Costly Verification and Money Burning

Rohit Patel* Can Urgun†

May 2023 ‡

Abstract

A principal designs a mechanism to allocate an indivisible, productivity increasing good to one of many agents. Monetary transfers are not allowed. Instead, we consider the interplay between two instruments studied only in isolation: "costly verification of the agent's type" and "money burning". We use a graph theoretic approach and characterize the optimal mechanism completely.

JEL Code: C72, D73, D82
Keywords: costly verification, money burning, allocation problem

*Meta, Data Science, rohit@rohit-patel.com

†Department of Economics, Princeton University, curgun@princeton.edu

‡We would also like to thank various seminar audiences, Rakesh Vohra, Eddie Dekel, Niko Matouschek, James Schummer, Alvaro Sandroni, Michael Powell, Navin Kartik, Alessandro Lizzeri, Slyvain Chassang, Leeat Yariv, Hector Chade, Andrew Ferdowsian, Ayca Kaya, Doruk Cetemen, Christopher Sandmann, Stephanie Chen and Pietro Ortoleva for helpful comments and suggestions.

1 Introduction

Consider a local branch manager (principal) of a large firm and salespeople (agents) under this manager. It is fairly common that when a salesperson makes a sale the principal and the salesperson each get a share (potentially different). Suppose the branch manager is tasked with the allocation of a productivity increasing object, say, a new workstation, among these salespeople. Each agent wants this object and has a privately known productivity increase associated with it. If an agent receives this object, both that agent and the principal will enjoy their respective shares of the increased productivity. It is reasonable to assume that the principal cannot use direct monetary transfers (i.e. sell the workstation) to elicit the agents' private information. One common approach is that the principal can at a cost audit the eligibility of these agents, say by hiring someone to go through the accounts, (costly verification) as Ben-Porath et al. (2014) analyzed. Another approach is considering various forms of money burning whether it is "buttering up the boss" as in Milgrom and Roberts (1988) or bureaucratic paperwork as in Ross (1986). In either case the agent spends time on ultimately unproductive endeavours instead of doing their work. This, in turn, is costly to both the agent and the principal. At a first glance, one might think only one of those tools is necessary: they are both costly to the principal and serve the same purpose so the principal should stick with the one that is less costly. Indeed, prior literature has also treated these instruments in isolation. However, there are some economically relevant settings such as capital financing within a firm where we see evidence of both costly verification (Harris et al. (1982)) and money burning (Ross (1986)).

In this paper, we study such allocation problems where both instruments are present and we explore whether and how they should be optimally combined. We characterize the optimal mechanism and show that it has a cutoff structure where types below the cutoff are pooled, causing allocative inefficiency among the lower types. If this cutoff is high enough, then the optimal mechanism has no money burning and only costly verification is necessary. If the cutoff is not high, then both tools are used *simultaneously for some types.* In particular, the mechanism has only money burning for the lower types. And as types increase this money burning increases at first. However, this increase plateaus when it becomes at least worthwhile to use verification on a type. For types where verification is not a guaranteed loss, the money burning remains present, but is constant across all such types. For these higher types, the verification probability is the one that is increasing instead. Thus, these tools are complementary - money burning is more effective to disincentivize lower types from lying, whereas, costly verification is more effective to disincentivize higher types.

Our mechanism has two instruments, thus our problem requires a novel approach. Instead of relying on envelope arguments, we interpret our mechanism as a directed graph. This enables us to characterize incentive constraints regardless of whether money burning, costly verification, or a combination of the two is used. We construct our graph by drawing a directed edge from type t to t' whenever the incentive constraint of type t regarding t' is tight (holding with equality). This approach also highlights the difference in incentive structures with either tool in isolation. In mechanisms with monetary transfers (e.g. standard auctions) or just money burning, each type only wants to mimic the type immediately adjacent to them, thus the graph connects each type to an adjacent one. In mechanisms with only costly verification, there is a cutoff type that wants to mimic all the types above. But for any two types above the cutoff, the incentive constraints are slack. The combination of two tools allows for more connected graphs, potentially lowering the cutoff and allowing edges between the types above the cutoff.

Research on costly verification in mechanism design is not new. Townsend (1979) models optimal contracts with costly verification. Border and Sobel (1987), Mookherjee and Png (1989) and Dunne and Loewenstein (1995) further study problems on costly state verification. Ben-Porath et al. (2014) and Erlanson and Kleiner (2019) studies allocation problem with multiple agents in the presence of costly verification techniques, and Vohra (2012) considers a discrete version of the same problem. Taking a different approach to the above Li (2020) considers a costly verification problem where there are limited punishments available to the designer. These papers, however, study the effects of costly verification in isolation. There is also considerable research on money burning in mechanism design. McAfee et al. (1992), Hartline and Roughgarden (2008), Yoon (2011), and Chakravarty and Kaplan (2013) all study optimal allocation problems with some form of money burning. Again, much of the literature in money burning deals with the problem in isolation.

2 Model and Results

2.1 Basic Setup

There are D agents denoted as $1, 2, \ldots, D$. We assume that each agent i has a type t_i drawn an $i.i.d.$ distribution F over a finite type space $t \in \{\hat{1}, \hat{2}, \ldots, \hat{n}\}$. Each type t is associated with a value $\hat{v}_t$ which captures the increase in their productivity if they are allocated the object with $\hat{v}_{t-1} < \hat{v}_t$ for all $\hat{t}$. If an agent's productivity increases by v they receive $0 < r < 1$ of the share of their increased productivity and the principal receives $0 < R < 1$ of the share, with $r + R \leq 1$. A simple interpretation is that both the agents themselves and the principal receive fixed commissions. With a slight loss in generality we will assume that each type under F has a probability that is a rational number. The agents and the principal are risk neutral. The objective of the principal is to maximize the expected value from allocation minus the costs of any assessments and the costs due to the money burning activities. Verification of a claim is perfect and costs the principal $\hat{k}$; irrespective of the type of the agent.

We assume burning p units of money by an agent is interpreted as time wasted in unproductive activities measured in the same manner as their productivity. Money burning activities can be of many forms – the principal can introduce a policy of extensive documentation for claims, or multiple rounds of meetings which would consume time and burn resources. This, we call bureaucracy. It is costly and does not have any direct output. It is dictated by the principal directly. Instead of bureaucracy, one can consider some expected lobbying or "buttering up the boss", where agents decide the amount of money burning instead, wasting productive time. In those cases, if the agents pick any level of money burning other than what the principal want, the principal can just rule them out of the running. Thus, the amount of money burned can be interpreted as either lobbying or bureaucracy or some combination of the two. The principal suffers these costs because the time wasted by the agent is directly costly in terms of foregone work: if an agent spends say ten hours, on some lobbying activity or bureaucratic paperwork, that is ten hours that the agent didn't do their job, regardless of whether they received the object or not. For example suppose the principal requires an agent with value v to burn p units of money and assigns the object after verifying their value. The agent will have a payoff of $r\hat{v} - rp$ and the principal will receive a payoff of $R\hat{v} - Rp - \hat{k}$.[1] We further assume there exists at least

[1]If money burning was costly only to the agent, clearly the principal would only use

one type t such that $R\hat{v}_t > \hat{k}$ so that verification is indeed potentially useful. Finally, without loss of generality we will assume that $\hat{v}_1 = 1$.[2]

The principal must devise a mechanism for optimal allocation of the object using costly verification (costly only to the principal) and money burning (costly to both sides). We consider the Bayes Nash Equilibrium of the mechanism. This mechanism in general could be fairly complex, involving multiple rounds of checks and money burning. Each agent wants to maximize the expected value they enjoy from having the object allocated to them minus the money burning costs. Agents have an outside option valued at 0 available to them throughout.

We focus on the following type of mechanisms:

1. The principal commits to a selection policy, an inspection policy and a money burning policy;
2. Agents report their types to the principal;
3. Following the selection policy, an agent is selected for Stage 4;
4. The selected agent burns the required money;
5. The selected agent's claim is inspected according to the inspection policy;
6. The selected agent receives the object if there is no inspection, or if the agent is found truthful upon inspection.

Both the optimality of such mechanisms and the revelation principle are simple extensions of argument in Ben-Porath et al. (2014).[3] It is easy to see that we only need truthful mechanisms as the principal can simply simulate the BNE of any complicated game after asking for the types. An agent's claim is inspected only if their project is being considered for allocation, since there is no point in spending money on inspecting an agent's claim otherwise. For maximum penalty, the agents that lie must have the object withheld from them. Money burning can't be a punishment as it is costly

money burning and the problem reduces to an auction where the objective of the designer is only maximizing allocative efficiency instead of maximizing revenue.

[2]Clearly this can be achieved by simply dividing all valuations and $\hat{k}$ by $\hat{v}_1$.

[3]A formal proof of the revelation principle and optimality can be found in the online appendix.

to the principal and agents required to burn money without a chance of allocation will simply leave. Finally, money burning must precede inspection in an optimal mechanism as inspection is fully revealing.

Before we formalize the problem further, for notational simplicity we will let $k = \frac{\hat{k}}{R}$ to denote the cost of verification in terms of the share of the principal and introduce the following transformation of the distribution and type space.

Definition 1 (Uniformization)**.** *Let F be a distribution over a type space $\{\hat{1}, \hat{2}, \dots, \hat{n}\}$ with valuations $v_{\hat{1}}, \dots, v_{\hat{n}}$. Find the smallest number, n, so that the probability of any type t can be represented as $f(t)/n$ for some $f(t)$ sharing the common denominator n. Now, replicate each type t, $f(t)$ times to construct the new type space with a total of n types. In this type space, the types $1, \dots f(1)$ have valuation $\hat{v}_1$, $f(1)+1, \dots f(2)$ have valuation $\hat{v}_2$ and so on until the types $f(\hat{n}-1) - f(\hat{n})$ have valuation $\hat{v}_{\hat{n}}$.*

Since we assumed F to have rational probabilities, a uniformization is always well defined. Clearly after uniformization we can assume the distribution over the new type space is uniform. The set of all valuations available has not changed, but in this new type space there are duplicate types if $f(t) > 1$. With a slight abuse of notation we will denote the type space after uniformization as $\{1, \dots, n\}$ and the values $V = \{v_1, \dots, v_n\}$ where the valuations are only weakly increasing.

A truthful mechanism is mapping from V^D to allocation probabilities, money burning and verification probabilities. Formally a mechanism is three functions: $p_i(t_i, t_{-i}) : V^D \mapsto \mathbb{R}_+$, $a_i(t_i, t_{-i}) : V^D \mapsto [0, 1]$, and $c_i(t_i, t_{-i}) : V^D \mapsto [0, 1]$. Here $p_i(t_i, t_{-i})$ is the amount of money burning by agent i in Stage 4, when the announced types are $\boldsymbol{t} = \{t_i, t_{-i}\}$. Since the amount of money burned cannot be negative, we impose a constraint $p_i(\boldsymbol{t}) \geq 0 \ \forall i \in D, \forall \boldsymbol{t} \in V^D$. $a_i(t_i, t_{-i})$ is the probability of the i^{th} agent being selected in Stage 3 for further review given the announced types $\boldsymbol{t} = (t_i, t_{-i})$, thus, we must have $\sum_{i=1}^{D} a_i(\boldsymbol{t}) \leq 1 \ \forall \boldsymbol{t} \in V^D$ and $a_i(\boldsymbol{t}) \geq 0 \ \forall \boldsymbol{t} \in V^D$. We have $(1 - c_i(t_i, t_{-i}))$ as the probability that a costly inspection of the selected agent is done in Stage 5, hence $0 \leq c_i(\boldsymbol{t}) \leq 1 \ \forall \boldsymbol{t} \in V^D$.

Within this setting the principal wishes to maximize:

$$R \times E_{\boldsymbol{t}} \left[\sum_{i=1}^{D} \left[v_{t_i} a_i(t_i, t_{-i}) - p_i(t_i, t_{-i}) - k \left(1 - c_i(t_i, t_{-i})\right) a_i(t_i, t_{-i}) \right] \right]$$

subject to:

$$rE_{t_{-i}}\left[v_{t_i}a_i(t_i,t_{-i})-p_i(t_i,t_{-i})\right]\geq rE_{t_{-i}}\left[v_{t_i}a_i(\hat{t}_i,t_{-i})c_i(\hat{t}_i,t_{-i})-p_i(\hat{t}_i,t_{-i})\right]\ \forall t_i,\hat{t}_i\in V$$
$$rE_{t_{-i}}\left[v_{t_i}a_i(t_i,t_{-i})-p_i(t_i,t_{-i})\right]\geq 0\ \forall t_i\in V$$

After the uniformization and rescalings of $\hat{k}$ to k clearly we can ignore R and r in the objective and the constraints. Furthermore, since the agent types are i.i.d., it is without loss to focus on symmetric mechanisms. Thus, we abuse notation to define $\mathcal{A}_t = E_{t_{-i}}\left[a_i(t_i,t_{-i})\right]$, $p_t = E_{t_{-i}}\left[p_i(t_i,t_{-i})\right]$ when $t_i = t$. Also, define $c_t = \frac{E_{t_{-i}}[c_i(t_i,t_{-i})a_i(t_i,t_{-i})]}{E_{t_{-i}}[a_i(t_i,t_{-i})]}$. Clearly, $0 \leq c_t \leq 1$ since $0 \leq c_i(t_i,t_{-i}) \leq 1\ \forall \boldsymbol{t} \in V^D$. Of course, we need to ensure the interim probabilities, $\mathcal{A}_t$, can be achieved using an allocation rule. Border (1991) provides us with a characterization of the set of feasible interim allocations. We use $\langle\mathcal{A}\rangle$ to denote the collection $\langle\mathcal{A}_t\rangle_{t\in V}$. The inequalities state that $\langle\mathcal{A}\rangle \in \mathbb{R}^n_+$ is feasible if and only if,

$$\frac{D}{n}\sum_{t\in S}\mathcal{A}_t \leq 1-\left(1-\frac{|S|}{n}\right)^D \qquad\qquad \forall S\subseteq V.$$

For any S, the inequality can be written as $\sum_{t\in S}\mathcal{A}_t \leq g(S)$ where we define $g:\ \mathcal{P}(V)\ \mapsto\ \mathbb{R}^+$ as $g(S) = \frac{n}{D}-\frac{n}{D}\left(1-\frac{|S|}{n}\right)^D$; $\mathcal{P}(\cdot)$ denotes the set of all subsets. $g(S)$ is essentially a function of the cardinality of the set S. One useful property of g is that it is strictly submodular, thus the feasible set of allocations is a *polymatroid*.[4] Furthermore, we have $p_t \geq 0\ \forall t \in V$ since $p_i(t_i,t_{-i}) \geq 0\ \forall i \in D, \forall \boldsymbol{t} \in V^n$. Let us denote by the set $\mathscr{A}$ the set of all feasible intermediate allocations that satisfy Border's inequalities. A mechanism is a tuple $M = \{\langle\mathcal{A}\rangle, \langle c\rangle, \langle p\rangle\}$ where,

$$\begin{aligned}\mathcal{A}_t &: \text{interim probability of selection of type } t\\ 1-c_t &: \text{probability of assessment if type } t \text{ is the selected type}\\ p_t &: \text{expected amount of money burning required for type } t\end{aligned}$$

[4]A formal proof is provided in the online appendix for completeness.

Using symmetry and uniformization we can write the problem as:

$$\max_{\langle \mathcal{A}\rangle, \langle c\rangle, \langle p\rangle} E_t \left[v_t \mathcal{A}_t - p_t + k\mathcal{A}_t c_t - k\mathcal{A}_t \right]$$

$$\begin{aligned}
&s.t.\\
&\langle \mathcal{A}\rangle \in \mathscr{A};\ \ 0 \le c_t p_t \le p_t\ \forall t \in V;\ c_t, p_t \in \mathbb{R}^+ && \text{(F)}\\
&v_t \mathcal{A}_t - p_t \ge v_t \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} \quad \forall t, \hat{t} \in V && \text{(IC)}\\
&v_t \mathcal{A}_t - p_t \ge 0 \qquad\qquad\quad \forall t \in V. && \text{(IR)}
\end{aligned}$$

We say that a mechanism $M = \{\langle \mathcal{A}\rangle, \langle c\rangle, \langle p\rangle\}$ is feasible if $\langle \mathcal{A}\rangle \in \mathscr{A}$ and $0 \le c_t p_t \le p_t\ \forall t \in V$; $c_t, p_t \in \mathbb{R}_+$. A mechanism is said to be incentive compatible if it satisfies all IC constraints.

