) dy + \sum_{\theta\in J(y_1,y_2)} l^a(\inf_{\theta'>\theta} y(\theta')) - l^a(\sup_{\theta'<\theta} y(\theta')). \tag{9}$$

The integral term captures the change in the amount of money burning that is accumulated during intervals on which the action scheme is continuous, while the second term adds up the discrete changes in money burning that are associated with points of discontinuity of the action scheme. It is obvious from the above expression that money burning is increasing as long as the prescribed action stays below the optimal point of the agent (that is, if there is no overshooting). If the action scheme is continuous, then the expression in the lemma simplifies to $\tilde{m}(y_2) - \tilde{m}(y_1) = \int\limits_{y\in D(y_1,y_2)} \left(-\frac{\partial l^a\left(\tilde{\theta}(y),y\right)}{\partial y}\right) dy$. This integral has a convenient graphical representation when the agent has a quadratic utility function. In this case $-\frac{\partial l^a\left(\tilde{\theta}(y),y\right)}{\partial y} = 2\left(\tilde{\theta}(y) + b - y\right)$, 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 the next figure.

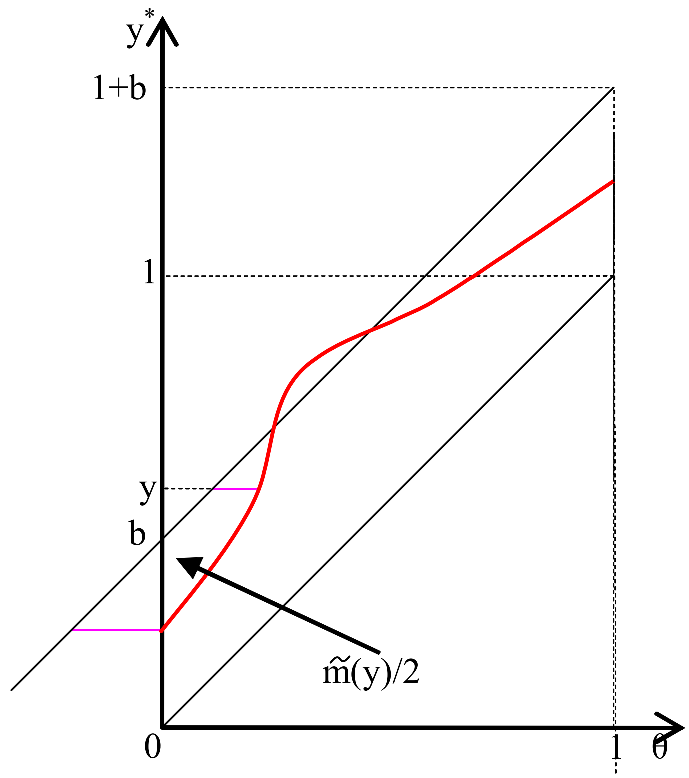

Figure 1: Representation of money burned as an integral

The next result states that the implemented policy is never below the ideal point of the principal (there is no undershooting in the optimal contract). The argument we use is that a deviation from the contract involving increasing the implemented action slightly over an interval in the undershooting region would increase the expected utility of the principal. Indeed, if the contract for types above the interval were modified by lowering the money burning of all these types by just the right amount so that exactly they choose exactly the same actions as before, then such an action makes the principal better off, since (i) the action profile scheme gets weakly closer to her ideal curve (strictly closer on the selected interval in the undershooting region); (ii) in every state the agent is weakly better off than before (strictly on the selected interval and in higher states), reducing the ex ante transfer. The only caveat is that if at some higher state there is overshooting, the prescribed money burning might get close enough to zero such that it cannot be decreased to the level required to maintain incentive compatibility of the modified contract, without violating the nonnegativity constraint on money burning. However, in these states the prescribed action can be lowered (all the way to the agent's ideal point at that state, if necessary) to restore incentives, which again makes both parties better off.

Claim 6 *If* (y^*, m^*) *is a solution to the problem (2) then* $y^*(\theta) \geq \theta$ *for every* $\theta \in \Theta$.

We can now use these results to establish the existence of an optimal contract.

Theorem 7 *There exists a solution to problem* $(2) - (4)$.

To summarize the results above, the following hold in general in our model: there exists an optimal contract; for any optimal contract the implemented action is weakly increasing in the state and never falls below the principal's ideal point; and the agent's ex post utility is Lipschitz-continuous in the state. The amount of money burning is bounded from above and it is not bounded away from zero.

To establish further properties of the optimal contract, in what follows we impose two regularity assumptions.

Assumption 1 $\frac{\frac{\partial l^p(\theta_0, y)}{\partial y}}{-\frac{\partial l^a(\theta_0, y)}{\partial\theta\partial y}}$ *is increasing in* y *for* $y > \theta_0$.

Assumption 2 $\frac{\frac{\partial l^p(\theta, \theta+b(\theta))}{\partial y} f(\theta)}{-\frac{\partial^2 l^a(\theta, \theta+b(\theta))}{\partial\theta\partial y}}$ *is non-decreasing in* θ.

For symmetric loss functions and constant bias (which is assumed in most of the literature), that is when $l^p(\theta, y) = l(y - \theta)$ and $l^a(\theta, y) = l(y - \theta - b)$, Assumption 1 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), Assumption 1 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 expression in Assumption 1 is constant, hence the strict convexity of l^p implies that the condition holds. This, together with the subsequent results, suggests that the important assumption for the qualitative conclusions from the popular uniform-quadratic example to remain valid is that the agent's utility function is quadratic (while the principal can have any strictly convex loss function).

A sufficient condition for Assumption 2 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 θ and $f(\theta)$ is non-decreasing in θ. For symmetric loss functions and constant bias the first condition holds automatically, therefore the simple condition that $f(\theta)$ is non-decreasing in θ is sufficient for Assumption 2 to hold.

For cases when the above regularity conditions do not hold, see the discussion in Section 5.

Next we show that Assumptions 1 and 2 imply that both $y()$ and $m()$ are continuous on the interior of Θ in an optimal contract.

Theorem 8 *Assume A1 holds. If* $(y^*(\theta), m^*(\theta))_{\theta\in\Theta}$ *is an optimal contract, then both* $y^*(\theta)$ *and* $m^*(\theta)$ *are continuous on* $(0,1)$. *Moreover, if* $(y^*(\theta), m^*(\theta))_{\theta\in\Theta}$ *is an optimal contract, then there exists another optimal contract* $\{y^{*\prime}(\theta), m^{*\prime}(\theta)\}_{\theta\in\Theta}$ *such that* $y^{*\prime}(\theta)$ *and* $m^{*\prime}(\theta)$ *are continuous on* $[0,1]$.

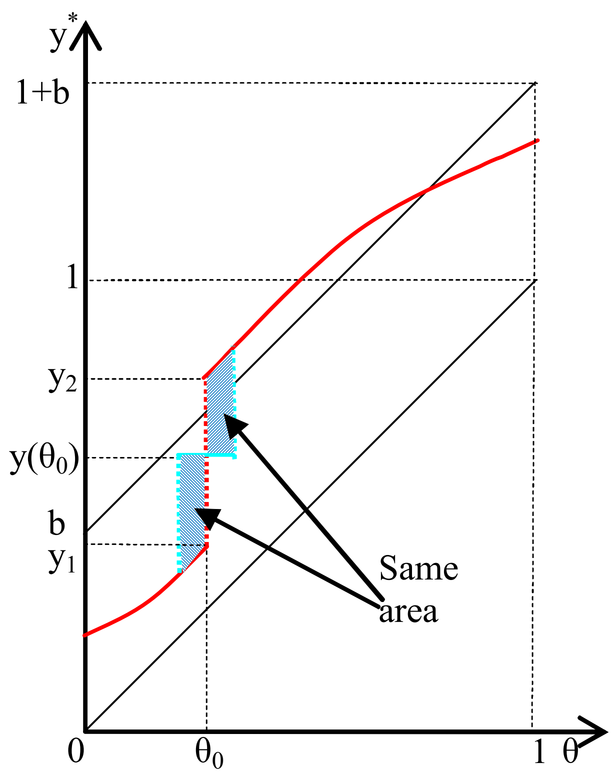

Figure 2: Continuity

For an intuition of the proof, consider Figure 2. The depicted action scheme has a jump at state θ_0. Note that at this state the agent has to be indifferent between the supremum of action-money burning pairs chosen by types below θ_0, and the infimum of action-money burning pairs chosen by types above θ_0. Let these pairs be (y_1, m_1) and (y_2, m_2), respectively. It is further true that for any action between y_1 and y_2 there is an amount of money burning that would make this action indifferent for the agent to the above two options. Now consider complementing the existing contract with an extra option that specifies some in-between action ($y(\theta_0)$ on the picture) with a slightly lower amount of money burning than the one above. Such a contract attracts not only θ_0, but also an interval of types around θ_0. The addition of the new possible choice to the contract weakly improves the agent's utility in all states, decreasing the ex ante transfer that the principal has to pay to satisfy the IR constraint. As for the implemented policy, the principal loses on types on the left of θ_0, but gains on the types on the right of θ_0 choosing the new option. If the principal's loss function is quadratic, these gains and losses are represented by the shaded areas to the left and to the right of θ_0. The sign of the welfare change for the principal depends on the relative magnitudes of $y_2 - y(\theta_0)$ versus $y(\theta_0) - y_1$, as well as on the relative mass of types choosing the new option on the left versus on the right of θ_0. We show that for money burning in the new option that is close enough to making θ_0 indifferent, the principal gains iff $\frac{l^p(\theta_0,y_2)-l^p(\theta_0,y(\theta_0))}{l^p(\theta_0,y(\theta_0))-l^p(\theta_0,y_1)} > \frac{\frac{\partial l^a(\theta_0,y(\theta_0))}{\partial\theta}-\frac{\partial l^a(\theta_0,y_2)}{\partial\theta}}{\frac{\partial l^a(\theta_0,y_1)}{\partial\theta}-\frac{\partial l^a(\theta_0,y(\theta_0))}{\partial\theta}}$. It turns out that this inequality always holds for strictly convex loss functions if y_1 is below the agent's ideal point (if the jump involves no overshooting).[9] Moreover, we show that Assumption 1 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.

Now we can show that if the regularity conditions hold then the optimal contract involves no overshooting, that is, the implemented action is never above the ideal point of the agent. Together with the no undershooting result, this implies that the optimal policy is always between the ideal points of the sender and the receiver. Unlike the no undershooting result, no overshooting does not hold in general though. In Section 5 we provide an example in which the optimal policy involves overshooting, and point out that overshooting and discontinuity of the action scheme are interrelated phenomena.

Theorem 9 *Assume A1 and A2 hold. If contract* $(y^*(\theta), m^*(\theta))$ *solves the problem* (2) *then for any* $\theta \in \Theta$ *we have:*

$$y^*(\theta) \leq \theta + b(\theta).$$

For the intuition behind the result, consider Figure 3. It is straightforward to show that

[9] For more on this, see Section 5.

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. Theorem 8 establishes that optimal action scheme is continuous. Let now θ_0 be the infimum of states with overshooting, as on the picture. Consider now a deviation which keeps the implemented action on the agent's ideal curve for a small interval on the right of θ_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. If the agent has a quadratic loss function then this loss is represented by the shaded area to the right of $(\theta_0, y(\theta_0))$ on the picture. In order to cancel out this increase in money burning, we also specify increasing the prescribed action on a small interval on the left of θ_0, to the agent's optimal curve. If the agent has a quadratic loss function, the resulting gain in from the reduction in money burning is represented by the shaded area to the left of $(\theta_0, y(\theta_0))$ on the picture. Therefore, for quadratic loss function on the agent's side, the two shaded areas are equal in the deviation we propose. Whether the well-being of the principal increases with the deviation depends on the relative magnitude of the loss imposed on the principal by increasing the prescribed action on the left of θ_0 versus the gain resulting from decreasing the prescribed action on the right of θ_0. We show that Assumption 2 (which in the case of symmetric loss functions and fixed biases only requires that $f(\theta)$ is nondecreasing) 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.

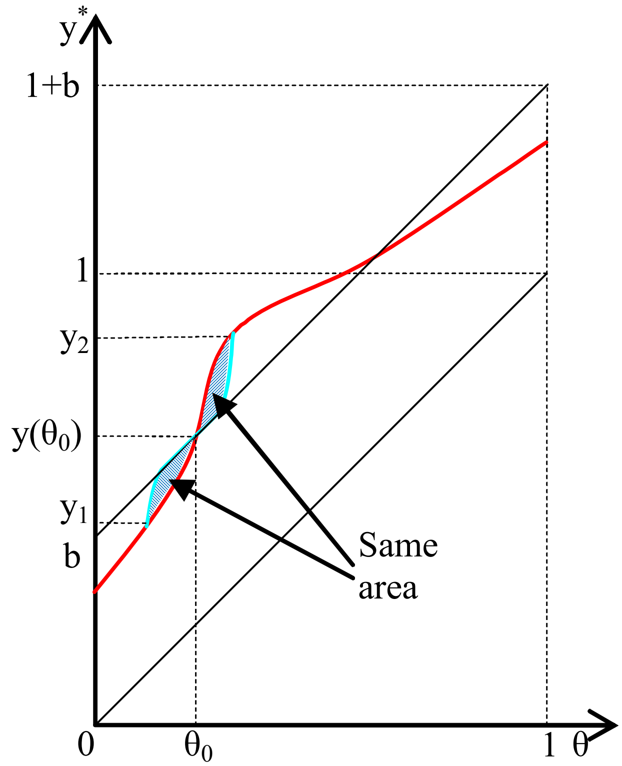

Figure 3: No overshooting

Note that Lemma (5) and Theorem (9) together imply that money burning is monotonically increasing in the optimal contract.

We conclude the section by rewriting the maximization problem, given the results obtained earlier. We have

$$\int_{\Theta} \left(l^{p}\left(\theta, y\left(\theta\right)\right) + l^{a}\left(\theta, y\left(\theta\right)\right) + m\left(\theta\right)\right) f\left(\theta\right) d\theta = \int_{\Theta} \left(l^{p}\left(\theta, y\left(\theta\right)\right) + L^{a}\left(\theta\right)\right) f\left(\theta\right) d\theta,$$

and, taking into account (7),

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

Since in the optimum $m(0) = 0$, then $L^a(0) = l^a(0, y(0))$, and thus the optimization problem is equivalent to the following one:

$$\min_{y(\cdot)} \left(l^a(0, y(0)) + \int_0^1 \left(l^p(\theta, y(\theta)) f(\theta) + \frac{\partial l^a(\theta, y(\theta))}{\partial \theta}(1 - F(\theta)) \right) d\theta \right)$$
$$\text{s.t. } y(\cdot) \text{ is non-decreasing and continuous,}$$
$$y(\theta) \leq \theta + b(\theta).$$

The rewritten form of the optimization problem has the advantage that the incentive constraints are incorporated in the objective function. The objective function indicates that there is 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 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.

