Money Burning in the Theory of Delegation*

Manuel Amador
Federal Reserve Bank of Minneapolis
University of Minnesota

Kyle Bagwell
Stanford University

November 15, 2016

Abstract

This paper uses a Lagrangian approach to provide general sufficient conditions under which money burning expenditures are used in an optimal delegation contract. To better understand the role of money burning, we also establish simple sufficient conditions for the optimality of interval delegation in the form of a cap allocation under a restricted set of preferences for a benchmark setting in which money burning is not allowed. We also apply our findings to a model of cooperation and delegation and to a model with quadratic preferences and general distribution functions.

1 Introduction

In the standard delegation model, a principal engages with a privately informed but biased agent, and contingent transfers are infeasible. The principal chooses a set of permissible actions, knowing that the agent will privately observe the state of the world and then pick his preferred action from this set. In choosing the optimal set, the principal balances the benefit of expanding the agent's choice set, so as to give the agent the flexibility to respond to the state of nature, against the benefit of restricting the agent's choice set, so as to limit the expression of the agent's bias.

In this paper, we extend the standard delegation model to allow for a two-dimensional action set. In particular, the principal may specify that a given action is permissible only when the agent also incurs a certain amount of wasteful or "money burning" expenditure. Put differently, we extend the standard delegation model to allow for contingent money burning. Under the assumption that the money burning expenditure lowers the welfare of

*Some of the results in this paper were originally circulated in an earlier version of Amador and Bagwell (2013). We thank Alex Frankel for very helpful comments. Manuel Amador is grateful for financial support from the NSF under grant number 0952816.

the principal and agent in a symmetric fashion, we provide general conditions under which money burning expenditures are used in the optimal delegation contract.

To establish our findings, we follow Amador and Bagwell (2013) and consider a general representation of the delegation problem. The agent's action is taken from an interval on the real line, the state has a continuous distribution over a bounded interval on the real line, and the preferences of the principal and agent are represented in a general fashion. In particular, unlike most of the previous delegation literature, we do not restrict the preferences of the principal and agent to take quadratic forms. The principal's welfare function is continuous in the action and the state of nature, and is twice differentiable and concave in the action. The agent's welfare function is twice differentiable and strictly concave in the action, and the state enters the agent's welfare function in a standard multiplicative fashion. We assume that the agent's preferred (or "flexible") action is interior and strictly increasing in the state.

The possibility of money burning can be interpreted in a variety of ways. For example, we can imagine a setting in which certain "exceptional" actions are permitted only when wasteful administrative costs are incurred. Ambrus and Egorov (2015) provide extensive discussion of this interpretation. Additionally, if preferences are quadratic, stochastic allocations can be interpreted as money burning. Goltsman, Hörner, Pavlov and Squintani (2009) and Kováč and Mylovanov (2009) provide related work. Also, in repeated games with privately observed and iid shocks, symmetric punishments can be understood as money burning. For related themes, see Athey, Bagwell and Sanchirico (2004), Athey, Atkeson and Kehoe (2005) and Bagwell and Lee (2012).[1] Finally, as Amador, Werning and Angeletos (2006) show, in a consumption-savings problem, money burning can be interpreted as arising when a consumption-savings bundle is selected that lies in the interior of the consumer's budget set.

To motivate our assumption that money burning symmetrically lowers the welfare of the principal and agent, we consider two economic examples. First, and following Amador and Bagwell (2013), we may consider a setting in which two players select a menu of permissible actions and associated money burning levels, with the objective of maximizing the ex ante sum of their welfares and with the understanding that one of the two players will later observe the state and make the selection from the menu that maximizes that player's welfare. In this example, money burning directly lowers the welfare of the "agent" (i.e., the player that makes the selection from the menu) and thereby lowers the welfare of the "principal" (i.e., the ex ante sum of their welfares) in a symmetric fashion. Second, and following Ambrus and Egorov (2015), we may consider a setting with a principal and agent, in which an ex-ante participation constraint arises since the agent has an outside option. If the

[1] For related static settings with multiple agents, see McAfee and McMillan (1992), Condorelli (2012) and Chakravarty and Kaplan (2013).

principal and agent can achieve non-contingent transfers at the ex ante stage, then the binding participation constraint implies that the principal must use an ex ante transfer to compensate the agent for expected money burning expenses. As a consequence, money burning again lowers the welfare of the principal and agent in a symmetric fashion.

We provide two general propositions that characterize conditions under which money burning expenditures are used in the optimal delegation contract. Proposition 1 offers a set of conditions under which optimal delegation requires that (i) for states below a critical level, the agent has full flexibility and selects the preferred action while incurring no money burning expenses; and (ii) for states above the critical level, higher states induce higher actions that are paired with higher levels of money burning. Proposition 2 provides a different set of conditions under which optimal delegation requires that (i) for states below a critical level, the agent pools at a common action with no money burning; and (ii) for states above the critical level, higher states induce higher actions that are paired with higher levels of money burning. In both cases, for higher states, the agent selects an action that is below the preferred level and incurs money burning expenses. Intuitively, for higher states, an agent resists deviating to a higher action, closer to the agent's preferred level, due to the higher money burning expense that would then be required.

To develop these results, we build on work by Amador et al. (2006) and Amador and Bagwell (2013) and use a general Lagrangian approach. This approach enables the analyst to develop general sufficient conditions for the optimality of a conjectured solution. Motivated by the work of Ambrus and Egorov (2015), we conjecture solutions with the features described in the previous paragraph. In a set up with quadratic utility and a uniform distribution for states of nature, Ambrus and Egorov (2015) provide an analytical characterization of optimal delegation contracts that feature money burning for certain parameters ranges. To our knowledge, Ambrus and Egorov (2015) thus provide a first explicit characterization in a standard quadratic-uniform model of optimal delegation contracts that feature actual money burning.[2] Using our Lagrangian approach, we provide a more general set of sufficient conditions for money burning. In particular, the sufficient conditions that we provide include but extend beyond the conditions provided by Ambrus and Egorov (2015) for the quadratic-uniform model.

To better understand the role of money burning, we also consider optimal delegation when money burning is not allowed. Our analysis of this setting builds on Amador and Bagwell (2013), who establish, among other results, sufficient conditions for a "cap allocation" (i.e.,

[2] Within the context of a consumption-savings problem, Amador et al. (2006) numerically obtained solutions to the delegation problem where money burning is used when the distribution of states of nature is discrete. See also the related work of Ambrus and Egorov (forthcoming).

a delegation set in which the set of permissible actions is defined by a maximal action and in which money burning (if feasible) is never required) to be optimal. For a restricted family of preferences, we establish here easier-to-check sufficient conditions for the optimality of a cap allocation in the setting where money burning in infeasible. This result implies that, under a slight strengthening of the assumption used to support Proposition 1, optimal delegation would take the form of a cap were money burning not allowed. Based on this comparison, we conclude that the use of money burning is attractive in this case, because it provides an incentive-compatible means for the agent that observes higher states to select actions that rise somewhat with the underlying state.

Our propositions provide general sufficient conditions for money burning and may be easily employed for applications. To illustrate the value of our general propositions, we explore in detail two specific settings.

First, we analyze a model of cooperation and delegation in which two players select a menu of permissible actions and associated money burning levels. As in the first economic example mentioned above, the players pick the menu with the objective of maximizing the ex ante sum of their welfares and with the understanding that one of the players, the agent, will later observe the state and make the selection from the menu that maximizes the agent's welfare. The specific cooperation model that we analyze allows for general (non-quadratic) preferences and distributions, but a key ingredient of the model is that the two players have welfare functions that take related, but distinct, forms. In particular, the players equally share the benefits generated by an action for any given state; however, the agent incurs a lower marginal cost and is thus biased toward actions that are too high from the other player's perspective. We find that the optimal menu uses money burning if the density is non-increasing and the cost asymmetry is not too large.

Second, and in line with the second economic example mentioned above, we follow Ambrus and Egorov (2015) and consider quadratic preferences while also extending their analysis to allow for general distribution functions. After showing that our two propositions capture Ambrus and Egorov's (2015) corresponding results on money burning when the distribution function is uniform, we generalize their results to allow for general families of distribution functions.[3] Our work here illustrates how our propositions can be easily applied to obtain new results. In addition, since this example fits within the restricted family of preferences mentioned above, we may easily characterize conditions under which a cap allocation would be optimal were money burning not allowed. We illustrate this point for

[3]At the same time, Ambrus and Egorov (2015) is more general in other respects. For example, they study the effects of a bound on the ex-ante transfer between the principal and the agent, and they also consider an extension that allows for limited state contingent transfers.

the uniform-distribution setting. For this setting, we identify conditions under which the optimal delegation contract when money burning is allowed requires for higher states that the agent engages in money burning and selects actions above the cap that would prevail were money burning not allowed. Whether or not money burning is allowed, the agent has full flexibility when lower states are realized; however, the agent enjoys such flexibility under a smaller subset of states when money burning is allowed.

Our work relates to a rich literature on delegation theory. Most of this work develops conditions under which "interval delegation" is optimal, where interval delegation is defined by a delegation set in which the set of permissible actions takes the form of a single interval and money burning (if feasible) is never required.[4] Holmstrom (1977) first defined and analyzed the delegation problem. For a setting without the possibility of money burning, he provides sufficient conditions for the existence of an optimal solution to the delegation problem. He also establishes for a specific example with quadratic preferences that interval delegation is optimal. Alonso and Matouschek (2008) provide necessary and sufficient conditions for the optimality of interval delegation when preferences are quadratic and money burning is infeasible.[5] Analyzing a consumption-savings problem, Amador et al. (2006) provide conditions under which interval delegation (in the form of a minimum savings rule) is optimal when money burning is feasible. In the quadratic-uniform model, Ambrus and Egorov (2015) identify a range of parameters under which interval delegation is optimal when money burning is feasible. Finally, Amador and Bagwell (2013) consider a general representation of the delegation problem that includes as special cases previous findings regarding the optimality of interval delegation. Using Lagrangian techniques, they establish necessary and sufficient conditions for the optimality of interval delegation, both when money burning is infeasible and when money burning is feasible.[6]

In comparison to the interval delegation literature, an important distinction of the present

[4] A cap allocation is thus a special form of interval delegation, in which the upper bound of the interval is defined by the cap and the lower bound is defined by the preferred action of the lowest type.

[5] In addition, Alonso and Matouschek (2008) provide comparative-statics results and characterize the value of delegation. They also obtain a characterization when interval delegation is not optimal.

[6] For other work on the theory of delegation, see, for example, Armstrong and Vickers (2010), Martimort and Semenov (2006), Melumad and Shibano (1991), and Mylovanov (2008). For models with quadratic preferences, Koessler and Martimort (2012) and Frankel (2012) consider optimal delegation with multi-dimensional actions; however, they do not consider money burning, and their results differ significantly from those found here. See also Frankel (2014) for analysis of optimal max-min mechanisms in a multiple-decision delegation problem in which the principal is uncertain about the agent's payoffs. Finally, some recent work also uses Lagrangian methods related to those used by Amador et al. (2006) and Amador and Bagwell (2013) to study optimal delegation in other settings. Amador and Bagwell (2014) consider a regulatory setting with a participation constraint but without the possibility of money burning, Burkett (2016) examines two regimes of which one is a delegation setting with a participation constraint in which money burning is allowed but not optimal, and Guo (forthcoming) and Krähmer and Kováč (2016) study optimal delegation in dynamic settings.

paper is that we assume that money burning is feasible and then provide general conditions under which money burning expenditures are actually used in the optimal delegation contract. We also provide simple sufficient conditions for the optimality of a cap allocation for a restricted family of preferences when money burning is not allowed. We illustrate the value of our findings in two specific settings and expect that they will be useful for future applications as well.

The paper is organized as follows. The basic model is presented in Section 2. In Section 3, we analyze optimal delegation with money burning; specifically, we develop our general approach and state and prove our two general propositions. Next, in Section 4, we consider the setting without money burning and establish simple sufficient conditions for the optimality of a cap allocation. In Section 5, we develop our specific model of cooperation and delegation, and we apply our general propositions to develop findings for this application. We focus in Section 6 on the case of quadratic welfare functions and general distribution functions. Here we illustrate the value of our general propositions and generalize the findings of Ambrus and Egorov (2015) on money burning in optimal delegation contracts. We also use our findings to compare optimal delegation when money burning is feasible and when it is not. Section 7 concludes. Remaining proofs and findings are found in the Appendix.

2 Basic Set Up

We consider a delegation problem with one principal and one agent. Letting the values of π and γ respectively represent an action (or allocation) and a state (or shock), we denote the principal's welfare function as $w(\gamma, \pi) - t$ and the agent's welfare function as $\gamma\pi + b(\pi) - t$, where $t \geq 0$ is an additional action that reduces everyone's welfare. By including the action t, we extend the standard delegation problem to allow for the possibility of money burning.

We assume that γ is privately observed by the agent and has a continuous distribution $F(\gamma)$ with bounded support $\Gamma = [\underline{\gamma}, \overline{\gamma}]$ and with an associated continuous density $f(\gamma) > 0$. We assume further that $\pi \in \Pi$, where Π is an interval of the real line. Without loss of generality, for $\overline{\pi}$ in the extended reals, we may define $\Pi = [0, \overline{\pi}]$ if $\overline{\pi} < \infty$ or $\Pi = [0, \overline{\pi})$ if $\overline{\pi} = \infty$. Henceforth, and following Amador and Bagwell (2013), we also maintain the following basic conditions:[7]

Assumption 1. *The following holds:* (i) $w : \Gamma \times \Pi \to \mathbb{R}$ *is continuous on* $\Gamma \times \Pi$*;* (ii) $w(\gamma, \pi)$ *is twice differentiable and concave in* π*;* (iii) $b : \Pi \to \mathbb{R}$ *is twice differentiable and strictly*

[7]Note that Assumption 1 is slightly different from Assumption 1 in Amador and Bagwell (2013), because, differently from that paper, we do not need to incorporate cases where the payoff functions have infinite derivatives at the boundaries of Π. Similar modifications can be undertaken in the present paper.

concave in π*;* (iv) $\gamma\pi + b(\pi)$ *has a unique interior maximum over* $\pi \in \Pi$ *for all* $\gamma \in \Gamma$*, denoted by* $\pi_f(\gamma)$ *with* π_f *twice differentiable and* $\pi_f'(\gamma) > 0$*;* (v) $w_\pi(\gamma, \pi)$ *is continuous on* $\Gamma \times \Pi$*; and* (vi) *if* $\bar{\pi} < \infty$ *then* $b'(\bar{\pi})$ *and* $w_\pi(\gamma, \bar{\pi})$ *are finite for all* γ*.*

Conditions (i)-(v) in Assumption 1 are standard. Condition (vi) simplifies the analysis in some places but is straightforward to relax.

Notice that $\pi_f(\gamma)$ denotes the agent's preferred or flexible action, given the state γ. We may thus understand that $w_\pi(\gamma, \pi_f(\gamma))$ indicates the direction of the *bias* of the agent. For example, if $w_\pi(\gamma, \pi_f(\gamma)) > 0$, then the principal would prefer a higher action than the one most preferred by the agent of type γ.

The problem is to choose an action allocation, $\pi : \Gamma \to \Pi$, and an amount of money burned, $t : \Gamma \to \mathbb{R}$, so as to maximize the principal's expected welfare:

$$\begin{aligned} \max \int_\Gamma \Big(w(\gamma, \pi(\gamma)) - t(\gamma) \Big) dF(\gamma) & \quad \text{subject to:} \\ \gamma \in \arg\max_{\tilde{\gamma} \in \Gamma} \Big\{ \gamma\pi(\tilde{\gamma}) + b(\pi(\tilde{\gamma})) - t(\tilde{\gamma}) \Big\}, & \quad \text{for all } \gamma \in \Gamma \\ t(\gamma) \geq 0\,; \ \forall \gamma & \end{aligned} \tag{P}$$

where the first constraint arises from γ being private information of the agent and the second constraint ensures that money can be burned but not printed.[8]

To motivate our analysis, we describe two economic examples that fit into our framework.