2.2 Graph Theoretic Formulation

In much of standard mechanism design, tight incentive constraints is enough to simplify the problem to an allocation problem and the usage of tools (such as required payments) are simply deduced from these allocation probabilities. Suppose the constraint between types s and t are tight. Indeed, in standard auctions, with a slight abuse of notation letting $p_{\hat{t}}$ denote the expected payment for type $\hat{t}$, the allocation probabilities pin down the payments entirely up to a constant, via the incentive condition, $v_s \mathcal{A}_s - p_s = v_s \mathcal{A}_t - p_t$. So the standard approach boils down to first narrowing down which set of constraints are tight via an assumption of monotone hazard ratio and proceed to solve an optimal allocation problem given these tight constraints. Similarly in the problem with only costly verification explored in Ben-Porath et al. (2014) and Vohra (2012) the allocation probabilities pins down the probability of verification entirely via the incentive condition, $v_s \mathcal{A}_s = v_s \mathcal{A}_t c_t$. In our setting however, even if the allocation probabilities and the set of tight incentive constraints are known, the verification or money burning cannot be simply backed out. More importantly even narrowing down the set of tight incentive constraints ex ante (such as limiting them to only the local ones) seems difficult.

We tackle these difficulties by interpreting a mechanism M as a directed graph. Let G_M be the directed graph that represents the mechanism M. Let $E(G_M)$ be the set of edges where $(s, t) \in E(G_M)$ means that there is a directed edge originating at s and pointing towards t in G_M. Since all graphs we consider have the same vertex set, we simply use V to denote the set of vertices. We say that $(s, t) \in E(G_M)$ if the incentive constraint for type s

pretending to be type t is tight. That is $(s,t) \in E(G_M)$ if and only if

$$v_s \mathcal{A}_s - p_s = v_s \mathcal{A}_t c_t - p_t$$

We call a graph G_M associated with an optimal mechanism M an optimal graph. Certainly such a graph can be constructed for any mechanism design problem. For example, in case of an optimal auction with *i.i.d.* types, we know that allocation probabilities are increasing and each type t only has a binding constraint for the type immediately above it $t+1$ as depicted below.

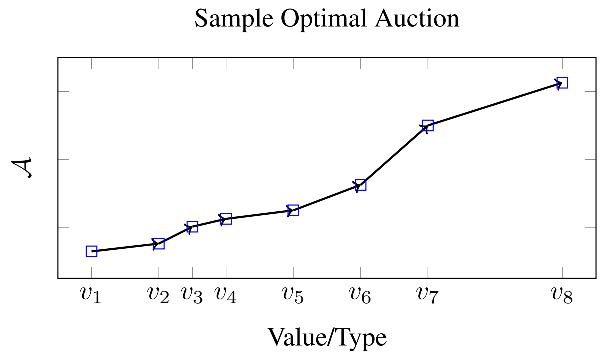

Figure 1: Sample optimal auction with no reservation price

Similarly let us depict the problem with only costly verification, explored in Ben-Porath et al. (2014) and Vohra (2012). With *i.i.d.* types, the optimal mechanism involves a cutoff, below which the types are pooled and not subject to verification and all the binding constraints involve this cutoff. Notably, there are no binding constraints between higher types as depicted below.

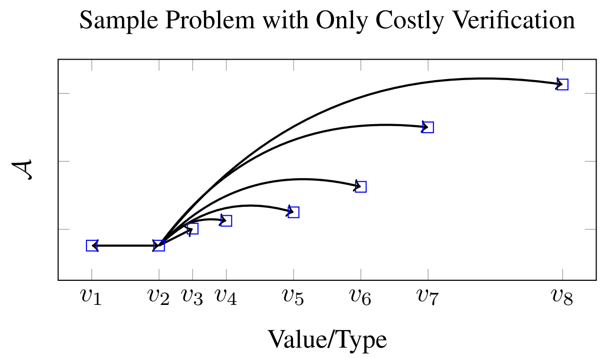

Figure 2: Sample optimal allocation mechanism with only costly verification

However, in the general problem as we will soon show, the set of edges could be quite complicated. Relying only on incentive constraints directly will not help us to proceed even under stringent assumptions on the distribution of the types and valuations. Instead, we will rely on *mutually binding incentive constraints* and optimality. A set of types S is said to have mutually binding constraints if for any all pairs $s, t \in S$, t is indifferent between telling their true type and s and s is indifferent between telling their true type and t. This does not rule out some types outside S that might want to mimic types in S, or the types in S wanting to mimic types outside S, but in such cases such constraints are one-sided. In environments with just money burning (also in much of standard mechanism design such as auctions) mutually binding incentive constraints is equivalent to pooling; the types are treated identically. With only costly verification, the types need to have the same allocation probability and have no verification, thus pooling is necessary but not sufficient to conclude mutually binding incentive constraints.[5] With two instruments, types might want to mimic each other despite being treated differently, thus pooling is neither necessary nor sufficient. A set of types with mutually binding incentive constraints between them correspond to a special structure on a graph called a *web*.

Definition 2. *We say that $W \subseteq V$ is a* web *in G_M iff $(s,t) \in E(G_M)$ $\forall s, t \in W$ and there do not exist r, s such that $r \in V \backslash W$; $s \in W$ and $\{(r,s),(s,r)\} \subseteq E(G_M)$. A web W is* non-trivial *if $|W| > 1$.*

In essence, a web is a set of types where every node points to every other node in W, and no more nodes can be added to it while still having every node pointing to every other node. A web necessarily contains directed cycles, but directed cycles need not contain webs.

2.3 Results

Our first result says that by using *optimality* and *mutually binding incentive constraints* we can drop all the downward facing constraints from the problem and it remains unchanged – that we can assume no agent will ever devalue their claim unless they are pooled *and* in a web.

[5] Indeed one can have two types with the same allocation probability and the same, strictly positive probability of verification, these types are treated identically hence they are pooled, but do not wish to mimic each other.

Proposition 1. *In an optimal mechanism, agents are strictly better off by announcing their true value than any other lower value unless the types are pooled in a web, i.e., if $(s,t) \in E(G_M)$ then either $v_s < v_t$ or there exists a web W such that $\{s,t\} \subseteq W$.*

The proof of this result relies on a couple of Lemmas that capture the properties of an optimal mechanism and their relation to webs. In the appendix, we first show that optimality will require that a directed cycle in an optimal mechanism must be a web and in a web with more than one type all the types have the *same amount of money burning and same chance of allocation and no verification.* (Lemmas 4, 5). Furthermore, an optimal graph must be connected and contain exactly one web W^* - the types(nodes) in the web requiring no money burning. (Lemmas 5, 6). Now, if there was a type that was downward pointing, then it would have to be part of some directed cycle, which would mean it would be in a web where the types are pooled.

Our next goal is to understand a very basic amount of monotonicity in the mechanism. Money burning can signal higher values, but it is commonsensical that if the true value will be learned by inspection, that takes away the need to burn money. This suggests that money burning activities must be related to higher probability of allocation, but without assessment as we next show.

Lemma 8. *In an optimal mechanism, the amount of money burning is strictly monotonic in the probability of project implementation without an assessment, i.e., in an optimal M, $\mathcal{A}_t c_t > \mathcal{A}_s c_s \iff p_t > p_s$ and $\mathcal{A}_s c_s = \mathcal{A}_t c_t \iff p_s = p_t$.*

The second part of the lemma is worthy of special mention when money burning is not allowed in the mechanism. Indeed, in the problem with no money burning for every pair of types t, s we have $\mathcal{A}_s c_s = \mathcal{A}_t c_t$. For the problem without money burning indeed just this observation is enough to tease out the rest of the mechanism.

The next bit of monotonicity is about allocation probabilities. Higher types have higher allocation probabilities. This is intuitive as the principal also benefits from allocative efficiency. Furthermore, since we have shown that only the upward constraints are binding, we would expect higher types to have higher incentive rents, hence higher expected utilities.

Proposition 2. *In an optimal mechanism higher types have higher interim allocation probabilities, i.e., $v_t > v_{\hat{t}} \implies \mathcal{A}_t \geq \mathcal{A}_{\hat{t}}$. Furthermore, higher types also have higher expected payoffs, i.e., $v_t > v_{\hat{t}} \implies v_t \mathcal{A}_t - p_t \geq v_{\hat{t}} \mathcal{A}_{\hat{t}} - p_{\hat{t}}$.*

For proving these result we show that if the proposition is not true, then the higher type has a positive inspection probability which eventually leads to a contradiction of *optimality.* It is noteworthy that monotonicity results such as the ones above only requires *incentive compatibility* in much of the classical mechanism design literature. However, with two tools incentive compatibility is not enough and we must try to characterize what is the optimal use of the incentive tools available to the principal.

Propositions 1, 2 are very desirable and intuitive results, but still more structure is necessary to back out verification probabilities and money burning. Notably, the binding constraints are always upward looking, but they need not be local, and in fact there might be multiple binding constraints involving a single type and the only information we have is that the payoffs are weakly increasing. Additionally we still have not addressed whether money burning is a substitute or a complement to verification.

Proposition 3. *Every optimal mechanism M is of the form where $\exists\ \underline{t} \in V$ such that for all types with $v_t \leq v_{\underline{t}}$ there is no verification and no money burning. Furthermore, for all the types above this cutoff, i.e. $v_t > v_{\underline{t}}$, there is either a positive probability of verification, or some money burning, or both. Also, all the types below this cutoff have the lowest probability of implementation, i.e., $\mathcal{A}_t = \min_{r \in V} \mathcal{A}_r\ \forall t \leq \underline{t}$.*

Since lower types have a lower chance of allocation we do not expect very low value types typically to attract an assessment or a lot of money burning. This intuition is formalized as a cutoff and types below the cutoff have a fixed probability of allocation. No money burning is required for these low value types, furthermore, no verification is also used on these types.

The proof of this result is also closely tied to the proof of proposition 1 and the Lemmas leading to it. Notably, it was necessary to show that in every optimal mechanism there was a unique web, W^* which had money burning. But we also know that any set of nodes grouped in a web must have the same allocation probability, and no verification. In the web W^* there is also no money burning, hence it is commonsensical that these types will be grouped at the at the bottom identifying a cutoff structure.

Here the structure of the problem is worth noting. In principle our problem is a polymatroid optimization problem with incentive constraints. The issue is that incentive constraints intrinsically have the allocation inside them. Thus we will first focus on finding the optimal use of the incentives tools for a given allocation then solve for the optimal allocation. Such a step is clearly not necessary in much of the standard mechanism design literature as the local incentive constraints would pin down the usage of the tools directly from the allocation.

We have so far shown that an optimal mechanism includes an endogenous cutoff type $\underline{t}$. However, in our problem set-up there is also an exogenous cutoff, the types that are not "worth" verifying, which is tied to the cost of verification.

Definition 3. *Efficient verification level, t^{EF} is the smallest type where the value exceeds the cost of verification. Formally, $t^{EF} = \min\{t \in \{1, \ldots, n\} : v_t \geq k\}$.*

How these two cutoffs (one endogenous, one exogenous) interact turn out identify the use of the optimal tools, for any allocation probabilities that satisfies the properties we have so far identified. Going forward we will refer to the types equal or below t^{EF} as "lower types" and types above t^{EF} as "higher types".

Proposition 4. *Let $\underline{t}$ be the cutoff type in an optimal mechanism and let $v_{\underline{t}} > t^{EF}$. Then $p_t = 0$ for all t and $c_t = \frac{\mathcal{A}_{\underline{t}}}{\mathcal{A}_t}$ for all $t \geq \underline{t}$.*

To glean some intuition first observe that if there is a type that has money burning, no type below the cutoff can point to this type. The reasoning is simple, if that was the case then the cutoff type's incentive constraint would be violated. Since we consider the case where the cutoff type itself is a higher type, then the only incentive constraints are between higher types. These types are willing to burn considerably more compared to lower types money since their valuation is already high. Hence, verification (and thus the threat of withholding) is a much more efficient method to induce higher types to be truthful.

This result is also in line with the results of Ben-Porath et al. (2014) and Vohra (2012). Indeed when money burning is not allowed at all the optimal mechanism is identified by a cutoff type and the verification probabilities are exactly pinned down by the ratio of allocation probabilities as we had seen in Lemma 8.

On the other hand when the cutoff is below t^{EF} money burning must be part of an optimal mechanism. In fact, as we will show, for all types below t^{EF} that is above the cutoff that is the only tool that is used. This is intuitive; trying to verify these types is not worth the cost. Additionally, the required money burning increases monotonically until it reaches a maximum level at t^{EF}. Formally we have:

Proposition 5. *Let $\underline{t}$ be the cutoff type in an optimal mechanism and let $v_{\underline{t}} < t^{EF}$. For all $t \leq t^{EF}$, $c_t = 1$, $p_t = v_{t-1}(\mathcal{A}_t - \mathcal{A}_{t-1}) + p_{t-1}$, with $p_{\underline{t}} = 0$. Additionally, if $v_{t-r} = v_{t-1}$ for any $r \leq t-1$, then, $p_t = v_{t-r}(\mathcal{A}_t - \mathcal{A}_{t-r}) + p_{t-r}$.*

Building upon the above result our next aim is to consider the usage of incentive tools for types above t^{EF}. For these types verification is not a guaranteed loss as the cost to verify is below their value. Indeed, the principal starts using verification for these types. One might potentially consider whether the principal would like to substitute away from money burning completely for these types since verification is perfect. However, as we show below, the types above t^{EF} have both money burning and costly verification.

Proposition 6. *Let $\underline{t}$ be the cutoff type in an optimal mechanism and let $v_{\underline{t}} < t^{EF}$. Then for all $t > t^{EF}$ $p_t = p_{t^{EF}}$ and $c_t = \frac{\mathcal{A}_{t^{EF}}}{\mathcal{A}_t}$*

Beside providing a complete characterization of the usage of tools for a given low enough cutoff value, the above result also delivers surprising monotonicity properties. Money burning caps out, it is fixed at the level of t^{EF} but the verification probabilities increase monotonically. For higher types money burning is relatively insignificant compared to the loss from withholding resulting from a failed verification. Thus for maintaining incentive constraints between high types it is the more efficient option. However, lower types still need to be deterred from declaring higher types thus the money burning remains present but does not increase. Notably, instead of substituting away from money burning, verification is used to complement money burning.

The last parts of Propositions 5 and 6 provides the information that we need in order to go from our uniformization via replication back to the original distribution. Recall that we had the type space weakly ordered where there were potentially identical types. A relatively obvious observation is the following:

Lemma 12. *In an optimal mechanism if $v_t = v_{t-r}$ for any $1 < r < t \leq t^{EF}$ then $\mathcal{A}_t = \mathcal{A}_r$. In addition there exists an optimal mechanism where if $v'_t = v_{t'-r'}$ for any $1 < r' < t'$ then $\mathcal{A}_{t'} = \mathcal{A}_{r'}$.*

The implication here is that any duplicate type will indeed be treated identically when they are only subject to money burning. Recall that types in a web have the same probability of allocation, no verification and the same amount of money burning. It turns out pooling these duplicate types when they belong to a web is necessary. Remarkably, it is without loss to pool higher duplicated types as well, but it is not necessary since these types do not belong to a web.

Since it is without loss to gather higher types, going forward we will assume that is indeed the case. Since we have fully characterized the usage of the instruments, the remaining part of the problem is the characterization of the allocation probabilities. Given Lemma 12 it is easier to focus on the allocation probabilities in the original type space without duplication.

Notice that in the first best, where the principal could see the realized type, the highest realized type would get the object. In terms of interim probabilities, this corresponds to a greedy solution to the polymatroid problem starting from the highest type ending in the lowest type with no withholding.

Proposition 7. *Let $\{\hat{1}, \ldots, \hat{n}\}$ denote the type space with a type $\hat{t}$ having $f(\hat{t})$ replications. In any optimal mechanism, for a given cutoff $\underline{\hat{t}}$, the probability of allocation for all types represented in the original type space are given by*

$$\begin{aligned}
\mathcal{A}_{\hat{n}} &= g(f(\hat{n})) \\
\mathcal{A}_{\hat{n}-1} &= g\left(f(\hat{n}) + f(\hat{n}-1)\right) - g(f(\hat{n})) \\
&\vdots \\
\mathcal{A}_{\hat{t}} &= g\left(\sum_{t=\hat{t}}^{\hat{n}} f(t)\right) - g\left(\sum_{t=\hat{t}+1}^{\hat{n}} f(t)\right) \\
\mathcal{A}_{\underline{\hat{t}}} &= \frac{g\left(\sum_{t=\hat{1}}^{\hat{n}} f(t)\right) - g\left(\sum_{t=\underline{\hat{t}}+1}^{\hat{n}} f(t)\right)}{\sum_{t=\hat{1}}^{\underline{\hat{t}}} f(t)}
\end{aligned}$$

With monetary transfers from the agents to the principal withholding and thus allocative inefficiency is clearly not necessary. The result above shows that a combination of costly verification and money burning is not enough to compensate for the lack of transfers. There is allocative inefficiency, restricted only to the types below the cutoff. Technically, the solution above the cutoff remains "greedy" despite the incentive constraints.