4 The optimal contract in uniform-quadratic settings

In this section we explicitly solve for the optimal contract, using the simplified form of the principal's minimization problem contained in the previous section, for a class of models in which both the principal and the agent have quadratic utility functions.

For the remainder of the section, assume that θ is distributed uniformly on $[0, 1]$. Moreover, assume that loss functions are given by:

$$\begin{aligned} l^p(\theta, y) &= A(y - \theta)^2 \\ l^a(\theta, y) &= (y - \theta - b)^2 \end{aligned}$$

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 in monetary terms (recall that utilities are quasilinear in money) the principal's loss from the chosen policy differing by a given amount from the principal's ideal point is smaller than the agent's loss from the chosen policy differing from his ideal point by the same amount. Values $A > 1$ imply the opposite. This latter case is particularly realistic if the principal represents a large organization (or the state) and deviations from the optimal policy can lead to large financial losses (or large social welfare losses), while the agent's preferences over the policy outcomes come from relatively small private benefits/rents. As we show below, the qualitative features of the optimal contract, including whether money burning is used in equilibrium, depend crucially on the value of this parameter.

In the above setting the principal's problem 3 becomes:

$$\min_{y(\cdot)} \left((y(0) - b)^2 + \int_0^1 \left(A (y(\theta) - \theta)^2 - 2 (y - \theta - b)(1 - \theta) \right) d\theta \right)$$
$$\text{s.t. } y(\cdot) \text{ is non-decreasing and continuous,}$$
$$y(\theta) \leq \theta + b.$$

Note that the integrand in the objective function is quadratic in y, and is minimized at:

$$x(\theta) = \frac{A-1}{A}\theta + \frac{1}{A} = \theta + \frac{1-\theta}{A}.$$

The next lemma shows that an optimal contract should coincide with the minimum of $\theta + b$ and $x(\theta)$, except for that it may reach a floor or a ceiling. The intuition behind this is that if at some non-extreme state θ_0 the optimal contract specifies a strictly lower (respectively, higher) action than $\min(\theta + b, x(\theta))$ then there is an interval of states around θ_0 such the prescribed action in this interval can be increased (respectively, decreased) in a way that decreases the objective function in problem (4) while respecting the constraints in the problem.

Claim 10 *Suppose $y(\cdot)$ solves problem (4). Then if θ is such that $y(0) < y(\theta) < y(1)$, then $y(\theta) = \min\{x(\theta), \theta + b\}$.*

We can now give an explicit characterization of optimal contracts. Before proceeding, note that $x(1) = 1$ and $x(0) = \frac{1}{A}$. In addition, the solution to equation $x(\theta) = \theta + b$ is given by

$$\theta = 1 - Ab.$$

We first characterize the solution under the assumption $b < 1$.

We start with the case when $A > 1$, hence $x(\theta) = \theta + \frac{1-\theta}{A}$. Then $x(\theta)$ is increasing. If $\frac{1}{A} \geq b$, so $x(0) \geq b$, the optimal contract is given by

$$y^*(\theta) = \begin{cases} \theta + b & \text{if } \theta \leq 1 - Ab; \\ \theta + \frac{1-\theta}{A} & \text{if } \theta > 1 - Ab. \end{cases}$$

Indeed, for this contract, both terms in the objective function of problem (4) are minimized, subject to the constraints.

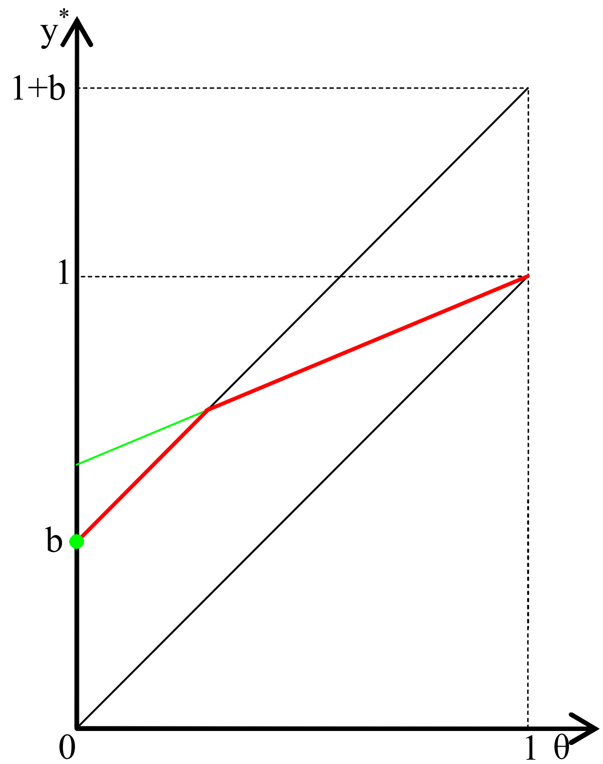

Figure 4: Optimal contract if $1 < A < 1/b$

Now suppose again that $A > 1$, but $\frac{1}{A} < b$, so $x(0) < b$. It is easy to see that $y(0)$ must lie between 0 and b, so

$$y^*(\theta) = \begin{cases} \theta^* + \frac{1-\theta^*}{A} & \text{if } \theta \le \theta^*; \\ \theta + \frac{1-\theta}{A} & \text{if } \theta > \theta^*, \end{cases}$$

where $\theta^* \in \left[0, \frac{Ab-1}{A-1}\right]$ (at the latter point, $x(\theta) = b$). Writing out $V(y(\cdot))$:

$$\begin{aligned} V(y(\cdot)) &= \int_0^{\theta^*} \left(A\left(\theta^* + \frac{1-\theta^*}{A}\right)^2 - 2\left(\theta^* + \frac{1-\theta^*}{A}\right)(1-\theta+A\theta) + A\theta^2 + (b+\theta)^2 \right) d\theta \\ &\quad + \int_{\theta^*}^1 \left(A\left(\theta + \frac{1-\theta}{A}\right)^2 - 2\left(\theta + \frac{1-\theta}{A}\right)(1-\theta+A\theta) + A\theta^2 + (b+\theta)^2 \right) d\theta \\ &\quad -2b\left(\theta^* + \frac{1-\theta^*}{A}\right) + \left(\theta^* + \frac{1-\theta^*}{A}\right)^2. \end{aligned}$$

The above expression needs to be minimized with respect to θ^*. After differentiating and algebraic manipulations, we find that θ^* is given by

$$\theta^* = \frac{1}{A}\left(\sqrt{1+2\frac{A}{A-1}\left(Ab-1\right)}-1\right).$$

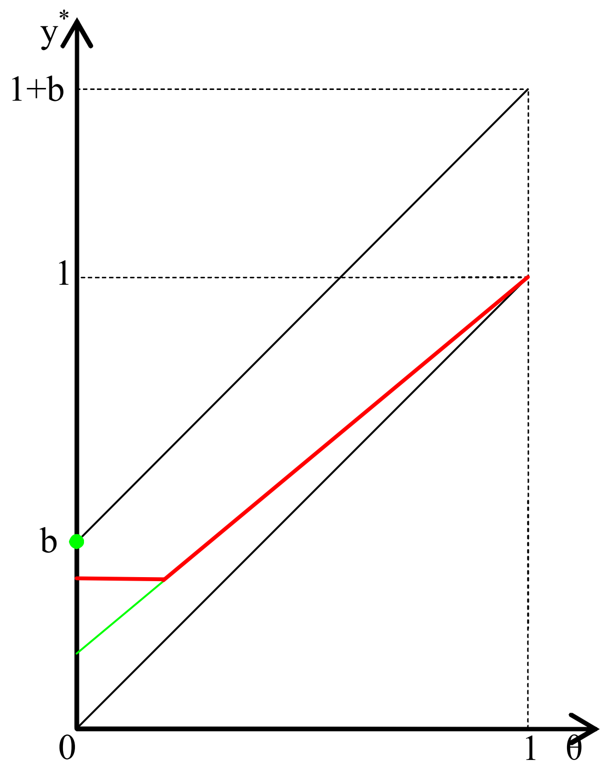

Figure 5: Optimal contract if $1/b < A$

Consider next the case $A = 1$. In this case $x\left(\theta\right)$ is a constant, and therefore it is easy to show that the optimal contract is given by:

$$y^*\left(\theta\right) = \begin{cases} \theta + b & \text{if } \theta \leq 1-b; \\ 1 & \text{if } \theta > 1-b. \end{cases}$$

Finally, consider the case $A < 1$. Then $x\left(\theta\right)$ is decreasing. Moreover, $x(0) > b$ implies that $y\left(0\right) = b$. Therefore the optimal contract is of the form:

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

Substituting this into the objective function yields:

$$
\begin{aligned}
V\left(y\left(\cdot\right)\right) &= \int_0^{\theta^*}\left(A\left(\theta+b\right)^2-2\left(\theta+b\right)\left(1-\theta+A\theta\right)+A\theta^2+\left(b+\theta\right)^2\right)d\theta \\
&\quad +\int_{\theta^*}^1\left(A\left(\theta^*+b\right)^2-2\left(\theta^*+b\right)\left(1-\theta+A\theta\right)+A\theta^2+\left(b+\theta\right)^2\right)d\theta \\
&\quad -2b^2+b^2.
\end{aligned}
$$

This expression has to be minimized in θ^*. After differentiating and manipulations, we get

$$\theta^* = 1-\frac{2Ab}{A+1}.$$

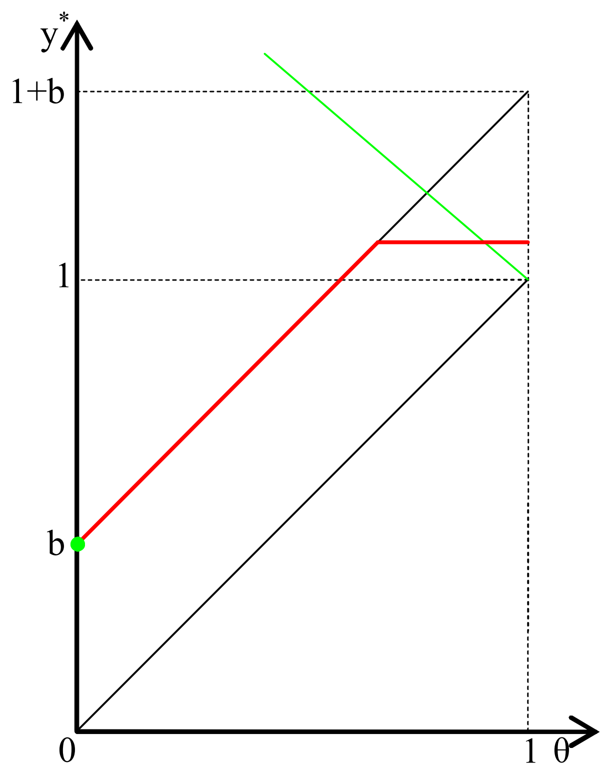

Figure 6: Optimal contract if $A<1$

For completeness, we also consider the optimal contract when $b \geq 1$. We can prove the following. As before, if $A>1$, then optimal contract is

$$y^*\left(\theta\right)=\begin{cases}\theta^*+\frac{1-\theta^*}{A} & \text{if } \theta\leq\theta^*;\\ \theta+\frac{1-\theta}{A} & \text{if } \theta>\theta^*,\end{cases}$$

where

$$\theta^* = \frac{1}{A}\left(\sqrt{1 + 2\frac{A}{A-1}(Ab-1)} - 1\right).$$

If $A \leq 1$, we get

$$y(\theta) = \min\left(\frac{1}{2} + \frac{b}{A+1}, \theta + b\right).$$

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 IR constraint, and symmetric quadratic loss functions (corresponding to the $A = 1$ case above). 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. One qualitative difference is that in KM a large bias implies that the implemented action is always strictly below the agent's ideal line (the decision is never "fully delegated" to the agent), while in our model there is a region (for $A \leq 1$) at which the agent's ideal point is implemented no matter how large the bias is. This difference results from the fact that as opposed to our model, there is no IR constraint in KM.

5 Properties of the optimal contract when the regularity conditions do not hold

Under the regularity conditions A1 and A2 the optimal contract has the intuitive feature that the implemented action is always between the ideal points of the principal and the agent. Moreover, both the action scheme and the amount of money burning are continuous and weakly increasing functions of the state. Below we show that if the regularity conditions do not hold, the optimal contract might not have any of the above features (besides the implemented action being weakly increasing in the state, which is a general property by Lemma 1. Moreover, we show that the violations of these properties are interrelated.

The next theorem establishes that if the optimal contract involves no overshooting (that is, if the implemented policy is always between the players' ideal points) then both the implemented action and money burning are continuous and increasing in the state.

Theorem 11 *Assume that* $(y^*(\cdot), m^*(\cdot))$ *is an optimal contract, and* $\theta \leq y^*(\theta) \leq \theta + b(\theta)$ *for every* $\theta \in (0,1)$. *Then both* $y^*(\cdot)$ *and* $m^*(\cdot)$ *are continuous and weakly increasing on* $(0,1)$.

We prove the above result by showing that the type of deviation considered in the proof of Theorem 8, that is making the jump more gradual by offering an in-between option to types around the jump point, increases the expected utility of the principal for arbitrary convex loss functions, as long as the jump is in between the ideal points of the players.

Next, we construct an example in which the optimal contract indeed involves overshooting and discontinuities, as well as non-monotonicity of money burning. We also provide an intuitive explanation why overshooting is optimal for the principal.