Example 1: Cooperation and delegation. Following Amador and Bagwell (2013), we may consider a setting in which two players choose a menu of permissible actions and associated money burning levels with the goal of maximizing the ex ante sum of their welfares and with the understanding that one of the two players will later observe the state and make a selection from the menu. In this case, we may define the welfare functions of the two players as $\gamma\pi(\gamma) + b(\pi(\gamma)) - t(\gamma)$ and $v(\gamma, \pi(\gamma))$, respectively, where the former welfare function is associated with the player that later observes the state. The sum of players' welfare functions then takes the form $w(\gamma, \pi(\gamma)) - t(\gamma) \equiv \gamma\pi(\gamma) + b(\pi(\gamma)) + v(\gamma, \pi(\gamma)) - t(\gamma)$. Here, money burning enters the objective function and the constraint in a symmetric way, since in each instance it is associated with a loss in one of the player's welfare.[9] We provide further discussion of this example in Section 5.

Example 2: Participation constraint and ex ante transfer. Following Ambrus and Egorov (2015), we may consider a setting with a principal and an agent, where the agent privately

[8] All integrals used in the paper are Lebesgue integrals.

[9] Amador and Bagwell (2013) use this setting to interpret the design of trade agreements.

observes the state and makes a selection from the menu selected by the principal. The agent has an outside option yielding zero utility, which results in an ex ante participation constraint. At the ex ante stage, the principal and agent can also achieve non-contingent transfers, where $p \in \Re$ denotes a transfer or payment from the principal to the agent. The agent's ex ante participation constaint then takes the following form: $\int_{\Gamma}(\gamma\pi(\gamma)+b(\pi(\gamma))-t(\gamma))dF(\gamma)+p \geq 0$.[10] The principal's objective is to maximize her expected welfare less the ex ante payment to the agent: $\int_{\Gamma} v(\gamma,\pi(\gamma))dF(\gamma)-p$. Since the participation constraint will bind, we may solve for p and substitute into the objective, so that the principal's objective function becomes $\int_{\Gamma}(v(\gamma,\pi(\gamma))+\gamma\pi(\gamma)+b(\pi(\gamma))-t(\gamma))dF(\gamma)$. This setting fits into our framework as well once we define $w(\gamma,\pi(\gamma)) \equiv \gamma\pi(\gamma)+b(\pi(\gamma))+v(\gamma,\pi(\gamma))$. Here, money burning appears symmetrically in the objective function and the constraint, since the participation constraint ensures that the principal uses an ex ante transfer to compensate the agent for expected money burning expenses. We discuss this example further in Section 6.

3 Optimal Delegation with Money Burning

Our goal in this section is to provide sufficient conditions for the optimality of money burning. Following Amador et al. (2006) and Amador and Bagwell (2013), we begin by re-stating Problem P and describing our general Lagrangian approach. Next, motivated by the work of Ambrus and Egorov (2015), we conjecture the existence of particular allocations that feature money burning and provide sufficient conditions for these allocations to be optimal. Ambrus and Egorov (2015) provide an analytical characterization of optimal allocations that feature strict money burning within a set up with quadratic utility and a uniform distribution for shocks. Using our Lagrangian approach, we are able to establish a more general set of sufficient conditions.

3.1 General Approach

We begin by re-stating Problem P. By expressing the incentive constraints (i.e., the first constraint in Problem P) in their standard integral form plus a monotonicity restriction, we

[10]The zero utility outside option is without loss of generality as long as the ex-ante transfer, p, is unrestricted. See Ambrus and Egorov (2015) for analysis of a case where p is restricted.

may re-write Problem P as:[11]

$$\max_{\left\{\substack{\pi:\Gamma\to\Pi,\\ t:\Gamma\to\mathbb{R}}\right\}} \int (w(\gamma,\pi(\gamma)) - t(\gamma))dF(\gamma) \quad \text{subject to:}$$

$$\gamma\pi(\gamma) + b(\pi(\gamma)) - t(\gamma) = \int_{\underline{\gamma}}^{\gamma} \pi(\tilde{\gamma})d\tilde{\gamma} + \underline{U}, \text{ for all } \gamma \in \Gamma$$

$$\pi \text{ non-decreasing, and } t(\gamma) \geq 0, \text{ for all } \gamma \in \Gamma$$

where $\underline{U} \equiv \underline{\gamma}\pi(\underline{\gamma}) + b(\pi(\underline{\gamma})) - t(\underline{\gamma})$.

We next solve the integral equation for $t(\gamma)$ and substitute into both the objective and the non-negativity constraint. We thereby generate the following equivalent problem:

$$\max_{\left\{\substack{\pi:\Gamma\to\Pi,\\ t(\underline{\gamma})\geq 0}\right\}} \int \Big(v(\gamma,\pi(\gamma))f(\gamma) + (1-F(\gamma))\pi(\gamma)\Big)d\gamma + \underline{U} \quad \text{subject to:} \tag{P'}$$

$$\gamma\pi(\gamma) + b(\pi(\gamma)) - \int_{\underline{\gamma}}^{\gamma} \pi(\tilde{\gamma})d\tilde{\gamma} - \underline{U} \geq 0; \quad \text{for all } \gamma \in \Gamma \tag{1}$$

$$\pi \text{ non-decreasing} \tag{2}$$

where v is defined such that $v(\gamma,\pi(\gamma)) \equiv w(\gamma,\pi(\gamma)) - b(\pi(\gamma)) - \gamma\pi(\gamma)$.

Note that Problem P' requires solving for $\pi(\gamma)$ and $t(\underline{\gamma})$. With these solutions in place, we can recover the entire function $t(\gamma)$ using the integral constraint as follows:

$$t(\gamma) = \gamma\pi(\gamma) + b(\pi(\gamma)) - \int_{\underline{\gamma}}^{\gamma} \pi(\tilde{\gamma})d\tilde{\gamma} - \underline{U} \tag{3}$$

To solve Problem P', we follow and extend the Lagrangian approach used by Amador et al. (2006) and Amador and Bagwell (2013). We thus first embed the monotonicity constraint (2) into the choice set of $\pi(\gamma)$. Letting the choice set for π be defined as $\Phi \equiv \{\pi | \pi : \Gamma \to \Pi;$ and π non-decreasing$\}$, the problem is then to choose a function $\pi \in \Phi$ and a number $t(\underline{\gamma}) \geq 0$ so as to maximize the objective of Problem P' subject to (1).

By assigning a cumulative multiplier function Λ to constraint (1), we can write the Lagrangian for the problem as follows:

[11]See Milgrom and Segal (2002).

$$\begin{aligned}\mathcal{L}(\pi, t(\underline{\gamma})|\Lambda) \equiv & \int_{\gamma\in\Gamma}\Big(v(\gamma,\pi(\gamma))f(\gamma)+(1-F(\gamma))\pi(\gamma)\Big)d\gamma+\underline{U} \\ & -\int_{\gamma\in\Gamma}\left(\int_{\underline{\gamma}}^{\gamma}\pi(\gamma')d\gamma'+\underline{U}-\gamma\pi(\gamma)-b(\pi(\gamma))\right)d\Lambda(\gamma).\end{aligned} \tag{4}$$

The Lagrange multiplier Λ is restricted to be a non-decreasing function.

To develop our sufficiency results, we rely on Theorem 1 stated in Appendix A, which is a slightly modified version of the sufficiency Theorem 1 of Section 8.4 in Luenberger (1969, p. 220). Our approach is to identify a candidate solution and construct an associated non-decreasing Lagrange multiplier function such that complementary slackness is satisfied and the candidate solution maximizes the resulting Lagrangian. To verify that the candidate solution maximizes the resulting Lagrangian, we utilize first order conditions as developed in Amador et al. (2006). In turn, to apply first order conditions, we must ensure that our constructed Lagrange multiplier Λ is also such that the resulting Lagrangian is concave in π when evaluated at this specific multiplier.[12] Amador and Bagwell (2013) follow a similar approach; however, differently from that paper, we develop here propositions under which the solution involves actual money burning.

3.2 Money Burning: Case 1

We now look for an allocation with the following properties: from $\underline{\gamma}$ to some γ_x, the allocation provides full flexibility where no money is burned; and from γ_x to $\overline{\gamma}$, the allocation burns money but still provides some flexibility. The next assumption imposes conditions that turn out to be sufficient for such an allocation to be optimal.

Assumption 2. *The following holds:* (i) $v(\gamma,\pi)$ *is concave in* π *for all* $\gamma\in\Gamma$*;* (ii) $v_\pi(\underline{\gamma},\pi_f(\underline{\gamma})) \leq 0$*;* (iii) *there exists a continuous function* $x(\gamma)$ *and a value for* $\gamma_x\in\Gamma$ *such that:* (a) $v_\pi(\gamma,x(\gamma)) f(\gamma)=F(\gamma)-1$ *for all* $\gamma\geq\gamma_x$*;* (b) $x(\gamma)\leq\pi_f(\gamma)$ *for all* $\gamma>\gamma_x$ *and with equality at* γ_x*;* (c) $x(\gamma)$ *is non-decreasing for all* $\gamma\geq\gamma_x$*;* (d) $F(\gamma)-v_\pi(\gamma,\pi_f(\gamma))f(\gamma)$ *is non-decreasing for all* $\gamma<\gamma_x$*.*

Our proposed allocation in this case is:

$$\pi_x(\gamma)=\begin{cases}\pi_f(\gamma) & ;\ \text{for } \gamma\in[\underline{\gamma},\gamma_x] \\ x(\gamma) & ;\ \text{for } \gamma\in(\gamma_x,\overline{\gamma}]\end{cases}$$

[12] Observe that the Lagrangian is linear in $t(\underline{\gamma})$.

with $t_x(\underline{\gamma}) = 0$ and where $x(\gamma)$ and γ_x are defined by Assumption 2.

Let us now briefly discuss the conditions in Assumption 2. Assumption 2i guarantees that problem (P') is a convex problem, and hence that the Lagrangian is concave. We note that the function v is concave in π in the economic examples mentioned above if the uninformed player has concave preferences (in the first example) or if the principal has concave preferences (in the second example). Assumption 2ii guarantees that the bias at the bottom of the distribution is positive, which must hold if the principal does not want to pool allocations at the lowest states. Assumption 2iii describes the conditions that point γ_x and the allocation in the money burning region, $x(\gamma)$, must satisfy, where $x(\gamma)$ is defined by $v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1$ and γ_x is defined by $x(\gamma_x) = \pi_f(\gamma_x)$. We note that the function x is defined to maximize point-wise the integrand in Problem (P'), with part (a) providing the corresponding first order condition. For our purposes, we need only ensure that this function is defined for $\gamma \geq \gamma_x$. In the proof below, part (d) ensures that the constructed multiplier function is non-decreasing and the Lagrangian is concave for $\gamma < \gamma_x$; part (c) guarantees monotonicity of the proposed allocation; part (b) ensures that the allocation is continuous and burns non-negative amounts of money for all types above γ_x; and part (a) guarantees that the proposed allocation satisfies the first order conditions of the Lagrangian.[13]

The following is a sufficiency result:

Proposition 1. *The proposed allocation $(\pi_x(\gamma), t_x(\underline{\gamma}))$ solves Problem P' if Assumption 2 holds.*

Proof. We begin by taking the expression for the Lagrangian in (4) and integrating by parts.[14] Setting $\Lambda(\overline{\gamma}) = 1$ without loss of generality, we get:

$$\begin{aligned}\mathcal{L}(\pi, t(\underline{\gamma})|\Lambda) \equiv & \int_{\gamma\in\Gamma} \Big(v(\gamma, \pi(\gamma))f(\gamma) + (\Lambda(\gamma) - F(\gamma))\pi(\gamma)\Big)d\gamma \\ & + \int_{\gamma\in\Gamma} (\gamma\pi(\gamma) + b(\pi(\gamma)))\, d\Lambda(\gamma) + \Lambda(\underline{\gamma})\underline{U}.\end{aligned} \quad (5)$$

[13]Note that part (c) of Assumption 2iii is sure to hold if, for all $\gamma \in \Gamma$ and $\pi \in \Pi$, $v_{\pi\pi}(\gamma, \pi) < 0$ and $v_\pi(\gamma, \pi) + (1 - F(\gamma))/f(\gamma)$ is non-decreasing. For distributions with a non-decreasing hazard rate, the latter condition requires $v_{\pi\gamma}(\gamma, \pi) \geq 0$. Similar observations apply for Assumption 3 below.

[14]Observe that $h(\gamma) \equiv \int_{\underline{\gamma}}^{\gamma} \pi(\gamma')d\gamma'$ exists (since π is bounded and measurable by monotonicity) and is absolutely continuous; and $\Lambda(\gamma)$ is a function of bounded variation (as it is non-decreasing over $[\underline{\gamma}, \overline{\gamma}]$). We conclude that $\int_{\underline{\gamma}}^{\overline{\gamma}} h(\gamma)d\Lambda(\gamma)$ exists as the Riemman-Stieltjes integral and that integration by parts can be done as follows: $\int_{\underline{\gamma}}^{\overline{\gamma}} h(\gamma)d\Lambda(\gamma) = h(\overline{\gamma})\Lambda(\overline{\gamma}) - h(\underline{\gamma})\Lambda(\underline{\gamma}) - \int_{\underline{\gamma}}^{\overline{\gamma}} \Lambda(\gamma)dh(\gamma)$. Finally, since $h(\gamma)$ is absolutely continuous, we can use $\pi(\gamma)d\gamma$ in place of $dh(\gamma)$.

A proposed multiplier. Our Lagrange multiplier for this case takes the following form:[15]

$$\Lambda(\gamma) = \begin{cases} 0 & ; for\ \gamma = \underline{\gamma} \\ F(\gamma) - v_\pi(\gamma, \pi_f(\gamma))f(\gamma) & ; for\ \gamma \in (\underline{\gamma}, \gamma_x] \\ 1 & ; for\ \gamma \in (\gamma_x, \overline{\gamma}]. \end{cases}$$

Monotonicity of the Lagrange multiplier. Using Assumption 2iii parts (a), (d) and the equality condition of part (b), we have that Λ is non-decreasing for all $\gamma \in (\underline{\gamma}, \overline{\gamma})$. The Lagrange multiplier has a jump at $\underline{\gamma}$ which equals $-v_\pi(\underline{\gamma}, \pi_f(\underline{\gamma}))f(\underline{\gamma})$ which is non-negative by Assumption 2ii.

Concavity of the Lagrangian. Given our proposed multiplier, it follows that $\underline{U}$ drops from equation (5). Concavity of the Lagrangian in π follows from Λ non-decreasing and Assumption 2i which guarantees concavity of v in π.

Maximizing the Lagrangian. We now proceed to show that the proposed allocation $(\pi_x, t_x(\underline{\gamma}))$ maximizes the Lagrangian given our proposed multiplier function. We first note that $t_x(\underline{\gamma})$ does not appear in the resulting Lagrangian. We can thus restrict attention to maximizing the Lagrangian over $\pi(\gamma) \in \Phi \equiv \{\pi | \pi : \Gamma \to \Pi; \text{ and } \pi \text{ non-decreasing}\}$. For this, we may use the sufficiency part of Lemma A.2 in Amador et al. (2006) which concerns the maximization of concave functionals in a convex cone.