Many mechanism design problems beyond this setup fall into a category of problems called a polymatroid optimization with side constraints (see Border

(1991), Vohra (2012)) and the methodology of tackling them and how they relate to polymatroid optimization with *incentive constraints* might be of independent interest. Thus, let us briefly sketch the argument for the proof of Proposition 7. Proposition 1 shows that an optimal graph has a minimal web W^* that defines a cutoff type $\underline{t}$. Given a cutoff, we first take $\langle \mathcal{A}c, p\rangle$ (neither being necessarily the optimal ones we derived above) as given and ask what the optimal $\langle \mathcal{A}\rangle$ is. Notice changing $\langle \mathcal{A}\rangle$ while taking $\langle \mathcal{A}c\rangle$ as given is equivalent to changing c. Due to Proposition 1 for each potentially binding IC regarding $t' \geq t$, we can let $d_{t,t'} = \mathcal{A}_{t'}c_{t'} - \frac{p_{t'}}{v_t} + \frac{p_t}{v_t}$ and let $d_t = \max_{t' \geq t}\{d_{t,t'}\}$. Define the associated polytope $\mathcal{P}(\mathcal{A}c, p) = \{\sum_{i \in S} \mathcal{A}_i \leq g(S), S \subseteq V; \mathcal{A}_i \geq d_i,\ i \in V\}$. The optimal $\mathcal{A}'s$ given $\mathcal{A}c, p$'s solve the following problem

$$\max_{\langle \mathcal{A}\rangle \in P(\mathcal{A}c,p)} \sum_{t=1}^{n} v_t \mathcal{A}_t$$

This is a linear optimization over a polymatroid identified by the set function $g(S)$ and side constraints d_i. By Proposition 2 and feasibility we can restrict attention to $d_i \leq g(1)$'s that are weakly increasing. Clearly if $(d_1, \ldots, d_n) \notin \mathcal{P}(\mathcal{A}c, p)$ then the problem is not feasible. If the problem is feasible, by proposition 2.2 of Fujishige (1980) we can consider *the contraction of the polymatroid with respect to the constraints* d_i, which in turn also defines a polymatroid with the set function defined as $\rho(S) = \min_{S' \supseteq S}(g(S') - \sum_{i \in S'} d_i)$. In words, *a contraction of the polymatroid is a polymatroid itself.* This polymatroid consists of all the vectors $(\mathcal{B}_i)_{i \in \{1,\ldots,n\}}$ such that $(\mathcal{B}_i + d_i)_{i \in \{1,\ldots,n\}}$ is feasible in the original polymatroid. The monotonicity we acquired of the $d_i's$ being weakly increasing turns out to be enough to identify the structure of ρ. This new polymatroid problem then can be simply solved by a greedy algorithm starting from the highest type since there are no more side constraints. Given the form of the contraction from $\mathcal{B}$'s we can back out $\mathcal{A}$'s for the given $\mathcal{A}c, p$. We then show that if the resulting $\mathcal{A}$'s are not of the stated form $\mathcal{A}c, p$ can not be optimal.

At this point it is also important to remark that we could have pinned down the allocation problem using only the weak monotonicity of utilities (Proposition 2) and assuming that duplicate types are pooled. Indeed, directly jumping to this step would reduce the choice of incentive tools to an application of the Simplex algorithm. Nonetheless in the present setting the result of the Simplex would be hard to gauge hence we formalized the maximization earlier.

Since we have pinned down the mechanism completely depending on the cutoff, the final piece in the optimal mechanism is to deliver a characterization of the optimal cutoff. Towards this end it will be convenient to introduce a final bit of notation. We will let $\bar{\mathcal{A}}_{\hat{t}}$ denote the allocation probability for type $\hat{t}$, when $\hat{t}$ is the cutoff itself, as we want to change the cutoff.

Theorem 1. *For every original typeset $\{\hat{1}, \ldots, \hat{n}\}$ with the optimal cutoff with $\mathcal{A}_{\hat{t}}$ given in Proposition 7. And further let $\bar{\mathcal{A}}_{\hat{t}} = \frac{g\left(\sum_{t=\hat{1}}^{\hat{n}} f(t)\right) - g\left(\sum_{t=\hat{t}+1}^{\hat{n}} f(t)\right)}{\sum_{t=\hat{1}}^{\hat{t}} f(t)}$. Then the optimal cutoff solves the following:*

$$\arg\max_{\underline{\hat{t}}} \Bigg\{ \max_{\underline{\hat{t}} \geq t^{EF}} \bar{\mathcal{A}}_{\underline{\hat{t}}} \left(\sum_{t=1}^{\underline{t}} v_t - \sum_{\underline{\hat{t}}+1}^{n} k \right) - \frac{1}{D} \sum_{t=1}^{n} v_t + \sum_{t=\underline{\hat{t}}+1}^{n} \mathcal{A}_t v_t,$$

$$\max_{1<\underline{t}<t^{EF}} \bar{\mathcal{A}}_{\underline{t}} \Big(\sum_{t=1}^{\underline{t}-1} v_t + \sum_{t=\underline{t}}^{n} v_{\underline{t}} \Big) - \sum_{t=2}^{\underline{t}} \mathcal{A}_t (v_t - v_{t-1})(n-t+1) - \frac{\left(\frac{1}{n}\right)^{D-2}}{D} \Bigg\}$$

Notice that the above characterization doesn't impose any restriction on neither the state space nor the distribution. Indeed when the state space consists of consecutive values the second maximization simplifies dramatically. However as we have shown above such a restrictions are not necessary for identifying the characterizations of the mechanism in general.

2.3.1 Comparative Statics

As Theorem 1 suggest, broad comparative statics without stringent assumptions are not possible in this setup. There two main difficulties that limit the analysis. First, the presence of $v_l - v_{l-1}$ in the objective of Theorem 1 means that simple first order or second order stochastic dominance shifts are generally inconclusive. A second difficulty is the polymatroid base. One can show for example that $\mathcal{A}_t$ terms are all decreasing in D, but this decrease is also non-monotone. Hence, the effect of a change in $\underline{t}$ in the objective is hard to determine. The last parameter, cost of verification is amenable to some analysis, but also highlights the subtleties involved.

Proposition 8. *The optimal cutoff $\underline{t}$ is increasing in the cost of verification k. Consequently, the probability of verification, $1 - c_t$, is weakly decreasing in k for all t.*

The intuition for the result is slightly subtle; since the principal can simply substitute away from verification into money burning without changing allocation probabilities when the cost of verification increases. However, notice that an increase in k necessarily increases t^{EF}. Now suppose the principal does not change the cutoff value. If the original mechanism entailed money burning then by Propositions 5 and 6 an increase in t^{EF} results in an increase in the money burning required for all types that are now above t^{EF}, whereas the other types money burning remains the same. So the principal might want to reduce the money burning required to improve and the optimal way to achieve this is by increasing the cutoff. If the original mechanism had no money burning, then if it still has no money burning, just the cost of verification for all types have increased again the principal might want to reduce the costs related to verification and hence increase the cutoff. Finally if the original mechanism had no money burning, now it might be necessary to introduce money burning for all types above the cutoff level. Similar to the first case the principal aims to curb the money burning by increasing the cutoff. Once the cutoff being weakly increasing is established the verification probabilities decreasing follows simply from the characterizations in Propositions 4 and 6. So, we can uniformly say that as the cost of verification increases, the amount of verification in the mechanism weakly decreases.

In light of Propositions 5, 6 and 8. One might be inclined to think that the money burning is also weakly decreasing in k since the cutoff is weakly increasing. Indeed, for fixed k, a higher cutoff results in weakly lower payments for all types. Additionally, for all types below t^{EF} the money burning required also weakly decreases simply from Proposition 5. However for higher types Proposition 6 implies that their money burning is given by $\sum_{l=\underline{t}}^{t^{EF}-1} v_l(\mathcal{A}_{l+1} - \mathcal{A}_l)$. Since both t^{EF} and $\underline{t}$ is increasing, the whether the principal substitutes away from or substitutes into money burning might vary.

3 Conclusions and Discussion

Before we conclude let us consider a simple example that highlights the characteristics and the incentives at play.

Example 1. *Consider three agents that can take any one of the three types* $V = \{1, 6, 10\}$ *each with probability* $1/3$ *and let the cost of inspection be* $k = 2$. *Table 1 represents an optimal mechanism with both instruments on the first column, only verification in the second column and only money burning in the third column.*

	both			only verification		only money burning	
v_t	$\mathcal{A}_t$	$1-c_t$	p_t	$\mathcal{A}_t$	$1-c_t$	$\mathcal{A}_t$	p_t
1	0.04	0	0	0.15	0	0.04	0
6	0.26	0	0.22	0.15	0	0.26	0.22
10	0.70	0.63	0.22	0.70	0.78	0.70	2.86

Table 1: An Optimal Mechanism

In the example since the underlying distribution is already uniform we do not need to duplicate any types. Consider the type $t = 10$. We note that the type is inspected in many states of the world, and yet must burn money at the same time. This happens because the two tools act as complements for each other when enforcing the incentive compatibility constraints. If only money burning were to be used, a great deal of money will need to be burned to deter the closer type $t = 6$ from announcing it's type as type $t = 10$. Verification is a much cheaper alternative when trying to deter $t = 6$ from announcing $t = 10$. However, a very small amount of money burned can deter $t = 1$ from announcing $t = 10$. On the other hand if only verification were used then $t = 6$ needs to be pooled with $t = 1$ causing allocative inefficiency, as separating the three types would require very high verification probabilities, namely $1 - c_6 = 0.84$ and $1 - c_{10} = 0.94$. Therefore, a combination of the two proves to be the most efficient option.

We develop a model with minimal assumptions - we don't even require monotone hazard rate condition (and given the lack of an envelope argument it is not certain that it will be directly useful). We allow the principal to use costly inspection and observable money burning activities to allocate optimally. From a mechanism design perspective, we see that when money burning activities are not available there are tight incentive conditions between the lowest value projects and the highest value projects. Minimizing the cost of inspection regarding these conditions might cause some projects to be implemented with less than ideal probabilities. Money burning activities enable the principal to disincentivize the lowest types from mimicking very high types thus might help the principal to avoid unnecessarily low probabilities of implementation for valuable projects. Going back to our motivation the perception that these money burning activities are just that - money burning - and somehow despite destroying surplus, it might be a useful instrument in lowering incentive rents that is not directly substitutable.

On the technical front we tackle a mechanism design problem with multiple tools via a graph theoretic approach. The applicability of a similar approach to other mechanism design problems is of particular interest. Our results relied on no assumptions on the distributions of types, yet a full characterization was achievable. Traditionally mechanism design does not require complicated optimizations to tackle the use of incentive tools, as the envelope theorem simplifies the analysis greatly. We have shown that when such an envelope argument is not present objects that are previously not well studied such as mutually binding incentive constraints and their interaction with maximization might be a fruitful endeavour. Additionally, the translation of polymatroid with side constraints to an associated polymatroid without constraints is a potentially useful technique that can be applied to broader problems. Notably, allocation of multiple objects with incentive constraints have a similar polymatroid base and thus our techniques could be applicable. We hope to address this particular application in future work.

4 Appendix

Definition 4 (Transform of a Mechanism)**.** *M^* is a transform of the mechanism M denoted as $M^* = T_{M,r,s}(\mathcal{A}'_r, \mathcal{A}'_s, c'_r, c'_s, p'_r, p'_s)$. If $M^* = \{\langle \mathcal{A}^* \rangle, \langle c^* \rangle, \langle p^* \rangle\}$ then $\mathcal{A}^*_r = \mathcal{A}'_r$; $\mathcal{A}^*_s = \mathcal{A}'_s$; $c^*_r = c'_r$; $c^*_s = c'_s$; $p^*_r = p'_r$; $p^*_s = p'_s$; $\mathcal{A}^*_t = \mathcal{A}_t \ \forall t \notin \{r, s\}$; $c^*_t = c_t \ \forall t \notin \{r, s\}$; $p^*_t = p_t \ \forall t \notin \{r, s\}$.*

Definition 5 (In-Out Sets)**.** *For any $t \in V$ let $deg^-_M(t)$ denote the indegree of t and let $deg^+_M(t)$ denote the outdegree of t in G_M. The in-set of type t is $i_M(t) = \{r : (r, t) \in E(G_M)\}$ out-set of type t is $o_M(t) = \{r : (t, r) \in E(G_M)\}$. For any $S \subseteq V$ in-set of S is $i_M(S) = \{t \in V \backslash S : (t, s) \in E(G_M)$ for some $s \in S\}$.*

Lemma 1. *Let $\langle \mathcal{A} \rangle \in \mathscr{A}$ and $\mathcal{A}_t \geq \mathcal{A}_s$ for some s, t, then $\exists\ \delta > 0$ such that $\langle \mathcal{A}' \rangle \in \mathscr{A}$, where $\mathcal{A}'_t = \mathcal{A}_t - \delta$, $\mathcal{A}'_s = \mathcal{A}_s + \delta$ and $\mathcal{A}'_r = \mathcal{A}_r \ \forall r \notin \{s, t\}$.*

Proof. Consider an arbitrary $S \subseteq V \backslash \{s, t\}$. Let $S_t \equiv S \cup \{t\}$, $S_s \equiv S \cup \{s\}$. Since $\mathcal{A}_s \leq \mathcal{A}_t$ we can write

$$\sum_{r \in S_s} \mathcal{A}_r \leq \frac{1}{2} \left[\sum_{r \in S} \mathcal{A}_r + \sum_{r \in S_t \cup S_s} \mathcal{A}_r \right]$$

We need to show,

$$\sum_{r \in S'} \mathcal{A}'_r \leq g(S') \qquad \forall S' \subseteq V$$

For the sets that do not contain s, or contain both s, t the inequality holds trivially. Since S is arbitrary and δ can be arbitrarily small, we need to show:

$$\sum_{r \in S_s} \mathcal{A}_r < g(S_s)$$

Since g is strictly submodular

$$\begin{aligned}
& g(S_s \cup S_t) + g(S) < g(S_s) + g(S_t) \\
\implies \quad & \sum_{r \in S} \mathcal{A}_r + \sum_{r \in S_t \cup S_s} \mathcal{A}_r < g(S_s) + g(S_t) \\
\implies \quad & 2 \sum_{r \in S_s} \mathcal{A}_r < g(S_s) + g(S_t) \\
\implies \quad & \sum_{r \in S_s} \mathcal{A}_r < g(S_s) \qquad \because g(S_s) = g(S_t)
\end{aligned}$$

□

Lemma 2. *In an optimal M if $deg^-_M(t) = 0$, then $p_t = 0$; $c_t = 1$ and $\mathcal{A}_t = \min_{r \in V} \mathcal{A}_r$*

Proof. If $p_t > 0$ then define $M' = T_{M,s,t}(\mathcal{A}_s, \mathcal{A}_t, c_s, c_t, p_s, p_t - \delta)$ for small $\delta > 0$ then M' is feasible and a δ improvement. If $c_t < 1$ then define $M' = T_{M,s,t}(\mathcal{A}_s, \mathcal{A}_t, c_s, c_t + \delta, p_s, p_t)$ for small $\delta > 0$ then M' is feasible and a $k\mathcal{A}_t\delta$ improvement. If $\mathcal{A}_t > \mathcal{A}_s$ for some s we have $v_s\mathcal{A}_s - p_s < v_s\mathcal{A}_t$ since $p_s \geq 0$, thus violating the IC. $\square$

Lemma 3. *In an optimal M,* $\min_s \mathcal{A}_s c_s \geq \min_s \mathcal{A}_s$

Proof. Suppose not. Consider r, t be such that $\mathcal{A}_r c_r = \min_s \mathcal{A}_s c_s$ and $\mathcal{A}_t = \min_s \mathcal{A}_s$ then $\mathcal{A}_r c_r < \mathcal{A}_t$, and $c_r < 1$. Now, we must have $deg^-_M(r) = 0$ otherwise there must be an r' such that $(r', r) \in E(G_M)$ but then $v_{r'}\mathcal{A}_t > v_{r'}\mathcal{A}_r c_r - p_r = v_{r'}\mathcal{A}_{r'} - p_{r'}$ violating IC. However, if $deg^-_M(r) = 0$ then $c_r = 1$ due to Lemma 2 a contradiction. $\square$

Lemma 4. *In an optimal mechanism M if $(t, \hat{t}) \in E(G_M)$ then $\mathcal{A}_{\hat{t}} c_{\hat{t}} \geq \mathcal{A}_t$*