To start with, consider a specification of the model in which the loss functions are of the form: $l^p(\theta, y(\theta)) = 2(y(\theta) - \theta)^2$, and $l^a(\theta, y(\theta)) = (y(\theta) - \theta - 0.05)^2$. This is a special case of the class of loss functions considered in Section 4, with $A = 2$ and $b = 0.05$. Moreover, temporarily assume that $f(\theta) = \frac{3}{2}$ for $\theta \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ and $f(\theta) = 0$ for $\theta \in (\frac{1}{3}, \frac{2}{3})$ (below we change the example so that the density is strictly positive everywhere). It is easy to see that in this example the principal can solve its optimization separately for the regions $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$. Using the results obtained in Section 4, the optimal contract specifies:

$$y^*(\theta) = \begin{cases} \theta + 0.05 & \text{if } \theta \leq \frac{7}{30}; \\ \frac{1}{6} + \frac{\theta}{2} & \text{if } \frac{7}{30} < \theta < \frac{1}{3} \end{cases}$$

and

$$y^*(\theta) = \begin{cases} \theta + 0.05 & \text{if } \frac{2}{3} \leq \theta \leq \frac{9}{10}; \\ \theta + \frac{1-\theta}{A} & \text{if } \frac{9}{10} < \theta. \end{cases}$$

Using (9), the amount of money burning implied by $y^*()$ at state $\theta = \frac{1}{3}$ is $\frac{1}{200}$ (twice the area between $y^*(\theta)$ and the agent's ideal line $\theta + 0.05$). Note that at this state the agent prefers action $y^*\left(\frac{1}{3}\right) = \frac{1}{3}$ and money burning $\frac{1}{200}$ to action $y^*\left(\frac{2}{3}\right) = \frac{43}{60}$ and money burning 0, and therefore the above $y^*(\theta)$ is incentive-compatible on $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ as long as $m^*\left(\frac{2}{3}\right) = 0$. On interval $(\frac{1}{3}, \frac{2}{3})$ the prior density is 0, therefore the action scheme (and money burning) specified on this region does not influence the principal's expected payoff directly. Therefore in optimum the principal should set the action scheme in this region in order to cut money burning back to zero (to achieve $m^*\left(\frac{2}{3}\right) = 0$.). This requires the contract to overshoot for at least part this region. For example the following specification achieves this:

$$y^*(\theta) = \begin{cases} \frac{1}{20}\sqrt{2} + \frac{23}{60} & \text{if } \frac{1}{3} < \theta \leq 0.404\,04; \\ \theta + 0.05 & \text{if } 0.404\,04 < \theta < \frac{2}{3}. \end{cases}$$

It is easy to verify that the utility that the above contract yields to the principal is bounded away from any contract that does not specify overshooting at any point of Θ. Modify now the above example such that $f(\theta) = \frac{3}{2} - 2\varepsilon$ for $\theta \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ and $f(\theta) = \varepsilon$ for $\theta \in (\frac{1}{3}, \frac{2}{3})$. For small enough $\varepsilon > 0$ the above contract (which is still incentive-compatible, since the latter does not depend on the prior distribution) continues to yield a strictly higher payoff to the principal than any contract that does not specify overshooting. This establishes that the optimal contract requires overshooting (and hence a non-monotonic money burning scheme) in the modified example, too. We also establish that the optimal contract requires a discontinuity at some state in $(\frac{1}{3}, \frac{2}{3})$.

Claim 12 *Assume that* $l^p(\theta, y(\theta)) = 2(y(\theta) - \theta)^2$, $l^a(\theta, y(\theta)) = (y(\theta) - \theta - 0.05)^2$, $f(\theta) = \frac{3}{2} - 2\varepsilon$ *for* $\theta \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ *and* $f(\theta) = \varepsilon$ *for* $\theta \in (\frac{1}{3}, \frac{2}{3})$. *Then for small enough* $\varepsilon > 0$ *the optimal contract prescribes a discontinuous action scheme* $y()$.

The intuition behind the optimal contract involving overshooting and discontinuities is the following: if the implemented action is kept between the optimal points of the principal and the agent, the amount of prescribed money burning is increasing, and if the implemented action is kept strictly below the agent's ideal point the money burning is strictly increasing. The only way the principal can decrease money burning at some state in an incentive compatible way is if she prescribes an overshooting action. In the above example, this becomes optimal in the region $(\frac{1}{3}, \frac{2}{3})$, where the density of the prior is low. The optimal policy involves increasing money burning in low states, then a discontinuous jump and overshooting, and finally increasing money burning again in high states. Intuitively, the principal sacrifices utility in the unlikely states, in order to better align incentives in the more likely states and at the same time do not accumulate too high levels of money burning.

6 Delegation with both Conditional Transfers and Money Burning

In this section we investigate the case when besides conditional money burning, the principal can also specify conditional monetary transfers in the contract. Note that in this context we can ignore the ex ante transfer, and focus on ex post transfers, since any transfer scheme with ex ante transfer $T > 0$ is equivalent to a transfer scheme with ex ante transfer $T = 0$ and ex post transfers that are T higher in every state. The key restriction we impose is that there is a minimum amount of transfer that the principal has to pay to the agent in any state. The lower bound on transfers can correspond to a minimum wage requirement, or if it is negative then to the maximum amount of punishment/fee that the principal can impose on the agent. The latter can result either from legal restrictions or because the agent is liquidity constrained. Without loss of generality we normalize this lower bound to be 0, implying that only monetary transfers from the principal to the agent are possible. Below we show that given this requirement, despite money burning is a less efficient way to create incentives for the agent than monetary transfers, the optimal contract can specify either only monetary transfers, or only money burning, or monetary transfers in low states and money burning in high states. The primary factor determining which case applies is the outside option of the agent, relative to the minimum transfer amount.

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 principal to agent $t(\theta) \geq 0$. As before, the agent's utility is quasilinear both in the transfer and money burning, hence he cares about $m(\theta)$ and $t(\theta)$ only through the difference $m(\theta) - t(\theta)$. The agent's loss function is now given by

$$L^a(\theta) = l^a(\theta, y(\theta)) + m(\theta) - t(\theta). \tag{10}$$

The principal's problem may now be written as follows (we immediately write it as a minimization problem):

$$\min_{\{y(\theta), m(\theta), t(\theta)\}_{\theta \in \Theta}} \int_{\Theta} \left(l^p(\theta, y(\theta)) + t(\theta)\right) d\theta \tag{11}$$

$$\text{s.t.} \int_{\Theta} \left(l^a(\theta, y(\theta)) + m(\theta) - t(\theta)\right) \leq L, \tag{12}$$

$$\forall \theta, \theta' \in \Theta : l^a(\theta, y(\theta)) + m(\theta) - t(\theta) \leq l^a\left(\theta, y\left(\theta'\right)\right) + m\left(\theta'\right) - t\left(\theta'\right), \tag{13}$$

$$\forall \theta \in \Theta : m(\theta) \geq 0, t(\theta) \geq 0. \tag{14}$$

We start by establishing properties of the optimal contract that are analogous to properties obtained in the case without transfers.

Claim 13 *There exists a solution to problem* (11) *subject to constraints* (12), (13), (14)*. Moreover, if A1 and A2 hold the for any optimal contract* $(y^*(\cdot), m^*(\cdot), t^*(\cdot))$ *the following hold:*
(i) $y^*(\cdot)$ *is weakly increasing on* $[0,1]$ *and continuous on* $(0,1)$*;*
(ii) $\theta \leq y^*(\theta) \leq \theta + b(\theta)$ *for all* $\theta \in [0,1]$*;*
(iii) for any $\theta \in (0,1)$,

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

and

$$L^a(\theta_2) - L^a(\theta_1) = \int_{\theta_1}^{\theta_2} \frac{\partial l^a(\theta, y(\theta))}{\partial \theta} d\theta;$$

(iv) for any θ_1,$\theta_2 \in [0,1]$,

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

The proofs of these results follow closely similar proofs for the case without conditional transfers, and are omitted. The next set of results state that there is essentially no state at which there is both money burning and nonzero conditional transfer. An optimal contract specifies conditional transfers in low states (if there is a region with nonzero transfers). These transfers are decreasing in the state and at some point reach 0. At the right of this point an optimal contract might specify money burning, such that money burning is increasing in the state in this region.

Claim 14 *Suppose A1 and A2 hold. Then any two solutions to* (11) *subject to constraints* $(12), (13), (14)$ *specify the same* $(y(\theta), m(\theta), t(\theta))$ *at almost every* $\theta \in \Theta$*. Moreover, there exists a solution which satisfies the following:*
(i) Either $m^*(\theta) = 0$ *or* $t^*(\theta) = 0$*, for every* $\theta \in \Theta$*. Moreover,* $m^*(\theta)$ *is non-decreasing in* θ *and* $t^*(\theta)$ *is non-increasing in* θ*; in particular,* $m^*(0) = 0$*.*
(ii) Either there exists $\theta_0 \in \Theta$ *such that* $m^*(\theta_0) = 0$ *and* $t^*(\theta_0) = 0$*, or* $t^*(\theta) > 0$ *for all* θ*.*

From now on we turn attention to deriving the optimal contract in the uniform-quadratic setting of Section 4. Recall that loss functions in this setting are given by:

$$l^p(\theta, y) = A(y - \theta)^2$$
$$l^a(\theta, y) = (y - \theta - b)^2$$

In light of result (ii) of Claim 14, there are two possibilities to consider. The next result shows that if $t^*(\theta) > 0$ for all θ, then the contract must achieve the weighted average between the principal's and the agent's ideal points that maximizes the joint surplus of the principal and the agent at the given point. That is, if the constraint on the minimal amount of transfer never binds then the optimal contract achieves an efficient outcome in every state. In essence, the principal and the agent form a partnership, and in every state the action maximizing the joint welfare of the partnership is chosen.

Claim 15 *If* $(y^*(\cdot), m^*(\cdot), t^*(\cdot))$ *is the optimal contract such that* $t^*(\theta) > 0$ *for all* $\theta \in [0,1]$, *then*

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

If transfers are not positive in every state, the principal's problem may be rewritten as

$$\min_{y(\theta)} \left\{ -(y(\theta_0) - \theta_0 - b)^2 \theta_0 + \int_0^{\theta_0} \left((y(\theta) - \theta - b)^2 - 2(y(\theta) - \theta - b)\theta \right) d\theta + \int_0^1 A(y(\theta) - \theta)^2 d\theta \right. \tag{15}$$

$$\left. + \lambda \left((y(\theta_0) - \theta_0 - b)^2 + \int_0^{\theta_0} 2(y(\theta) - \theta - b)\theta d\theta - \int_{\theta_0}^1 2(y(\theta) - \theta - b)(1-\theta) d\theta - L \right) \right\}$$

$$\text{s.t. } y(\cdot) \text{ is non-decreasing and continuous,}$$

$$y(\theta) \le \theta + b(\theta),$$

where θ_0 satisfies $m^*(\theta_0) = t^*(\theta_0) = 0$ and $\lambda \in [0,1]$ is the Lagrange multiplier. Moreover, $\lambda = 0$ only if $t^*(\theta) = 0$ for all θ, and $\lambda = 1$ only if $m^*(\theta) = 0$ for all θ.

The solution to this problem is difficult to describe explicitly. In order to describe the solution to the problem indirectly, we introduce two auxiliary functions

$$\begin{aligned} x(\theta) &= \frac{1}{A+1}(b + 2\theta + A\theta - \theta\lambda); \\ z(\theta) &= \frac{1}{A}(\lambda + A\theta - \theta\lambda). \end{aligned}$$

Note that functions $x(\theta)$ and $z(\theta)$ may been obtained by minimizing the partial Lagrangian with respect to $y(\theta)$ for $\theta < \theta_0$ and $\theta > \theta_0$, respectively.

We now formulate the solution, as a function of λ, for different cases.

Claim 16 *Suppose* $A \le \lambda$. *Then* $z(\theta)$ *is nonincreasing. In the optimal contract, there is no money-burning. If* $Ab(A + 2 - \lambda) > (1 - \lambda)(A + \lambda)$, *then the solution to* (15) *is given by*

$$y^*(\theta) = \begin{cases} x(\theta) & \text{if } \theta < \theta^* \\ x(\theta^*) & \text{if } \theta \ge \theta^* \end{cases},$$

where

$$\theta^* = \frac{A+\lambda}{A+1} - \sqrt{\frac{(1-\lambda)\left(2Ab - (1-\lambda)\frac{A+\lambda}{A+1}\right)}{(A+2-\lambda)(A+1)}}.$$

Otherwise, the solution to (15) *is given by*

$$y^*(\theta) = \begin{cases} x(\theta) & \text{if } \theta < \frac{Ab}{1-\lambda} \\ \theta + b & \text{if } \frac{Ab}{1-\lambda} \le \theta < \theta^{**} \\ \theta^{**} + b & \text{if } \theta \ge \theta^{**} \end{cases},$$

where

$$\theta^{**} = 1 - \frac{2Ab}{A+\lambda}.$$

Claim 17 *Suppose $A > \lambda$. Then $z(\theta)$ is increasing. In the optimal contract, there are transfers, except for the case $\lambda = 0$. If $Ab \le \lambda(1-\lambda)$, then the solution to (15) is*

$$y^*(\theta) = \begin{cases} x(\theta) & \text{if } \theta < \frac{Ab}{1-\lambda} \\ \theta + b & \text{if } \frac{Ab}{1-\lambda} \le \theta < 1 - \frac{Ab}{\lambda} \\ z(\theta) & \text{if } \theta \ge 1 - \frac{Ab}{\lambda} \end{cases};$$

if $\lambda(1-\lambda) < Ab < A(1-\lambda)(A+1-\lambda)$, then

$$y^*(\theta) = \begin{cases} x(\theta) & \text{if } \theta < \lambda \\ \frac{\lambda(A+2-\lambda)+b}{A+1} & \text{if } \lambda \le \theta \le \frac{\lambda\left(A - A\lambda + A^2 - 1\right) + Ab}{(A+1)(A-\lambda)} \\ z(\theta) & \text{if } \theta > \frac{\lambda\left(A - A\lambda + A^2 - 1\right) + Ab}{(A+1)(A-\lambda)} \end{cases};$$

if $Ab \ge A(1-\lambda)(A+1-\lambda)$, then

$$y^*(\theta) = \begin{cases} x(\theta) & \text{if } 0 \le \theta \le \lambda \\ \frac{\lambda(A+2-\lambda)+b}{A+1} & \text{if } \theta > \lambda \end{cases}.$$

In the first two cases, there are both transfers and money-burning in equilibrium; in the last case, there are transfers only.

Let us now recall that λ is not a parameter of the model, but a variable that is determined from the agent's IR constraint. However, it is straightforward to show that λ is a continuous and decreasing function of L (recall that the agent's outside option is $-L$), and that there are $\underline{L}$ and $\overline{L}$ such that $L \le \underline{L}$ implies that $\lambda = 0$, and $L \ge \overline{L}$ implies that $\lambda = 1$. The case $L \ge \overline{L}$ corresponds to such low levels of the outside option $-L$ that the principal can achieve her optimal policy at every state only through money burning, without paying transfers to the agent. The case $L \le \underline{L}$ corresponds to such high levels of the outside option that even after achieving the jointly efficient policy scheme purely through transfers as incentives, the agent had to be paid more money. More generally, for low enough L the optimal contract only specifies transfers, and for high enough L the optimal contract only specifies money burning. However, for any fixed A, there is a range of L for which the optimal contract specifies both transfers in low states and money-burning at high states. Typically for these values of L there is also an interval of states at which both money burning and transfers are specified to be zero.

For a fixed L, the dependence of λ on A is more complex. However, in the region where there is no money-burning, a higher A would make the agent better off for any fixed λ, and therefore increasing A must decrease λ in this region. We can then conclude the following: If A is sufficiently low, then the optimal contract specifies no money-burning, and $y^*(\theta)$ is constant in some neighborhood of 1. As A increases, the amount of transfers increases. As A increases further, there is both transfers for low types and money-burning for high types. For extremely high values of A, the contract line $y^* (\theta)$ will uniformly tend to the line $y = \theta$.