Given that $\mathcal{L}$ is a concave functional on π, if $\overline{\pi} = \infty$, then Φ is a convex cone, and so we may appeal to this lemma directly. We can then say that if

$$\partial \mathcal{L}(\pi_x; \pi_x | \Lambda) = 0, \tag{6}$$

$$\partial \mathcal{L}(\pi_x; y | \Lambda) \leq 0, \text{ for all } y \in \Phi, \tag{7}$$

then π_x maximizes the Lagrangian, where we have suppressed notation for the money burning variable $t_x(\underline{\gamma})$ and where the first order condition is in terms of Gateaux differentials.[16] If instead $\overline{\pi} < \infty$, then a bit more care is needed, but after invoking Assumption 1 we obtain

[15] Alternatively, we could have set $\Lambda(\underline{\gamma}) = -v_\pi(\underline{\gamma}, \pi_f(\underline{\gamma}))f(\underline{\gamma})$, and the proof would have proceeded as in Proposition 2 below.

[16] Given a function $T : \Omega \to Y$, where $\Omega \subset X$ and X and Y are normed spaces, if for $x \in \Omega$ and $h \in X$ the limit

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha} [T(x + \alpha h) - T(x)]$$

exists, then it is called the Gateaux differential at x with direction h and is denoted by $\partial T(x; h)$.

the same result.[17]

Taking the Gateaux differential of the Lagrangian in the direction $y \in \Phi$ and using that $b'(\pi_f(\gamma)) = -\gamma$, we get that:[18]

$$\partial \mathcal{L}(\pi_x; y|\Lambda) = \int_{\gamma_x}^{\overline{\gamma}} \Big(v_\pi(\gamma, x(\gamma)) f(\gamma) + (1 - F(\gamma)) \Big) y(\gamma) d\gamma.$$

By Assumption 2iii part (a) it follows then that $\partial \mathcal{L}(\pi_x; y|\Lambda) = 0$ for all y, and thus the first order conditions for π_x to maximize the Lagrangian are satisfied.

Money burning is non-negative. We now check that $t(\gamma) \geq 0$ for all γ. Recall that

$$t(\gamma) = \gamma \pi_x(\gamma) + b(\pi_x(\gamma)) - \int_{\underline{\gamma}}^{\gamma} \pi_x(\tilde{\gamma}) d\tilde{\gamma} - \underline{U}$$

for all γ. Given that $t(\gamma) = 0$ for $\gamma \in [\underline{\gamma}, \gamma_x]$, as the flexible allocation is offered in that region, we have that

$$t(\gamma) = \begin{cases} 0 & ; \text{ for } \gamma \in [\underline{\gamma}, \gamma_x] \\ \gamma x(\gamma) - \gamma_x x(\gamma_x) + b(x(\gamma)) - b(x(\gamma_x)) - \int_{\gamma_x}^{\gamma} x(\widetilde{\gamma}) d\widetilde{\gamma} & ; \text{ for } \gamma \in (\gamma_x, \overline{\gamma}] \end{cases}$$

Integrating by parts the last integrand, we get that:

$$t(\gamma) = b(x(\gamma)) - b(x(\gamma_x)) + \int_{\gamma_x}^{\gamma} \widetilde{\gamma} dx(\widetilde{\gamma})$$

[17] In this case, we may follow closely the approach used in Amador and Bagwell (2013) to extend b and w to the entire positive ray of the real line in the following way:

$$\hat{w}(\gamma, \pi) = \begin{cases} w(\gamma, \pi) & ; \text{ for } \pi \in \Pi \\ w(\gamma, \overline{\pi}) + w_\pi(\gamma, \overline{\pi})(\pi - \overline{\pi}) & ; \text{ for } \pi > \overline{\pi} \end{cases}$$

for all $\gamma \in \Gamma$ and $\pi \in [0, \infty)$. And we similarly define $\hat{b}$. These extensions are possible from the boundedness of the derivatives as stated in part (vi) of Assumption 1. Then we let $\hat{\Phi} = \{\pi | \pi : \Gamma \to \mathbb{R}_+ \text{ and } \pi \text{ non-decreasing}\}$ and observe that $\hat{\Phi}$ is a convex cone. Observe also that $\hat{b}$ and $\hat{w}$ are continuous, differentiable, and concave. The next step is to define the extended Lagrangian, $\hat{\mathcal{L}}(\pi|\Lambda)$, as in (5) but using $\hat{w}$ and $\hat{b}$ instead of w and b. We can now use Lemma A.2 in Amador et al. (2006). According to this Lemma, the Lagrangian $\hat{\mathcal{L}}$ is maximized at π_x if $\hat{\mathcal{L}}$ is a concave functional defined in a convex cone $\hat{\Phi}$; $\partial \hat{\mathcal{L}}(\pi_x; \pi_x|\Lambda) = 0$; and $\partial \hat{\mathcal{L}}(\pi_x; y|\Lambda) \leq 0$; for all $y \in \hat{\Phi}$. Now note that if $y \in \hat{\Phi}$, then for all sufficiently small $\alpha > 0$, we have that $\alpha y \in \Phi$ and $\partial \hat{\mathcal{L}}(\pi_x; y|\Lambda) = \frac{1}{\alpha} \partial \hat{\mathcal{L}}(\pi_x; \alpha y|\Lambda)$. Hence, it is sufficient to check the above first order conditions for all y in Φ. From Assumption 1 part (iv), it follows that $\pi_x + \alpha y \in \Phi$ for all small enough $\alpha > 0$, as π_f is interior. Then, by the definition of the Gateaux differential, we have that $\partial \hat{\mathcal{L}}(\pi_x; y|\Lambda) = \partial \mathcal{L}(\pi_x; y|\Lambda)$ for all $y \in \Phi$, and conditions (6)-(7) are sufficient for optimality.

[18] Under Assumption 1, arguments similar to those in footnote 14 guarantee the existence of the integral in the right hand side, and Lemma A.1 in Amador et al. (2006) can be used to established the existence of the Gateaux differential.

for $\gamma \geq \gamma_x$. By Assumption 2iii part (c), we know that x is monotone for $\gamma \geq \gamma_x$. By Assumption 2iii part (b), and concavity of b, we know that $b'(x(\gamma)) + \gamma \geq 0$. Then, it follows that

$$t(\gamma) \geq b(x(\gamma)) - b(x(\gamma_x)) - \int_{\gamma_x}^{\gamma} b'(x(\widetilde{\gamma}))dx(\widetilde{\gamma}) = 0$$

for all $\gamma \geq \gamma_x$.[19]

Complementary Slackness. Now we check complementary slackness which requires that

$$\int_{\Gamma} \Big(\int_{\underline{\gamma}}^{\gamma} \pi_x(\gamma')d\gamma' + \underline{U} - \gamma\pi_x(\gamma) - b(\pi_x(\gamma)) \Big) d\Lambda(\gamma) = 0$$

which follows from $t(\gamma) = 0$ for $\gamma \in [\underline{\gamma}, \gamma_x]$ and $\Lambda(\gamma) = 1$ for $\gamma \in [\gamma_x, \overline{\gamma}]$.

Applying Luenberger's Sufficiency Theorem: Here we follow the exact mapping used in the proof of part (b) of Proposition 1 in Amador and Bagwell (2013) in order to use Theorem 1 in Appendix A and show that the proposed allocation $(\pi_x, 0)$ solves Problem (P'). Specifically, we set (i) $x_0 \equiv (\pi_x, 0)$; (ii) f to be given by the negative of the objective function, $f \equiv -\int_{\Gamma}(v(\gamma, \pi(\gamma))f(\gamma) + (1 - F(\gamma))\pi(\gamma))d\gamma - \underline{U}$; (iii) $X \equiv \{\pi, \underline{t} | \underline{t} \in \Re_+, \pi : \Gamma \to \Pi\}$; (iv) $Z \equiv \{z | z : \Gamma \to \Re \text{ with } z \text{ integrable}\}$; (v) $\Omega \equiv \{\pi, \underline{t} | \underline{t} \in \Re_+, \pi : \Gamma \to \Pi;$ and π non-decreasing$\}$; (vi) $P \equiv \{z | z \in Z \text{ such that } z(\gamma) \geq 0 \text{ for all } \gamma \in \Gamma\}$; (vii) G to be the mapping from Ω to Z given by by negative of the left hand side of inequality (1); (viii) $T(z)$ to be the linear mapping:

$$T(z) = \int_{\Gamma} z(\gamma)d\Lambda(\gamma),$$

where $T(z) \geq 0$ for $z \in P$ follows from Λ non-decreasing. We have that

$$T(G(x_0)) \equiv \int_{\Gamma} \Big(\int_{\underline{\gamma}}^{\gamma} \pi_x(\widetilde{\gamma})d\widetilde{\gamma} + \underline{U} - \gamma\pi_x(\gamma) - b(\pi_x(\gamma)) \Big) d\Lambda(\gamma) = 0,$$

which follows here since $t(\gamma) = 0$ for $\gamma \in [\underline{\gamma}, \gamma_x]$ and $\Lambda(\gamma) = 1$ for $\gamma \in [\gamma_x, \overline{\gamma}]$. We have found conditions under which the proposed allocation, $x_0 \equiv (\pi_x, 0)$, minimizes $f(x) + T(G(x))$ for $x \in \Omega$. Given that $T(G(x_0)) = 0$, the conditions of Theorem 1 hold and it follows that $(\pi_x, 0)$ solves $\min_{x \in \Omega} f(x)$ subject to $-G(x) \in P$, which is Problem (P). □

Figure 1 illustrates the money burning solution for Case 1. For lower states (i.e., for $\gamma \leq \gamma_x$), the agent selects the flexible action and no money burning is required. For higher states (i.e., for $\gamma > \gamma_x$), the agent's action is determined by the function x, which for any

[19] The last equality follows from the continuity and monotonicity of x. (See Carter and Brunt, 2000, Theorem 6.2.1 and the subsequent discussion.)

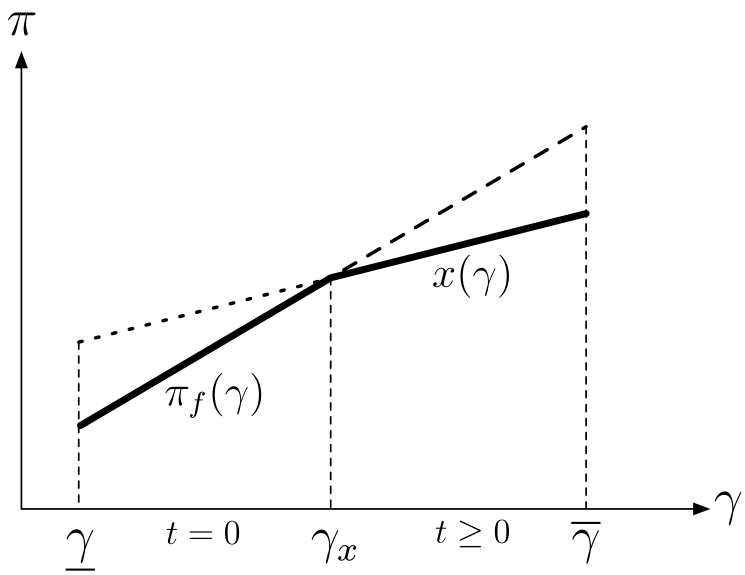

Figure 1: Money burning, Case 1.

given state lies below the agent's flexible action. To prevent the agent from deviating and selecting a higher action intended for a higher state, the optimal delegation contract utilizes money burning, where higher actions are paired with higher levels of money burning.

3.3 Money Burning: Case 2

We next look for an allocation with the following properties: from $\underline{\gamma}$ to some γ_y, the allocation provides no flexibility; and from γ_y to $\overline{\gamma}$, the allocation burns money while providing some flexibility. The next assumption imposes conditions that turn out to be sufficient for such an allocation to be optimal.

Assumption 3. *The following holds:* (i) $v(\gamma, \pi)$ *is concave in* π *for all* $\gamma \in \Gamma$*;* (ii) *there exists a continuous function* $x(\gamma)$ *and a value* $\gamma_y \in \Gamma$ *such that* (a) $v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1$*, for all* $\gamma \geq \gamma_y$*;* (b) $x(\gamma) \leq \pi_f(\gamma)$*, for all* $\gamma \geq \gamma_y$*;* (c) $x(\gamma)$ *is non-decreasing for all* $\gamma \geq \gamma_y$*;* (d) *the following holds:*

$$\underline{\gamma} + b'(x(\gamma_y)) + \int_{\underline{\gamma}}^{\gamma} [v_\pi(\tilde{\gamma}, x(\gamma_y))f(\tilde{\gamma}) + 1 - F(\tilde{\gamma})]d\tilde{\gamma} \geq 0,$$

for all $\gamma \in [\underline{\gamma}, \gamma_y]$ *with equality at* γ_y*.*

Our proposed allocation in this case is:

$$\pi_y(\gamma) = \begin{cases} x(\gamma_y) & ; \text{ for } \gamma \in [\underline{\gamma}, \gamma_y] \\ x(\gamma) & ; \text{ for } \gamma \in (\gamma_y, \overline{\gamma}] \end{cases}$$

with $t_y(\underline{\gamma}) = 0$ and where $x(\gamma)$ and γ_y are defined by Assumption 3.

We may interpret the conditions in Assumption 3 as follows. As discussed above in the context of Assumption 2, part (i) of Assumption 3 may be directly interpreted in terms of our economic examples. This assumption again guarantees that problem (P') is a convex problem, and hence that the Lagrangian is concave. Part (ii) of Assumption 3 provides conditions that the point γ_y and the allocation in the money burning region, $x(\gamma)$, must satisfy, where $x(\gamma)$ is defined by $v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1$. We recall that the function x maximizes pointwise the integrand in Problem (P'), with part (a) providing the corresponding first order condition, where for our present purposes we only require that the function x is defined for $\gamma \geq \gamma_y$. In the proof below, parts (b) and (c) guarantee monotonicity of the proposed allocation and that the allocation is continuous and burns non-negative amounts for money for all types above γ_y; and parts (a) and (d) guarantee that the proposed allocation satisfies the first order conditions of the Lagrangian.

The next proposition is an alternative sufficiency result for money burning:

Proposition 2. *The proposed allocation* $(\pi_y(\gamma), t_y(\underline{\gamma}))$ *solves Problem (P') if Assumption 3 holds.*

Proof. Just as in the proof of Proposition 1, we begin by deriving the Lagrangian:

$$\mathcal{L}(\pi, t(\underline{\gamma})|\Lambda) = \int_{\gamma\in\Gamma} \Big[v(\gamma, \pi(\gamma))f(\gamma) + (\Lambda(\gamma) - F(\gamma))\pi(\gamma) \Big] d\gamma \\ + \int_{\gamma\in\Gamma} \big(\gamma\pi(\gamma) + b(\pi(\gamma)) \big) d\Lambda(\gamma) + \Lambda(\underline{\gamma})\underline{U}.$$

where we used without loss of generality that $\Lambda(\overline{\gamma}) = 1$. Recall that the Lagrange multiplier, Λ, must be non-decreasing.