Proof. Suppose not, and let $\mathcal{A}_{\hat{t}} c_{\hat{t}} < \mathcal{A}_t$, then it must be that $p_t > p_{\hat{t}} \geq 0$. Consider

$$M' = T_{M,t,\hat{t}}(\mathcal{A}_{\hat{t}} c_{\hat{t}}, \mathcal{A}_{\hat{t}}, c_t, c_{\hat{t}}, p_{\hat{t}}, p_{\hat{t}})$$

M' is feasible. Incentive constraints not involving the type t remain unaffected. For $r \in o_M(t)$, we have

$$\begin{aligned}
v_t\mathcal{A}'_t - p'_t &= v_t\mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} \\
&= v_t\mathcal{A}_t - p_t && \because (t, \hat{t}) \in E(G_M) \\
&\geq v_t\mathcal{A}_r c_r - p_r \quad \forall r \neq t \\
&= v_t\mathcal{A}'_r c'_r - p'_r
\end{aligned}$$

For $r \in i_M(t)$ we have

$$v_r\mathcal{A}'_r - p'_r = v_r\mathcal{A}_r - p_r \geq v_r\mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} \geq v_r\mathcal{A}_{\hat{t}} c_{\hat{t}} c_t - p_{\hat{t}} = v_r\mathcal{A}'_t c'_t - p'_t$$

Thus, all incentive constraints are satisfied. The objective is increased by,

$$Obj(M') - Obj(M) = k(1 - c_t)(\mathcal{A}_t - \mathcal{A}_{\hat{t}} c_{\hat{t}})$$

This is positive if $c_t < 1$ giving a contradiction. If $c_t = 1$ then we have $deg^-_M(t) > 0$ otherwise Lemma 2 and Lemma 3 jointly give us a contradiction since $\mathcal{A}_t > \mathcal{A}_{\hat{t}} c_{\hat{t}}$. Since $deg^-_M(t) > 0$ let $r' \in i_M(t)$ and $r' < t$ then

$$\begin{aligned} & v_t \mathcal{A}_t - p_t = v_t \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} \\ \implies \quad & v_{r'}(\mathcal{A}_t - \mathcal{A}_{\hat{t}} c_{\hat{t}}) < p_t - p_{\hat{t}} && \because \mathcal{A}_t > \mathcal{A}_{\hat{t}} c_{\hat{t}};\ t > r' \\ \implies \quad & v_{r'} \mathcal{A}_{r'} - p_{r'} < v_{r'} \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} && \because c_t = 1;\ (r, t) \in E(G_M) \end{aligned}$$

Again, giving us a contradiction. So we must have $r' \geq t$ for all r' such that $(r', t) \in E(G_M)$.

Given above define $M' = T_{M,t,\hat{t}}(\mathcal{A}_t - \delta, \mathcal{A}_{\hat{t}}, c_t, c_{\hat{t}}, p_t - v_t\delta, p_{\hat{t}})$. M' is feasible for small δ since $p_t > 0$. For any r' such that $(r', t) \in E(G_M)$ we have

$$\begin{aligned} & v_{r'} \mathcal{A}_{r'} - p_{r'} = v_{r'} \mathcal{A}_t - p_t \\ \implies \quad & v_{r'} \mathcal{A}_{r'} - p_{r'} \geq v_{r'} \mathcal{A}_t - v_{r'}\delta - p_t + v_t \delta \\ \implies \quad & v_{r'} \mathcal{A}_{r'} - p_{r'} \geq v_{r'}(\mathcal{A}_t - \delta) - (p_t - v_t \delta) \\ \implies \quad & v_{r'} \mathcal{A}_{r'} - p_{r'} \geq v_{r'} \mathcal{A}'_t - p'_t \end{aligned}$$

Hence M' is incentive compatible and optimal(since M was). Moreover, $deg^-_{M'}(t) = 0$ and hence we have $\mathcal{A}'_t \leq \mathcal{A}_{\hat{t}} c_{\hat{t}}$ due to lemma 2 and lemma 3 but for δ small enough we still must have $\mathcal{A}'_t \leq \mathcal{A}_{\hat{t}} c_{\hat{t}} < \mathcal{A}'_t = \mathcal{A}_t - \delta$ which is a contradiction. □

Lemma 5. *In an optimal M, if W' is a directed cycle then all the types in W' must have same probability of allocation, and the same amount of money burning with no chance of assessment. Equivalently, $W' \subseteq W$ for some web W.*

Proof. Let $W' = \{t_1, t_2, \ldots, t_n\}$ be the directed cycle then from Lemma 4 we have $\mathcal{A}_{t_1} \leq \mathcal{A}_{t_2} c_{t_2} \leq \mathcal{A}_{t_2} \leq \mathcal{A}_{t_3} c_{t_3} \leq \cdots \leq \mathcal{A}_{t_1} c_{t_1}$ thus giving us an equality throughout. From IC, we have $v_{t_1} \mathcal{A}_{t_1} - p_{t_1} = v_{t_1} \mathcal{A}_{t_2} c_{t_2} - p_{t_2} \implies p_{t_1} - p_{t_2} = 0$, hence all the agents in the cycle must have the same requirements. Since $\mathcal{A}_{t_i} = \mathcal{A}_{t_j};\ p_{t_i} = p_{t_j};\ c_{t_i} = c_{t_j} = 1\ \forall t_i, t_j \in W'$ we have $t_i \in i_M(t_j)\ \forall t_i, t_j \in W'$ and W' must be part of a web. □

Lemma 6. *Every optimal G_M is weakly connected and contains exactly one web W^* with $i_M(W^*) = \emptyset$.*

Proof. Let G_M be optimal, we will first show that there must exist a web W in G_M with $i_M(W) = \emptyset$. If $\exists t \in V$ such that $i_M(t) = \emptyset$ then $\{t\}$ is the desired web. So, $deg^-_M(t) > 0\ \forall t \in V$, and G_M must contain directed cycles. Due to Lemma 5 these cycles must belong in a web thus G_M must contain webs. Let $\mathscr{W}$ be the set of webs in G_M.

Let $W^* = \arg\min_{W\in\mathscr{W}}\{\mathcal{A}_t : t \in W\}$, thus for any $t^* \in W^*$ and any $t \in W \in \mathscr{W}$, $W \neq W^*$, $\mathcal{A}_{t^*} < \mathcal{A}_t$.

It must be that $i_M(W^*) = \emptyset$, for otherwise let $(t_1, t^*) \in E(G_M)$; $t^* \in W^*$; $t_1 \in V\backslash W^*$ then $\mathcal{A}_{t^*} \geq \mathcal{A}_{t_1}$ due to Lemma 4. Now, since $deg^-_M(t_1) > 0$ we can find a t_2 such that $\mathcal{A}_{t_1} \geq \mathcal{A}_{t_2}$ and so on until we end up in another cycle with $\mathcal{A}_{t_k} \leq \mathcal{A}_{t^*}$ which is a contradiction to $W^* = \arg\min_{W\in\mathscr{W}}\{\mathcal{A}_t : t \in W\}$.

W^* is the only web with $i_M(W^*) = \emptyset$. Suppose not, let W be another such web. $deg^-_M(W) = 0 \implies p_t = 0\ \forall t \in W$, for if not then we can decrease money burning for all the types in the web by small $\delta > 0$ and improve the objective function without violating any IC constraints. Similarly, $deg^-_M(W^*) = 0$ mean that there is no money burning in W^*. However, let $t \in W, t^* \in W^*$, since $p_t = p_{t^*} = 0$ and $\mathcal{A}_t > \mathcal{A}_{t^*}$, $v_{t^*}\mathcal{A}_{t^*} < v_{t^*}\mathcal{A}_t$ thus violating the IC constraint and giving a contradiction.

If the graph is not connected, then repeating the steps above every connected component must contain a web W with $deg^-_M(W^*) = 0$. However, this gives us a contradiction since there can only be one such web. □

Lemma 7. *If M is optimal, $(s,t) \in E(G_M)$ and $\mathcal{A}_s = \mathcal{A}_t c_t$ then $c_s = 1$.*

Proof. If s has indegree zero then Lemma 2 delivers the result. Thus, assume $\exists r$ such that (r,s) bind, and let $c_s < 1$. Since $\mathcal{A}_s = \mathcal{A}_t c_t$ and $v_s\mathcal{A}_s - p_s = v_s\mathcal{A}_t c_t - p_t$ we have $p_s = p_t$. Since (r,s) bind,

$$\begin{aligned} v_r\mathcal{A}_r - p_r &= v_r\mathcal{A}_s c_s - p_s \\ &= v_r\mathcal{A}_t c_t c_s - p_s \\ &< v_r\mathcal{A}_t c_t - p_t \end{aligned}$$

contradicting the IC constraint. □

Proof of Proposition 1. Suppose not and let $v_s > v_t$ but s points to t. Note that Lemma 4 gives us $\mathcal{A}_t \geq \mathcal{A}_s$. We consider four cases,
Case I: $c_t < 1$; $o_M(t) \neq \emptyset$
Let $r \in o_M(t)$. We have,

$$\begin{aligned} & v_t\mathcal{A}_t - p_t = v_t\mathcal{A}_r c_r - p_r \\ \implies\quad & p_r - p_t \leq v_s\left(\mathcal{A}_r c_r - \mathcal{A}_t\right) \quad \because v_s > v_t;\ \mathcal{A}_r c_r \geq \mathcal{A}_t \text{ due to Lemma 4} \\ \implies\quad & p_r - p_t < v_s\left(\mathcal{A}_r c_r - \mathcal{A}_t c_t\right) \quad \because c_t < 1 \\ \implies\quad & v_s\mathcal{A}_s - p_s < v_s\mathcal{A}_r c_r - p_r \quad \because (s,t) \in E(G_M) \end{aligned}$$

Violating the IC constraint.

Case II: $c_t < 1$; $o_M(t) = \emptyset$
Consider the mechanism

$$M' = T_{M,s,t}\left(\mathcal{A}_s + \delta, \mathcal{A}_t - \delta, \frac{\mathcal{A}_s c_s}{\mathcal{A}_s + \delta}, \frac{\mathcal{A}_t c_t}{\mathcal{A}_t - \delta}, p_s, p_t\right)$$

then M' is feasible due to Lemma 1 and the fact that $c_t < 1$. IC's not involving s, t remain unchanged. $\mathcal{A}'_t c'_t = \mathcal{A}_t c_t$ and $\mathcal{A}'_s c'_s = \mathcal{A}_s c_s$ and so no other type has an incentive to pretend to be s or t. Surplus of s has increased so it has no incentive to deviate in M' and since t has outdegree zero, for small enough δ, t will not deviate. M', however, is an improvement of $\delta(v_s - v_t)$ in the objective function.

Case III: $c_t = 1$; $\mathcal{A}_t = \mathcal{A}_s$
From Lemma 7 we have that $c_s = 1$. Since $v_s\mathcal{A}_s - p_s = v_s\mathcal{A}_t - p_t$ we get $p_s = p_t$ and so $s \in i_M(t)$; $t \in i_M(s)$ and they must belong in a web.

Case IV: $c_t = 1$; $\mathcal{A}_t > \mathcal{A}_s$
We have,

$$\begin{aligned}
& & p_t - p_s &= v_s(\mathcal{A}_t - \mathcal{A}_s) & \\
&\implies & p_t - p_s &> v_t(\mathcal{A}_t - \mathcal{A}_s) & \because v_t < v_s \text{ and } \mathcal{A}_s < \mathcal{A}_t \\
&\implies & v_t\mathcal{A}_s - p_s &> v_t\mathcal{A}_t - p_t &
\end{aligned}$$

Consider $M' = T_{M,s,t}(\mathcal{A}_t, \mathcal{A}_s, c_t, c_s, p_t, p_s)$. s, t have no incentives to deviate in M' since their surplus is not decreasing. No other type has any incentive to pretend to be s or t either since they would otherwise have done that in M. M' improves the objective by $(v_t\mathcal{A}_s - p_s) - (v_t\mathcal{A}_t - p_t) > 0$. $\square$

Proof of Lemma 8. If $deg^-_M(s) \cdot deg^-_M(t) = 0$ then the result follows from Lemma 2 and Lemma 7. Let s', t' be such that $(s', s) \in E(G_M)$ and $(t', t) \in E(G_M)$ then we have

$$v_{s'}\mathcal{A}_s c_s - p_s \geq v_{s'}\mathcal{A}_t c_t - p_t$$
$$v_{t'}\mathcal{A}_t c_t - p_t \geq v_{t'}\mathcal{A}_s c_s - p_s$$

Combining the two we get

$$v_{t'}(\mathcal{A}_t c_t - \mathcal{A}_s c_s) \geq p_t - p_s \geq v_{s'}(\mathcal{A}_t c_t - \mathcal{A}_s c_s)$$

Thus concluding the proof. $\square$

Proof of Proposition 2. The second part of the proposition is proved separately as a Lemma below. Let $\hat{t}, t$ be such that $v_t > v_{\hat{t}}$ and $\mathcal{A}_t < \mathcal{A}_{\hat{t}}$ in an optimal mechanism M. If $deg^-_M(\hat{t}) = 0$ then the proposition is trivially true due to Lemma 2, so $i_M(\hat{t}) \neq \emptyset$. The proof consists of two parts. First, we show that for this mechanism M we must have $c_t < 1$ and $p_{\hat{t}} > 0$.

Suppose $c_t = 1$. If $c_{\hat{t}} = c_t = 1$ then from the IC constraints of $t, \hat{t}$ we have

$$v_{\hat{t}}(\mathcal{A}_{\hat{t}} - \mathcal{A}_t) \geq p_{\hat{t}} - p_t \geq v_t(\mathcal{A}_{\hat{t}} - \mathcal{A}_t)$$

which is not possible since $\mathcal{A}_{\hat{t}} > \mathcal{A}_t$ and $v_t > v_{\hat{t}}$, so:

$$c_{\hat{t}} c_t < 1 \tag{1}$$

If $deg^+_M(\hat{t}) = 0$ and $c_{\hat{t}} < 1$ then consider the mechanism M'' with

$$M' = T_{M,\hat{t},t}\left(\mathcal{A}_{\hat{t}} - \delta, \mathcal{A}_t + \delta, \frac{\mathcal{A}_{\hat{t}} c_{\hat{t}}}{\mathcal{A}_{\hat{t}} - \delta}, \frac{\mathcal{A}_t c_t}{\mathcal{A}_t + \delta}, p_{\hat{t}}, p_t\right)$$

By Lemma 1, for small δ this is feasible. All IC constraints can be verified to be satisfied for small δ and the objective function improvement is $\delta(v_t - v_{\hat{t}})$. And so if $deg^+_M(\hat{t}) = 0 \implies c_{\hat{t}} = 1$. However, this implies that $p_{\hat{t}} > 0$, since otherwise $\hat{t}$ must have lowest allocation probability and $\mathcal{A}_{\hat{t}} \leq \mathcal{A}_t$ contradicting the supposition. It also implies $c_t < 1$ by (1). Thus, we have $\neg(c_t < 1 \& p_{\hat{t}} > 0) \implies deg^+_M(\hat{t}) > 0$.

Now, since $deg^+_M(\hat{t}) > 0$ (i.e. $o_M(\hat{t}) \neq \emptyset$), let $\hat{t}' \in o_M(\hat{t})$. We have:

$$\begin{aligned}
& v_t\mathcal{A}_t - p_t \geq v_t\mathcal{A}_{\hat{t}'} c_{\hat{t}'} - p_{\hat{t}'} \\
\implies \quad & p_{\hat{t}'} - p_t \geq v_t(\mathcal{A}_{\hat{t}'} c_{\hat{t}'} - \mathcal{A}_t) \\
\implies \quad & p_{\hat{t}'} - p_t > v_{\hat{t}}(\mathcal{A}_{\hat{t}'} c_{\hat{t}'} - \mathcal{A}_t) \qquad \because \text{Lemma 4} : \mathcal{A}_{\hat{t}'} c_{\hat{t}'} \geq \mathcal{A}_{\hat{t}} > \mathcal{A}_t \\
\implies \quad & v_{\hat{t}}\mathcal{A}_t - p_t > v_{\hat{t}}\mathcal{A}_{\hat{t}'} c_{\hat{t}'} - p_{\hat{t}'}
\end{aligned}$$

$$\implies \quad v_{\hat{t}}\mathcal{A}_t - p_t > v_{\hat{t}}\mathcal{A}_{\hat{t}} - p_{\hat{t}} \tag{2}$$

Consider $M' = T_{M,\hat{t},t}(\mathcal{A}_t, \mathcal{A}_t, c_t, c_t, p_t, p_t)$, which is IC by (2). Objective changes by

$$\begin{aligned}
& v_{\hat{t}}\mathcal{A}_t - p_t + k\mathcal{A}_t c_t - k\mathcal{A}_t - [v_{\hat{t}}\mathcal{A}_{\hat{t}} - p_{\hat{t}} + k\mathcal{A}_{\hat{t}} c_{\hat{t}} - k\mathcal{A}_{\hat{t}}] \\
= & (v_{\hat{t}}\mathcal{A}_t - p_t - v_{\hat{t}}\mathcal{A}_{\hat{t}} + p_{\hat{t}}) + k(\mathcal{A}_{\hat{t}} - \mathcal{A}_t) + k(\mathcal{A}_t c_t - \mathcal{A}_{\hat{t}} c_{\hat{t}})
\end{aligned}$$

Which is strictly positive if $c_t = 1$, or $\mathcal{A}_t c_t \geq \mathcal{A}_{\hat{t}} c_{\hat{t}}$. Thus $c_t < 1$ and $\mathcal{A}_{\hat{t}} c_{\hat{t}} > \mathcal{A}_t c_t$ implying $p_{\hat{t}} > p_t \geq 0$ due to Lemma 8.