The economic lesson from this is that the optimal contract only achieves efficiency if the agent has a high enough outside option (relative to the minimal wage). Lower levels of outside option result in two sources of inefficiency: (i) the implemented action scheme gets distorted from the jointly efficient scheme; (ii) the principal uses socially inefficient money burning (at least in some states) to distort the action choices of the agent. If the agent's outside option is very low (the minimum wage that the principal has to pay to the agent is very high relative to the utility the agent could obtain outside the relationship) then this inefficiency is particularly severe, and wasteful money burning is prescribed at almost every state.

7 Appendix

Proof of Claim 1: If $\theta_2 = \theta_1$, the statement is trivial, so assume $\theta_2 > \theta_1$. Denote for brevity $y_i = y(\theta_i)$, $m_i = m(\theta_i)$ for $i = 1, 2$. By (3), we have

$$\begin{aligned} l^a(\theta_1, y_1) + m_1 &\leq l^a(\theta_1, y_2) + m_2; \\ l^a(\theta_2, y_2) + m_2 &\leq l^a(\theta_2, y_1) + m_1. \end{aligned}$$

These may be rewritten as

$$\begin{aligned} m_2 - m_1 &\geq l^a(\theta_1, y_1) - l^a(\theta_1, y_2); \\ m_2 - m_1 &\leq l^a(\theta_2, y_1) - l^a(\theta_2, y_2). \end{aligned} \tag{16}$$

Consequently,

$$l^a(\theta_2, y_1) - l^a(\theta_2, y_2) \geq l^a(\theta_1, y_1) - l^a(\theta_1, y_2), \tag{17}$$

Suppose that $\theta_2 > \theta_1$, but $y_2 < y_1$. Then

$$\begin{aligned} l^a(\theta_2, y_1) - l^a(\theta_2, y_2) &= \int_{y_2}^{y_1} \frac{\partial l^a(\theta_2, y)}{\partial y} dy \\ &< \int_{y_2}^{y_1} \frac{\partial l^a(\theta_1, y)}{\partial y} dy \\ &= l^a(\theta_1, y_1) - l^a(\theta_1, y_2), \end{aligned}$$

where the inequality follows from the assumption that $y_2 < y_1$ and the single-crossing condition: $\frac{\partial l^a(\theta_2,y)}{\partial y} - \frac{\partial l^a(\theta_1,y)}{\partial y} = \int_{\theta_1}^{\theta_2} \frac{\partial l^a(\theta,y)}{\partial \theta \partial y} d\theta < 0$. But this contradicts (17). ■

Proof of Claim 2: If $\theta_2 = \theta_1$, the statement is trivial. Without loss of generality assume $\theta_2 > \theta_1$, in which case (16) holds. We immediately get

$$\begin{aligned} m_1 - m_2 &\leq |y_2 - y_1| \max_{y_1 \leq y \leq y_2} \left| \frac{\partial l^a(\theta_1, y)}{\partial y} \right| \leq |y_2 - y_1| \Delta_y, \\ m_2 - m_1 &\leq |y_2 - y_1| \max_{y_1 \leq y \leq y_2} \left| \frac{\partial l^a(\theta_2, y)}{\partial y} \right| \leq |y_2 - y_1| \Delta_y. \end{aligned}$$

This implies

$$|m_2 - m_1| \leq |y_2 - y_1| \Delta_y,$$

which establishes (5). Now bounded variation of $m(\theta)$ follows immediately: take any subdivision of $[0, 1]$, say,

$$0 = \theta_0 < \theta_1 < \ldots < \theta_{n-1} < \theta_n = 1.$$

For any $k : 0 \leq k \leq n-1$, we have $y(\theta_k) \leq y(\theta_{k+1})$, as follows from part (i). Therefore,

$$\sum\nolimits_{k=0}^{n-1} |m(\theta_{k+1}) - m(\theta_k)| \leq \sum\nolimits_{k=0}^{n-1} (y(\theta_{k+1}) - y(\theta_k)) \Delta = (y(1) - y(0)) \Delta_y,$$

which proves part (ii). ■

Proof of Claim 3. Suppose (y^*, m^*) minimizes V^p subject to (3)–(4), and $\inf_{\theta\in\Theta} m^*(\theta) = \varepsilon > 0$. Consider m' such that $m'(\theta) = m^*(\theta) - \varepsilon \ \forall\ \theta \in \Theta$. By construction, (y^*, m') satisfies condition (4). Moreover, since (y^*,m^*) satisfies condition (4), and hence for every $\theta \in \Theta$ we have $l^a(\theta, y^*(\theta)) + m^*(\theta) \le l^a(\theta, y^*(\theta')) + m^*(\theta')$, we also have $l^a(\theta, y^*(\theta)) + m'(\theta) \le l^a(\theta, y^*(\theta')) + m'(\theta')$. This implies that (y^*, m') satisfies condition (4), too. Note now that $V^p(y^*(\cdot), m'(\cdot)) = V^p(y^*(\cdot), m^*(\cdot)) - \varepsilon$, which contradicts that (y^*,m^*) minimizes V^p subject to (3)–(4). ■

Proof of Claim 4. By (3), we have

$$\begin{aligned}
L^a(\theta_1) &\le l^a(\theta_1, y_2) + m_2 \\
&\le L^a(\theta_2) + (l^a(\theta_1, y_2) - l^a(\theta_2, y_2)) \\
&\le L^a(\theta_2) + |\theta_2 - \theta_1| \Delta_\theta,
\end{aligned}$$

so

$$L^a(\theta_1) - L^a(\theta_2) \le |\theta_2 - \theta_1| \Delta_\theta.$$

Similarly,

$$L^a(\theta_2) - L^a(\theta_1) \le |\theta_2 - \theta_1| \Delta_\theta,$$

which imply the statements in the lemma.

(ii) Note that we have for sufficiently small ε

$$\begin{aligned}
L^a(\theta_0) &\le l^a(\theta_0, y(\theta_0+\varepsilon)) + m(\theta_0+\varepsilon) \\
&\le L^a(\theta_0+\varepsilon) + l^a(\theta_0, y(\theta_0+\varepsilon)) - l^a(\theta_0+\varepsilon, y(\theta_0+\varepsilon)),
\end{aligned}$$

and similarly

$$\begin{aligned}
L^a(\theta_0+\varepsilon) &\le l^a\left(\theta_0+\varepsilon, y\left(\theta_0+\varepsilon^2\right)\right) + m\left(\theta_0+\varepsilon^2\right) \\
&\le L^a\left(\theta_0+\varepsilon^2\right) + l^a\left(\theta_0+\varepsilon, y\left(\theta_0+\varepsilon^2\right)\right) - l^a\left(\theta_0+\varepsilon^2, y\left(\theta_0+\varepsilon^2\right)\right) \\
&\le L^a(\theta_0) + \varepsilon^2\Delta_\theta + l^a\left(\theta_0+\varepsilon, y\left(\theta_0+\varepsilon^2\right)\right) - l^a\left(\theta_0+\varepsilon^2, y\left(\theta_0+\varepsilon^2\right)\right).
\end{aligned}$$

Hence,

$$\frac{l^a(\theta_0+\varepsilon, y(\theta_0+\varepsilon)) - l^a(\theta_0, y(\theta_0+\varepsilon))}{\varepsilon} \le \frac{L^a(\theta_0+\varepsilon) - L^a(\theta_0)}{\varepsilon}$$

$$\le \frac{\varepsilon^2\Delta_\theta + l^a\left(\theta_0+\varepsilon, y\left(\theta_0+\varepsilon^2\right)\right) - l^a\left(\theta_0+\varepsilon^2, y\left(\theta_0+\varepsilon^2\right)\right)}{\varepsilon}$$

It is now trivial to check that both the left-hand side and the right-hand side have the same limit $\frac{\partial l^a(\theta_0, \lim_{\theta\to\theta_0+} y(\theta))}{\partial\theta}$ as $\varepsilon \to 0+$, hence

$$\frac{d^r L^a(\theta_0)}{d\theta} = \lim_{\varepsilon\to 0+} \frac{L^a(\theta_0+\varepsilon) - L^a(\theta_0)}{\varepsilon} = \frac{\partial l^a(\theta_0, \lim_{\theta\to\theta_0+} y(\theta))}{\partial\theta}.$$

We can prove the formula for the left derivative similarly. Now, since $\frac{\partial l^a(\theta_0,y)}{\partial\theta}$ is strictly monotonic in y, we have $\frac{\partial l^a(\theta_0, \lim_{\theta\to\theta_0-} y(\theta))}{\partial\theta} = \frac{\partial l^a(\theta_0, \lim_{\theta\to\theta_0+} y(\theta))}{\partial\theta}$ if and only if $\lim_{\theta\to\theta_0-} y(\theta) = \lim_{\theta\to\theta_0+} y(\theta)$. Since $y(\theta)$ is monotonic, this is equivalent to continuity of $y(\theta)$ at θ_0, so the result on continuity follows. This completes the proof. ■

Lemma 1 *Suppose that $(y^*(\cdot), m^*(\cdot))$ satisfies (3). Assume that $y \in R(y^*)$ is such that $y^*(\cdot)$ is continuous at every $\theta \in \Theta$ for which $y^*(\theta) = y$. Then function $\tilde{m}(\cdot)$ given by (8) satisfies the following conditions:*

(i) $\tilde{m}(\cdot)$ is continuous at y; its left derivative exists at y if $y > y^*(0)$, and its right derivative exists at y if $y < y^*(1)$ and are equal to:

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

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

(ii) $\tilde{m}(y)$ is differentiable at $y \in (y(\theta_1), y(\theta_2))$ if and only if $(y^*)^{-1}(y)$ is a singleton.

Proof of Lemma 1. (i) We prove the result for the left derivative; then the result for the right derivative may be proved similarly, and continuity will follow. Take any $y \in (y(0), y(1)]$ and sufficiently small $\varepsilon > 0$. Applying (3) to types $\tilde{\theta}(y)$ and $\tilde{\theta}(y-\varepsilon)$ which choose y and $y-\varepsilon$, respectively, we can write (since $\tilde{m}(y) = m\left(\tilde{\theta}(y)\right)$ for each y)

$$\begin{gathered} l^a\left(\tilde{\theta}(y-\varepsilon), y-\varepsilon\right) + \tilde{m}(y-\varepsilon) \le l^a\left(\tilde{\theta}(y-\varepsilon), y\right) + \tilde{m}(y); \\ l^a\left(\tilde{\theta}(y), y\right) + \tilde{m}(y) \le l^a\left(\tilde{\theta}(y), y-\varepsilon\right) + \tilde{m}(y-\varepsilon). \end{gathered}$$

These inequalities imply

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

Since this holds for any function $\tilde{\theta}(\cdot)$ that satisfies $y^*\left(\tilde{\theta}(y)\right) = y$, we can take that $\tilde{\theta}_{min}(y) = \min_{y^*(\theta)=y}\theta \in (\theta_1, \theta_2]$; this limit exists due to continuity of y^* at every $\theta \in \Theta$ for which

$y^*(\theta) = y$. Dividing all parts by ε, we notice that the leftmost and the rightmost parts tend to $-\frac{\partial l^a(\tilde{\theta}_{min}(y),y)}{\partial y}$, this shows that $\frac{d^l \tilde{m}(y)}{dy}$ exists and is given by the formula

(ii) Trivially follows from (i). ■

Proof of Claim 5. Let $D(y_1, y_2) = \{y \in [y_1, y_2] \cap R(y^*)| - \exists\ \theta \in J(y_1, y_2)$ such that $y = \sup_{\theta'<\theta} y(\theta')$ or $y = \inf_{\theta'>\theta} y(\theta')\}$. Monotonicity of $y^*()$ implies that $J(y_1, y_2)$ is a countable set, which in turn implies that $D(y_1, y_2)$ is a countable set. Moreover, $\sum_{\theta \in J(y_1,y_2)} |l^a(\inf_{\theta'>\theta} y(\theta')) - l^a(\sup_{\theta'<\theta} y(\theta'))| \leq \int_{y_1}^{y_2} \max(|\frac{\partial l^a(0,y)}{\partial y}|, |\frac{\partial l^a(1,y)}{\partial y}|) < \infty$. Hence, the total increment of $\widetilde{m}$ at points $D(y_1, y_2)$ is well-defined and given by $\sum_{\theta \in J(y_1,y_2)} l^a(\inf_{\theta'>\theta} y(\theta')) - l^a(\sup_{\theta'<\theta} y(\theta'))$. Part (i) of Lemma 1 implies that on $([y_1, y_2] \cap R(y^*) \backslash D(y_1, y_2))$ function $\widetilde{m}$ is absolutely continuous, hence the total change in $\widetilde{m}$ over $([y_1, y_2] \cap R(y^*) \backslash D(y_1, y_2))$ is given by $\int_{y \in ([y_1,y_2] \cap R(y^*) \backslash D(y_1,y_2))} \left(-\frac{\partial l^a(\tilde{\theta}(y),y)}{\partial y}\right) dy$. Since $D(y_1, y_2)$ is countable, the latter integral is equal to $\int_{y \in ([y_1,y_2] \cap R(y^*)} \left(-\frac{\partial l^a(\tilde{\theta}(y),y)}{\partial y}\right) dy$. ■

Proof of Claim 6**:** First suppose that $y^*(0) < 0$ and $y^*(\theta) \geq \theta\ \forall\ \theta \in \Theta$. If $\lim_{\theta \searrow 0} m^*(\theta) = 0$ then incentive compatibility implies that there is $\theta_1 > 0$ and $K \geq \theta_1 + b(\theta_1)$ such that (i) either $\theta_1 = 1$ or $K \geq \theta_1 + b(\theta_1)$; (ii) $y^*(\theta) = K$ and $m^*(\theta) = 0$ for every $\theta \in (0, \theta_1]$. Changing $y^*()$ to specify $\theta + b(\theta)$ for $\theta \in [0, \theta_1]$ and leaving other aspects of the delegation scheme unchanged is then clearly a profitable deviation for the principal. If $\lim_{\theta \searrow 0} m^*(\theta) > 0$ then consider the following scheme $y^{**}(), m^{**}()$ by the principal: let $m^{**}(0) = 0$ and $y^{**}(0) = \min(b(0), \lim_{\theta \searrow 0} y^*(\theta))$. For $\theta \in (0, 1]$, let $y^{**}(\theta) = \min(y^*(\theta), \theta + b(\theta))$, and define $m^{**}(\theta)$ such that $y^{**}(), m^{**}()$ satisfies (3) over $\theta \in [0, \theta_1)$. Note that, given a fixed $m^{**}(0)$ and $y^{**}(0)$, by Claim 5 there is exactly one scheme $y^{**}(), m^{**}()$ satisfying the above requirements. Note that $\lim_{\theta \searrow 0} m^*(\theta) > 0$ implies that there is some $\theta_1 > 0$ such that at all states $[0, \theta_1]$ the agent prefers scheme $y^{**}(), m^{**}()$ to scheme $y^*(), m^*()$. If the agent prefers $y^{**}(), m^{**}()$ to $y^*(), m^*()$ at all states, then scheme $y^{**}(), m^{**}()$ is a profitable deviation for the principal, since it requires a smaller amount of T to satisfy (1b), and since by construction $l^p(\theta, y^{**}(\theta)) \leq l^p(\theta, y^*(\theta))$ for every $\theta \in (0, 1]$. Suppose now that there is some $\theta > \theta_1$ such that at θ the agent prefers $y^{**}(\theta), m^{**}(\theta)$ to $y^*(\theta), m^*(\theta)$. Then, since by Claim 4 the agent's ex post utility is continuous both at scheme $y^*(), m^*()$ and at scheme $y^{**}(), m^{**}()$, there exists $\theta_2 > \theta_1$ such that the agent strictly prefers $y^{**}(\theta), m^{**}(\theta)$ to $y^*(\theta), m^*(\theta)$ at every $\theta \in (0, \theta_2)$, and exactly indifferent between $y^{**}(\theta_2), m^{**}(\theta_2)$ to $y^*(\theta_2), m^*(\theta_2)$ at θ_2. Consider now a deviation by the principal that specifies scheme $y^{**}(\theta), m^{**}(\theta)$ for states $\theta \in [0, \theta_2)$, and scheme $y^*(\theta), m^*(\theta)$ for states $\theta \in [\theta_2, 1]$. By construction $y^*(\theta_2) \geq y^{**}(\theta_2)$, and both $y^*(), m^*()$ and $y^{**}(), m^{**}()$ satisfy (3), therefore the proposed scheme also satisfies (3). The deviation is profitable, since in all states both the principal and the agent are better off (the agent strictly over $[0, \theta_2)$), and $l^p(\theta, y^{**}(\theta)) \leq l^p(\theta, y^*(\theta))$ for every $\theta \in (0, 1]$.