A proposed multiplier and monotonicity. Our proposed Lagrange multiplier function is $\Lambda(\gamma) = 1$ for all $\gamma \in \Gamma$. This function is trivially non-decreasing.[20]

Concavity of the Lagrangian. When this multiplier function is imposed, the second

[20] It seems from this multiplier that the non-negativity constraint is not binding for all γ. If this were true, then clearly the solution should make t as large and negative as possible! However, recall that there is still a non-negativity constraint that is imposed through the choice set, that is, that $t(\underline{\gamma}) \geq 0$.

integral disappears, permitting the Lagrangian to be expressed as

$$\mathcal{L}(\pi, t(\underline{\gamma})|\Lambda) = \int_{\gamma\in\Gamma} \Big[v(\gamma, \pi(\gamma))f(\gamma) + (1 - F(\gamma))\pi(\gamma)\Big] d\gamma + \underline{U}. \tag{8}$$

Given that v is concave under Assumption 3i and that $\underline{U} \equiv \underline{\gamma}\pi(\underline{\gamma}) + b(\pi(\underline{\gamma})) - t(\underline{\gamma})$ is concave (via b) in $\pi(\underline{\gamma})$ and linear in $t(\underline{\gamma})$, we see that $\mathcal{L}(\pi, t(\underline{\gamma})|\Lambda)$ is concave in the policy functions.

Maximizing the Lagrangian. Our next step then is to consider the Gateaux differentials of the Lagrangian:[21]

$$\partial\mathcal{L}(\pi, \underline{t}; y, \underline{t}_0|\Lambda) = \int_{\gamma\in\Gamma} \Big[v_\pi(\gamma, \pi(\gamma))f(\gamma) + (1 - F(\gamma))\Big] y(\gamma) d\gamma + \underline{\gamma} y(\underline{\gamma}) + b'(\pi(\underline{\gamma}))y(\underline{\gamma}) - \underline{t}_0$$

for any $y : \Gamma \to \Pi$, y non decreasing and $\underline{t}_0 \geq 0$. Evaluating the Gateaux differential at the proposed allocation:

$$\begin{aligned}\partial\mathcal{L}(\pi_y, 0; \pi_y, 0|\Lambda) &= \int_{\gamma\in\Gamma} \Big[v_\pi(\gamma, \pi_y(\gamma))f(\gamma) + (1 - F(\gamma))\Big] \pi_y(\gamma) d\gamma + \underline{\gamma}\pi_y(\underline{\gamma}) + b'(\pi_y(\underline{\gamma}))\pi_y(\underline{\gamma}) \\ &= x(\gamma_y) \left\{ \int_{\underline{\gamma}}^{\gamma_y} \Big[v_\pi(\gamma, x(\gamma_y))f(\gamma) + (1 - F(\gamma))\Big] d\gamma + \underline{\gamma} + b'(x(\gamma_y)) \right\} = 0\end{aligned}$$

where we have used Assumption 3ii part (a) to obtain the second equality. The last equality follows from part (d).

The differential in any other direction is:

$$\begin{aligned}\partial\mathcal{L}(\pi_y, 0; y, \underline{t}_0|\Lambda) &= \int_{\underline{\gamma}}^{\gamma_y} \Big[v_\pi(\gamma, x(\gamma_y))f(\gamma) + (1 - F(\gamma))\Big] y(\gamma) d\gamma + \underline{\gamma} y(\underline{\gamma}) + b'(x(\gamma_y))y(\underline{\gamma}) - \underline{t}_0 \\ &= y(\gamma_y) \int_{\underline{\gamma}}^{\gamma_y} \Big[v_\pi(\gamma, x(\gamma_y))f(\gamma) + (1 - F(\gamma))\Big] d\gamma \\ &\qquad - \int_{\underline{\gamma}}^{\gamma_y} \left\{ \int_{\underline{\gamma}}^{\gamma} \Big[v_\pi(\gamma, x(\gamma_y))f(\gamma) + (1 - F(\gamma))\Big] d\gamma \right\} dy(\gamma) \\ &\qquad\qquad + \underline{\gamma} y(\underline{\gamma}) + b'(x(\gamma_y))y(\underline{\gamma}) - \underline{t}_0 \\ &= -\int_{\underline{\gamma}}^{\gamma_y} \left\{ \int_{\underline{\gamma}}^{\gamma} \Big[v_\pi(\gamma, x(\gamma_y))f(\gamma) + (1 - F(\gamma))\Big] d\gamma \right\} dy(\gamma) \\ &\qquad - \big(y(\gamma_y) - y(\underline{\gamma})\big)\big(\underline{\gamma} + b'(x(\gamma_y))\big) - \underline{t}_0\end{aligned}$$

[21] The existence of the Gateaux differential follows the same argument as in footnote 18.

$$= -\int_{\underline{\gamma}}^{\gamma_y} \left\{ \int_{\underline{\gamma}}^{\gamma} \Big[v_\pi(\gamma, x(\gamma_y)) f(\gamma) + (1 - F(\gamma)) \Big] d\gamma + \underline{\gamma} + b'(x(\gamma_y)) \right\} dy(\gamma) - \underline{t}_0 \leq 0$$

where the second equality follows from integration by parts (which can be done, as one of the functions involved is continuous and the other one is monotonic); and the third equality follows from using Assumption 3ii part (d). The fourth equality follows by just noticing that $\underline{\gamma} + b'(x(\gamma_y))$ is a constant, and the last inequality follows from Assumption 3ii part (d) and $\underline{t}_0 \geq 0$. Taken together, the above first order conditions imply that $(\pi_y, 0)$ maximizes the Lagrangian.[22]

Money burning is non-negative. Our next step is to show that the induced value for $t(\gamma)$ is nonnegative. Using (3), we have that

$$t(\gamma) = \begin{cases} 0 & ; \text{ for } \gamma \in [\underline{\gamma}, \gamma_y] \\ \gamma x(\gamma) - \gamma_y x(\gamma_y) + b(x(\gamma)) - b(x(\gamma_y)) - \int_{\gamma_y}^{\gamma} x(\gamma) d\gamma & ; \text{ for } \gamma \in (\gamma_y, \overline{\gamma}] \end{cases}$$

The proof that $t(\gamma)$ is non-negative follows exactly the same steps as in the proof of Proposition 1, where we now use Assumption 3ii part (b) and (c).

Complementarity Slackness. Complementarity slackness follows from $\Lambda(\gamma) = 1$ for all γ.

Applying Luenberger's Sufficiency Theorem. Here we follow the exact mapping used in the proof of Proposition 1 to use Theorem 1 in Appendix A and show that proposed allocation $(\pi_y, 0)$ solves Problem (P'). □

In Appendix B, we provide a set of sufficient conditions that, although less general, are easier to check than those in Assumption 3.[23] Note also that Assumption 3ii part (d) implies that $\underline{\gamma} + b'(x(\gamma_y)) \geq 0$ and thus $x(\gamma_y) \leq \pi_f(\underline{\gamma})$.

Figure 2 illustrates the money burning solution for Case 2. For lower states (i.e., for $\gamma \leq \gamma_y$), the agent pools at the flexible action for type γ_y and no money burning is required. For higher states (i.e., for $\gamma > \gamma_y$), the agent's action is determined by the function x, which for any given state lies below the agent's flexible action. As before, to prevent the agent from deviating and selecting a higher action intended for a higher state, the optimal delegation contract utilizes money burning, where higher actions are paired with higher levels of money burning.

[22] This follows from a formal argument analogous to that in the proof of Proposition 1, which applies Lemma A.2 of Amador et al. (2006), augmented to allow for the presence of $\underline{t}_0$ in the Lagrangian.

[23] The easier-to-check conditions do not involve γ_y and are stated only in terms of the function $x(\gamma)$ which is defined for all $\gamma \in \Gamma$.

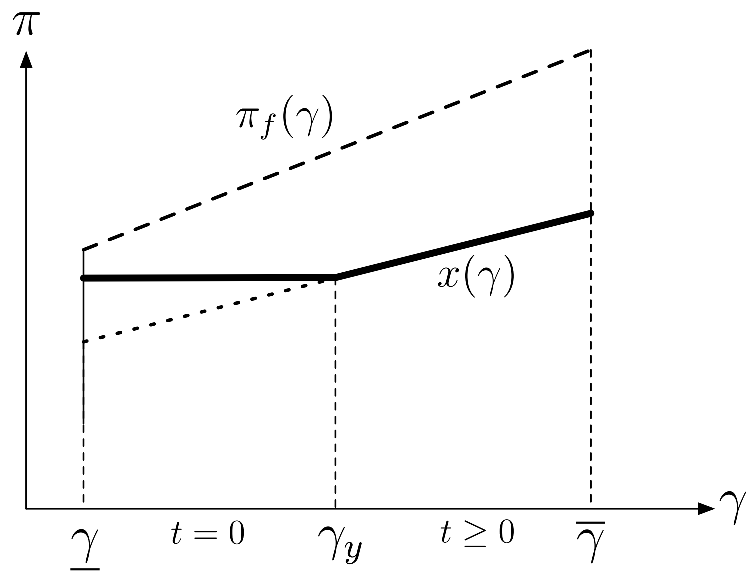

Figure 2: Money burning, Case 2.

4 Optimal Delegation without Money Burning

In the previous section, we extend the standard delegation problem to allow for the possibility of money burning and establish conditions under which optimal delegation entails money burning. To better understand the role of money burning, we consider in this section optimal delegation when money burning is not allowed. Our analysis builds on Amador and Bagwell (2013), who provide sufficient (and necessary) conditions for interval delegation to be optimal, with and without money burning. Here, we assume that money burning is not possible, consider a restricted family of preferences and establish easier-to-check sufficient conditions for the optimality of interval delegation in the form of a cap allocation. For the preference family, this result implies that, under a slight strengthening of Assumption 2, optimal delegation would take the form of a cap were money burning not allowed. We thereby provide a useful benchmark for understanding the role of money burning in Case 1 above. We expect that this result will be useful for other applications as well.

Formally, we consider in this section a restricted family of preferences, where the principal's and the agent's utilities satisfy the following equality:

$$w(\gamma, \pi) = A\big(b(\pi) + B(\gamma) + C(\gamma)\pi\big) \tag{9}$$

for all γ and π, and for some functions B and C and some constant A. Amador and Bagwell (2013) also consider this family of preferences when establishing necessary conditions for the optimality of interval delegation. As they discuss, this restricted class of preferences includes the standard quadratic preferences used in the delegation literature as well as the preferences that Amador et al. (2006) analyze.

To define the candidate interval allocation, we assume that there exists a $\gamma_H \in (\underline{\gamma}, \overline{\gamma})$ that is the solution to the following equation:

$$\int_{\gamma_H}^{\overline{\gamma}} w_\pi(\gamma, \pi_f(\gamma_H)) f(\gamma) d\gamma = 0 \tag{10}$$

We now define the interval allocation, $\pi^\star$, that offers flexibility below γ_H and imposes a cap above γ_H:

$$\pi^\star(\gamma) = \begin{cases} \pi_f(\gamma_H) & \text{for } \gamma > \gamma_H \\ \pi_f(\gamma) & \text{for } \gamma \leq \gamma_H \end{cases}$$

We note that $\pi^\star$ is an interval allocation that takes the particular form of a cap allocation.

The delegation problem when money burning is not allowed is formally defined by Problem P under the further requirement that $t(\gamma) \equiv 0$. For a general family of preferences, Amador and Bagwell (2013) establish sufficient conditions under which the interval allocation $\pi^\star$ is optimal for this problem. Under the preference restriction (9), Amador and Bagwell (2013)'s sufficient conditions for the optimality of $\pi^\star$ for the delegation problem when money burning is not allowed take the following form:

(c1) $AF(\gamma) - v_\pi(\gamma, \pi_f(\gamma))f(\gamma)$ is non-decreasing for all $\gamma \in [\underline{\gamma}, \gamma_H]$.

(c2) if $\gamma_H < \overline{\gamma}$,

$$(\gamma - \gamma_H)A \geq \int_{\gamma}^{\overline{\gamma}} w_\pi(\tilde{\gamma}, \pi_f(\gamma_H)) \frac{f(\tilde{\gamma})}{1 - F(\gamma)} d\tilde{\gamma}\ ,\ \forall \gamma \in [\gamma_H, \overline{\gamma}]$$

with equality at γ_H.

(c3) $v_\pi(\underline{\gamma}, \pi_f(\underline{\gamma})) \leq 0$.

In writing (c1), (c2) and (c3) in this way, we have used the fact that κ, as defined in Amador and Bagwell (2013), equals A, given the preference restriction (9). We have also used the definition for $v(\gamma, \pi(\gamma))$ given above and under which $w_\pi(\gamma, \pi_f(\gamma)) = v_\pi(\gamma, \pi_f(\gamma))$.

We now have the following lemma, showing that if condition (c1) holds globally for the entire domain of $\gamma \in [\underline{\gamma}, \overline{\gamma}]$, then condition (c2) holds as well:

Lemma 1. *Under the preference restriction (9), if $AF(\gamma) - v_\pi(\gamma, \pi_f(\gamma))f(\gamma)$ is non-decreasing for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$, then condition (c2) holds.*

Proof. In Appendix C. □

For applications, it may be easier to confirm that condition (c1) holds globally than to verify directly that condition (c2) holds.[24] Lemma 1 then offers an easier-to-check sufficient condition for the optimality of interval delegation when money burning is not allowed.

We now provide a no-money-burning benchmark finding that may be compared with our Case 1 characterization for the setting with money burning. From Assumption 2 we have that Assumption 2ii is the same as condition (c3). We also have that if Assumption 2i holds and Assumption 2iii part (d) holds globally, then condition (c1) holds for all γ for the preference specification (9), as $A \geq 1$ in this case. We may now use Lemma 1 to obtain the following result:

Lemma 2. *Suppose that Assumption 2i holds, Assumption 2ii holds, and Assumption 2iii part (d) holds for all γ. Suppose as well that (9) holds and that γ_H is given by (10). Then the interval allocation, $\pi^\star$ is optimal for the delegation problem when money burning is not allowed.*

Proof. In Appendix D. □

Referring to Proposition 1, we see that the interval allocation $\pi^\star$ is similar to the money-burning allocation $(\pi_x(\gamma), t_x(\underline{\gamma}))$, in that both allocations grant full flexibility (with no money burning) for low values of γ. For higher values of γ, however, the money-burning allocation $(\pi_x(\gamma), t_x(\underline{\gamma}))$ provides for higher actions that are paired with money-burning expenditures whereas the interval allocation simply imposes a hard cap against higher actions. Based on this comparison, we conclude that the use of money burning is attractive, because it provides an incentive-compatible means for the agent that observes higher values for γ to select actions that rise somewhat with γ.

5 Cooperation and Delegation

As discussed in Example 1, our results can be used to analyze cooperation among two players. The players select a menu of permissible actions and associated money burning levels with the objective of maximizing the ex ante sum of their welfares and with the understanding that one of the players will later observe the state of nature and select an action from the menu. The player who observes the state and takes the action is the agent, and for convenience we refer to the other player as the "partner." The principal's objective in this context is to

[24] We exploit a similar idea in Amador and Bagwell (2014) in the specific context of a regulation problem under logarithmic demand.

maximize the ex ante sum of the agent's and partner's welfare functions. In this section, we use the propositions above to characterize the optimal delegation contract in this setting while allowing for general (non-quadratic) welfare functions and distribution functions. A key ingredient in our analysis is the assumption that the players' welfare functions take related, but distinct, forms; specifically, the agent and partner agree on the benefits generated by an action for any given state, but the partner experiences a higher marginal cost from the action than does the agent.

5.1 The Cooperation Model

Formally, we assume that the agent's welfare function is given as $\beta(\pi) - \alpha c(\pi) + \gamma\pi - t$, from which it follows that the function $b(\pi)$ is captured here as

$$b(\pi) = \beta(\pi) - \alpha c(\pi).$$

The partner's welfare function is related, but distinct, and takes the form

$$v(\gamma, \pi) = \beta(\pi) - (1 - \alpha)c(\pi) + \gamma\pi.$$

Summing the agent and partner's welfare functions, we obtain the joint welfare function as $2\beta(\pi) - c(\pi) + 2\gamma\pi - t$. It follows that the principal's welfare function, $w(\gamma, \pi) - t$, is represented here as

$$w(\gamma, \pi) - t = 2\beta(\pi) - c(\pi) + 2\gamma\pi - t.$$

We refer to this model as the *cooperation model*.