Now given $c_t < 1$ & $p_{\hat{t}} > 0$, consider $M^* = T_{M,\hat{t},t}(\mathcal{A}^*_{\hat{t}}, \mathcal{A}^*_t, c^*_{\hat{t}}, c^*_t, p^*_{\hat{t}}, p^*_t)$ where,

$$\mathcal{A}^*_{\hat{t}} = \mathcal{A}_{\hat{t}} - \varepsilon;\ c^*_{\hat{t}} = c_{\hat{t}} - \varepsilon \left(\frac{v_t(\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_t c_t) - c_{\hat{t}}(p_{\hat{t}} - p_t)}{(p_{\hat{t}} - p_t)(\mathcal{A}_{\hat{t}} - \varepsilon)} \right);\ p^*_{\hat{t}} = p_{\hat{t}} - v_t \varepsilon$$

$$\mathcal{A}^*_t = \mathcal{A}_t + \varepsilon;\ c^*_t = c_t + \varepsilon \left(\frac{v_t(\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_t c_t) - c_t(p_{\hat{t}} - p_t)}{(p_{\hat{t}} - p_t)(\mathcal{A}_t + \varepsilon)} \right);\ p^*_t = p_t + v_t \varepsilon$$

This part of the proof proceeds in 3 steps.

Step 1 - M^* is feasible: Lemma 1 shows that $\mathcal{A}^*_{\hat{t}}$ and $\mathcal{A}^*_t$ are feasible for small ε. $p_{\hat{t}} > 0$ and so $p_{\hat{t}}, p_t$ are feasible. c^*_t is feasible for small enough ε, since $c_t < 1$. Let us show $0 < c^*_{\hat{t}} \leq c_{\hat{t}}$. The first inequality holds for ε sufficiently small. For the second inequality we must show

$$\left(\frac{v_t(\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_t c_t) - c_{\hat{t}}(p_{\hat{t}} - p_t)}{(p_{\hat{t}} - p_t)(\mathcal{A}_{\hat{t}} - \varepsilon)} \right) \geq 0 \tag{3}$$

Consider $\underline{\hat{t}} = \inf i_{M^*}(\hat{t})$, then $t > \hat{t} > \underline{\hat{t}}$ due to Proposition 1. Thus:

$$\begin{aligned}
& & v_{\underline{\hat{t}}} \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} &\geq v_{\underline{\hat{t}}} \mathcal{A}_t c_t - p_t & \\
&\implies & v_{\underline{\hat{t}}}(\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_t c_t) &\geq p_{\hat{t}} - p_t & \\
&\implies & v_t(\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_t c_t) &\geq p_{\hat{t}} - p_t & \because \mathcal{A}_{\hat{t}} c_{\hat{t}} > \mathcal{A}_t c_t \\
&\implies & v_t(\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_t c_t) - c_{\hat{t}}(p_{\hat{t}} - p_t) &\geq 0 &
\end{aligned}$$

Proving (3).

Step 2 - M^* is IC: The only changes concern type $\hat{t}$ and t. We have $v_{\hat{t}} \mathcal{A}^*_{\hat{t}} - p^*_{\hat{t}} \geq v_{\hat{t}} \mathcal{A}_{\hat{t}} - p_{\hat{t}}$ and $v_t \mathcal{A}^*_t - p^*_t \geq v_t \mathcal{A}_t - p_t$ and so $\hat{t}, t$ have no incentive to lie. Define

$$\delta = \frac{v_t \varepsilon}{p_{\hat{t}} - p_t}$$

Then we see

$$\begin{aligned}
p^*_{\hat{t}} &= (1 - \delta) p_{\hat{t}} + \delta p_t \\
p^*_t &= (1 - \delta) p_t + \delta p_{\hat{t}} \\
\mathcal{A}^*_{\hat{t}} c^*_{\hat{t}} &= (1 - \delta) \mathcal{A}_{\hat{t}} c_{\hat{t}} - \delta \mathcal{A}_t c_t \\
\mathcal{A}^*_t c^*_t &= (1 - \delta) \mathcal{A}_t c_t - \delta \mathcal{A}_{\hat{t}} c_{\hat{t}}
\end{aligned}$$

Consider any arbitrary type r:

$$\begin{aligned}
v_r \mathcal{A}^*_t c^*_t - p^*_t &= v_r \left((1 - \delta) \mathcal{A}_t c_t + \delta \mathcal{A}_{\hat{t}} c_{\hat{t}} \right) - \left((1 - \delta) p_t + \delta p_{\hat{t}} \right) \\
&= (1 - \delta) \left[v_r \mathcal{A}_t c_t - p_t \right] + \delta \left[v_r \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} \right] \\
&\leq v_r \mathcal{A}_r - p_r
\end{aligned}$$

And so r doesn't mimic t. Similar argument shows that r doesn't mimic $\hat{t}$ either.

Step 3 - M^* is an improvement:

$$\begin{aligned}
Obj(M^*) - Obj(M) &= v_{\hat{t}}\mathcal{A}^*_{\hat{t}} - p^*_{\hat{t}} + k\mathcal{A}^*_{\hat{t}}c^*_{\hat{t}} - k\mathcal{A}^*_{\hat{t}} - (v_{\hat{t}}\mathcal{A}_{\hat{t}} - p_{\hat{t}} + k\mathcal{A}_{\hat{t}}c_{\hat{t}} - k\mathcal{A}_{\hat{t}}) \\
&\quad + v_t\mathcal{A}^*_t - p^*_t + k\mathcal{A}^*_t c^*_t - k\mathcal{A}^*_t - (v_t\mathcal{A}_t - p_t + k\mathcal{A}_t c_t - k\mathcal{A}_t) \\
&= (v_t - v_{\hat{t}})\varepsilon - k\delta(\mathcal{A}_{\hat{t}}c_{\hat{t}} - \mathcal{A}_t c_t) + k\varepsilon \\
&\quad + k\delta(\mathcal{A}_{\hat{t}}c_{\hat{t}} - \mathcal{A}_t c_t) - k\varepsilon \\
&= (v_t - v_{\hat{t}})\varepsilon \\
&> 0
\end{aligned}$$

□

Lemma 9. $v_t\mathcal{A}_t - p_t \geq v_{\hat{t}}\mathcal{A}_{\hat{t}} - p_{\hat{t}}$ *if* $v_t \geq v_{\hat{t}}$.

Proof. Suppose not and let $v_{\hat{t}} < v_t$ with $v_{\hat{t}}\mathcal{A}_{\hat{t}} - p_{\hat{t}} > v_t\mathcal{A}_t - p_t$ then we must have $c_{\hat{t}} < 1$ otherwise $v_t\mathcal{A}_t - p_t < v_t\mathcal{A}_{\hat{t}} - p_{\hat{t}}$ violating the IC constraint. We must also have $o_M(\hat{t}) = \emptyset$ otherwise if $r \in o_M(\hat{t})$, $v_t\mathcal{A}_t - p_t < v_{\hat{t}}\mathcal{A}_{\hat{t}} - p_{\hat{t}} = v_{\hat{t}}\mathcal{A}_r c_r - p_r < v_t\mathcal{A}_r c_r - p_r$ violating the IC constraint. Now, consider

$$M' = T_{M,\hat{t},t}\left(\mathcal{A}_{\hat{t}} - \delta, \mathcal{A}_t + \delta, \frac{\mathcal{A}_{\hat{t}}c_{\hat{t}}}{\mathcal{A}_{\hat{t}} - \delta}, \frac{\mathcal{A}_t c_t}{\mathcal{A}_t + \delta}, p_{\hat{t}}, p_t\right)$$

M' is IC and is feasible due to Lemma 1 and the fact that $c_{\hat{t}} < 1$. The improvement in the objective is $\delta(v_t - v_{\hat{t}})$. □

Proof of Proposition 3. Let M^* be an optimal mechanism and consider the web W^* with $i_{M^*}(W^*) = \emptyset$. Let $\underline{t} = maxW^*$. Since $i_{M^*}(W^*) = \emptyset$ we must have $p_t = 0\ \forall t \in W^*$, otherwise the mechanism M' defined by decreasing money burning for all the projects in W^* by small ε is an improvement over M^* and is feasible and incentive compatible. Since $p_t = 0;\ c_t = 1\ \forall t \in W^*$ we must have $\mathcal{A}_t = \min_{r\in V}\mathcal{A}_r\ \forall t \in V$ for if not and let $\mathcal{A}_{\hat{t}} > \mathcal{A}_t$ then $v_{\hat{t}}\mathcal{A}_{\hat{t}}c_{\hat{t}} - p_{\hat{t}} < v_{\hat{t}}\mathcal{A}_t$, thus violating the IC constraint. Due to Proposition 2 we must have $t \in W^*\ \forall t < \underline{t}$, since for all such $t, \mathcal{A}_t = \mathcal{A}_{\underline{t}}$ thus implying $c_t = 1;\ p_t = 0$ and $\{(t, \underline{t}), (\underline{t}, t)\} \subseteq E(G_M)$. Moreover, for any $t > \underline{t}$ we must have $\mathcal{A}_t > \mathcal{A}_{\underline{t}}$ by definition. Thus by IC every $t > \underline{t}$ must either have $c_t < 1, p_t > 0$ or both. □

Lemma 10. *Let* $\underline{t}$ *be the cutoff, let* $\hat{t}$ *be such that* $p_{\hat{t}} > 0$. *Then* $\nexists t' < \underline{t}$ *with* $v_{t'} < v_{\underline{t}}$ *and* $t' \in i_M(\hat{t})$.

Proof. Suppose not, $\exists t' < \underline{t}$ with $t' \in i_M(\hat{t})$. By lemma 3 we have $\mathcal{A}_{\underline{t}} = \mathcal{A}_{t'}$ Since $t' \in i_M(\hat{t})$ and $p_{\hat{t}} > 0$ we have:

$$p_{\hat{t}} = (\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_{t'})\, v_{t'}$$

But since $v_{\underline{t}} > v_{t'}$.

$$p_{\hat{t}} < (\mathcal{A}_{\hat{t}} c_{\hat{t}} - \mathcal{A}_{\underline{t}})\, v_{\underline{t}}$$

Contradicting the IC of type $\underline{t}$. □

Proof of Proposition 4. Suppose not, let t be such that $p_t > 0$ in the optimal mechanism. By lemma 10 any type in $i_M(t)$ must be weakly larger than $\underline{t}$. Let $t' = \min i_M(t)$, and reduce p_t by ε and c_t by $\frac{\varepsilon}{\mathcal{A}_t v_{t'}}$. This leads to $i_M(t') = \{t'\}$ and the objective is also increased $1/n\varepsilon\mathcal{A}_t(t' - k) > 0$ delivering the desired contradiction. Since there is no money burning in the mechanism the c_t's simply follow from Lemma 8. □

Proof of Proposition 5. We are going to prove the proposition by induction.

Basis Step: Let $\underline{t}$ be the cutoff. If $v_{\underline{t}} < k$ then $p_{\underline{t}+1} = v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \mathcal{A}_{\underline{t}})$ and $c_{\underline{t}+1} = 1$.

Proof of Basis Step. By definition p, c must be optimal for the optimal $\mathcal{A}$. Let $\mathcal{A}$ be given by the optimal mechanism, then using Proposition 2 and Lemmas 9, 10, we can equivalently write the problem of choosing c and p follows:

$$\begin{aligned}
\max_{\langle c\rangle,\langle p\rangle} & \sum_{t=\underline{t}+1}^{n} [-p_t + k\mathcal{A}_t c_t] \\
& s.t. \\
& c_t \le 1 \ \forall t > \underline{t};\ c_t, p_t \in \mathbb{R}^+ \\
& v_t \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} + p_t \le v_t \mathcal{A}_t \quad \forall t, \hat{t} \in \{\underline{t}, \ldots n\}, \hat{t} > t
\end{aligned}$$

Consider the dual of this LP problem, letting $y_{i,j} \ge 0$ denote the coefficient of i pretending to be j. Similarly let $r_j \ge 0$ denote the coefficient for the constraint $c_j \le 1$. The coefficients are all weakly positive since the constraints are all required to be less than equal to their respective bounds. The dual is:

$$\min_{\langle y_{i,j}\rangle,\langle r\rangle} \sum_{i\in\{\underline{t},n\}}\sum_{j>i} y_{i,j}v_i\mathcal{A}_i + \sum_{j\in\{\underline{t}+1,n\}} r_j$$

$$\sum_{i=\underline{t}}^{j-1} v_i\mathcal{A}_j y_{i,j} + r_j \geq k\mathcal{A}_j, \quad \forall j\in\{\underline{t}+1,n\}$$

$$-\sum_{i=\underline{t}}^{j-1} y_{i,j} + \sum_{l=j+1}^{n} y_{j,l} \geq -1 \quad \forall j\in\{\underline{t}+1,n\}$$

Towards a contradiction, assume $p_{\underline{t}+1} = 0$. Then the constraint $-y_{\underline{t},\underline{t}+1} + \sum_{l=\underline{t}+2}^{n} y_{\underline{t}+1,l} \geq -1$ is slack. But then we must have $v_{\underline{t}}\mathcal{A}_{\underline{t}+1}y_{\underline{t},\underline{t}+1} = k\mathcal{A}_{\underline{t}+1}$ and $r_{\underline{t}+1} = 0$ as otherwise we would have $c_{\underline{t}+1} = 1$, which contradicts $\underline{t}$ being the cutoff. This in turn implies $\sum_{l=\underline{t}+2}^{n} y_{\underline{t}+1,l} > \frac{k}{v_{\underline{t}}} - 1 > 0$. Which further implies there is at least one l with $y_{\underline{t}+1,l} > 0$. Since $-y_{\underline{t},\underline{t}+1} + \sum_{l=\underline{t}+2}^{n} y_{\underline{t}+1,l} > -1$ we can reduce $y_{\underline{t}+1,l}$ by $\varepsilon > 0$ that is small enough and increase $y_{\underline{t},l}$ by $\delta = \frac{\varepsilon v_{\underline{t}+1}}{v_{\underline{t}}}$. Then the dual decreases by $\varepsilon v_{\underline{t}+1}(\mathcal{A}_{\underline{t}+1} - \frac{\mathcal{A}_{\underline{t}}}{v_{\underline{t}}}) > 0$ contradicting $p_{\underline{t}+1} = 0$ minimizing the dual (equivalently maximizing the primal). Hence we must have $-y_{\underline{t},\underline{t}+1} + \sum_{l=\underline{t}+2}^{n} y_{\underline{t}+1,l} = -1$, which by complementary slackness implies $p_{\underline{t}+1} > 0$.

To see $c_{\underline{t}+1} = 1$, for a contradiction assume $0 < c_{\underline{t}+1} < 1$ which implies $r_{\underline{t}+1} = 0$. But then we must have $v_{\underline{t}}\mathcal{A}_{\underline{t}+1}y_{\underline{t},\underline{t}+1} = k\mathcal{A}_{\underline{t}+1}$ as otherwise complementary slackness would imply $c_{\underline{t}+1} = 0$, which in turn would imply that there can be no types that want to mimic $\underline{t}+1$ (since all claims are checked with certainty). But then increasing $c_{\underline{t}+1}$ by a small ε would be a strict improvement of the objective so we must have $v_{\underline{t}}\mathcal{A}_{\underline{t}+1}y_{\underline{t},\underline{t}+1} = k\mathcal{A}_{\underline{t}+1}$. Consider decreasing $y_{\underline{t},\underline{t}+1} = \frac{k}{v_{\underline{t}}} > 1$ by $\epsilon > 0$ such that $y_{\underline{t},\underline{t}+1} - \epsilon > 1$ and setting $r_{\underline{t}+1} = \epsilon v_{\underline{t}}\mathcal{A}_{\underline{t}+1}$. Then the constraint regarding $p_{\underline{t}+1}$, $-y_{\underline{t},\underline{t}+1} + \sum_{l=\underline{t}+2}^{n} y_{\underline{t}+1,l} = -1$ is now relaxed but the objective in the dual is unchanged. Thus we can find at least one l with $y_{\underline{t}+1,l} > 0$ reduce $y_{\underline{t}+1,l}$ by ε and increase $y_{\underline{t},l}$ by $\delta = \frac{\varepsilon v_{\underline{t}+1}}{v_{\underline{t}}}$. Then the dual decreases by by $\varepsilon v_{\underline{t}+1}(\mathcal{A}_{\underline{t}+1} - \frac{\mathcal{A}_{\underline{t}}}{v_{\underline{t}}}) > 0$ contradicting the optimality of $c_{\underline{t}+1} < 1$.