Suppose now that there is $\theta' \in (0, 1]$ for which $y^*(\theta) < \theta$. By Claim 1 then there exist $\theta'' < \theta'$ and $\varepsilon > 0$ such that $y^*(\theta) < \theta - \varepsilon$ for every $\theta \in [\theta'', \theta']$. Moreover, there exist $\theta_1 \leq \theta''$ and $\theta_2 \geq \theta'$ such that: (i) $y^*(\theta) < \theta$ for every $\theta \in (\theta_1, \theta_2)$; (ii) either $\theta_1 = 0$ or $y^*(\theta_1) = \theta_1$; (iii) either $\theta_2 = 1$ or $y^*(\theta_2) \geq \theta_2$. Consider now the following scheme $y^{**}(), m^{**}()$: let $y^{**}(\theta) = y^*(\theta)$ and $m^{**}(\theta) = m^*(\theta)$ for every $\theta \in [0, \theta_1)$; let $y^{**}(\theta) = \theta$ for every $\theta \in [\theta_1, \theta_2)$; let $y^{**}(\theta) = \min(y^*(\theta), \theta + b(\theta))$ for every $\theta \in [\theta_2, 1]$; and let $m^{**}()$ over $[\theta_1, 1]$ be defined such that $y^{**}(), m^{**}()$ satisfies (3). Note that by Claim 4 there exists only one scheme $y^{**}(), m^{**}()$ satisfying the above requirement. Further note that the agent weakly prefers $y^{**}(\theta), m^{**}(\theta)$ to $y^*(\theta), m^*(\theta)$ for $\theta \leq \theta_2$, and strictly for $\theta \in (\theta_1, \theta_2)$. Consider first the case that the agent weakly prefers $y^{**}(\theta), m^{**}(\theta)$ to $y^*(\theta), m^*(\theta)$ for every $\theta \in \Theta$. Then scheme $y^{**}(), m^{**}()$ is unambiguously a profitable deviation for the principal, since it both reduces the amount of T needed to satisfy (1b), and since by construction $l^p(\theta, y^{**}(\theta)) \leq l^p(\theta, y^*(\theta))$ for every $\theta \in (0, 1]$. Consider next the case that there is some $\theta_3 > \theta_2$ such that the agent prefers $y^{**}(\theta_3), m^{**}(\theta_3)$ to $y^*(\theta_3), m^*(\theta_3)$ at θ_3. Claim 4 implies that ex post utilities are continuous for both $y^*(), m^*()$ and $y^{**}(), m^{**}()$. Then there exists some θ_0 such that the agent weakly prefers $y^{**}(\theta), m^{**}(\theta)$ to $y^*(\theta), m^*(\theta)$ at every $\theta \in [0, \theta_0]$, and the agent is exactly indifferent between $y^{**}(\theta_0), m^{**}(\theta_0)$ to $y^*(\theta_0), m^*(\theta_0)$ at θ_0. Then a scheme which prescribes $y^*(\theta), m^*(\theta)$ for $\theta \in [0, \theta_0)$ and $y^{**}(\theta), m^{**}(\theta)$ for $\theta \in [\theta_0, 1]$ is unambiguously a profitable deviation for the principal. Note that by construction $y^{**}(\theta_0) \leq y^*(\theta_0)$, hence the fact that both $y^*(), m^*()$ and $y^{**}(), m^{**}()$ satisfy (3) imply that the above deviation scheme satisfies (3), too. ■

Lemma 2 *Suppose that for each $k \in \mathbb{N}$, $y_k : \Theta \to Y$ (where Y is compact) is a weakly increasing function. Then there exists a strictly increasing sequence of natural numbers $\{k_n\}_{n \in \mathbb{N}}$ and weakly increasing function $y : \Theta \to Y$ such that subsequence of functions y_{k_n} converges to y pointwisely, or, formally,*

$$\forall \theta \in \Theta : \lim_{n \to \infty} y_{k_n}(\theta) = y(\theta).$$

Proof of Lemma 2. Take a countable dense subset of Θ which includes $\theta_L = 0$, for example, the rational numbers, and enumerate its elements as $\{\theta_r\}_{r \in \mathbb{N}}$. Since Y is a compact, there is a subsequence $\{y_{l_n^1}\}_{n \in \mathbb{N}}$ (where $\{l_n^1\}_{n \in \mathbb{N}}$ is a strictly increasing sequence of natural numbers) such that sequence $y_{l_n^1}(\theta_1)$ converges to some element of set Y; denote it by q_1. From sequence of functions $\{y_{l_n^1}\}_{n \in \mathbb{N}}$ take a subsequence $\{y_{l_n^2}\}_{n \in \mathbb{N}}$ (so $\{l_n^2\}_{n \in \mathbb{N}} \subset \{l_n^1\}_{n \in \mathbb{N}}$) such that $y_{l_n^2}(\theta_2)$ converges to some $q_2 \in Y$. We then proceed likewise and construct subsequences $y_{l_n^r}$ such that

$y_{l_n^r}(\theta_j)$ converges to some q_r for each $r \in \mathbb{N}$. Now consider the "diagonal" sequence of functions $\{y_{l_n}\}_{n\in\mathbb{N}}$, where $y_{l_n} = y_{l_n^n}$; evidently, sequence $\{y_{l_n}(\theta_j)\}_{n\in\mathbb{N}}$ converges to q_j for any $j \in \mathbb{N}$.

Define define function $z : \Theta \to Y$ by

$$z(\theta) = \sup_{r:\theta_r \leq \theta} q_r. \tag{18}$$

Since θ_L belongs to $\{\theta_r\}_{r\in\mathbb{N}}$ by construction, function $z(\theta)$ is well-defined; it is also weakly increasing. Note that if $\theta_r > \theta_j$, then $q_r \geq q_j$ (this is easy to prove by contradiction); this immediately implies that $z(\theta_r) = q_r$ for all $r \in \mathbb{N}$.

One can show that if z is continuous at θ, then $\lim_{n\to\infty} y_{l_n}(\theta) = z(\theta)$. Indeed, take any $\varepsilon > 0$ and let r_1 and r_2 be such that $\theta_{r_1} \leq \theta \leq \theta_{r_2}$, $|z(\theta_{r_1}) - z(\theta)| < \varepsilon/2$, and $|z(\theta_{r_2}) - z(\theta)| < \varepsilon/2$. There exists $N \in \mathbb{N}$ such that for $n > N$ we have $|y_{l_n}(\theta_{r_1}) - z(\theta_{r_1})| < \varepsilon/2$ and $|y_{l_n}(\theta_{r_2}) - z(\theta_{r_2})| < \varepsilon/2$. Now for $n > N$ we have, since both y_{l_n} and z are weakly increasing,

$$\begin{aligned} y_{l_n}(\theta) - z(\theta) &\leq y_{l_n}(\theta_{r_2}) - z(\theta) \\ &= (y_{l_n}(\theta_{r_2}) - z(\theta_{r_2})) + (z(\theta_{r_2}) - z(\theta)) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{aligned}$$

and, similarly,

$$\begin{aligned} z(\theta) - y_{l_n}(\theta) &\leq z(\theta) - y_{l_n}(\theta_{r_1}) \\ &= (z(\theta_{r_1}) - y_{l_n}(\theta_{r_1})) + (z(\theta) - z(\theta_{r_1})) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

This implies that $|z(\theta) - y_{l_n}(\theta)| < \varepsilon$ for $n > N$, so $\lim_{n\to\infty} y_{l_n}(\theta) = z(\theta)$ whenever z is continuous at θ.

Our final step is to pick a subsequence of sequence $\{y_{l_n}(\theta_j)\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} y_{l_n}(\theta)$ exists even if z is discontinuous at θ. As z is weakly increasing, the set of such points is empty, finite, or countable. If it is empty, then we can let $y(\theta) = z(\theta)$ for all $\theta \in \Theta$ and finish the proof. Otherwise, enumerate the points of discontinuity and denote them by $\left\{\theta'_r\right\}$. As before, we first take a subsequence $\left\{y_{k_n^1}\right\}_{n\in\mathbb{N}}$ such that $\left\{y_{k_n^1}\left(\theta'_1\right)\right\}_{n\in\mathbb{N}}$ converges to some q'_1, then sequence $\left\{y_{k_n^2}\right\}_{n\in\mathbb{N}}$ etc. If the set of points of discontinuity is finite, we will be done in a finite number of steps, otherwise we again take the diagonal subsequence. In any case, we end up with a subsequence of functions $\{y_{k_n}\}_{n\in\mathbb{N}}$ such that $\{y_{k_n}(\theta)\}_{n\in\mathbb{N}}$ has a limit for all $\theta \in \Theta$; denote this limit function by y. Evidently, the set of points where y is continuous coincides with the set where z is continuous, and for such points $y(\theta) = z(\theta)$. Again, it is easy to prove by contradiction that y is weakly increasing. This finishes the proof of Lemma 2. ■

Proof of Theorem 7. First, we observe that there is a pair of functions $(y(\theta), m(\theta))_{\theta\in\Theta}$ such that conditions (3) and (4) are satisfied and $m(\theta) \in [0, M]$ $\forall$ $\theta \in \Theta$: take, for example, $y(\theta) = b(\theta) = \min l^a(\theta, y)$ and $m(\theta) = 0$. On the other hand, $V^p(y(\cdot), m(\cdot)) \geq 0$, provided that (4) holds. Hence, the function $V(y(\cdot), m(\cdot))$ has a well-defined infimum $z \geq 0$ on the set of pairs of functions $(y(\theta), m(\theta))_{\theta\in\Theta}$ which satisfy (3) and (4). Denote the set of pairs of such (measurable) functions by B. Our goal is to show that $\exists\,(y(\cdot), m(\cdot)) \in B$ such that $V^p(y(\cdot), m(\cdot)) = z$.

Suppose, to obtain a contradiction, that $\forall\,(y(\cdot), m(\cdot)) \in B$ we have $V(y(\cdot), m(\cdot)) > z$. Take a sequence of pairs $\{(y_k(\cdot), m_k(\cdot))\}_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} V(y_k(\cdot), m_k(\cdot)) = z$. Without loss of generality, assume that for each k, $y_k(\theta) \in Y$ for all $\theta \in \Theta$, where Y is some compact, and also that for each k, $\inf_{\theta\in\Theta} m_k(\theta) = 0$ (the latter doesn't entail loss of generality because decreasing $m_k(\cdot)$ by a constant decreases $V(y_k(\cdot), m_k(\cdot))$). By Lemma 1, $y_k(\cdot)$ is weakly increasing for any k, while $m_k(\cdot)$ is a function of bounded variation with variation not exceeding $(y_H - y_L)\Delta_y$ for each k. The last condition implies that each $m_k(\cdot)$ may be represented as a difference of two weakly increasing functions $m_k^+(\cdot), m_k^-(\cdot) : \Theta \to [0, 2(\max_{y\in Y} y - \min_{y\in Y} y)\Delta_y]$, so that $m_k(\theta) = m_k^+(\theta) - m_k^-(\theta)$ for all θ. By Lemma 2 we can choose a subsequence k_n such that each of the sequences of functions $\{y_{k_n}(\cdot)\}_{n\in\mathbb{N}}$, $\{m_{k_n}^+(\cdot)\}_{n\in\mathbb{N}}$, and $\{m_{k_n}^-(\cdot)\}_{n\in\mathbb{N}}$ pointwisely converge, respectively, to some functions $y^*(\cdot)$, $(m^+)^*(\cdot)$, and $(m^-)^*(\cdot)$ (to prove this, we need to apply Lemma 2 three times, each time to the subsequence obtained in the previous step). Now define function $m^*(\cdot)$ by

$$m^*(\theta) = (m^+)^*(\theta) - (m^-)^*(\theta) \text{ for all } \theta \in \Theta; \tag{19}$$

obviously, $m_{k_n}(\cdot) = m_{k_n}^+(\cdot) - m_{k_n}^-(\cdot)$ pointwisely converges to $m^*(\cdot)$.

We now show that delegation plan $(y^*(\cdot), m^*(\cdot))$ minimizes V subject to (3) and (4). Condition (3) trivially holds, since $m^*(\cdot)$ is defined by (19) and $m_{k_n}^+(\theta) - m_{k_n}^-(\theta) = m_{k_n}(\theta) \geq 0$ for all $n \in \mathbb{N}$ and for all $\theta \in \Theta$. Now take any $\theta, \theta' \in \Theta$; for all n, we have

$$l^a(\theta, y_{k_n}(\theta)) + m_{k_n}(\theta) \leq l^a(\theta, y_{k_n}(\theta')) + m_{k_n}(\theta').$$

Taking the limit with $n \to \infty$, we obtain

$$l^a(\theta, y^*(\theta)) + m^*(\theta) \leq l^a(\theta, y^*(\theta')) + m^*(\theta'),$$

meaning that (3) holds. Finally,

$$\begin{aligned} 0 &\leq l^p(\theta, y_{k_n}(\theta)) + l^a(\theta, y_{k_n}(\theta)) + m_{k_n}(\theta) \\ &\leq \max_{\theta\in\Theta, y\in Y} l^p(\theta, y) + \max_{\theta\in\Theta, y\in Y} l^a(\theta, y) + M, \end{aligned}$$

where the expression on the right is a constant. Hence, by the dominated convergence theorem, we have

$$V^p(y^*(\cdot), m^*(\cdot)) = \lim_{n\to\infty} V^p(y_{k_n}(\cdot), m_{k_n}(\cdot)) = z.$$

This proves that contract $(y^*(\cdot), m^*(\cdot))$ solves (2) and is therefore optimal. ■

Lemma 3 *Suppose that* $\frac{\frac{\partial l^a(\theta,y_1)}{\partial\theta} - \frac{\partial l^a(\theta,y_0)}{\partial\theta}}{l^p(\theta,y_0) - l^p(\theta,y_1)} > \frac{\frac{\partial l^a(\theta,y_1)}{\partial\theta} - \frac{\partial l^a(\theta,y_2)}{\partial\theta}}{l^p(\theta,y_2) - l^p(\theta,y_1)}$, *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* $(0,1)$.