In our analysis of the cooperation model, we assume that $\alpha \in (0, 1/2)$ and that the functions $\beta(\pi)$ and $c(\pi)$ are twice continuously differentiable with $c'(\pi) > 0, \beta''(\pi) < 0$ and $c''(\pi) > 0$. To ensure that the function $\pi_f(\gamma)$ is interior, we also assume that

$$\beta'(0) - \alpha c'(0) + \underline{\gamma} > 0 > \beta'(\overline{\pi}) - \alpha c'(\overline{\pi}) + \overline{\gamma}, \tag{11}$$

where for simplicity we assume that $\overline{\pi} < \infty$ and thus $\Pi = [0, \overline{\pi}]$.

For both the agent and partner, we may interpret $\beta(\pi) + \gamma\pi$ as capturing the benefit of the action π for a given state γ. The marginal cost of π, however, is higher for the partner, since $\alpha \in (0, 1/2)$. The state variable γ is privately observed by the agent at the time that an action is selected. The state may be observed by the partner as well before payoffs are realized, but in any case we follow the delegation literature and assume that is infeasible to write contracts that are contingent on the state. For example, γ may not be verifiable.

We may easily verify that the cooperation model satisfies Assumption 1. Notice that $b(\pi)$ and $w(\gamma, \pi)$ are strictly concave functions of π. The flexible allocation $\pi_f(\gamma)$ solves

$$b'(\pi) + \gamma \equiv \beta'(\pi) - \alpha c'(\pi) + \gamma = 0.$$

Under (11), the flexible schedule is interior. It is also strictly increasing:

$$\pi_f'(\gamma) = \frac{-1}{\beta''(\pi_f) - \alpha c''(\pi_f)} > 0.$$

Thus, Assumption 1 holds.

We next consider further the partner's welfare function. We observe first that $v(\gamma, \pi)$ is also strictly concave in π. The cooperation model thus satisfies part (i) of Assumptions 2 and 3.[25] Observe next that

$$\begin{aligned} v_\pi(\gamma, \pi_f(\gamma)) &= \beta'(\pi_f(\gamma)) - (1-\alpha)c'(\pi_f(\gamma)) + \gamma \\ &= (2\alpha - 1)c'(\pi_f(\gamma)) < 0. \end{aligned}$$

This means that, from the partner's perspective, the agent is biased toward actions that are too high. The cooperation model thus also satisfies part (ii) of Assumption 2.

Before proceeding with our formal analysis, we consider a possible interpretation of the cooperation model.[26] We may imagine two players contemplating the level of a symmetric investment, π, for a joint project. One of the players, the agent, privately observes a state variable, γ, related to the potential promise of the project. The state variable ultimately plays a role in determining the benefit enjoyed by the players; in particular, a higher value for the state variable indicates a greater marginal benefit from investment for the players. The agent in this scenario, however, also incurs less than half of the cost of the project, which is captured in the model by the assumption that $\alpha \in (0, 1/2)$. The agent is thus biased toward actions that are too high from the partner's perspective. Money burning in this context could correspond to persuasion or lobbying costs that the agent must incur as an expression of conviction before the agent is allowed to select a higher investment level.

As one example of such a situation, we can imagine a co-author pair that seeks to de-

[25] Since $v(\gamma, \pi)$ is strictly concave in π, part (i) of Assumption 4 in the Appendix is also satisfied.

[26] The cooperation model considered here is similar to the trade-agreement applications considered by Amador and Bagwell (2013), in the broad sense that in both cases the objective is to maximize the ex ante sum of welfares for two players. The welfare functions in the trade applications considered by Amador and Bagwell (2013), however, are significantly different from those considered here. In the trade applications that Amador and Bagwell (2013) explore, the function v depends on π but is independent of γ, and the function v is also strictly convex in π. The trade applications that Amador and Bagwell (2013) examine thus do not satisfy Assumptions 2 and 3 above and hence are not captured by the cooperation model.

termine an amount of time to devote to a new research project. One of the co-authors, the agent, has better information about the potential promise of the project, but the other co-author, the partner, has technical expertise and expects to provide the majority of the effort for any investment. As another example, this time in the commercial context, the agent may be an entrepreneur with a start-up idea. The entrepreneur has superior information as to the potential value of the idea but requires the assistance of a talented engineer. The engineer, who is the partner in this case, may have a higher opportunity cost of time and thus anticipate bearing a disproportionate cost for any investment in the project. In both of these examples, the partner may be reluctant to "buy in" to plans for a big investment unless the agent undertakes sufficient persuasion or lobbying costs.

5.2 Results

We now develop the applications of our propositions to the cooperation model. To begin, we consider the conditions for our cooperation model under which Assumption 2 holds so that Proposition 1 may be applied. The corresponding Case 1 solution is depicted in Figure 1 above. We note above that parts (i) and (ii) of Assumption 2 hold. The remaining task is to address part (iii) of Assumption 2.

Before stating our finding, we require a definition. For any $\gamma \in \Gamma$, we define $x(\gamma)$ by

$$v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1. \tag{12}$$

We may now report the following result:

Corollary 1. *Consider the cooperation model with $c'(\pi_f(\underline{\gamma}))[1-2\alpha] < 1/f(\underline{\gamma})$. Let $x(\gamma)$ satisfy (12). Then, there exists $\gamma_x \in (\underline{\gamma}, \overline{\gamma})$ such that $x(\gamma) \leq \pi_f(\gamma)$ for all $\gamma > \gamma_x$ with equality at γ_x. If $f'(\gamma) \leq 0$ for $\gamma \in [\gamma_x, \overline{\gamma}]$ and*

$$f(\gamma)\left(\frac{\beta''(\pi_f(\gamma)) - (1-\alpha)c''(\pi_f(\gamma))}{\beta''(\pi_f) - \alpha c''(\pi_f)}\right) \geq (2\alpha - 1)c'(\pi_f(\gamma))f'(\gamma) \text{ for } \gamma \in [\underline{\gamma}, \gamma_x),$$

then the allocation $(\pi_x(\gamma), t_x(\underline{\gamma}))$ identified in Proposition 1 is optimal.

Proof. In Appendix E.

The conditions in Corollary 1 are sure to hold if the density is non-increasing and α is not too far below 1/2.[27] Thus, when the density is non-increasing, if the partner does not bear

[27] As an instructive limiting case, if we were to allow $\alpha = 1/2$, then $x(\overline{\gamma}) = \pi_f(\overline{\gamma})$ would follow. Intuitively, in this case, the agent would have no bias and should have discretion across the full support. This outcome obtains when $\gamma_x = \overline{\gamma}$.

too much of an additional cost burden, then the optimal allocation gives the agent flexibility for states below a critical level, γ_x, and otherwise allows the agent to select actions higher than that preferred under the critical state provided that money burning occurs. We note that strictly positive money burning is used if the density is strictly decreasing for $\gamma > \gamma_x$, so that $x(\gamma)$ is strictly increasing. Actions strictly above the critical action, $\pi_f(\gamma_x)$, then may occur, with strictly higher actions requiring strictly higher money burning expenses.

We consider next the allocation featured in Proposition 2 and depicted as the Case 2 solution in Figure 2 above. This solution is optimal if Assumption 3 holds, where we know from above that part (i) of this assumption holds. As noted following the proof of Proposition 2, we provide in Appendix B a set of sufficient conditions that are easier to check, but less general, than those in Assumption 3. Using these conditions, we find that:

Corollary 2. *Consider the cooperation model. Let $x(\gamma)$ satisfy (12) and assume $x(\underline{\gamma}) > 0$. If $(1-2\alpha)c'(x(\overline{\gamma})) < 2(\overline{\gamma} - E(\gamma))$ and, for all $\gamma \in \Gamma$, $f'(\gamma) < 0$ and $(1-F(\gamma))/f(\gamma) < (1-2\alpha)c'(x(\gamma))$, then there exists $\gamma_y \in (\underline{\gamma}, \overline{\gamma})$ such that the allocation $(\pi_y(\gamma), t_y(\underline{\gamma}))$ identified in Proposition 2 is optimal.*

Proof. In Appendix F.

The conditions in Corollary 2 cannot hold when α is too close to $1/2$ and thus direct attention to settings in which the partner bears a significantly higher cost burden. Under the conditions given in Corollary 2, the optimal allocation imposes a cap on actions unless money burning occurs. An agent that observes a state below a critical level, γ_y, selects the capped action and avoids money burning, while an agent that observes a state above this critical level selects a higher action and engages in money burning. Above the critical state, strictly higher actions require strictly higher money burning expenses.

We make two further points about the conditions in Corollary 2. First, our assumption that $x(\underline{\gamma}) > 0$ can be stated in terms of model primitives and holds if

$$v_\pi(\underline{\gamma}, 0) = \beta'(0) - (1-\alpha)c'(0) + \underline{\gamma} > -\frac{1}{f(\underline{\gamma})}.$$

Second, given that $c''(\pi) > 0$, the conditions in Corollary 2 are mutually consistent only if $(1-F(\gamma))/f(\gamma) < 2(\overline{\gamma} - E(\gamma))$ for all $\gamma \in \Gamma$. A simple example that satisfies this inequality is $f(\gamma) = a + 2(1-a)\gamma$, where $\gamma \in [0,1]$ and $a \in (1,2)$.

Finally, let us briefly discuss the case when money burning is not allowed. As a general matter, the welfare functions for the cooperation model need not satisfy the preference restriction (9); however, if there exist constants C_1, C_2 and C_3 such that the relationship $c(\pi) = C_1\beta(\pi) + C_2\pi + C_3$ holds, then it can be shown that the preference restriction (9) is satisfied. So, when the welfare functions for the cooperation model satisfy this relationship, it is possible to apply the results of Lemma 2. One example for which this relationship holds is when $c(\pi)$ and $\beta(\pi)$ are quadratic, a case that we explore in detail next.

6 Quadratic Welfare and General Distributions

As discussed in Example 2 above, Ambrus and Egorov (2015) analyze a delegation problem with a principal and a privately informed agent, where the contract specifies incentive-compatible actions and money burning levels for the agent. The model includes an initial transfer payment between the principal and agent that is used to satisfy the agent's ex ante participation constraint. The contract must then compensate the agent for expected money burning expenses under the proposed contract.[28] Their model thus fits into the framework that we analyze here, since the principal incurs the money burning expense as well.

Ambrus and Egorov (2015) explicitly solve for optimal contracts in the *quadratic-uniform model*, in which the principal and agent have quadratic utility functions and the type is distributed uniformly over $[0, 1]$. The quadratic-uniform model is commonly used in the delegation literature, and Ambrus and Egorov (2015) generalize this model to include an extra parameter that determines the relative importance in monetary terms of the action choice for the principal and agent. Ambrus and Egorov (2015) then provide a first explicit characterization of solutions to the optimal delegation problem in a continuous setting that include actual money burning.

Our main goal in this section is to illustrate the value of our general propositions by capturing and then generalizing the corresponding results on optimal delegation with money burning by Ambrus and Egorov (2015). Specifically, we generalize the results to families of distributions functions. An additional goal is to apply the simple sufficient conditions developed in Section 4 to characterize optimal delegation when money burning is infeasible and compare this solution with that which occurs in Case 1 when money burning is feasible.

6.0.1 Ambrus and Egorov (2015)'s Quadratic-Uniform Model

[28] We focus on the case where the initial transfer payments are unrestricted. Transfers between the principal and agent are otherwise infeasible. Ambrus and Egorov (2015) also allow for the possibility that the intial transfer is restricted. In an extension, they consider as well a situation where limited contingent transfers are allowed.

In the quadratic-uniform model considered by Ambrus and Egorov (2015), the welfare function for the agent is given as

$$\widetilde{u}(\gamma, \pi) \equiv -[\pi - (\gamma + \beta)]^2 - \widetilde{t}(\gamma),$$

where $\widetilde{t}(\gamma)$ is a money burning variable and $0 < \beta < 1$.[29] The welfare function for the principal is defined as

$$\widetilde{v}(\gamma, \pi) \equiv -\alpha[\pi - \gamma]^2,$$

where $\alpha > 0$ is a parameter that determines the relative importance of the selected action to the principal and agent. The delegation problem thus features positive bias, in that the agent's flexible action, $\pi_f(\gamma) = \gamma + \beta$, exceeds the principal's ideal action, $\pi = \gamma$. As discussed in Example 2, if an ex ante transfer payment p is available and the agent faces an ex ante participation constraint, then we may follow Ambrus and Egorov (2015) and specify a program in the form of (P) in which the principal seeks to maximize the expected value of

$$\widetilde{w}(\gamma, \pi) - \widetilde{t}(\gamma) \equiv -\{\alpha[\pi - \gamma]^2 + [\pi - (\gamma + \beta)]^2\} - \widetilde{t}(\gamma).$$

Dividing by two, ignoring constants and defining $t(\gamma) \equiv \widetilde{t}(\gamma)/2$, we can now map the quadratic-uniform model studied by Ambrus and Egorov (2015) into our problem (P) by specifying that

$$w(\gamma, \pi) = -\frac{\alpha + 1}{2}\left(\pi - \gamma - \frac{\beta}{\alpha + 1}\right)^2; \quad u(\gamma, \pi) = \gamma\pi + b(\pi); \quad \text{and} \quad b(\pi) = \beta\pi - \frac{\pi^2}{2}$$

where the function u denotes the agent's utility function and where, as noted above, $\alpha > 0$ and $0 < \beta < 1$.[30] Recall further that

$$v(\gamma, \pi) \equiv w(\gamma, \pi) - (\gamma\pi + b(\pi)).$$

Using $\pi_f(\gamma) = \gamma + \beta$, we may easily verify that $v_\pi(\gamma, \pi_f(\gamma)) = -\alpha\beta < 0$. It is also straightforward to confirm that $v(\gamma, \pi)$ is strictly concave in π. Finally, for all $\gamma \in \Gamma$, we may define $x(\gamma)$ by $v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1$, which under the quadratic preferences considered here yields the closed-form solution

$$x(\gamma) = \frac{1}{\alpha}\left(\frac{1 - F(\gamma)}{f(\gamma)} + \alpha\gamma\right). \tag{13}$$

[29] Ambrus and Egorov (2015) highlight cases in which $\beta < 1$; however, they also discuss the case where $\beta \geq 1$. We assume here that $\beta < 1$.

[30] This representation is also described in the Online Appendix of Amador and Bagwell (2013).

Using our propositions, we can capture Ambrus and Egorov (2015)'s corresponding characterizations for the quadratic-uniform model. To this end, we now follow Ambrus and Egorov (2015) and further specify that the distribution is uniform, with $F(\gamma) = \gamma$, $\underline{\gamma} = 0$ and $\overline{\gamma} = 1$. When $\alpha \leq 1$, Amador and Bagwell (2013) use part b of their Proposition 1 to show the optimality of a cap allocation with full flexibility up to the critical type $\gamma_H = 1 - 2\alpha\beta/(1+\alpha)$. For $\alpha > 1$, we use our propositions to characterize optimal delegation contracts, where we note using (13) that the function x takes the simple form $x(\gamma) = \gamma + (1-\gamma)/\alpha$ under the assumed uniform distribution.