Finally to see that $p_{\underline{t}+1} = v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \mathcal{A}_{\underline{t}})$, observe that the constraint regarding $\underline{t}$ mimicking $\underline{t}+1$ must bind as there has to be at least one type that wants to mimic $\underline{t}+1$, but since $c_{\underline{t}+1} = 1$ we must have $p_{\underline{t}+1} = v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \mathcal{A}_{\underline{t}})$. □

Inductive Step: Let $\underline{t}$ be the cutoff and let $v_{t'} < k$ and $t' > \underline{t}$. By the inductive hypothesis we have $p_{t'-l} - p_{t'-l-1} = v_{t'-l-1}(\mathcal{A}_{t'-l} - \mathcal{A}_{t'-l-1})$ and $c_{t'-l} = 1$ for all $t' - \underline{t} > l \geq 0$. We need to show $p_{t'+1} - p_{t'} = v_{t'}(\mathcal{A}_{t'+1} - \mathcal{A}_{t'})$ and $c_{t'+1} = 1$.

Proof of Inductive Step. By the inductive hypothesis the money burning and verification probabilities of all types including type t' is known and fixed. Similar to the basis step we can consider the LP that pins down the rest of the money burning and verification for the optimal $\mathcal{A}$. The problem is given by

$$\begin{aligned}
&\max_{\langle c\rangle,\langle p\rangle} \sum_{t=t'+1}^{n} [-p_t + k\mathcal{A}_t c_t] \\
&\quad s.t. \\
&\quad c_t \leq 1 \ \forall t > t';\ c_t, p_t \in \mathbb{R}^+ \\
&\quad v_t \mathcal{A}_{\hat{t}} c_{\hat{t}} - p_{\hat{t}} + p_t \leq v_t \mathcal{A}_t \quad \forall \hat{t} \in \{t', \ldots n\}, t \in \{\underline{t}, \ldots n\}, \hat{t} > t
\end{aligned}$$

The only difference between this LP and the one in the basis step is that we need to keep track of the incentive constraints regarding types below t' mimicking types above t' as well, since a choice of money burning and verification probability for some type $t > t'$ still needs to satisfy those constraints. Similar to the basis step let us consider the dual letting $y_{i,j} \geq 0$ denote the coefficient of i pretending to be j. Similarly let $r_j \geq 0$ denote the coefficient for the constraint $c_j \leq 1$. The dual is:

$$\begin{aligned}
&\min_{\langle y_{i,j}\rangle,\langle r\rangle} \sum_{i\in\{\underline{t},\ldots n\}} \sum_{j\in\{t'+1,\ldots,n\}, j>i} y_{i,j} v_i \mathcal{A}_i + \sum_{j\in\{t'+1,n\}} r_j \\
&\quad \sum_{i=\underline{t}}^{j-1} v_i \mathcal{A}_j y_{i,j} + r_j \geq k\mathcal{A}_j, \quad \forall j \in \{t'+1, n\} \\
&\quad -\sum_{i=\underline{t}}^{j-1} y_{i,j} + \sum_{l=j+1}^{n} y_{j,l} \geq -1 \quad \forall j \in \{t'+1, n\}
\end{aligned}$$

Towards a contradiction, assume $p_{t'+1} = 0$. Then the constraint $-\sum_{i=\underline{t}}^{t'} y_{i,t'+1}$ $+\sum_{l=t'+2}^{n} y_{t'+1,l} \geq -1$ is slack. But then we must have $\sum_{i=\underline{t}}^{t'} v_i \mathcal{A}_{t'+1} y_{i,t'+1} = k\mathcal{A}_{t'+1}$ and $r_{t'+1} = 0$ as otherwise we would have $c_{t'+1} = 1$, which contradicts t' being above the cutoff. We can cancel the $A_{t'+1}$'s and reorganize the left handside term $\sum_{i=\underline{t}}^{t'} v_i y_{i,t'+1} = \sum_{i=\underline{t}}^{t'} v_{t'} y_{i,t'+1} - \sum_{i=\underline{t}}^{t'} (v_{t'} - v_i) y_{i,t'+1} = k$. Since all $(v_{t'} - v_i) y_{i,t'+1}$'s are positive we thus must have $\sum_{i=\underline{t}}^{t'} y_{i,t'+1} > \frac{k}{v_{t'}} > 1$.

This in turn implies $\sum_{l=t'+2}^{n} y_{t'+1,l} > 0$. Which further implies there is at least one l with $y_{t'+1,l} > 0$. Since $-\sum_{i=\underline{t}}^{t'} y_{i,t'+1} + \sum_{l=t'+2}^{n} y_{t'+1,l} > -1$ we can reduce $y_{t'+1,l}$ by $\varepsilon > 0$ that is small enough and increase $y_{t,l}$ by $\delta = \frac{\varepsilon v_{t'+1}}{v_{t'}}$. Then the dual decreases by $\varepsilon v_{t'+1}(\mathcal{A}_{t'+1} - \frac{\mathcal{A}_{t'}}{v_{t'}}) > 0$ contradicting $p_{t'+1} = 0$ minimizing the dual (equivalently maximizing the primal). Hence we must have $p_{t'+1} > 0$.

To see $c_{t'+1} = 1$, for a contradiction assume $0 < c_{t'+1} < 1$ which implies $r_{t'+1} = 0$. But then we must have $\sum_{i=\underline{t}}^{t'} v_i \mathcal{A}_{t'+1} y_{i,t'+1} = k\mathcal{A}_{t'+1}$ as otherwise complementary slackness would imply $c_{t'+1} = 0$ which cannot be optimal similar to the basis case. Consider decreasing $y_{\underline{t},t'+1}$ by $\epsilon > 0$ such that $\sum_{i=\underline{t}}^{t'} y_{i,t'+1} - \epsilon > 1$ and setting $r_{t'+1} = \epsilon v_{\underline{t}} \mathcal{A}_{t'+1}$. Then similar to the basis step the constraint regarding $p_{t'+1}$, is now relaxed but the objective in the dual is unchanged. Thus similar to above we can find at least one l with $y_{t'+1,l} > 0$ reduce $y_{t'+1,l}$ by $\varepsilon > 0$ that is small enough and increase $y_{t,l}$ by $\delta = \frac{\varepsilon v_{t'+1}}{v_{t'}}$. Then the dual decreases by $\varepsilon v_{t'+1}(\mathcal{A}_{t'+1} - \frac{\mathcal{A}_{t'}}{v_{t'}}) > 0$ contradicting the optimality of $c_{t'+1} < 1$.

Finally, to see that $p_{t'+1} = v_{t'}(\mathcal{A}_{t'+1} - \mathcal{A}_{t'}) + p_{t'}$ observe that again there has to be at least one type that wants to mimic $t'+1$. For a contradiction suppose some type l, $\underline{t} \leq l < t'$ has a binding incentive constraint regarding $t'+1$. Then we must have $p_{t'+1} = v_l(\mathcal{A}_{t'+1} - \mathcal{A}_l) + p_l$ and p_l is fixed by the inductive hypothesis. Let us now consider the type $l+1$. By the inductive hypothesis we also must have $p_{l+1} = v_l(\mathcal{A}_{l+1} - \mathcal{A}_l) + p_l$. But then, incentive constraint for type $l+1$ regarding $t'+1$ would imply

$$
\begin{aligned}
v_{l+1}\mathcal{A}_{l+1} - p_{l+1} &\geq v_{l+1}\mathcal{A}_{t'+1} - p_{t+1},\\
v_{l+1}\mathcal{A}_{l+1} - p_{l+1} &\geq v_{l+1}\mathcal{A}_{t'+1} - v_l(\mathcal{A}_{t'+1} - \mathcal{A}) - p_l,\\
v_l(\mathcal{A}_{t'+1} - \mathcal{A}_l) &\geq v_{l+1}(\mathcal{A}_{t'+1} - \mathcal{A}_{l+1}) + p_{l+1} - p_l,\\
v_l(\mathcal{A}_{t'+1} - \mathcal{A}_l) &\geq v_{l+1}(\mathcal{A}_{t'+1} - \mathcal{A}_{l+1}) + v_l(\mathcal{A}_{l+1} - \mathcal{A}_l),\\
v_l(\mathcal{A}_{t'+1} - \mathcal{A}_{l+1}) &\geq v_{l+1}(\mathcal{A}_{t'+1} - \mathcal{A}_{l+1});
\end{aligned}
$$

Now, if $v_l < v_{l+1}$ then we reach the desired contradiction as the IC of type v_{l+1} will be violated. If not, then we must have $v_l = v_{l+1}$, and v_{l+1} also has a binding constraint $p_{t'+1} = v_{l+1}(\mathcal{A}_{t'+1} - \mathcal{A}_{l+1}) + p_{l+1}$. If $l+1 = t'$ then we are done. If not we consider the IC for type $l+2$. Again, either we would have a violation of IC for type $l+2$ if $v_{l+1} < v_{l+2}$ or we must have $v_{l+2} = v_{l+1} = v_l$ with $l+2$ having a binding constraint. Proceeding inductively we either will reach a contradiction or all types starting from l has the same value until t' and the IC of t' gives the desired identity $p_{t'+1} = v_{t'}(\mathcal{A}_{t'+1} - \mathcal{A}_{t'}) + p_{t'}$.

Notice, in the case that all types have the same value, that is $v_l = v_{l+1} \dots v_{t'}$ then we have $\{(l, t'+1), (l+1, t'+1), \dots (t', t'+1)\} \subset E(G_M)$. $\square$

$\square$

Lemma 11. *Let $\underline{t}$ be the cutoff and let $t^{EF} = \min\{t \in \{\underline{t}, \dots, n\} : v_t \geq k\}$. Then for all $t > t^{EF}$, $c_t < 1$.*

Proof. By proposition 7 $\mathcal{A}$'s are all fixed for a given cutoff $\underline{t}$. And by proposition 5 we know that for all $\hat{t} \leq t^{EF}$, $c_{\hat{t}} = 1$ and $p_{\hat{t}} = v_{\hat{t}-1}(\mathcal{A}_{\hat{t}} - \mathcal{A}_{\hat{t}-1}) + p_{\hat{t}-1}$ starting with $p_{\underline{t}} = 0$. Then the principal's problem can be written as a standard LP

$$\begin{aligned}
\max_{\langle c \rangle, \langle p \rangle} & \sum_{t=t^k+1}^{n} [-p_t + k\mathcal{A}_t c_t] \\
& s.t. \\
& c_t \leq 1 \ \forall t \in \{t^{EF}+1, \dots, n\}; \ c_t, p_t \in \mathbb{R}^+ \\
& v_{\hat{t}}\mathcal{A}_{\hat{t}} \geq v_{\hat{t}}\mathcal{A}_t c_t - p_t + p_{\hat{t}} \quad \forall t \in \{t^{EF}+1, \dots, n\}, \hat{t} \in \{\underline{t}, \dots, n\}, \hat{t} < t
\end{aligned}$$

Notice in this formulation we still need to account for lower types (whose c and p is fixed) potentially mimicking higher types. The lemma concerns c_t's so instead of directly using Simplex method, we apply the upper bounding technique detailed in Eiselt and Sandblom (2007) (pp 212-218). That is we first seperate the upper bounds on c_t, and consider the constraint matrix, defined by just the IC constraints. The constraint matrix including the slack variables for the IC constraints has the following representation

$$\left[\begin{array}{cccccccccc:c}
\mathrm{V}_{t^{EF}+1} & 0 & \cdots & 0 & 0 & -\mathbf{1}^{t^{EF}+1-\underline{t}+1} & 0 & \cdots & 0 & 0 & \\
0 & \mathrm{V}_{t^{EF}+2} & \cdots & 0 & 0 & e_1^{t'+2-\underline{t}+1} & -\mathbf{1}^{t^{EF}+2-\underline{t}+1} & \cdots & 0 & 0 & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & I \\
0 & 0 & \cdots & \mathrm{V}_{n-1} & 0 & e_1^{n-1-\underline{t}+1} & e_2^{n-1-\underline{t}+1} & \cdots & -\mathbf{1}^{n-1-\underline{t}+1} & 0 & \\
0 & 0 & \cdots & 0 & \mathrm{V}_n & e_1^{n-\underline{t}+1} & e_2^{n-\underline{t}+1} & \cdots & e_{n-\underline{t}+1}^{n-\underline{t}+1} & -\mathbf{1}^{n-\underline{t}+1} &
\end{array}\right]$$

$\hat{\mathrm{V}}_j$ is the column vector of size $j - \underline{t} + 1$ with entries $(v_{\underline{t}}\mathcal{A}_j, v_{\underline{t}+1}\mathcal{A}_j, \dots, v_{j-1}\mathcal{A}_j)$. $-\mathbf{1}^j$ is the column vector of size j with all entries -1 and e_l^j is the unit column vector of size j, that is, it has zero in every entry except the lth entry which has a value 1. Finally I denotes the identity matrix of size $\sum_{r=t'+1}^{n}(r - \underline{t} + 1)$. In words each column vector $\hat{\mathrm{V}}_j$ corresponds to all the types that might potentially want to mimic type j, which are all the types starting from $\underline{t}$ and ending in $j-1$. Any type that wants to mimic type j would need to pay p_j hence there are $j - \underline{t} + 1$ constraints which has p_j with coefficient -1. If type $j > t'$ wants to mimic other, higher types, then we account for those with the entries e since p_j is still a choice variable. For types less than or equal to t' the p's are fixed by proposition 5. For each j we construct the following column vector $\nu_j = (v_{\underline{t}}\mathcal{A}_t, \dots v_{t^{EF}}\mathcal{A}_{t^{EF}} - p_{t^{EF}}, v_{t^{EF}+1}\mathcal{A}_{t^{EF}+1}, \dots, v_{j-1}\mathcal{A}_{j-1})$. And the column vector of size $\sum_{r=t^{EF}+1}^{n}(r - \underline{t} + 1)$, $\nu = (\nu_{t^{EF}+1}, \dots, \nu_n)$ which corresponds to the RHS of all the IC's.

Given that the p's have a negative value in the objective, we need to start pivoting from any of the c variables first. Suppose we start from the variable c_j, since all the entries in the column $(0, \dots, 0, \hat{\mathrm{V}}_j, 0, \dots, 0)$ is weakly positive and all the entries in the vector ν are positive upper bounding yields $v_{\underline{t}\mathcal{A}_j}$ as the pivot variable. On the other hand if we applied standard simplex for our constraint matrix will have $n - \underline{t} + 1$ additional rows and columns, rows corresponding to the upper bound constraints for c variables, and columns corresponding to their associated slack variables. In particular since the objective hasn't changed we would start pivoting from any of the c variables first again. Suppose we start from the column corresponding to the variable c_j, the associated column is now longer by $n - \underline{t} + 1$ entries where all the new entries are 0 except one 1 corresponding to the equation $c_j + s_{cj} = 1$ where s_{cj} is the slack variable introduced to capture the constraint $c_j \le 1$. That is the column will now look like $(0 \dots 0, 1, 0 \dots 0; 0, \dots, 0, \hat{\mathrm{V}}_j, 0, \dots, 0)$. However, similar to before the pivot variable would be $v_{\underline{t}\mathcal{A}_j}$ since the quotient for the new constraint is just 1 and we have $\mathcal{A}_{\underline{t}}/\mathcal{A}_j < 1$. But since both algorithms pivot the same way with the same quotient for the departing variable c_j, the constraint $c_j \le 1$ cannot be binding. Proceeding with the rest of the c variables in an analogous manner, we conclude that for all $t > t^E F$, $c_t < 1$. □

Proof of Proposition 6. By proposition 7 $\mathcal{A}$'s are all fixed for a given cutoff $\underline{t}$. And by proposition 5 we know that for all $\hat{t} \le t^{EF}$, $c_{\hat{t}} = 1$ and $p_{\hat{t}} = v_{\hat{t}-1}(\mathcal{A}_{\hat{t}} - \mathcal{A}_{\hat{t}-1}) + p_{\hat{t}-1}$ starting with $p_{\underline{t}} = 0$. Finally by Lemma 11 we know that for all $t > t^{EF}$, $c_t < 1$.