Proof: Note that Claim 5 implies that m^* is discontinuous at θ iff y^* is discontinuous at θ. The proof below is by contradiction. Suppose that for some $\theta_0 \in (0,1)$, y^* is discontinuous at θ_0. Denote

$$\hat{y}_1 = \sup\nolimits_{\theta\in[0,\theta_0)} y^*(\theta),$$
$$\hat{y}_2 = \inf\nolimits_{\theta\in(\theta_0,1]} y^*(\theta).$$

Note that, since $y^*(\theta)$ is monotonic, it is true that $\hat{y}_1 = \lim_{\theta\to\theta_0-} y^*(\theta)$, $\hat{y}_2 = \lim_{\theta\to\theta_0+} y^*(\theta)$. Define $\hat{m}_1 = \lim_{\theta\to\theta_0-} m^*(\theta) \geq 0$ and $\hat{m}_2 = \lim_{\theta\to\theta_0+} m^*(\theta) \geq 0$; these limits exist by continuity of loss function $L^a(\theta)$: $\hat{m}_1 = \lim_{\theta\to\theta_0-} L^a(\theta) - l^a(\theta_0, \hat{y}_1)$, and similarly $\hat{m}_2 = \lim_{\theta\to\theta_0+} L^a(\theta) - l^a(\theta_0, \hat{y}_2)$. It is evident that an agent of type θ_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 (3).

The idea of the proof is to perturb the optimal contract $(y^*(\theta), m^*(\theta))_{\theta\in\Theta}$ around the point of discontinuity θ_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

$$\frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial\theta} = a\frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial\theta} + (1-a)\frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial\theta}; \tag{20}$$

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, (20) 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 θ_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 (3), 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 θ does, and choosing $(\hat{y}_0, \hat{m}_0)$ if he had such an option. From Claim 4 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 θ such that $\theta_1(\varepsilon)$ is decreasing and $\theta_2(\varepsilon)$ is increasing in ε. As $\varepsilon \to 0$, $\theta_1(\varepsilon) \to \theta_0$ and $\theta_2(\varepsilon) \to \theta_0$. 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 Lemma 1 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

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

Therefore,

$$\begin{aligned} \lim_{\varepsilon \to 0+} \frac{\theta_0 - \theta_1(\varepsilon)}{\theta_2(\varepsilon) - \theta_0} &= \frac{\lim_{\varepsilon \to 0+} \frac{\varepsilon}{\theta_2(\varepsilon) - \theta_0}}{-\lim_{\varepsilon \to 0+} \frac{-\varepsilon}{\theta_0 - \theta_1(\varepsilon)}} \\ &= \frac{\frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_2)}{\partial \theta}}{\frac{\partial l^a(\theta_0, \hat{y}_1)}{\partial \theta} - \frac{\partial l^a(\theta_0, \hat{y}_0)}{\partial \theta}} = \frac{a}{1-a}. \end{aligned} \tag{21}$$

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 principal's payoff. This change equals

$$\int_{\theta_1(\varepsilon)}^{\theta_2(\varepsilon)} (l^p(\theta, y(\theta)) - l^p(\theta, \hat{y}_0)) f(\theta) d\theta = \int_{\theta_1(\varepsilon)}^{\theta_2(\varepsilon)} \int_{\hat{y}_0}^{y(\theta)} \frac{\partial l^p(\theta, y)}{\partial y} dy f(\theta) d\theta \tag{22}$$

$$= \int_{\theta_0}^{\theta_2(\varepsilon)} (l^p(\theta, y(\theta)) - l^p(\theta, \hat{y}_0)) f(\theta) d\theta - \int_{\theta_1(\varepsilon)}^{\theta_0} (l^p(\theta, \hat{y}_0) - l^p(\theta, y(\theta))) f(\theta) d\theta.$$

To check that this expression is positive, it is sufficient, given the continuity of $f(\theta)$ at θ_0 and existence of limits $\lim_{\theta \to \theta_0 -} \frac{\partial l^p(\theta, y)}{\partial y} = \frac{\partial l^p(\theta_0, y)}{\partial y}$ and $\lim_{\theta \to \theta_0 +} \frac{\partial l^p(\theta, y)}{\partial y} = \frac{\partial l^p(\theta_0, y)}{\partial y}$, to prove that

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

In light of (21), it suffices to prove that

$$(1-a)(l^p(\theta, \hat{y}_2) - l^p(\theta, \hat{y}_0)) > a(l^p(\theta, \hat{y}_0) - l^p(\theta, \hat{y}_1)). \tag{23}$$

By Claim 6, $l^p(\theta, \hat{y}_2) > l^p(\theta, \hat{y}_0)$ and $l^p(\theta, \hat{y}_0) > l^p(\theta, \hat{y}_1)$, and (23) is equivalent to

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

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

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

Now, rearranging (this is safe since the denominators are positive) and changing the sign, we get

$$\frac{\frac{\partial l^a(\theta_0,\hat{y}_1)}{\partial\theta}-\frac{\partial l^a(\theta_0,\hat{y}_0)}{\partial\theta}}{l^p(\theta_0,\hat{y}_0)-l^p(\theta_0,\hat{y}_1)} > \frac{\frac{\partial l^a(\theta_0,\hat{y}_1)}{\partial\theta}-\frac{\partial l^a(\theta_0,\hat{y}_2)}{\partial\theta}}{l^p(\theta_0,\hat{y}_2)-l^p(\theta_0,\hat{y}_1)}. \tag{24}$$

Claims 1 and 6 imply that $\theta \leq \hat{y}_1 < \hat{y}_0 < \hat{y}_2$, hence the assumption of the lemma implies that (24) holds. This implies that the proposed deviation is profitable, contradicting y^* is discontinuous at θ_0. ■

Proof of Theorem 8 Note that A1 implies that for any $\theta \leq y_1 < y_0 < y_2$ the following holds:

$$\frac{\frac{\frac{\partial l^a(\theta_0,\hat{y}_1)}{\partial\theta}-\frac{\partial l^a(\theta_0,y_0)}{\partial\theta}}{\hat{y}_0-\hat{y}_1}}{\frac{l^p(\theta_0,y_0)-l^p(\theta_0,y_1)}{y_0-y_1}} > \frac{\frac{\frac{\partial l^a(\theta_0,y_0)}{\partial\theta}-\frac{\partial l^a(\theta_0,y_2)}{\partial\theta}}{y_2-y_0}}{\frac{l^p(\theta_0,y_2)-l^p(\theta_0,y_0)}{y_2-y_0}}.$$

This is equivalent to $\frac{\frac{\partial l^a(\theta_0,\hat{y}_1)}{\partial\theta}-\frac{\partial l^a(\theta_0,\hat{y}_0)}{\partial\theta}}{l^p(\theta_0,\hat{y}_0)-l^p(\theta_0,\hat{y}_1)} > \frac{\frac{\partial l^a(\theta_0,\hat{y}_1)}{\partial\theta}-\frac{\partial l^a(\theta_0,\hat{y}_2)}{\partial\theta}}{l^p(\theta_0,\hat{y}_2)-l^p(\theta_0,\hat{y}_1)}$, therefore Lemma 3 implies the claim in the theorem. ■

Lemma 4 *Suppose* $(y(\theta), m(\theta))$ *satisfies* (3) *and* $y(\theta)$ *is continuous on* Θ. *Then for any* $\theta_1, \theta_2 \in \Theta$, *we have*

$$m(\theta_2) - m(\theta_1) = l^a(\theta_1, y(\theta_1)) - l^a(\theta_2, y(\theta_2)) + \int_{\theta_1}^{\theta_2} \left(\frac{\partial l^a(\theta, y(\theta))}{\partial\theta}\right) d\theta. \tag{25}$$

Proof of Lemma 4. From (7), we have

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

Rearranging, we obtain (25). ■

Proof of Theorem 9. By Theorem 8, $y^*(\theta)$ 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 $\theta_1 = \inf\{\theta : y^*(\theta) < \theta + b(\theta)\}$ if such θ exists; otherwise, let $\theta_1 = 1$. Define function $y(\theta)$ by

$$y(\theta) = \begin{cases} y^*(\theta) \text{ if } \theta > \theta_1, \\ \theta + b \text{ if } \theta \leq \theta_1; \end{cases}$$

Note that by continuity, $y^*(\theta_1) = \theta_1 + b(\theta_1)$, hence the above function is continuous. Furthermore, let

$$m(\theta) = \begin{cases} m^*(\theta) \text{ if } \theta > \theta_1, \\ m^*(\theta_1) \text{ if } \theta \leq \theta_1. \end{cases}$$

Given that scheme $y^*(), m^*()$ satisfies (3) and (4), it is straightforward to verify that scheme $(y(), m())$ also satisfies (3) and (4). In the modified scheme, the utility of the agent at $\theta \geq \theta_0$ is unchanged. If $\theta < \theta_0$, then, by (7)

$$\begin{aligned} L^a(\theta, y(\theta), m(\theta)) &= L^a(\theta_1, y(\theta_1), m(\theta_1)) - \int_\theta^{\theta_1} \frac{\partial l^a(\xi, y(\xi))}{\partial \theta} d\xi \\ &< L^a(\theta_1, y^*(\theta_1), m^*(\theta_1)) - \int_\theta^{\theta_1} \frac{\partial l^a(\xi, y^*(\xi))}{\partial \theta} d\xi = L^a(\theta, y^*(\theta), m^*(\theta)); \end{aligned}$$

this holds because at θ_1 the contract is unchanged, and

$$\int_\theta^{\theta_1} \frac{\partial l^a(\xi, y^*(\xi))}{\partial \theta} d\xi - \int_\theta^{\theta_1} \frac{\partial l^a(\xi, y(\xi))}{\partial \theta} d\xi = \int_\theta^{\theta_1} \int_{y(\xi)}^{y^*(\xi)} \frac{\partial^2 l^a(\xi, y)}{\partial \theta \partial y} d\xi < 0,$$

since $y(\xi) < y^*(\xi)$ whenever $\xi < \theta_1$ is close to θ_1. Consequently, all types of agent are at least weakly better off. The principal, is obviously better off, since for some θ, $y(\theta)$ became closer to θ than $y^*(\theta)$. This contradicts that contract $(y^*(\theta), m^*(\theta))$ solves the problem (2).

Now suppose that 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, $\hat{\theta}_1$ and $\hat{\theta}_2$ 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 $y^*(\theta_2 + \varepsilon) > \theta_1 + \varepsilon + b(\theta_1 + \varepsilon)$. This implies, in particular, that $m^*(\theta)$ is bounded away from 0 on $[\theta_1, \theta_2]$ (this is a trivial corollary of Claim 4, since $m^*(\theta_1 - \varepsilon)$ is non-negative).

Let us construct an alternative $y(\theta)$ as follows. We take ε_1 and ε_2 to be such small positive numbers such that

$$\int_{\theta_1 - \varepsilon_1}^{\theta_1} \int_{y^*(\theta)}^{\theta + b(\theta)} \left(-\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y}\right) dy d\theta = \int_{\theta_2}^{\theta_2 + \varepsilon_2} \int_{\theta + b(\theta)}^{y^*(\theta)} \left(-\frac{\partial^2 l^a(\theta, y)}{\partial \theta \partial y}\right) dy d\theta, \tag{26}$$

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

$$\int_{\theta_1-\varepsilon_1-\varepsilon_0}^{\theta_1}\int_{y^*(\theta)}^{y(\theta)}\left(-\frac{\partial^2 l^a(\theta,y)}{\partial\theta\partial y}\right)dyd\theta=\int_{\theta_2}^{\theta_2+\varepsilon_2+\varepsilon_0}\int_{y^*(\theta)}^{y(\theta)}\left(-\frac{\partial^2 l^a(\theta,y)}{\partial\theta\partial y}\right)dyd\theta. \tag{27}$$

Now, if we define $m(\theta)$ to be such that the agent's loss function $L^a(\theta,y(\theta),m(\theta))$ satisfies (7) 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 (3) and (4).

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 (26). The change in the principal's utility is given by

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

It suffices to show that

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

Dividing this by (27), we are to prove

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

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

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

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

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

However, this follows from A2. This completes the proof. ■

Proof of Claim 10 Suppose, to obtain a contradiction, that this does not hold. Then there is θ_0 such that $y(0)<y(\theta_0)<y(1)$ (which means, in particular, that $0<\theta_0<1$) and

$y(\theta_0) \neq \min\{x(\theta_0), \theta_0 + b\}$. First, consider the case where $x(\theta)$ is increasing or constant (note that it is a linear function of θ). Suppose $y(\theta_0) < \min\{x(\theta_0), \theta_0 + b\}$. Then, by continuity of $y(\theta)$ (a constraint in the optimization problem) and the assumption that $y(\theta_0) < y(1)$, there exists $\theta' > \theta_0$ such that $y(\theta') < \min\{x(\theta'), \theta' + b\}$ and $y(\theta_0) < y(\theta')$. But then slightly increasing $y(\theta)$ for $\theta \in (\theta_0, \theta')$ while preserving $y(\theta_0)$ and $y(\theta')$ would decrease the value function without violating the constraints. Now suppose $y(\theta_0) > \min\{x(\theta_0), \theta_0 + b\}$; since $y(\theta_0) \leq \theta_0 + b$ (a constraint in the optimization problem), we must have $x(\theta_0) < y(\theta_0) \leq \theta_0 + b$. Since $x(\theta)$ is increasing or constant and $y(\cdot)$ is continuous, we can choose $\theta' < \theta_0$ such that $x(\theta') < y(\theta') < y(\theta_0)$. Then if we slightly decrease $y(\theta)$ for $\theta \in (\theta', \theta_0)$ while preserving $y(\theta')$ and $y(\theta_0)$ would decrease the value function without violating the constraints. So, if $x(\theta)$ is not decreasing, we get to a contradiction.