Specifically, when $1 < \alpha \leq \frac{1}{\beta}$ we can use our Proposition 1 to show that the money burning allocation

$$\pi_x(\gamma) = \begin{cases} \pi_f(\gamma) & ; \gamma \leq 1 - \alpha\beta \\ \gamma + (1-\gamma)/\alpha & ; \gamma > 1 - \alpha\beta \end{cases}$$

is optimal as Assumption 2 holds for $x(\gamma) = \gamma + (1-\gamma)/\alpha$ and $\gamma_x = 1 - \alpha\beta$. Similarly, when $\alpha > \frac{1}{\beta}$, we can use Proposition 2 to show that the money burning allocation

$$\pi_y(\gamma) = \begin{cases} \gamma_y + (1-\gamma_y)/\alpha & ; \gamma \leq \gamma_y \\ \gamma + (1-\gamma)/\alpha & ; \gamma \geq \gamma_y \end{cases}$$

where $\gamma_y = \frac{1}{\alpha}\left(\sqrt{1 + \frac{2\alpha}{\alpha-1}(\alpha\beta - 1)} - 1\right) \in (0,1)$ is optimal as Assumption 3 holds for $x(\gamma) = \gamma + (1-\gamma)/\alpha$.

6.0.2 Applying Lemma 2: Simple Sufficient Conditions

Since the quadratic-uniform model satisfies the preference restriction given in (9), $v_{\pi\pi}(\gamma,\pi) < 0$ and $v_\pi(\underline{\gamma}, \pi_f(\underline{\gamma})) < 0$, we can now apply the simple sufficient conditions identified in Lemma 2. This application confirms the value of Lemma 2: for the uniform-quadratic model, once we solve the simple first-order condition (10) for $\gamma_H = 1 - 2\alpha\beta/(1+\alpha)$, we can *immediately* identify conditions under which the cap allocation $\pi^\star$ is optimal for the delegation problem when money burning is infeasible.[31] Specifically, using $\alpha > 0$, $\beta > 0$ and that $F(\gamma) - v_\pi(\gamma, \pi_f(\gamma))f(\gamma) = F(\gamma) + \alpha\beta f(\gamma)$ is obviously non-decreasing for all $\gamma \in \Gamma$ when the distribution function is uniform with $F(\gamma) = \gamma$, $\underline{\gamma} = 0$ and $\overline{\gamma} = 1$, we can employ Lemma 2 to confirm the optimality of $\pi^\star$ for for the delegation problem when money burning is not allowed. We observe that $\gamma_H \in (0,1)$ under the further constraint that $2\alpha\beta < 1 + \alpha$.

[31] As noted previously, Amador and Bagwell (2013) establish that the quadratic-uniform model satisfies the preference restriction given in (9). The specific values for A, $B(\gamma)$ and $C(\gamma)$ are $A = 1 + \alpha$, $B(\gamma) = -[\gamma + \beta/(1+\alpha)]^2/2$ and $C(\gamma) = \gamma - \alpha\beta/(1+\alpha)$. Amador and Bagwell (2013) consider the case where $\alpha \leq 1$, for which money burning is not used even when it is allowed.

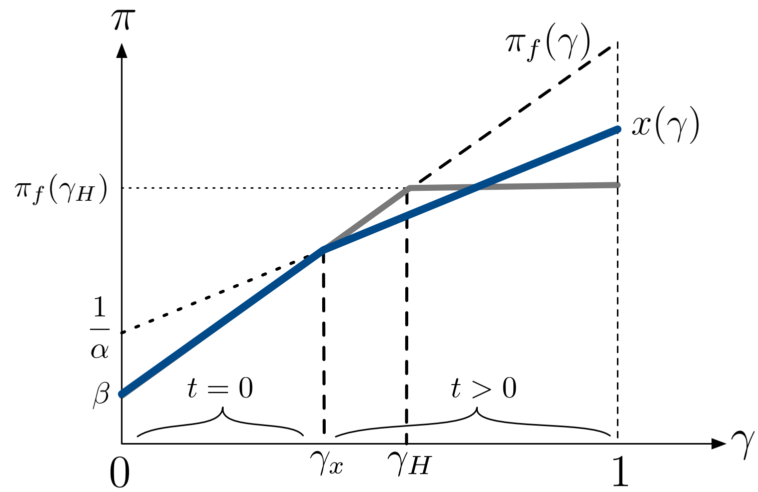

Figure 3: Comparison for the Quadratic-Uniform model: Case 1

With this benchmark established, we may now compare the optimal delegation contract when money burning is allowed with that when it is not. Following the results described just above, when $1 < \alpha < \frac{1}{\beta}$ we can use our Proposition 1 to show that the allocation $\pi_x(\gamma)$ is optimal when money burning is allowed, as Assumption 2 holds for $x(\gamma) = \gamma + (1-\gamma)/\alpha$ and $\gamma_x = 1 - \alpha\beta \in (0, 1)$.[32] Notice, too, that $1 < \alpha < \frac{1}{\beta}$ implies $2\alpha\beta < 1 + \alpha$, so that the further constraint mentioned in the previous paragraph holds automatically. Simple calculations now yield that $\gamma_H - \gamma_x = \alpha\beta(\alpha-1)/(1+\alpha) > 0$ and that $x(\overline{\gamma}) - \pi_f(\gamma_H) = \beta(\alpha-1)/(1+\alpha) > 0$.

For the quadratic-uniform model with $1 < \alpha < \frac{1}{\beta}$, Figure 3 illustrates the optimal delegation contracts when money burning is allowed and when it is not. For the lowest states, both contracts give full flexibility to the agent. For intermediate and high states, however, the use of money burning enables the agent to select actions that rise with the state but are less than the agent's preferred action. Furthermore, when money burning is allowed, the agent's action for intermediate (high) states is below (above) the optimal cap that obtains when money burning is infeasible.

6.0.3 Generalizations

We now further illustrate how our propositions can be easily applied to obtain new results. In what follows, we maintain the quadratic preferences of Ambrus and Egorov (2015) but relax their uniform distribution assumption. We refer to the resulting model as the *generalized Ambrus and Egorov (2015) quadratic model.*

When $\alpha \leq 1$, Amador and Bagwell (2013) use part (b) of their Proposition 1 to establish

[32]It is straightfoward to show that in this case, $t(\gamma)$ is strictly positive and strictly increasing for $\gamma > \gamma_x$.

a general family of distribution functions under which a cap allocation is optimal. When $1 < \alpha \leq \frac{1}{\beta f(\underline{\gamma})}$, we can use Proposition 1 to obtain conditions under which a money burning allocation is optimal:

Corollary 3. *Consider the generalized Ambrus and Egorov (2015) quadratic model with $1 < \alpha \leq \frac{1}{\beta f(\underline{\gamma})}$. For all $\gamma \in \Gamma$, let $x(\gamma)$ be defined by (13). Then, there exists $\gamma_x \in [\underline{\gamma}, \overline{\gamma})$ such that $x(\gamma_x) = \pi_f(\gamma_x)$. If* (i) $\frac{1-F(\gamma)}{f(\gamma)} \leq \alpha\beta$ *for $\gamma \in [\gamma_x, \overline{\gamma}]$,* (ii) $\frac{1-F(\gamma)}{f(\gamma)} + \alpha\gamma$ *is non-decreasing for $\gamma \in [\gamma_x, \overline{\gamma}]$, and* (iii) $F(\gamma) + \alpha\beta f(\gamma)$ *is non-decreasing for $\gamma \in [\underline{\gamma}, \gamma_x)$, then the allocation π_x:*

$$\pi_x(\gamma) = \begin{cases} \pi_f(\gamma) & ; \forall \gamma \leq \gamma_x \\ x(\gamma) & ; \forall \gamma > \gamma_x \end{cases}$$

is optimal.

Proof. In Appendix G. □

The conditions in Corollary 3 are satisfied for the uniform distribution, but also include a large family of non-uniform distributions as well. To illustrate, we consider a family of linear densities: $f(\gamma) = a + 2(1-a)\gamma$ where γ is distributed over $[0, 1]$ and $a \in (0, 2)$. This family includes the uniform density ($a = 1$), increasing densities ($a < 1$) and decreasing densities ($a > 1$). As we show in Appendix I, if $\alpha > 1$ and $1 \geq a\alpha\beta$, then all of the conditions of Corollary 3 are satisfied.

We conclude with the characterization of the last parameter region, $\alpha > \frac{1}{\beta f(\underline{\gamma})}$, where we can use Proposition 2 to show that an allocation with money burning is optimal:

Corollary 4. *Consider the generalized Ambrus and Egorov (2015) quadratic model with $\alpha > \frac{1}{\beta f(\underline{\gamma})}$. For all $\gamma \in \Gamma$, let $x(\gamma)$ be defined by (13). If* (i) $\frac{1-F(\gamma)}{f(\gamma)} + \alpha\gamma$ *is strictly increasing for $\gamma \in \Gamma$,* (ii) $\frac{1-F(\gamma)}{f(\gamma)} \leq \frac{1}{f(\underline{\gamma})}$ *for $\gamma \in \Gamma$, and* (iii) $(1+\alpha)(\overline{\gamma} - \mathbb{E}[\gamma]) > \beta$*, then an allocation of the form*

$$\pi_y(\gamma) = \begin{cases} x(\gamma_y) & ; \forall \gamma \leq \gamma_y \\ x(\gamma) & ; \forall \gamma > \gamma_y \end{cases}$$

for some $\gamma_y \in (\underline{\gamma}, \overline{\gamma})$ is optimal.

Proof. In Appendix H. □

As an illustration of a non-uniform distribution that satisfies the assumptions of Corollary 4, consider the linear density $f(\gamma) = a + 2(1-a)\gamma$ where γ is distributed over $[0, 1]$ and $a \in [1, 2)$. The assumption that $\alpha > 1/(\beta f(\underline{\gamma}))$ and assumption (iii) of Corollary 4 hold if

$1 < a\alpha\beta$ and $a > 6\beta/(1+\alpha) - 2$, respectively. These inequalities can simultaneously hold if $\alpha \geq 1/2$. As we show in the Appendix I, assumptions (i) and (ii) of Corollary 4 hold for this distribution when $a \in [1, 2)$.

7 Conclusions

We consider a general representation of the delegation problem with money burning, and we characterize optimal solutions to the delegation problem in which actual money burning occurs. The solutions that we identify are motivated by the findings that Ambrus and Egorov (2015) develop for a setting with quadratic payoffs and a uniform distribution. Using and extending the Lagrangian methods developed by Amador et al. (2006) and Amador and Bagwell (2013), we are able to establish general conditions on welfare and distribution functions under which money burning is optimal. To better understand the role of money burning, we also establish simple sufficient conditions for the optimality of interval delegation in the form of a cap allocation under a restricted set of preferences for a benchmark setting in which money burning is not allowed. In addition, we illustrate the value of our propositions by applying them to a specific model of cooperation and delegation and by generalizing the corresponding findings of Ambrus and Egorov (2015) for the quadratic-uniform model to include families of distribution functions. For one case in the quadratic-uniform model, we also apply our simple sufficient conditions to characterize optimal delegation when money burning is infeasible and compare this solution with that which occurs when money burning is feasible. We expect that our techniques and solutions will also be useful for other studies of applied mechanism design when contingent transfers are infeasible and wasteful actions may be used to provide incentives.

A A Modified Version of Luenberger's Sufficiency Theorem

Here we restate, from Amador and Bagwell (2013), a slightly modified version of Theorem 1 in section 8.4 of Luenberger (1969) that makes explicit the complementary slackness condition:

Theorem 1. *Let f be a real valued functional defined on a subset Ω of a linear space X. Let G be a mapping from Ω into the linear space Z having non-empty positive cone P. Suppose that* (i) *there exists a linear function $T : Z \to \mathbb{R}$ such that $T(z) \geq 0$ for all $z \in P$,* (ii) *there* *is an element* $x_0 \in \Omega$ *such that*

$$f(x_0) + T(G(x_0)) \leq f(x) + T(G(x)), \text{ for all } x \in \Omega,$$

(iii) $-G(x_0) \in P$, *and* (iv) $T(G(x_0)) = 0$. *Then* x_0 *solves:*

$$\min f(x) \text{ subject to: } -G(x) \in P, \quad x \in \Omega$$

Proof. For completeness, we restate the proof done in Amador and Bagwell (2013). Note that from (ii) and (iii), x_0 is in the constraint set of the minimization problem. Suppose that there exists an $x_1 \in \Omega$ with $f(x_1) < f(x_0)$ and $-G(x_1) \in P$, so that x_0 is not a minimizer. Then, by hypothesis (i), $T(-G(x_1)) \geq 0$. Linearity implies that $T(G(x_1)) \leq 0$. Using this together with (iv), it follows that $f(x_1) + T(G(x_1)) < f(x_0) = f(x_0) + T(G(x_0))$, which contradicts hypothesis (ii). □

B Easier-to-Check Assumptions that imply Assumption 3

Proposition 2 is a useful proposition, but one may wonder when Assumption 3ii part (d) would hold. To address these concerns, in this section we obtain easier-to-check conditions that deliver Assumption 3ii part (d). To this end, we state the following alternative assumption:

Assumption 4. *The following holds:*

(i) $v(\gamma, \pi)$ *strictly concave in* π*;*

(ii) *there exists a function* $x(\gamma)$ *such that:* (a) $v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1$*, for all* $\gamma \in \Gamma$*;* (b) $x(\gamma) \leq \pi_f(\gamma)$*, for all* $\gamma \in \Gamma$ *and strictly at* $\underline{\gamma}$*;* (c) $x(\gamma)$ *is strictly increasing, for all* $\gamma \in \Gamma$*;* (d) $\mathbb{E}[\gamma] + b'(x(\overline{\gamma})) + \int_{\underline{\gamma}}^{\overline{\gamma}} v_\pi(\tilde{\gamma}, x(\overline{\gamma}))f(\tilde{\gamma})d\tilde{\gamma} < 0$.

This assumption is stated only in terms of the function $x(\gamma)$, which is now defined for all $\gamma \in \Gamma$. In effect, Assumption 4ii strengthens slightly parts (a), (b) and (c) of Assumption 3ii, while replacing part (d). Notice that γ_y does not appear in this assumption. We have the following result:

Proposition 3. *Suppose that Assumption 4 holds. Then there exists a unique value $\gamma_g \in (\underline{\gamma}, \overline{\gamma})$ such that:*

$$\gamma_g + b'(x(\gamma_g)) + \int_{\underline{\gamma}}^{\gamma_g} \big[v_\pi(\gamma, x(\gamma_g))f(\gamma) - F(\gamma)\big] d\gamma = 0$$

If $v_\pi(\gamma, x(\gamma_g))f(\gamma) - F(\gamma)$ is non-decreasing for all $\gamma \in [\underline{\gamma}, \gamma_g]$, then the allocation $(\pi_y(\gamma), 0)$ with $\gamma_y = \gamma_g$ is optimal.