We first show $p_{t^{EF}+1} = p_{t^{EF}}$. For a contradiction, suppose not. There are two cases to consider

Case I: $p_{t^{EF}+1} < p_{t^{EF}}$

Recall that $\mathcal{A}_t c_t = \mathcal{A}_t$ for all $t \le t^{EF}$ by Lemma 4. 8, and Proposition 1 together imply that the only types that can have a binding constraint regarding type $t^{EF} + 1$ are the types $\hat{t}$ such that $p_{\hat{t}} \le p_{t^{EF}+1}$. First, suppose that multiple such types with binding constraints exist and let $\tilde{t}$ be the minimal type. From the IC of type $\tilde{t}$ we have

$$v_{\tilde{t}}(\mathcal{A}_{t^{EF}+1}c_{t^{EF}+1} - \mathcal{A}_{\tilde{t}}) = p_{t^{EF}+1} - p_{\tilde{t}}$$

On the other hand by Proposition 5 we also have

$$v_{\tilde{t}}\mathcal{A}_{\tilde{t}} - p_{\tilde{t}} = v_{\tilde{t}}\mathcal{A}_{\tilde{t}+1} - p_{\tilde{t}+1}$$

which together imply that

$$v_{\tilde{t}+1}(\mathcal{A}_{t^{EF}+1}c_{t^{EF}+1} - \mathcal{A}_{\tilde{t}}) > p_{t^{EF}+1} - p_{\tilde{t}}$$

A violation of the IC constraint for type $\tilde{t}+1$. A similar violation also occurs if there is a single type $\tilde{t}$ but $p_{\tilde{t}+1} \leq p_{t^{EF}+1}$. Thus, the only plausible case is if the only type that is mimicking type $t^{EF}+1$ is type $t^{EF}-1$ so that we have $p_{t^{EF}-1} \leq p_{t^{EF}+1} < p_{t^{EF}}$. But then consider increasing $p_{t^{EF}+1}$ by δ and $c_{t^{EF}+1}$ by ϵ where $\mathcal{A}_{t^{EF}+1}\epsilon v_{t^{EF}-1} = \delta$. Such an increase means $t^{EF}-1$ is still indifferent whereas all the other constraints remain unaffected, but the increase in the objective is $\epsilon\mathcal{A}_{t^{EF}+1}(k - v_{t^{EF}-1}) > 0$, a contradiction.

Case II: $p_{t^{EF}+1} > p_{t^{EF}}$

In this case first suppose some type $\tilde{t} < t^{EF}$ has a binding constraint with type $t^{EF}+1$. Then similar to above we would have

$$v_{\tilde{t}}(\mathcal{A}_{t^{EF}+1}c_{t^{EF}+1} - \mathcal{A}_{\tilde{t}}) = p_{t^{EF}+1} - p_{\tilde{t}} \quad \text{and; } v_{\tilde{t}}\mathcal{A}_{\tilde{t}} - p_{\tilde{t}} = v_{\tilde{t}}\mathcal{A}_{\tilde{t}+1} - p_{\tilde{t}+1}$$

which together imply that

$$v_{\tilde{t}+1}(\mathcal{A}_{t^{EF}+1}c_{t^{EF}+1} - \mathcal{A}_{\tilde{t}}) > p_{t^{EF}+1} - p_{\tilde{t}}$$

A violation of the IC constraint for type $\tilde{t}+1$. Therefore the only types that can mimic type $t^{EF}+1$ are types above t^{EF}. But then consider decreasing $p_{t^{EF}+1}$ by δ and $c_{t^{EF}+1}$ by ϵ where $\mathcal{A}_{t^{EF}+1}\epsilon v_{t^{EF}} = \delta$. Then any IC pertaining to type $t^{EF}+1$ from any type weakly above t^{EF} will continue to hold, potentially strictly. The increase in the objective is $\epsilon\mathcal{A}_{t^{EF}+1}(v_{t^{EF}} - k) > 0$, a contradiction.

Proceeding inductively with identical cases we must have $p_{t^{EF}+2} = p_{t^{EF}}$, which in turn implies $p_{t^{EF}+3} = p_{t^{EF}}$ and so on until we have $p_n = p_{n-1} \ldots = p_{t^{EF}}$. Since the payments are the same by Lemma 8 we have

$$\mathcal{A}_n c_n = \mathcal{A}_{n-1}c_{n-1} = \ldots = \mathcal{A}_{t^{EF}+1}c_{t^{EF}+1} = \mathcal{A}_{t^{EF}}.$$

$\square$

Proof of Lemma 12. Suppose not. Then under the optimal mechanism there with cutoff $\underline{t}$ exists a pair of types $\underline{t} < t' < t'+1$ such that $v_{t'} = v_{t'+1}$ but $\mathcal{A}_{t'} < \mathcal{A}_{t'+1}$. Let us construct an improvement. Wlog assume that $\mathcal{A}_{t'} + \mathcal{A}_{t'+1} = A$ for some $A < 1$. Recall that the total money burning bill is given by

$$\sum_{t=\underline{t}+1}^{n} \sum_{s=\underline{t}}^{t-1} v_s(\mathcal{A}_{s+1} - \mathcal{A}_s)$$

Thus, holding the rest of the mechanism constant minimizing the required payments leads to an improvement.

$$\begin{aligned} \min_{\mathcal{A}_{t'},\mathcal{A}_{t'+1}} \; & v_{t'}(\mathcal{A}_{t'+1} - \mathcal{A}_{t'}) + v_{t'-1}(\mathcal{A}_{t'} - \mathcal{A}_{t'-1}) \\ \text{subject to } & \mathcal{A}_{t'} + \mathcal{A}_{t'+1} = A, \quad \mathcal{A}_{t'+1} \geq \mathcal{A}_{t'}, \end{aligned}$$

Clearly this simple linear program is minimized at $\mathcal{A}_{t'} = \mathcal{A}_{t'+1}$. This covers all types below and including t^{EF} since there is no verification for these types. For the second part of the lemma we need to consider types strictly above t^{EF}. For these types, regardless of allocation probability their required payments are the same, but their allocation probability also determines the probability of verification. Therefore again consider two types $v_{t'} = v_{t'+1}$ and $\mathcal{A}_{t'} < \mathcal{A}_{t'+1}$ with $t' > t^{EF}$. Again let $\mathcal{A}_{t'} + \mathcal{A}_{t'+1} = A$ for some $A < 1$. Since types above t^{EF} have no binding constraints amongst each other their contribution and the costs associated can be maximized in isolation. Namely we must maximize:

$$\begin{aligned} \max_{\mathcal{A}_{t'},\mathcal{A}_{t'+1}} \; & \mathcal{A}_{t'} v_{t'} + \mathcal{A}_{t'+1} v_{t'+1} - 2k\mathcal{A}_{t^{EF}} \\ \text{subject to } & \mathcal{A}_{t'} + \mathcal{A}_{t'+1} = A, \quad \mathcal{A}_{t'+1} \geq \mathcal{A}_{t'}. \end{aligned}$$

Clearly, every feasible option, including $\mathcal{A}_{t'} = \mathcal{A}_{t'+1}$ is a solution. □

Proof of Proposition 7. If $v_{\underline{t}} > k$, then Proposition 4 implies there is no money burning at all in the mechanism and the result is the one that coincides with Vohra (2012) which satisfies the proposition. Consider the case where $v_{\underline{t}} < k$. Following the construction in the main text the polymatroid problem is:

$$\begin{aligned} \max_{\langle \mathcal{B} \rangle} & \sum_{t=1}^{n} v_t(\mathcal{B}_t + d_t) \\ \text{s.t.} & \sum_{s \in S} \mathcal{B}_s \leq \rho(S) \quad \forall S \subset V \end{aligned}$$

Since g is submodular and $d_i's$ are increasing, the ρ function for any set $\{n, n-1, n-2, \ldots t\}$ with $t \geq \underline{t}+1$ must be of the form $\rho(\{n, n-1, \ldots, t\}) = g(n-r+1) - \sum_{i=r}^{n} d_i$ for some $t \geq r \geq \underline{t}$. From lemma 12 we know for that if i, j are duplicates types then we must have $d_i = d_j$. The only way to "skip" some terms is by skipping a duplicate type and replacing it with another. However, recall that g is submodular thus since every duplicate has the same d_i a minimizing superset contains all duplicates together. It is not possible to consider subsets that "skip" non-duplicate terms t and r as the term $\sum_{i=t}^{n} d_t$ is maximized by choosing consecutive elements due to weakly increasing d_i's. If $\rho(\{n, n-1, \ldots, t\}) = g(n-t+1) - \sum_{i=t}^{n} d_i$ for all sets of the form $\{n, n-1, n-2, \ldots t\}$ then the greedy algorithm assigns $\mathcal{B}_t = g\left(\sum_{t=\hat{t}}^{\hat{n}} f(t)\right) - g\left(\sum_{t=\hat{t}+1}^{\hat{n}} f(t)\right) - f(t)d_t$ to each t which results in $\mathcal{A}_t = g\left(\sum_{t=\hat{t}}^{\hat{n}} f(t)\right) - g\left(\sum_{t=\hat{t}+1}^{\hat{n}} f(t)\right)$, the claim of the proposition. Towards a contradiction, suppose there exists some $\tilde{t}$ (WLOG assume $\tilde{t}$ is the lowest index among the duplicates) such that the set $\rho(\{n, n-1, \ldots, \tilde{t}\}) = g(n-r+1) - \sum_{i=r}^{n} d_i < g(n-\tilde{t}+1) - \sum_{i=\tilde{t}}^{n} d_i$ for some $\tilde{t} > r \geq \underline{t}$. Then the greedy algorithm necessarily assigns $\mathcal{B}_{\tilde{t}} = g(n-r+1) - g(n-\tilde{t}+1) - \sum_{j=r}^{\tilde{t}} d_j > d_{\tilde{t}}$, and sets $\mathcal{B}_{t'} = 0$ for any t' such that $\tilde{t} > t' \geq r$, hence those types are at their constraint level. The first inequality is strict since $\rho(\{n, n-1, \ldots, \tilde{t}\}) = g(n-r+1) - \sum_{i=r}^{n} d_i < g(n-\tilde{t}+1) - \sum_{i=\tilde{t}}^{n} d_i$. But $\mathcal{B}_{\tilde{t}} > d_{\tilde{t}}$ implies that type $\tilde{t}$ does not want to mimic any type, their allocated payoff is strictly larger than even the best type of lie they can tell. Since their constraint is strict it also must be the case that they must be subject to some verification as otherwise higher types would want to mimic type $\tilde{t}$. Thus we must have $o_M(\tilde{t}) = \emptyset$, and has $c_{\tilde{t}} < 1$. Furthermore, the resulting $\mathcal{A}_{\tilde{t}} = \mathcal{B}_{\tilde{t}} + d_{\tilde{t}} = g(n-r+1) - g(n-\tilde{t}+1) - \sum_{i=r}^{\tilde{t}-1} d_i < g(n-\tilde{t}) - g(n-\tilde{t}+1)$, that is the constraints arising from the lower types necessitates some withholding of type $\tilde{t}$. Let's construct an improvement. By proposition 5 in $\mathcal{A}c, p$ we must have $p_{\underline{t}+1} > 0$ and $c_{\underline{t}+1} = 1$ so $\tilde{t} > \underline{t}+1$. Consider $M^* = T_{M,\underline{t}+1,t}(\mathcal{A}^*_{\underline{t}+1}, \mathcal{A}^*_{\tilde{t}}, c^*_{\underline{t}+1}, c^*_{\tilde{t}}, p^*_{\underline{t}+1}, p^*_{\tilde{t}})$ where,

$$\mathcal{A}^*_{\underline{t}+1} = \mathcal{A}_{\underline{t}+1} - \varepsilon;\ c^*_{\underline{t}+1} = c_{\underline{t}} - \varepsilon\left(\frac{v_{\tilde{t}}(\mathcal{A}_{\tilde{t}}c_{\tilde{t}} - \mathcal{A}_{\underline{t}+1}c_{\underline{t}+1}) - c_{\underline{t}+1}(p_{\tilde{t}} - p_{\underline{t}+1})}{(p_{\tilde{t}} - p_{\underline{t}+1})(\mathcal{A}_{\underline{t}+1} - \varepsilon)}\right);$$

$$p^*_{\underline{t}+1} = p_{\underline{t}+1} - v_{\tilde{t}}\varepsilon$$

$$\mathcal{A}^*_{\tilde{t}} = \mathcal{A}_{\tilde{t}} + \varepsilon;\ c^*_{\tilde{t}} = c_{\tilde{t}} + \varepsilon\left(\frac{v_{\tilde{t}}(\mathcal{A}_{\tilde{t}}c_{\tilde{t}} - \mathcal{A}_{\underline{t}+1}c_{\underline{t}+1}) - c_t(p_t - p_{\underline{t}+1})}{(p_t - p_{\underline{t}+1})(\mathcal{A}_t + \varepsilon)}\right);$$

$$p^*_t = p_{\tilde{t}} + v_{\tilde{t}}\varepsilon$$

Improving $\mathcal{A}_{\tilde{t}}$ this way is feasible since $\mathcal{A}_{\tilde{t}} < g(N-\tilde{t}) - \sum_{j=\tilde{t}+1}^{n} \mathcal{A}_j$ so there is slack for small enough ε. Thus, noting $c_{\tilde{t}} < 1$ and replicating the steps 1-2-3 of the proof of proposition 2 shows that M^* is feasible, IC and an improvement. □

Proof of Theorem 1. Let $F(\underline{t})$ denote the total payoff from the mechanism when the cutoff is $\underline{t}$. We will now formulate the change in payoff when the cutoff is decreased from $\underline{t}+1$ to $\underline{t}$. $F(\underline{t})-F(\underline{t}+1)$. To facilitate the comparison let us introduce the notation $\bar{\mathcal{A}}_{\underline{t}}$ to denote the allocation probability of the type $\underline{t}$ when $\underline{t}$ is the cutoff and we use $\mathcal{A}_{\underline{t}}$ to denote the allocation probability of type $\underline{t}$ when the cutoff is below $\underline{t}$. Similarly let $p_t(\underline{t})$ denote the money burning required of type t when the cutoff is $\underline{t}$. **Case I:** $\underline{t} \geq t^{EF}$

In this case there is no money burning in the mechanism

$$\begin{aligned}
F(\underline{t}) &= \bar{\mathcal{A}}_{\underline{t}} \sum_{t=1}^{\underline{t}} v_t + \sum_{\underline{t}+1}^{n} \mathcal{A}_t v_t - \sum_{\underline{t}+1}^{n} \mathcal{A}_t c_t k \\
&= \bar{\mathcal{A}}_{\underline{t}} \sum_{t=1}^{\underline{t}} v_t + \sum_{\underline{t}+1}^{n} \mathcal{A}_t v_t - \sum_{\underline{t}+1}^{n} \bar{\mathcal{A}}_{\underline{t}} k \\
F(\underline{t}-1) &= \bar{\mathcal{A}}_{\underline{t}-1} \sum_{t=1}^{\underline{t}-1} v_t + \sum_{\underline{t}}^{n} \mathcal{A}_t v_t - \sum_{\underline{t}}^{n} \mathcal{A}_t c_t k \\
&= \bar{\mathcal{A}}_{\underline{t}-1} \sum_{t=1}^{\underline{t}-1} v_t + \sum_{\underline{t}}^{n} \mathcal{A}_t v_t - \sum_{\underline{t}}^{n} \bar{\mathcal{A}}_{\underline{t}-1} k
\end{aligned}$$

Hence

$$\begin{aligned}
F(\underline{t}-1) - F(\underline{t}) &= (\bar{\mathcal{A}}_{\underline{t}-1} - \bar{\mathcal{A}}_{\underline{t}}) \sum_{t=1}^{\underline{t}-1} v_t - \bar{\mathcal{A}}_{\underline{t}} v_{\underline{t}} + \mathcal{A}_{\underline{t}} v_{\underline{t}} - (\sum_{\underline{t}}^{n} \bar{\mathcal{A}}_{\underline{t}-1} - \sum_{\underline{t}+1}^{n} \bar{\mathcal{A}}_{\underline{t}}) k \\
&= \bar{\mathcal{A}}_{\underline{t}-1} \left(\sum_{t=1}^{\underline{t}-1} v_t - \sum_{\underline{t}}^{n} k \right) - \bar{\mathcal{A}}_{\underline{t}} \left(\sum_{t=1}^{\underline{t}} v_t - \sum_{\underline{t}+1}^{n} k \right) + \mathcal{A}_{\underline{t}} v_{\underline{t}}
\end{aligned}$$