Now suppose that $x(\theta)$ is strictly decreasing. Let us first suppose that $y(\theta_0) > \min\{x(\theta_0), \theta_0 + b\}$, i.e., $x(\theta_0) < y(\theta_0) \leq \theta_0 + b$. Then $x(1) < y(1) \leq 1 + b$, 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 $x(\theta)$ on $(\theta_0, 1]$. This means, in particular that in this case, if for some θ', $y(\theta') = x(\theta')$ 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 $x(\theta) < y(\theta) < y(1)$, which is, as we just proved, impossible. Now consider the remaining case, $y(\theta_0) < \min\{x(\theta_0), \theta_0 + b\}$. There are two possibilities. If $x(1) \geq y(1)$, then, as before, there exists $\theta' > \theta_0$ such that $y(\theta') < \min\{x(\theta'), \theta' + b\}$ and $y(\theta_0) < y(\theta')$, and slightly increasing $y(\theta)$ for $\theta \in (\theta_0, \theta')$ while preserving $y(\theta_0)$ and $y(\theta')$ would decrease the value function. If, however, $x(1) < y(1)$, then there is some $\theta' \in (\theta_0, \theta_H)$ for which $y(\theta') = x(\theta')$. But then, as we argued above, $y(\theta') = y(\theta_H) > y(\theta_0)$. 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 value function, again contradicting that the contract given by $y(\cdot)$ is optimal. This contradiction completes the proof of Claim 10. ■

Proof of Theorem 11: Note that the requirement that $\frac{\frac{\partial l^a(\theta,y_1)}{\partial\theta} - \frac{\partial l^a(\theta,y_0)}{\partial\theta}}{l^p(\theta,y_0) - l^p(\theta,y_1)} > \frac{\frac{\partial l^a(\theta,y_1)}{\partial\theta} - \frac{\partial l^a(\theta,y_2)}{\partial\theta}}{l^p(\theta,y_2) - l^p(\theta,y_1)}$ for every $\theta \in (0,1)$ and $\inf_{\theta\in(\theta_0,1]} y^*(\theta) \geq y_2 > y_0 > y_1 \geq \theta$ holds whenever:

$$\frac{\frac{\partial l^a(\theta_0,\hat{y}_1)}{\partial\theta} - \frac{\partial l^a(\theta_0,y)}{\partial\theta}}{l^p(\theta_0, y) - l^p(\theta_0, \hat{y}_1)} \tag{28}$$

is decreasing in y for $\inf_{\theta\in(\theta_0,1]} y^*(\theta) \geq y > \hat{y}_1$, for every $\theta_0 \in (0,1)$ and $\hat{y}_1 \geq \theta_0$. If $y^*(\theta) \leq \theta + b(\theta)$ for every $\theta \in (0,1)$, then $\inf_{\theta\in(\theta_0,1]} y^*(\theta) \leq \theta + b(\theta)$ for every $\theta \in (0,1)$. Then for every $\inf_{\theta\in(\theta_0,1]} y^*(\theta) \geq y > \hat{y}_1$, the numerator of (28) is decreasing in y, while the denominator of (28) is increasing in y, implying that (28) is decreasing in y. Lemma 3 then implies the theorem.
■

Proof of Claim 12: Analogously to the proof of Theorem , it can be established that if A2 holds in a neighborhood of $\theta' \in (0,1)$ and $y^*()$ is continuous at θ' then $y^*(\theta) \leq \theta + b(\theta)$ for all θ in a neighborhood of θ'. This implies that for the model specification in the Claim, if $y^*()$ is continuous then either $y^*(\theta) \leq \theta + 0.05$ for every $\theta \in \Theta$, or $y^*(\theta) \leq \theta + 0.05$ for $\theta \leq \frac{2}{3}$ and there is $\delta > 0$ such that $y^*(\theta) > \theta + 0.05$ for $\theta \in (\frac{2}{3} + \delta)$. In both of these cases, for small enough $\varepsilon > 0$ $y^*()$ yields a lower utility for the principal than the contract given by:

$$y(\theta) = \begin{cases} \theta + 0.05 \text{ if } \theta \leq \frac{7}{30} \\ \frac{1}{6} + \frac{\theta}{2} \text{ if } \frac{7}{30} < \theta < \frac{1}{3} \\ \theta + 0.05 \text{ if } \frac{2}{3} \leq \theta \leq \frac{9}{10} \\ \theta + \frac{1-\theta}{A} \text{ if } \frac{9}{10} < \theta \\ \frac{1}{20}\sqrt{2} + \frac{23}{60} \text{ if } \frac{1}{3} < \theta \leq 0.404\,04 \\ \theta + 0.05 \text{ if } 0.404\,04 < \theta < \frac{2}{3}. \end{cases}$$

This contradicts that $y^*()$ is optimal, hence $y^*()$ has to be discontinuous. ■

Proof of Claim 13. The proofs of these results follow closely the proofs of similar results in the case without conditional transfers and are omitted. ■

Proof of Claim 14. (i) Suppose that $m^*(\theta) > 0$ and $t^*(\theta) > 0$ for a positive measure of θ. Then there exists $\varepsilon > 0$ such that the measure of the set $\{\theta : m^*(\theta) > \varepsilon, t^*(\theta) > \varepsilon\}$ is positive. For all such θ's, let $m'(\theta) = m^*(\theta) - \varepsilon$ and $t'(\theta) = t^*(\theta) - \varepsilon$; in other cases, let $m'(\theta) = m^*(\theta)$ and $t'(\theta) = t^*(\theta)$. Then the contract $(y^*(\cdot), m'(\cdot), t'(\cdot))$ would satisfy all constraints and yield a higher payoff to the principal than $(y^*(\cdot), m^*(\cdot), t^*(\cdot))$, which is impossible. This contradiction proves that either $m^*(\theta) = 0$ or $t^*(\theta) = 0$ for almost all θ. Now, from Claim 13 we get that function $m^*(\theta) - t^*(\theta)$ is nondecreasing, and may without loss of generality assumed to be continuous. Therefore, we must have $t^*(\theta) = 0$ whenever $m^*(\theta) - t^*(\theta) > 0$ and $m^*(\theta) = 0$ whenever $m^*(\theta) - t^*(\theta) < 0$. Consequently, $m^*(\theta) = \max\{m^*(\theta) - t^*(\theta), 0\}$ and is therefore nondecreasing, while $t^*(\theta) = \max\{t^*(\theta) - m^*(\theta), 0\}$ is nonincreasing.

(ii) Without loss of generality, we may restrict attention to contracts with $m^*(0) = 0$. If this were not the case, we could take $m'(\theta) = m^*(\theta) - m(0) \geq 0$ since $m^*(\theta)$ is nondecreasing, and we would get a contract $(y^*(\cdot), m'(\cdot), t^*(\cdot))$ which would satisfy all constraints and have $m'(0) = 0$. Given that, consider function $t^*(\theta)$. If $t^*(\theta) = 0$ for some, then consider the supremum θ_0 of such points. By continuity of $t^*(\cdot)$ and $m^*(\cdot)$, we must have that $t^*(\theta_0) = m^*(\theta_0) = 0$. This completes the proof. ■

Proof of Claim 15: Let us start with the case where $t^*(\theta_0) = m^*(\theta_0) = 0$. We then use the integral formulas to get the following. For any $\theta < \theta_0$,

$$t^*(\theta) = l^a(\theta, y^*(\theta)) - l^a(\theta_0, y^*(\theta_0)) + \int_\theta^{\theta_0} \left(\frac{\partial l^a(\xi, y^*(\xi))}{\partial \theta} \right) d\xi;$$

therefore, the total amount of transfers the agent receives is

$$\begin{aligned} T &= \int_0^{\theta_0} \left(l^a\left(\theta, y^*\left(\theta\right)\right) - l^a\left(\theta_0, y^*\left(\theta_0\right)\right) + \int_\theta^{\theta_0} \left(\frac{\partial l^a\left(\xi, y^*\left(\xi\right)\right)}{\partial \theta} \right) d\xi \right) f\left(\theta\right) d\theta \\ &= -l^a\left(\theta_0, y^*\left(\theta_0\right)\right) F\left(\theta_0\right) + \int_0^{\theta_0} l^a\left(\theta, y^*\left(\theta\right)\right) f\left(\theta\right) d\theta + \int_0^{\theta_0} \int_\theta^{\theta_0} \left(\frac{\partial l^a\left(\xi, y^*\left(\xi\right)\right)}{\partial \theta} \right) d\xi d\theta \\ &= -l^a\left(\theta_0, y^*\left(\theta_0\right)\right) F\left(\theta_0\right) + \int_0^{\theta_0} \left(l^a\left(\theta, y^*\left(\theta\right)\right) f\left(\theta\right) + \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} F\left(\theta\right) \right) d\theta. \end{aligned}$$

The loss of agent of type θ equals

$$L^a\left(\theta\right) = L^a\left(0\right) + \int_0^\theta \frac{\partial l^a\left(\xi, y^*\left(\xi\right)\right)}{\partial \theta} d\xi,$$

and the expected loss of agent is therefore

$$\begin{aligned} L^a &= \int_0^1 L^a\left(\theta\right) f\left(\theta\right) d\theta \\ &= \int_0^1 \left(L^a\left(0\right) + \int_0^\theta \frac{\partial l^a\left(\xi, y^*\left(\xi\right)\right)}{\partial \theta} d\xi \right) f\left(\theta\right) d\theta \\ &= L^a\left(0\right) + \int_0^1 \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} \left(1 - F\left(\theta\right)\right) d\theta \\ &= L^a\left(\theta_0\right) - \int_0^{\theta_0} \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} d\theta + \int_0^1 \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} \left(1 - F\left(\theta\right)\right) d\theta \\ &= l^a\left(\theta_0, y^*\left(\theta_0\right)\right) - \int_0^{\theta_0} \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} F\left(\theta\right) d\theta + \int_{\theta_0}^1 \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} \left(1 - F\left(\theta\right)\right) d\theta. \end{aligned}$$

Now we can substitute these formulas to the partial Lagrangian

$$\int_0^1 l^p\left(\theta, y\left(\theta\right)\right) f\left(\theta\right) d\theta - T + \lambda\left(L^a - L\right)$$

and get that the principal's problem may be rewritten as (15). If $\lambda < 0$, then uniformly increasing $m^*\left(\theta\right)$ would increase the value of the partial Lagrangian; similarly, if $\lambda > 1$ then uniformly increasing $t^*\left(\theta\right)$ would do the same. Since the value of the partial Lagrangian must be maximized at the optimal contract, then $\lambda \in [0, 1]$. Moreover, $\lambda = 0$ implies $t^*\left(\theta\right) = 0$ for all θ (otherwise instead of increasing $m^*\left(\theta\right)$ as for the case $\lambda < 0$ we could decrease $t^*\left(\theta\right)$ for the values of θ where $t^*\left(\theta\right) > 0$ and still increase the value of the partial Lagrangian; similarly, $\lambda = 1$ implies $m^*\left(\theta\right) = 0$ for all θ (otherwise we could decrease $m^*\left(\theta\right)$ for some values of θ).

Suppose now that $t^*\left(\theta\right) > 0$ for all θ. Denote $t^*\left(1\right) = t_1$; like we did before, we can compute the total amount of transfers and the amount of agent's losses and obtain the partial Lagrangian

$$\begin{aligned} -t_1 - l^a\left(1, y^*\left(1\right)\right) + \int_0^1 \left(l^a\left(\theta, y^*\left(\theta\right)\right) f\left(\theta\right) + \frac{\partial l^a\left(\theta, y^*\left(\theta\right)\right)}{\partial \theta} F\left(\theta\right) \right) d\theta + \int_0^1 l^p\left(\theta, y\left(\theta\right)\right) f\left(\theta\right) d\theta \\ + \lambda \left(t_1 + l^a\left(1, y\left(1\right)\right) - \int_0^1 \frac{\partial l^a\left(\theta, y\left(\theta\right)\right)}{\partial \theta} F\left(\theta\right) d\theta \le L \right) \Bigg\}, \end{aligned}$$

which coincides with (15) for the case $\theta_0 = 1$ except for the two new terms $-t_1 + \lambda t_1$. If $\lambda \neq 1$, then we could increase the value of the partial Lagrangian by increasing or decreasing t_1. Consequently, $\lambda = 1$. The formula for $y^*(\theta)$ follows by plugging $\lambda = 1$ in the solutions for the other cases (if $\lambda = 1$, then terms with t_1 vanish, and the partial Lagrangian coincides with the one for the case where $t^*(1) = m^*(1) = 0$. ■

Proof of Claim 16. Trivially, $z(\theta)$ is nonincreasing in this case. Let us show that $t^*(\theta) = 0$ for all θ. Indeed, if this were not the case, there would exist some interval (θ_1, θ_2) on which $y^*(\theta) < \theta + b$, $t^*(\theta) > 0$, and $y^*(\theta_1) < y^*(\theta_2)$; this follows from the integral formulas for $t^*(\theta) - m^*(\theta)$. Since $z(\theta)$ is nonincreasing, we have that either $z(\theta_1) > y^*(\theta_1)$ or $z(\theta_2) < y^*(\theta_2)$ (or both). In the first case, it would be profitable to slightly increase $y^*(\theta)$ on a small interval $(\theta_1, \theta_1 + \varepsilon)$, while in the second case it would be profitable to slightly decrease $y^*(\theta)$ on $(\theta_2 - \varepsilon, \theta_2)$. Hence, $t^*(\theta) = 0$ for all θ, which means that we can choose $\theta_0 = 1$ in the partial Lagrangian for the purpose of optimization.

The next step is to show that if $y^*(\theta) < y^*(1)$, then $y^*(\theta) = \min\{x(\theta), \theta + b\}$. Indeed, if $y^*(\theta) < \min\{x(\theta), \theta + b\}$, then we can increase $y^*(\cdot)$ in some neighborhood of θ and by doing that increase the value of the partial Lagrangian. If $y^*(\theta) > \min\{x(\theta), \theta + b\}$, then $x(\theta) < y^*(\theta) \leq \theta + b$ since from Claim 13 we know that $y^*(\theta) \leq \theta + b$, and hence we could decrease $y^*(\cdot)$ in some neighborhood of θ at the same time increasing the value of the partial Lagrangian (in these perturbations, it is important that the value of $y^*(\theta_0)$ does not change). Since the slope of $x(\theta)$ equals $\frac{A - \lambda + 2}{A+1} = \frac{1-\lambda}{A+1} \geq 1$, and $x(0) = \frac{b}{A+1} < b$, we have that $\min\{x(\theta), \theta + b\} = x(\theta)$ for small θ. Note that the solution to the equation $x(\theta') = \theta' + b$ (if it exists) is $\theta' = \frac{Ab}{1-\lambda}$; this may or may not lie on $[0, 1]$.