Before stating the proof, let us first define

$$G(\gamma) \equiv \gamma + b'(x(\gamma)) + \int_{\underline{\gamma}}^{\gamma} [v_\pi(\tilde{\gamma}, x(\gamma))f(\tilde{\gamma}) - F(\tilde{\gamma})] d\tilde{\gamma}.$$

We now have the following Lemma:

Lemma 3. *Under Assumption 4, there exists a unique $\gamma_g \in (\underline{\gamma}, \overline{\gamma})$ such that $G(\gamma_g) = 0$.*

Proof. We find that:

$$G(\underline{\gamma}) = \underline{\gamma} + b'(x(\underline{\gamma})) = -b'(\pi_f(\underline{\gamma})) + b'(x(\underline{\gamma})) > 0,$$

where the inequality follows since $x(\underline{\gamma}) < \pi_f(\underline{\gamma})$ by Assumption 4ii part (b) and since $b'' < 0$. Next, we find that

$$\begin{aligned} G(\overline{\gamma}) &= \overline{\gamma} + b'(x(\overline{\gamma})) + \int_{\underline{\gamma}}^{\overline{\gamma}} [v_\pi(\tilde{\gamma}, x(\overline{\gamma}))f(\tilde{\gamma}) - F(\tilde{\gamma})] d\tilde{\gamma} \\ &= \mathbb{E}[\gamma] + b'(x(\overline{\gamma})) + \int_{\underline{\gamma}}^{\overline{\gamma}} v_\pi(\tilde{\gamma}, x(\overline{\gamma}))f(\tilde{\gamma}) d\tilde{\gamma} < 0, \end{aligned}$$

where the inequality follows from Assumption 4ii part (d). Finally, we find that, for all $\gamma \in \Gamma$,

$$G(\gamma) = \underline{\gamma} + b'(x(\gamma)) + \int_{\underline{\gamma}}^{\gamma} \big[v_\pi(\tilde{\gamma}, x(\gamma))f(\tilde{\gamma}) + 1 - F(\tilde{\gamma})\big] d\tilde{\gamma}$$

Now note that $\underline{\gamma} + b'(x(\gamma))$ is strictly decreasing in γ given that x is strictly increasing and b is strictly concave. Note as well that

$$\int_{\underline{\gamma}}^{\gamma} \big[v_\pi(\tilde{\gamma}, x(\gamma))f(\tilde{\gamma}) + 1 - F(\tilde{\gamma})\big] d\tilde{\gamma} = \int_{\underline{\gamma}}^{\gamma} \big[v_\pi(\tilde{\gamma}, x(\gamma)) - v_\pi(\tilde{\gamma}, x(\tilde{\gamma}))\big] f(\tilde{\gamma}) d\tilde{\gamma} \tag{14}$$

where we have used Assumption 4ii part (a). Given that x is strictly increasing and v strictly concave, we have that $v_\pi(\tilde{\gamma}, x(\gamma)) - v_\pi(\tilde{\gamma}, x(\tilde{\gamma})) < 0$ for all $\tilde{\gamma} < \gamma$. This difference strictly decreases as $x(\gamma)$ increases by strict concavity of v. Hence an increase in γ strictly increases $x(\gamma)$, and thus strictly reduces (14) by both increasing the support of the integral and by making the integrand more negative at all points.

Taken together, it follows that G is strictly decreasing in γ as it is the sum of two strictly decreasing functions. Assumption 4i guarantees that x is continuous, and hence it follows that G is continuous as well. Hence there exists a unique $\gamma_g \in (\underline{\gamma}, \overline{\gamma})$ so that $G(\gamma_g) = 0$. □

Now we are ready to prove Proposition 3:

Proof. Let us define

$$KT(\gamma) \equiv \gamma + b'(x(\gamma_g)) + \int_{\underline{\gamma}}^{\gamma} [v_\pi(\tilde{\gamma}, x(\gamma_g))f(\tilde{\gamma}) - F(\tilde{\gamma})]d\tilde{\gamma}.$$

Observe that $KT(\gamma_g) = G(\gamma_g) = 0$ by Lemma 3. Note that:

$$KT'(\gamma) = 1 + v_\pi(\gamma, x(\gamma_g))f(\gamma) - F(\gamma)$$

which is non-decreasing by the hypothesis of the Proposition. Hence KT is convex. We also know that $KT'(\gamma_g) = 0$ by the definition of x. It follows then that γ_g is a minimizer of KT and thus $KT(\gamma) \geq KT(\gamma_g) = 0$ for all γ. Now let $\gamma_y = \gamma_g$. Then we have that Assumption 3ii part (a) holds. Assumption 3ii part (b) follows from Assumption 4ii part (b). Assumption 3ii part (c) follows from Assumption 4ii part (c). And finally, Assumption 3ii part (d) follows from $KT(\gamma) \geq 0$ for all γ and with equality at γ_g.

Hence Assumption 3 holds, and we can apply Proposition 2. QED. □

C Proof of Lemma 1

It is convenient to define

$$G(\gamma) \equiv \int_{\gamma}^{\overline{\gamma}} w_\pi(\widetilde{\gamma}, \pi_f(\gamma_H))f(\widetilde{\gamma})d\widetilde{\gamma} - (\gamma - \gamma_H)A(1 - F(\gamma)).$$

Now notice that (c2) can be rewritten as

(c2) $G(\gamma) \leq 0$ for $\gamma \geq \gamma_H$, with $G(\gamma_H) = 0$.

We next note that $G(\gamma_H) = 0 = G(\overline{\gamma})$, and

$$G'(\gamma) = -[w_\pi(\gamma, \pi_f(\gamma_H)) - (\gamma - \gamma_H)A]f(\gamma) - A(1 - F(\gamma)). \tag{15}$$

Note that under the preference specification (9), $w_\pi(\gamma, \pi_f(\gamma_H) - (\gamma - \gamma_H)A = w_\pi(\gamma, \pi_f(\gamma))$. To see this, we use $b'(\pi_f(\gamma)) + \gamma \equiv 0$, to get that

$$\begin{aligned} w_\pi(\gamma, \pi_f(\gamma_H)) - (\gamma - \gamma_H)A &= A[b'(\pi_f(\gamma_H)) + C(\gamma)] - (\gamma - \gamma_H)A \\ &= A[C(\gamma) - \gamma_H] - (\gamma - \gamma_H)A \\ &= A[C(\gamma) - \gamma] \end{aligned}$$

and likewise that

$$\begin{aligned} w_\pi(\gamma, \pi_f(\gamma)) &= A[b'(\pi_f(\gamma)) + C(\gamma)] \\ &= A[C(\gamma) - \gamma]. \end{aligned}$$

So we have that

$$G'(\gamma) = AF(\gamma) - w_\pi(\gamma, \pi_f(\gamma))f(\gamma) - A. \tag{16}$$

Given $w_\pi(\gamma, \pi_f(\gamma)) = v_\pi(\gamma, \pi_f(\gamma))$, it follows that from (16) that, if (c1) holds globally, then $G'(\gamma)$ is non-decreasing for the entire domain.

Using this result and the fact that $G(\gamma_H) = 0 = G(\overline{\gamma})$, we conclude that $G(\gamma) \leq 0$ for $\gamma \geq \gamma_H$. Condition (c2) thus holds.

D Proof of Lemma 2

We already argue in the text that conditions (c1), (c2) and (c3) follow from the premise of the lemma. As a result, we can apply Proposition 1(a) from Amador and Bagwell (2013) to obtain the optimality of the interval allocation.

E Proof of Corollary 1

For any $\gamma \in \Gamma$, we define $x(\gamma)$ by (12). We seek a condition under which $x(\gamma)$ defined by (12) is interior (i.e., $x(\gamma) \in (0, \overline{\pi})$ for all $\gamma \in \Gamma$) and there exists $\gamma_x \in (\underline{\gamma}, \overline{\gamma})$ such that $x(\gamma) \leq \pi_f(\gamma)$ for all $\gamma > \gamma_x$ with equality at γ_x. Under such a condition, Assumption 2(iii)(a) and (b) would hold. Since $\pi_f(\gamma)$ is interior under our assumptions above, it is sufficient to find a condition under which the range of $x(\gamma)$ is contained within the range of $\pi_f(\gamma)$ for $\gamma \in \Gamma$.

Our first step is thus to establish that $x(\overline{\gamma}) < \pi_f(\overline{\gamma})$. Observe that $x(\overline{\gamma})$ and $\pi_f(\overline{\gamma})$ respectively solve

$$\begin{aligned} \beta'(x(\overline{\gamma})) - (1-\alpha)c'(x(\overline{\gamma})) + \overline{\gamma} &= 0 \\ \beta'(\pi_f(\overline{\gamma})) - \alpha c'(\pi_f(\overline{\gamma})) + \overline{\gamma} &= 0. \end{aligned}$$

Subtracting, we get

$$\beta'(x(\overline{\gamma})) - \beta'(\pi_f(\overline{\gamma})) = (1-\alpha)c'(x(\overline{\gamma})) - \alpha c'(\pi_f(\overline{\gamma})). \tag{17}$$

Now suppose to the contrary that $x(\overline{\gamma}) \geq \pi_f(\overline{\gamma})$. By concavity of $\beta(\pi)$, the LHS of (17) must then be non-positive. Since $c(\pi)$ is convex, $c'(x(\overline{\gamma})) \geq c'(\pi_f(\overline{\gamma}))$ then follows. Given that $\alpha \in (0, 1/2)$, we may conclude that the RHS of (17) is positive. This is a contradiction, and so we conclude that $x(\overline{\gamma}) < \pi_f(\overline{\gamma})$.

Our second step is to provide a condition under which $x(\underline{\gamma}) > \pi_f(\underline{\gamma})$. Observe that $x(\underline{\gamma})$ and $\pi_f(\underline{\gamma})$ respectively solve

$$\begin{aligned} \beta'(x(\underline{\gamma})) - (1-\alpha)c'(x(\underline{\gamma})) &= -1/f(\underline{\gamma}) - \underline{\gamma} \\ \beta'(\pi_f(\underline{\gamma})) - \alpha c'(\pi_f(\underline{\gamma})) &= -\underline{\gamma}. \end{aligned}$$

Subtracting, we obtain

$$[\beta'(x(\underline{\gamma})) - \beta'(\pi_f(\underline{\gamma}))] - [(1-\alpha)c'(x(\underline{\gamma})) - \alpha c'(\pi_f(\underline{\gamma}))] = -1/f(\underline{\gamma}). \tag{18}$$

Now, the RHS of (18) is clearly negative. Suppose to the contrary that $x(\underline{\gamma}) \leq \pi_f(\underline{\gamma})$. Using the concavity of $\beta(\pi)$ and the convexity of $c(\pi)$, respectively, it would then follow that

$$\begin{aligned} & [\beta'(x(\underline{\gamma})) - \beta'(\pi_f(\underline{\gamma}))] - [(1-\alpha)c'(x(\underline{\gamma})) - \alpha c'(\pi_f(\underline{\gamma}))] \\ \geq \;& -[(1-\alpha)c'(x(\underline{\gamma})) - \alpha c'(\pi_f(\underline{\gamma}))] \\ \geq \;& -[(1-\alpha)c'(\pi_f(\underline{\gamma})) - \alpha c'(\pi_f(\underline{\gamma}))] \\ = \;& -c'(\pi_f(\underline{\gamma}))[1-2\alpha]. \end{aligned}$$

We may thus conclude that a contradiction occurs if $-c'(\pi_f(\underline{\gamma}))[1-2\alpha] > -1/f(\underline{\gamma})$. Equivalently, under the condition that

$$c'(\pi_f(\underline{\gamma}))[1-2\alpha] < 1/f(\underline{\gamma}), \tag{19}$$

we may conclude that $x(\underline{\gamma}) > \pi_f(\underline{\gamma})$ and hence that $x(\gamma)$ is interior. Assumption 2(iii)(a) and (b) are thus satisfied when (19) holds.

We turn next to Assumption 2(iii)(c), which requires $x(\gamma)$ non-decreasing for all $\gamma \geq \gamma_x$. With $x(\gamma)$ defined by (12), and using $v_{\pi\gamma} = 1$, we obtain

$$x'(\gamma) = \frac{-[1 + \frac{d}{d\gamma}(\frac{1-F(\gamma)}{f(\gamma)})]}{v_{\pi\pi}(\gamma, \pi)} = \frac{\frac{f'(\gamma)(1-F(\gamma))}{f^2(\gamma)}}{v_{\pi\pi}(\gamma, \pi)}. \tag{20}$$

Thus, $x'(\overline{\gamma}) = 0$, and $sign(x'(\gamma)) = sign(-f'(\gamma))$ for $\gamma < \overline{\gamma}$. We conclude that Assumption 2(iii)(c) holds if $f'(\gamma) \leq 0$ for $\gamma \in [\gamma_x, \overline{\gamma})$.

The remaining task is to provide a condition under which $F(\gamma) - v_\pi(\gamma, \pi_f(\gamma))f(\gamma)$ is non-decreasing for $\gamma < \gamma_x$, so that Assumption 2(iii)(d) holds. Differentiating and gathering terms, we find that Assumption 2(iii)(d) holds under the following condition:

$$f(\gamma)(\frac{\beta''(\pi_f(\gamma)) - (1-\alpha)c''(\pi_f(\gamma))}{\beta''(\pi_f(\gamma)) - \alpha c''(\pi_f(\gamma))}) \geq (2\alpha - 1)c'(\pi_f(\gamma))f'(\gamma) \text{ for } \gamma < \gamma_x. \tag{21}$$

Notice that the LHS of (21) is positive, while the sign of the RHS is opposite of the sign of $f'(\gamma)$ (since $\alpha < 1/2$). So, (19) and (21) can both hold if α is not too far below 1/2. $\square$

F Proof of Corollary 2

We observe first that $v(\gamma, \pi)$ is strictly concave in π; thus, Assumption 4(i) holds. Next, we establish that Assumption 4(ii)(a) holds when, for all $\gamma \in \Gamma$, we define $x(\gamma)$ by (12). To this end, we show that $x(\gamma)$ defined by (12) is interior (i.e., $x(\gamma) \in (0, \overline{\pi})$ for all $\gamma \in \Gamma$). We assume $x(\underline{\gamma}) > 0$. We also know from the proof of Corollary 1 that $x(\overline{\gamma}) < \pi_f(\overline{\gamma})$. Finally, referring to (20) and using $f'(\gamma) < 0$ for all $\gamma \in \Gamma$, we also know that $x'(\gamma) > 0$ for all $\gamma < \overline{\gamma}$ with $x'(\overline{\gamma}) = 0$. Interiority is thus established, and so Assumption 4(ii)(a) holds. Having shown that $x'(\gamma) > 0$ for all $\gamma < \overline{\gamma}$, we also now know that Assumption 4(ii)(c) holds. To confirm that Assumption 4 holds, we thus have left to show that Assumption 4(ii)(b) and (d) hold.

We begin with Assumption 4(ii)(d), which requires that

$$E[\gamma] + b'(x(\overline{\gamma})) + \int_{\underline{\gamma}}^{\overline{\gamma}} v_\pi(\widetilde{\gamma}, x(\overline{\gamma}))f(\widetilde{\gamma})d\widetilde{\gamma} < 0. \tag{22}$$

To evaluate this condition, we observe that

$$v_\pi(\widetilde{\gamma}, x(\overline{\gamma})) = \beta'(x(\overline{\gamma})) - (1-\alpha)c'(x(\overline{\gamma})) + \widetilde{\gamma} = -\overline{\gamma} + \widetilde{\gamma},$$

where the second equality follows since $x(\overline{\gamma})$ satisfies $v_\pi(\overline{\gamma}, x(\overline{\gamma}))f(\overline{\gamma}) = F(\overline{\gamma}) - 1 = 0$ and thus $v_\pi(\overline{\gamma}, x(\overline{\gamma})) = \beta'(x(\overline{\gamma})) - (1-\alpha)c'(x(\overline{\gamma})) + \overline{\gamma} = 0$. We may now also observe that

$$b'(x(\overline{\gamma})) = \beta'(x(\overline{\gamma})) - \alpha c'(x(\overline{\gamma})) = (1-2\alpha)c'(x(\overline{\gamma})) - \overline{\gamma}.$$

Substituting, we may rewrite (22) as

$$(1-2\alpha)c'(x(\overline{\gamma})) < 2(\overline{\gamma} - E[\gamma]).$$

Since this inequality is assumed in Corollary 2, we have that Assumption 4(ii)(d) holds.