Summing across starting from n until t by shifting the last index we get

$$F(\underline{t}) - F(n) = \bar{\mathcal{A}}_{\underline{t}} \left(\sum_{t=1}^{\underline{t}} v_t - \sum_{\underline{t}+1}^{n} k \right) - \frac{1}{D} \sum_{t=1}^{n} v_t + \sum_{\underline{t}+1}^{n} \mathcal{A}_t v_t$$

Thus the optimal cutoff above and including t^{EF} solves

$$\max_{\underline{t} \geq t^{EF}} \bar{\mathcal{A}}_{\underline{t}} \left(\sum_{t=1}^{\underline{t}} v_t - \sum_{\underline{t}+1}^{n} k \right) - \frac{1}{D} \sum_{t=1}^{n} v_t + \sum_{\underline{t}+1}^{n} \mathcal{A}_t v_t$$

Case II: $\underline{t} < t^{EF}$ In this case there is no effect on the types above t^{EF} beside their money burning requirements. Hence;

$$F(\underline{t}) = \bar{\mathcal{A}}_{\underline{t}} \sum_{t=1}^{\underline{t}} v_t + \sum_{\underline{t}+1}^{n} \mathcal{A}_t v_t - \sum_{t^{EF}+1}^{n} \mathcal{A}_{t^{EF}} k - \sum_{\underline{t}+1}^{n} p_t(\underline{t})$$

$$F(\underline{t}-1) = \bar{\mathcal{A}}_{\underline{t}-1} \sum_{t=1}^{\underline{t}-1} v_t + \sum_{\underline{t}}^{n} \mathcal{A}_t v_t - \sum_{t^{EF}+1}^{n} \mathcal{A}_{t^{EF}} k - \sum_{\underline{t}}^{n} p_t(\underline{t}-1)$$

Recall that

$$p_t(\underline{t}) = v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \bar{\mathcal{A}}_{\underline{t}}) + v_{\underline{t}+1}(\mathcal{A}_{\underline{t}+2} - \mathcal{A}_{\underline{t}+1}) + v_{\underline{t}+2}(\mathcal{A}_{\underline{t}+3} - \mathcal{A}_{\underline{t}+2}) \ldots + v(t-1)(\mathcal{A}_t - \mathcal{A}_{t-1})$$

$$p_t(\underline{t}-1) = v_{\underline{t}-1}(\mathcal{A}_{\underline{t}} - \bar{\mathcal{A}}_{\underline{t}-1}) + v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \mathcal{A}_{\underline{t}}) + v_{\underline{t}+1}(\mathcal{A}_{\underline{t}+2} - \mathcal{A}_{\underline{t}+1}) \ldots + v(t-1)(\mathcal{A}_t - \mathcal{A}_{t-1})$$

therefore

$$p_t(\underline{t}-1) - p_t(\underline{t}) = v_{\underline{t}-1}(\mathcal{A}_{\underline{t}} - \bar{\mathcal{A}}_{\underline{t}-1}) + v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \mathcal{A}_{\underline{t}}) - v_{\underline{t}}(\mathcal{A}_{\underline{t}+1} - \bar{\mathcal{A}}_{\underline{t}})$$

$$p_t(\underline{t}-1) - p_t(\underline{t}) = v_{\underline{t}-1}(\mathcal{A}_{\underline{t}} - \bar{\mathcal{A}}_{\underline{t}-1}) + v_{\underline{t}}(\bar{\mathcal{A}}_{\underline{t}} - \mathcal{A}_{\underline{t}})$$

Hence

$$\begin{aligned} F(\underline{t}-1) - F(\underline{t}) &= \bar{\mathcal{A}}_{\underline{t}-1} \sum_{t=1}^{\underline{t}-1} v_t - \bar{\mathcal{A}}_{\underline{t}} \sum_{t=1}^{\underline{t}} v_t + \mathcal{A}_{\underline{t}} v_{\underline{t}} - \sum_{\underline{t}+1}^{n} (v_{\underline{t}-1}(\mathcal{A}_{\underline{t}} - \bar{\mathcal{A}}_{\underline{t}-1}) + v_{\underline{t}}(\bar{\mathcal{A}}_{\underline{t}} - \mathcal{A}_{\underline{t}})) \\ &- v_{\underline{t}-1}(\mathcal{A}_{\underline{t}} - \bar{\mathcal{A}}_{\underline{t}-1}) \\ &= \bar{\mathcal{A}}_{\underline{t}-1} \left(\sum_{t=1}^{\underline{t}-1} v_t + \sum_{\underline{t}}^{n} v_{\underline{t}-1} \right) - \bar{\mathcal{A}}_{\underline{t}} \left(\sum_{t=1}^{\underline{t}} v_t + \sum_{\underline{t}+1}^{n} v_{\underline{t}} \right) \\ &+ \mathcal{A}_{\underline{t}} \sum_{\underline{t}}^{n} (v_{\underline{t}} - v_{\underline{t}-1}) \end{aligned}$$

Starting from state 1 and adding the differences we get

$$F(1) - F(\underline{t}) = \bar{\mathcal{A}}_1 n - \bar{\mathcal{A}}_{\underline{t}} (\sum_{t=1}^{\underline{t}-1} v_t + \sum_{t=\underline{t}}^{n} v_{\underline{t}}) + \sum_{t=2}^{\underline{t}} \mathcal{A}_t (v_t - v_{t-1})(n-t+1)$$

Thus the optimal cutoff below t^{EF} maximizes

$$\max_{1<\underline{t}<t^{EF}} \bar{\mathcal{A}}_{\underline{t}}(\sum_{t=1}^{\underline{t}-1} v_t + \sum_{t=\underline{t}}^{n} v_{\underline{t}}) - \sum_{t=2}^{\underline{t}} \mathcal{A}_t(v_t - v_{t-1})(n-t+1) - \frac{\left(\frac{1}{n}\right)^{D-2}}{D}$$

$\square$

Proof of Proposition 8. Notice that a change in k also changes t^{EF}, so with a slight abuse of notation we will define the function $t^{EF}(k)$, where $t^{EF}(k) = \min\{t \in \{1,\dots,n\} : v_t \geq k\}$. In order to prove the proposition let us introduce the function $F(\underline{t},k,k')$ capturing the value of the problem when the cutoff is $\underline{t}$, cost of verification k, the cutoff for money burning is given by $t^{EF}(k')$ and allocation probabilities are given by Proposition 7. Formally:

$$\begin{aligned} F(\underline{t},k,k') = {} & I_{\underline{t}\geq t^{EF}(k')}\left(\sum_{t=1}^{n}\mathcal{A}_t v_t - \sum_{l=\underline{t}+1}^{n} k\left(\mathcal{A}_l - \mathcal{A}_{\underline{t}}\right)\right) \\ & + I_{\underline{t}<t^{EF}(k')}\left(\sum_{t=1}^{t^{EF}(k')}(n-t+1)\mathcal{A}_t(v_t - v_{t-1}) + VK^{EF}(k,k')\right) \end{aligned}$$

Where $VK^{EF}(k,k')\sum_{t=t^{EF}(k')+1}^{n}(\mathcal{A}_t v_t - \mathcal{A}_{t^{EF}(k')}v_{t^{EF}(k')}) - \sum_{t=t^{EF}(k')+1}^{n} k(\mathcal{A}_t - \mathcal{A}_{t^{EF}(k')})$. Our goal is to show that when parametrized by k, the function $F(\cdot,k,k)$ dominates $F(\cdot,k',k')$ for $k > k'$ in the interval order (see Quah and Strulovici (2009)). For a contradiction suppose that this is not true, then there must exist a pair $\underline{t}'' > \underline{t}'$ and for all $\underline{t} \in \{\underline{t}',\dots,\underline{t}''\}$ we have $F(\underline{t}'',k',k') \geq F(\underline{t},k',k')$ but $F(\underline{t}'',k,k) < F(\underline{t}',k,k)$. There are four cases we need to consider.

Case I: $\underline{t}'' > \underline{t}' \geq t^{EF}(k)$

In this case we necessarily also have $\underline{t}'', \underline{t}' > t^{EF}(k')$ and there is no money burning in both mechanisms either with k or k'. But then $F(\underline{t}'',k',k') \geq F(\underline{t}',k',k')$ implies that

$$\sum_{t=1}^{\underline{t}''}\mathcal{A}_{\underline{t}''}v_t - \sum_{t=1}^{\underline{t}'}\mathcal{A}_{\underline{t}'}v_t - \sum_{\underline{t}'+1}^{\underline{t}''}\mathcal{A}_t v_t \geq k'\left(\sum_{l=\underline{t}''+1}^{n}(\mathcal{A}_l - \mathcal{A}_{\underline{t}''}) - \sum_{l=\underline{t}'+1}^{n}(\mathcal{A}_l - \mathcal{A}_{\underline{t}'})\right)$$

Notice that the term in parenthesis is negative as both $\mathcal{A}_{\underline{t}''} > \mathcal{A}_{\underline{t}'}$ and the first summation is over fewer terms. But then the inequality continues to hold when we replace k' with k thus implying $F(\underline{t}'', k, k) \geq F(\underline{t}', k, k)$ delivering a contradiction.

Case II: $t^{EF}(k') > \underline{t}'' > \underline{t}'$

In this case we necessarily have money burning in both mechanisms with k or k'. In particular, costs associated with k only corresponds to $VK^{EF}(\cdot, \cdot)$ in both mechanisms and hence cancel each other out. So we have $F(\underline{t}'', k', k') \geq F(\underline{t}', k', k')$ directly implying $F(\underline{t}'', k, k) \geq F(\underline{t}', k, k)$ delivering the desired contradiction.

Case III: $\underline{t}'' \geq t^{EF}(k) > \underline{t}' > t^{EF}(k')$

By definition we have $F(\underline{t}'', k', k') = F(\underline{t}'', k', k) = F(\underline{t}'', k, k) + (k - k') \sum_{\underline{t}''+1}^{n} \mathcal{A}_t - \mathcal{A}_{\underline{t}''}$. On the other hand we have

$$F(\underline{t}', k', k') \geq F(\underline{t}', k', k) = F(\underline{t}', k, k) + (k - k')(\sum_{t=t^{EF}(k)+1}^{n} \mathcal{A}_t - \mathcal{A}_{t^{EF}(k)})$$

But we already know $F(\underline{t}'', k', k') \geq F(\underline{t}', k', k')$ and hence we have

$$F(\underline{t}'', k, k) + (k - k') \sum_{\underline{t}''+1}^{n} \mathcal{A}_t - \mathcal{A}_{\underline{t}''} \geq F(\underline{t}', k, k) + (k - k')(\sum_{t=t^{EF}(k')+1}^{n} \mathcal{A}_t - \mathcal{A}_{t^{EF}(k')})$$

But, clearly we have $(k - k') \sum_{\underline{t}''+1}^{n} \mathcal{A}_t - \mathcal{A}_{\underline{t}''} \leq (k - k')(\sum_{t=t^{EF}(k')+1}^{n} \mathcal{A}_t - \mathcal{A}_{t^{EF}(k')})$ implying $F(\underline{t}'', k, k) > F(\underline{t}', k, k)$, the desired contradiction.

Case IV: $t^{EF}(k) > \underline{t}'' \geq t^{EF}(k') > \underline{t}'$

By definition we have $F(\underline{t}'', k', k') = F(\underline{t}'', k, k') + (k - k') \sum_{\underline{t}''+1}^{n} \mathcal{A}_t - \mathcal{A}_{\underline{t}''}$. Furthermore by optimality we have $F(\underline{t}'', k, k') \leq F(\underline{t}'', k, k)$ thus we have $F(\underline{t}'', k', k') \leq F(\underline{t}'', k, k) + (k - k') \sum_{\underline{t}''+1}^{n} \mathcal{A}_t - \mathcal{A}_{\underline{t}''}$ On the other hand we have

$$F(\underline{t}', k', k') \geq F(\underline{t}', k', k) = F(\underline{t}', k, k) + (k - k')(\sum_{t=t^{EF}(k)+1}^{n} \mathcal{A}_t - \mathcal{A}_{t^{EF}(k)})$$

But we already know $F(\underline{t}'', k', k') \geq F(\underline{t}', k', k')$ and hence we have

$$F(\underline{t}'', k, k) + (k - k') \sum_{\underline{t}''+1}^{n} \mathcal{A}_t - \mathcal{A}_{\underline{t}''} \geq F(\underline{t}', k, k) + (k - k')(\sum_{t=t^{EF}(k)+1}^{n} \mathcal{A}_t - \mathcal{A}_{t^{EF}(k)})$$

But, clearly we have $(k-k')\sum_{\underline{t}''+1}^{n}\mathcal{A}_t - \mathcal{A}_{\underline{t}''} \leq (k-k')(\sum_{t=t^{EF}(k')+1}^{n}\mathcal{A}_t - \mathcal{A}_{t^{EF}(k')})$ implying $F(\underline{t}'', k, k) > F(\underline{t}', k, k)$, the desired contradiction.

Since all cases lead to a contradiction we must have that $F(\cdot, k, k)$ dominates $F(\cdot, k', k')$ for $k > k'$ in the interval order. Hence by Theorem 1 of Quah and Strulovici (2009), we have that the optimal cutoff is weakly increasing in k. Now, with a slight abuse of notation let $\underline{t}(k)$ denote the cutoff level for cost k. By the above argument we know that $\underline{t}(k)$ is weakly increasing. Recall that $t^{EF}(k)$ is also increasing. Then the types that are subject to verification are all the types $t > \max\{\underline{t}(k), t^{EF}(k)\}$. Thus as k increases, the set of types that are subject to verification weakly decreases. Furthermore and their verification probabilities are given by $c_t = \frac{\max\{\mathcal{A}_{t^{EF}(k)}, \mathcal{A}_{\underline{t}(k)}\}}{\mathcal{A}_t}$. Hence the probability of verification $1-c_t$ for all the types that are subject to verification is also increasing in k. $\square$

References

Ben-Porath, E., E. Dekel, and B. L. Lipman (2014). Optimal allocation with costly verification. *The American Economic Review 104*(12), 3779–3813.

Border, K. C. (1991). Implementation of reduced form auctions: A geometric approach. *Econometrica: Journal of the Econometric Society 59*(4), 1175–1187.

Border, K. C. and J. Sobel (1987). Samurai accountant: A theory of auditing and plunder. *The Review of economic studies 54*(4), 525–540.

Chakravarty, S. and T. R. Kaplan (2013, January). Optimal allocation without transfer payments. *Games and Economic Behavior 77*(1), 1–20.

Dunne, S. A. and M. A. Loewenstein (1995). Costly verification of cost performance and the competition for incentive contracts. *The RAND Journal of Economics 26*(4), 690–703.

Eiselt, H. A. and C.-L. Sandblom (2007). *Linear programming and its applications.* Springer Science & Business Media.

Erlanson, A. and A. Kleiner (2019). A note on optimal allocation with costly verification. *Journal of Mathematical Economics 84*, 56–62.

Fujishige, S. (1980). Lexicographically optimal base of a polymatroid with respect to a weight vector. *Mathematics of Operations Research 5*(2), 186–196.

Harris, M., C. H. Kriebel, and A. Raviv (1982). Asymmetric information, incentives and intrafirm resource allocation. *Management Science*.

Hartline, J. J. D. and T. Roughgarden (2008). Optimal mechanism design and money burning. *Proceedings of the fourtieth annual ACM symposium on Theory of computing - STOC 08* (April), 1–23.

Li, Y. (2020). Mechanism design with costly verification and limited punishments. *Journal of Economic Theory 186*, 105000.

McAfee, P. R., J. McMillan, and R. P. McAfee (1992, February). Bidding Rings. *American Economic Review 82*(3), 579–599.

Milgrom, P. and J. Roberts (1988). An Economic Approach to Influence Activities in Organizations. *American Journal of Sociology 94*(s1), S154–S179.

Mookherjee, D. and I. Png (1989). Optimal auditing, insurance, and redistribution. *The Quarterly Journal of Economics 104*(2), 399–415.

Quah, J. K.-H. and B. Strulovici (2009). Comparative statics, informativeness, and the interval dominance order. *Econometrica 77*(6), 1949–1992.

Ross, M. (1986). Capital budgeting practices of twelve large manufacturers. *Financial Management 15*(4), 15–22.

Townsend, R. M. (1979, October). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory 21*(2), 265–293.

Vohra, R. V. (2012). Optimization and mechanism design. *Mathematical programming 134*(1), 283–303.

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