We are now trying to solve for the optimal value of $y^*(1)$ (i.e., for the position of the "cap") under two possible conditions: that $y^*(\theta)$ does or does not contain a part where $y^*(\theta) = \theta + b$. If it does (which means $\theta' = \frac{Ab}{1-\lambda} < 1$), we have the following minimization problem for the value s starting from which $y^*(\theta) = y^*(1)$ (note that the position of the cap does not have an impact on transfers as they are determined by $y^*(\theta)$ for $\theta < \theta'$ only):

$$\min_{\theta^*} \left(\int_{\theta'}^{\theta^*} Ab^2 d\theta + \int_{\theta^*}^{1} A(\theta^* + b - \theta)^2 d\theta - \lambda \int_{\theta^*}^{1} 2(k - \theta)(1 - \theta) d\theta \right).$$

Taking the integrals, we get a function which reaches a local minimum at $\theta^* = 1 - \frac{2Ab}{A+\lambda}$ and a local maximum at 1; therefore, the minimum of the function on $[\theta', 1]$ is reached on $\max\{\theta', \theta^*\}$. Hence, if $\theta^* \geq \theta'$, i.e., $1 - \frac{2Ab}{A+\lambda} \geq \frac{Ab}{1-\lambda}$ or, equivalently, $Ab(A + 2 - \lambda) \leq (1 - \lambda)(A + \lambda)$, then the optimal cap starts at θ^* (below we check that the optimal cap cannot start at $\theta < \theta'$ in this case). This proves the formula for the case $Ab\left(A+2-\lambda\right) < \left(1-\lambda\right)\left(A+\lambda\right)$. In the opposite case, i.e., if $Ab\left(A+2-\lambda\right) > \left(1-\lambda\right)\left(A+\lambda\right)$, the optimal cap must start at point θ^{**} where $y^{*}\left(\theta^{**}\right) = x\left(\theta^{**}\right)$ (even though it may be the rightmost of such points). The problem now is to minimize

$$-\left(x\left(\theta^{**}\right)-\theta^{**}-b\right)^{2}\theta^{**}+\int_{0}^{\theta^{**}}\left(\left(x\left(\theta\right)-\theta-b\right)^{2}-2\left(x\left(\theta\right)-\theta-b\right)\theta\right)d\theta+\int_{0}^{\theta^{**}}A\left(x\left(\theta\right)-\theta\right)^{2}d\theta+\int_{\theta^{**}}^{1}A\left(x\left(\theta^{**}\right)\right.$$

$$+\lambda\left(\left(x\left(\theta^{**}\right)-\theta^{**}-b\right)^{2}+\int_{0}^{\theta^{**}}2\left(x\left(\theta\right)-\theta-b\right)\theta d\theta-\int_{\theta^{**}}^{1}2\left(x\left(\theta^{**}\right)-\theta-b\right)\left(1-\theta\right)d\theta-L\right).$$

This function reaches its local minimum at

$$\theta^{**}=\frac{A+\lambda}{A+1}-\sqrt{\frac{\left(1-\lambda\right)\left(2Ab-\left(1-\lambda\right)\frac{A+\lambda}{A+1}\right)}{\left(A+2-\lambda\right)\left(A+1\right)}}\in\left(0,1\right)$$

and a local maximum at

$$\frac{A+\lambda}{A+1}+\sqrt{\frac{\left(1-\lambda\right)\left(2Ab-\left(1-\lambda\right)\frac{A+\lambda}{A+1}\right)}{\left(A+2-\lambda\right)\left(A+1\right)}}>1.$$

One can prove that $\theta^{**} < \theta'$ if and only if $Ab\left(A+2-\lambda\right) > \left(1-\lambda\right)\left(A+\lambda\right)$. Indeed, if $Ab\left(A+2-\lambda\right) > \left(1-\lambda\right)\left(A+\lambda\right)$, then

$$\theta^{**}-\theta'=\frac{\left(1-\lambda\right)\left(A+\lambda\right)-Ab\left(A+1\right)}{\left(1-\lambda\right)\left(A+1\right)}-\sqrt{\frac{\left(1-\lambda\right)\left(2Ab-\left(1-\lambda\right)\frac{A+\lambda}{A+1}\right)}{\left(A+2-\lambda\right)\left(A+1\right)}};$$

the first term on the right-hand side is either negative (then $\theta^{**} < \theta'$ immediately) or positive; in the latter case, $\theta^{**}-\theta' < 0$ is equivalent to

$$\left(\frac{\left(1-\lambda\right)\left(A+\lambda\right)-Ab\left(A+1\right)}{\left(1-\lambda\right)\left(A+1\right)}\right)^{2}<\frac{\left(1-\lambda\right)\left(2Ab-\left(1-\lambda\right)\frac{A+\lambda}{A+1}\right)}{\left(A+2-\lambda\right)\left(A+1\right)},$$

which simplifies to $Ab\left(A+2-\lambda\right) > \left(1-\lambda\right)\left(A+\lambda\right)$. If, however, $Ab\left(A+2-\lambda\right) \leq \left(1-\lambda\right)\left(A+\lambda\right)$, then the first term is unambiguously positive as it exceeds $\left(1-\lambda\right)\left(A+\lambda\right) - Ab\left(A+2-\lambda\right) > 0$; hence, we again can carry the root to the right-hand side and take a square of both parts to obtain $Ab\left(A+2-\lambda\right) > \left(1-\lambda\right)\left(A+\lambda\right)$. However, in this case it does not hold, implying that for such values $\theta^{**} \geq \theta'$, and the minimum on $\left[0,\theta'\right]$ is reached at θ'.

On the one hand, this immediately gives the solution to the problem (15) if $Ab\left(A+2-\lambda\right) > \left(1-\lambda\right)\left(A+\lambda\right)$; on the other hand, it shows that the optimal cap does not start at $\theta < \theta'$ if $Ab\left(A+2-\lambda\right) \leq \left(1-\lambda\right)\left(A+\lambda\right)$, as then it would be better to have a cap at θ', and in that case we already know that the optimal cap starts at $\theta^* = 1 - \frac{2Ab}{A+\lambda}$. This completes the proof of Claim 16. ■

Proof of Claim 17. Suppose $A > \lambda$. Consider the point θ_0 that satisfies $m^*(\theta_0) = t^*(\theta_0) = 0$. By using the same reasoning as in the proof of Claim 16, we can show that whenever $\theta < \theta_0$ and $y^*(\theta) < y^*(\theta_0)$, then $y^*(\theta) = \min\{x(\theta), \theta + b\}$. Similarly, if $\theta > \theta_0$ and $y^*(\theta) > y^*(\theta_0)$, then $y^*(\theta) = \min\{z(\theta), \theta + b\}$. This already implies that the optimal contract $y^*(\theta)$ consists of several (not more than five) segments of straight lines.

Let

$$\theta_{xz} = \frac{1}{A+\lambda}(\lambda + A\lambda - Ab)$$

be the solution to the equation $x(\theta_{xz}) = z(\theta_{xz})$; then

$$x(\theta_{xz}) = z(\theta_{xz}) = \frac{\lambda(A + 2 - \lambda + b) - Ab}{A+\lambda},$$

and $x(\theta_{xz}) \geq \theta_{xz} + b$ if and only if $Ab \leq \lambda(1-\lambda)$. Hence, if $Ab \leq \lambda(1-\lambda)$, there must be a part of the optimal contract where $y^*(\theta) = \theta + b$, for otherwise there will have to be a jump, and we have proved that there is none at the optimum. It is easy to see that for such θ, $m^*(\theta) = t^*(\theta) = 0$, and hence it may be taken as θ_0. But now it is almost immediate that the optimal contract has no "flat" part, i.e., $y^*(\theta) = y^*(\theta_0)$ only if $\theta = \theta_0$, for otherwise there will have to be a jump, and we have proved that there is none at the optimum. This already implies that the optimal contract is as defined in the statement of the claim.

Let us now consider the case $Ab > \lambda(1-\lambda)$. This implies that $y^*(\theta) < \theta + b$ for all θ, and hence the contract may consist of at most three parts: where $y^*(\theta) = x(\theta)$, where $y^*(\theta) = y^*(\theta_0)$, and where $y^*(\theta) = z(\theta)$. Moreover, in this case $\theta_{xz} < \lambda$, as follows from

$$\lambda - \theta_{xz} = \lambda - \frac{\lambda + A\lambda - Ab}{A+\lambda} = \frac{Ab - \lambda(1-\lambda)}{A+\lambda}.$$

Let us now identify precisely the position of the flat part, if any.

Suppose that the segment where $y^*(\theta) = y^*(\theta_0)$ is $[\theta_1, \theta_2]$. We consider the following deviation: we keep the contract curve $y^*(\theta)$, but move the "reference" point θ_0 slightly to the left of θ_1 or to the right of θ_2. This naturally corresponds to uniformly increasing or decreasing the value of $m^*(\theta) - t^*(\theta)$. When doing so, we notice that the left derivative $\frac{dy^l(\theta_0)}{d\theta_0}|_{\theta_0=\theta_1} > 0$ and the right derivative $\frac{dy^r(\theta_0)}{d\theta_0}|_{\theta_0=\theta_2} > 0$ (these values are actually given by the slopes of $x(\theta)$ and $z(\theta)$, respectively). We must have that

$$-(y(\theta_1) - \theta_1 - b)^2 - 2(y(\theta_1) - \theta_1 - b)\frac{1-\lambda}{A+1}\theta_1 + (y(\theta_1) - \theta_1 - b)^2 - 2(y(\theta_1) - \theta_1 - b)\theta_1$$
$$+\lambda\left(2(y(\theta_1) - \theta_1 - b)\frac{1-\lambda}{A+1} + 2(y(\theta_1) - \theta_1 - b)\theta_1 + 2(y(\theta_1) - \theta_1 - b)(1-\theta_1)\right) \leq 0,$$

which simplifies to $\theta_1 \leq \lambda$. Similarly, we must have that

$$-\left(y\left(\theta_2\right)-\theta_2-b\right)^2-2\left(y\left(\theta_2\right)-\theta_2-b\right)\frac{1-\lambda}{A+1}\theta_2+\left(y\left(\theta_2\right)-\theta_2-b\right)^2-2\left(y\left(\theta_2\right)-\theta_2-b\right)\theta_2$$
$$+\lambda\left(2\left(y\left(\theta_2\right)-\theta_2-b\right)\frac{1-\lambda}{A+1}+2\left(y\left(\theta_2\right)-\theta_2-b\right)\theta_2+2\left(y\left(\theta_2\right)-\theta_2-b\right)\left(1-\theta_2\right)\right)\geq 0,$$

which simplifies to $\theta_2 \leq \lambda$. In either case, $\lambda \in [\theta_1, \theta_2]$, and we have identified at least one point that necessarily belongs to the "flat part", λ. In what follows, we can therefore take $\theta_0 = \lambda$.

To find the exact shape of optimal contract, it now suffices to compute $y(\theta_0) = y(\lambda)$. Let us vary $y = y(\theta_0)$; the derivative of the partial Lagrangian with respect to this value is given by

$$-2\left(y-\theta_0-b\right)\theta_0+\int_{\theta_1}^{\theta_0}\left(2\left(y-\theta-b\right)-2\theta\right)d\theta+\int_{\theta_1}^{\theta_2}2A\left(y-\theta\right)d\theta$$
$$+\lambda\left(2\left(y-\theta_0-b\right)+\int_{\theta_1}^{\theta_2}2\theta d\theta-\int_{\theta_0}^{\theta_2}2d\theta\right),$$

where θ_0 is a constant interior point of $[\theta_1, \theta_2]$ (such point exists except for the extreme case where $\theta_1 = \theta_2 = \theta_{xz}$; we also took into account that increasing $y(\theta_0)$ moves the segment $[\theta_1, \theta_2]$ to the right while decreasing moves it to the left). One can easily check that this expression does not actually depend on θ_0 as long as $\theta_0 \in [\theta_1, \theta_2]$, and is equal to

$$-2\left(y-\theta_2-b\right)\theta_2+\int_{\theta_1}^{\theta_2}\left(2\left(y-\theta-b\right)-2\theta\right)d\theta+\int_{\theta_1}^{\theta_2}2A\left(y-\theta\right)d\theta+\lambda\left(2\left(y-\theta_2-b\right)+\int_{\theta_1}^{\theta_2}2\theta d\theta\right),$$

which equals, after division by $\theta_2 - \theta_1$, to

$$2Ay-\left(\theta_1+\theta_2\right)\left(A-\lambda\right)-2\frac{\lambda-\theta_1}{\theta_2-\theta_1}\left(b-y+\theta_1\right)-2\lambda.$$

This is greater than

$$2Ay-\left(\theta_1+\theta_2\right)\left(A-\lambda\right)-2\lambda,$$

except when $\lambda = \theta_1$. But $y > x(\theta_{xz}) = \frac{\lambda(A+2-\lambda+b)-Ab}{A+\lambda}$, $\theta_1 = \frac{y(A+1)-b}{A-\lambda+2}$ and $\theta_2 = \min\left\{\frac{Ay-\lambda}{A-\lambda}, 1\right\} \leq \frac{Ay-\lambda}{A-\lambda}$, and for these values, the expression becomes zero. Consequently, it is non-positive for y, θ_1, θ_2, and therefore the partial Lagrangian is decreasing with respect to $y(\theta_0)$. Consequently, this value should be picked as high as possible, which implies $\theta_1 = \lambda$, $y(\lambda) = x(\lambda) = \frac{b+2\lambda+A\lambda-\lambda^2}{A+1}$, and $\theta_2 = \min\left\{\frac{-\lambda+A\lambda-A\lambda^2+A^2\lambda+Ab}{(A+1)(A-\lambda)}, 1\right\}$. Then $\theta_2 < 1$ if and only if $A - b - 2\lambda - A\lambda + \lambda^2 + 1$, i.e., if $(1-\lambda)(A+1-\lambda) > b$. From this we immediately obtain the formulas for $y^*(\theta)$. This finishes the proof. ■

8 References

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.

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

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

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

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," 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.

Garicano, L. and E. Rossi-Hansberg (2006): "Organization and inequality in a knowledge economy," Quarterly Journal of Economics, 121, 1383-1435.

Goltsman, M., J Hörner, G. Pavlov, and F. Squintani (2007): "Mediated cheap talk," Mimeo, Boston University.

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).

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

Kováč, E. and T. Mylovanov (2007): "Stochastic mechanisms in settings without monetary transfers," Journal of Economic Theory, forthcoming.

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," Mimeo Penn State University and UC Berkeley.

Leonard, B. (2000): “Worker productivity might be tangled in corporate bureacracy,” HR Magazine, November 2000.

Martimort, D. (1997): “A theory of bureaucratization based on reciprocity and collusive behavior,” Scandinavian Journal of Economics, 99, 555-579.

McCubbins, M. D., R. G. Noll, and B. R. 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.

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

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

Walsh, J. P., and R. D. Dewar (1987): “Formalization and the organizational life cycle,” Journal of Management Studies,24, 215-231.

Wilson, J. Q. (1989): “Bureaucracy: what government agencies do and why they do it,” Basic Books, NY.