Now consider Assumption 4(ii)(b). Recall that $x(\gamma)$ and $\pi_f(\gamma)$ respectively solve:

$$\begin{aligned} [\beta'(x(\gamma)) - (1-\alpha)c'(x(\gamma)) + \gamma]f(\gamma) + 1 - F(\gamma) &= 0 \\ \beta'(\pi_f(\gamma)) - \alpha c'(\pi_f(\gamma)) + \gamma &= 0. \end{aligned}$$

Recall that we assume in Corollary 2 that

$$1 - F(\gamma) < (1-2\alpha)c'(x(\gamma))f(\gamma). \tag{23}$$

Using (23), we have that

$$\begin{aligned} & [\beta'(x(\gamma)) - (1-\alpha)c'(x(\gamma)) + \gamma]f(\gamma) + 1 - F(\gamma) \\ < \; & [\beta'(x(\gamma)) - (1-\alpha)c'(x(\gamma)) + \gamma]f(\gamma) + (1-2\alpha)c'(x(\gamma))f(\gamma) \\ = \; & [\beta'(x(\gamma)) - \alpha c'(x(\gamma)) + \gamma]f(\gamma) \end{aligned}$$

and so

$$\beta'(x(\gamma)) - \alpha c'(x(\gamma)) + \gamma > 0 = \beta'(\pi_f(\gamma)) - \alpha c'(\pi_f(\gamma)) + \gamma.$$

It follows by concavity of β and convexity of c that $x(\gamma) < \pi_f(\gamma)$. Thus, under (23), Assumption 4(ii)(b) holds. Further, we get $x(\underline{\gamma}) < \pi_f(\underline{\gamma})$, which is specifically required by Assumption 4(ii)(b).

Having shown that Assumption 4 holds under the conditions in Corollary 2, we now utilize Proposition 3. Proposition 3 further requires that

$$v_\pi(\gamma, x(\gamma_g))f(\gamma) - F(\gamma) \text{ is nondecreasing over } [\underline{\gamma}, \gamma_g] \tag{24}$$

where $\gamma_g \in (\underline{\gamma}, \overline{\gamma})$ is uniquely defined in Proposition 3 given that Assumption 4 holds. Differentiating and using $v_{\pi\gamma} = 1$, we find that

$$\begin{aligned}
\frac{d}{d\gamma}[v_\pi(\gamma, x(\gamma_g))f(\gamma) - F(\gamma)] &= v_\pi(\gamma, x(\gamma_g))f'(\gamma) \\
&= [v_\pi(\gamma_g, x(\gamma_g)) + (\gamma - \gamma_g)]f'(\gamma) \\
&= [(F(\gamma_g) - 1)/f(\gamma_g) + (\gamma - \gamma_g)]f'(\gamma).
\end{aligned}$$

Thus, given $f'(\gamma) < 0$ for all $\gamma < \overline{\gamma}$, and given $\gamma_g \in (\underline{\gamma}, \overline{\gamma})$, we conclude that (24) holds. We can thus apply Proposition 3. □

G Proof of Corollary 3

As we discuss in the text, conditions i and ii of Assumption 2 hold for the quadratic preferences we are using in this section.

To show part (iii) of Assumption 2, note that the inequality $x(\gamma) \leq \pi_f(\gamma)$ can be written as

$$\frac{1}{\alpha}\left(\frac{1 - F(\gamma)}{f(\gamma)}\right) \leq \beta. \tag{25}$$

There exists a value of $\gamma_x \in [\underline{\gamma}, \overline{\gamma})$ such that the above holds with equality. This follows because the left hand side is zero when $\gamma_x = \overline{\gamma}$ and is weakly bigger than β when $\gamma_x = \underline{\gamma}$ by the hypothesis of Corollary 3. Continuity of $f(\gamma)$ then implies existence of such a γ_x. From hypothesis (i) of Corollary 3, it follows that (25) holds for all $\gamma > \gamma_x$. It follows also that for $\gamma > \gamma_x$, $v_\pi(\gamma, x(\gamma))f(\gamma) = F(\gamma) - 1$ by the definition of x. The function $x(\gamma)$ is non-decreasing for all $\gamma \geq \gamma_x$ given the hypothesis (ii) of Corollary 3. And finally, hypothesis (iii) of the corollary implies that that $F(\gamma) - v_\pi(\gamma, \pi_f(\gamma))f(\gamma)$ is non-decreasing. Hence Assumption 2 holds, and we can apply Proposition 1 to prove the corollary. □

H Proof of Corollary 4

As argued before, Assumption 3i holds. Note that $x(\gamma) = \frac{1}{\alpha}((1 - F(\gamma))/f(\gamma) + \alpha\gamma)$ satisfies 3iia and 3iic by definition of x and hypothesis (i) of Corollary 4. Note also that $x(\gamma) - \pi_f(\gamma) = \frac{1}{\alpha}\frac{1-F(\gamma)}{f(\gamma)} - \beta \leq \frac{1}{\alpha}\frac{1}{f(\underline{\gamma})} - \beta < 0$, by hypothesis (ii) of Corollary 4, so that 3iib holds. Finally note that the condition in 3iid is equivalent to:

$$p(\gamma, \gamma_y) \equiv (1 - F(\gamma))\gamma + \beta + (1 + \alpha)\mathbb{E}[\tilde{\gamma}|\tilde{\gamma} < \gamma]F(\gamma) - \left(\frac{1 - F(\gamma_y)}{\alpha f(\gamma_y)} + \gamma_y\right)(1 + \alpha F(\gamma)) \geq 0$$

for all $\gamma \in [\underline{\gamma}, \gamma_y]$ with equality at γ_y. Note that $pp(\gamma_y) \equiv p(\gamma_y, \gamma_y) = (1 - F(\gamma_y))\gamma_y + \beta + (1 + \alpha)\mathbb{E}[\tilde{\gamma}|\tilde{\gamma} \leq \gamma_y]F(\gamma_y) - (1 + \alpha F(\gamma_y))[\gamma_y + (1 - F(\gamma_y))/(\alpha f(\gamma_y))]$. Note also that $pp(\gamma_y)$ is continuous in γ_y and that $pp(\underline{\gamma}) = \beta - \frac{1}{\alpha f(\underline{\gamma})} > 0$ and $pp(\overline{\gamma}) = \beta + (1 + \alpha)(\mathbb{E}[\gamma] - \overline{\gamma}) < 0$ where this last holds by hypothesis (iii). It follows then that there exists $\gamma_y \in (\underline{\gamma}, \overline{\gamma})$ such that $pp(\gamma_y) = 0$.

Note as well that

$$\begin{aligned}\partial p(\gamma, \gamma_y)/\partial\gamma &= v_\pi(\gamma, x(\gamma_y))f(\gamma) + 1 - F(\gamma) \\ &= f(\gamma)\left\{\left[\frac{1 - F(\gamma)}{f(\gamma)} + \alpha\gamma\right] - \left[\frac{1 - F(\gamma_y)}{f(\gamma_y)} + \alpha\gamma_y\right]\right\} \leq 0\end{aligned}$$

for $\gamma \leq \gamma_y$, where the inequality follows from hypothesis (i) of Corollary 4. Given that $p(\gamma_y, \gamma_y) = 0$ and that $\partial p(\gamma, \gamma_y)/\partial\gamma \leq 0$ for all $\gamma < \gamma_y$, it follows that $p(\gamma, \gamma_y) \geq 0$ for all $\gamma < \gamma_y$. Thus Assumption 3iid holds. We now apply Proposition 2 to prove the corollary. □

I Examples of Densities for Corollaries 3 and 4

Lemma 4. *In the generalized Ambrus and Egorov (2015) quadratic model with $\alpha > 1$, if the density takes the linear form $f(\gamma) = a + 2(1 - a)\gamma$, with $\Gamma = [0, 1]$ and $a \in (0, 2)$, and if $1 \geq a\alpha\beta$, then all of the assumptions in Corollary 3 are satisfied.*

Proof. To begin, we show that $\frac{1}{\alpha f(\underline{\gamma})} \geq \beta$. Since $f(\underline{\gamma}) = a$, this inequality holds due to the assumption in the Lemma that $1 \geq a\alpha\beta$. By Corollary 3, it thus follows that there exists $\gamma_x \in [\underline{\gamma}, \overline{\gamma})$ such that $x(\gamma_x) = \pi_f(\gamma_x)$, where $x(\gamma) \equiv \frac{1}{\alpha}(\frac{1-F(\gamma)}{f(\gamma)} + \alpha\gamma)$. Letting $h(\gamma) \equiv \frac{1-F(\gamma)}{f(\gamma)}$, we thus have that there exists $\gamma_x \in [\underline{\gamma}, \overline{\gamma})$ satisfying $h(\gamma_x) = \alpha\beta$. We may now consider the assumptions in Corollary 3 in sequential fashion:

(i) $\frac{1-F(\gamma)}{f(\gamma)} \leq \alpha\beta$ for $\gamma \in [\gamma_x, \overline{\gamma}]$. This assumption holds, since calculations confirm that

$h'(\gamma) < 0$ for all $a \in (0,2)$ and $\gamma \in [0,1]$.

(ii) $\frac{1-F(\gamma)}{f(\gamma)} + \alpha\gamma$ is non-decreasing for $\gamma \in [\gamma_x, \overline{\gamma}]$. This assumption may be expressed as $h'(\gamma) \geq -\alpha$ for $\gamma \in [\gamma_x, \overline{\gamma}]$. If $a \leq 1$, then we find that $h''(\gamma) \leq 0$, and so in this case it suffices to show that $h'(1) \geq -\alpha$. Calculations confirm that $h'(1) \geq -\alpha$ if $\alpha \geq 1$. If instead $a \in (1,2)$, then we find that $h''(\gamma) > 0$, and so in this case it suffices to show that $h'(0) \geq -\alpha$. Calculations confirm that $h'(0) \geq -\alpha$ if $\alpha \geq 1$ and $a \in (1,2)$.

(iii) $F(\gamma) + \alpha\beta f(\gamma)$ is non-decreasing for $\gamma \in [\underline{\gamma}, \gamma_x)$. We find that

$$\frac{d}{d\gamma}[F(\gamma) + \alpha\beta f(\gamma)] = a + 2(1-a)[\gamma + \alpha\beta] \equiv z(a, \gamma + \alpha\beta).$$

Thus, we express this assumption as $z(a, \gamma + \alpha\beta) \geq 0$ for all $\gamma \in [\underline{\gamma}, \gamma_x)$. This inequality clearly holds if $a \leq 1$. If instead $a \in (1,2)$, then $z_2(a, \gamma + \alpha\beta) = 2(1-a) < 0$. When $a \in (1,2)$, it is thus sufficient to show that

$$z(a, \gamma_x + \alpha\beta) = z(a, \gamma_x + h(\gamma_x)) \geq 0.$$

The equality follows from the existence of $\gamma_x \in [\underline{\gamma}, \overline{\gamma})$ satisfying $h(\gamma_x) = \alpha\beta$. Next, since $f'(\gamma) < 0$ when $a > 1$, we may easily verify that $\gamma + h(\gamma)$ is strictly increasing for $\gamma \in [0,1)$. We may thus conclude that, when $a \in (1,2)$, $\gamma_x + h(\gamma_x) < 1 + h(1) = 1$, where we note that γ_x approaches 1 as β approaches zero. Thus, if $a \in (1,2)$, then

$$z(a, \gamma_x + h(\gamma_x)) > z(a, 1) = 2 - a > 0.$$

$\square$

Lemma 5. *In the generalized Ambrus and Egorov (2015) quadratic model with $\alpha > 1$, if the density takes the linear form $f(\gamma) = a + 2(1-a)\gamma$, with $\Gamma = [0,1]$ and $a \in [1,2)$, and if $1 < a\alpha\beta$, then asssumptions (i) and (ii) of Corollary 4 are satisfied.*

Proof. The proof is similar to the proof of parts (i) and (ii) of Lemma 4. $\square$

References

Alonso, Ricardo and Niko Matouschek, "Optimal Delegation," *The Review of Economic Studies*, 2008, *75* (1), pp. 259–293.

Amador, Manuel and Kyle Bagwell, "The Theory of Optimal Delegation with An Application to Tariff Caps," *Econometrica*, July 2013, *81* (4), pp. 1541–1600.

_ **and** _ , "Regulating a Monopolist with Unknown Costs without Transfers," Working paper, Stanford University 2014.

_ **, Iván Werning, and George-Marios Angeletos**, "Commitment vs. Flexibility," *Econometrica*, 2006, *74* (2).

Ambrus, Attila and Georgy Egorov, "Delegation and Nonmonetary Incentives," Working paper, Duke University 2015.

_ **and** _ , "A Comment on Commitment and Flexbility," *Econometrica*, forthcoming.

Armstrong, Mark and John Vickers, "A Model of Delegated Project Choice," *Econometrica*, January 2010, *78* (1), 213–244.

Athey, Susan, Andrew Atkeson, and Patrick J. Kehoe, "The Optimal Degree of Discretion in Monetary Policy," *Econometrica*, 2005, *73* (5), pp. 1431–1475.

_ **, Kyle Bagwell, and Chris Sanchirico**, "Collusion and Price Rigidity," *The Review of Economic Studies*, 2004, *71* (2), pp. 317–349.

Bagwell, Kyle and Gea M. Lee, "Advertising Collusion in Retail Markets," *The B.E. Journal of Economic Analysis and Policy*, August 2012, *10* (1).

Burkett, Justin, "Optimally Constraining a Bidder Using a Simple Budget," *Theoretical Economics*, 2016, *11* (1), 133–155.

Chakravarty, Surajeet and Todd R. Kaplan, "Optimal Allocation without Transfer Payments," *Games and Economic Behavior*, 2013, *77* (1), 1–20.

Condorelli, Daniele, "What Money Can't Buy: Efficient Mechanism Design with Costly Signals," *Games and Economic Behavior*, 2012, *75* (2), 613–24.

Frankel, Alexander, "Delegating Multiple Decisions," Working paper, Chicago Booth 2012.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani, "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 2009, *144* (4), 1397 – 1420.

Guo, Yingni, "Dynamic Delegation of Experimentation," *American Economic Review*, forthcoming.

Holmstrom, Bengt, "On Incentives and Control in Organizations." PhD dissertation, Stanford University 1977.

Koessler, Frederic and David Martimort, "Optimal delegation with multi-dimensional decisions," *Journal of Economic Theory*, 2012, *147* (5), 1850–1881.

Kováč, Eugen and Tymofiy Mylovanov, "Stochastic mechanisms in settings without monetary transfers: The regular case," *Journal of Economic Theory*, 2009, *144* (4), 1373 – 1395.

Krähmer, Daniel and Eugen Kováč, "Optimal sequential delegation," *Journal of Economic Theory*, 2016, *163*, 849–888.

Luenberger, David .G., *Optimization by vector space methods* Series in decision and control, Wiley, 1969.

Martimort, David and Aggey Semenov, "Continuity in mechanism design without transfers," *Economics Letters*, 2006, *93* (2), 182 – 189.

McAfee, R. Preston and John McMillan, "Bidding Rings," *The American Economic Review*, 1992, *82* (3), pp. 579–599.

Melumad, Nahum D. and Toshiyuki Shibano, "Communication in Settings with No Transfers," *The RAND Journal of Economics*, 1991, *22* (2), pp. 173–198.

Milgrom, Paul and Ilya Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, March 2002, *70* (2), 583–601.

Mylovanov, Tymofiy, "Veto-based delegation," *Journal of Economic Theory*, January 2008, *138* (1), 297–307.