*American Economic Review 2014, 104(12): 3779–3813*
*http://dx.doi.org/10.1257/aer.104.12.3779*

# Optimal Allocation with Costly Verification†

*By* Elchanan Ben-Porath, Eddie Dekel, and Barton L. Lipman*

*A principal allocates an object to one of I agents. Each agent values receiving the object and has private information regarding the value to the principal of giving it to him. There are no monetary transfers, but the principal can check an agent's information at a cost. A* favored-agent mechanism *specifies a value $v^*$ and an agent $i^*$. If all agents other than $i^*$ report values below $v^*$, then $i^*$ receives the good and no one is checked. Otherwise, whoever reports the highest value is checked and receives the good if and only if her report is confirmed. All optimal mechanisms are essentially randomizations over optimal favored-agent mechanisms.* (*JEL* D82)

Consider a principal with a good to allocate among a number of agents, each of whom wants the good. Each agent $i$ knows the value the principal receives if he gives the good to $i$, but the principal does not. (This value need not coincide with the value to $i$ of getting the good.) The principal can verify the agents' private information at a cost, but cannot use transfers. There are a number of economic environments of interest that roughly correspond to this scenario; we discuss a few below. How does the principal maximize the expected gain from allocating the good less the costs of verification?

We characterize optimal mechanisms for such settings. We construct an optimal mechanism with a particularly simple structure which we call a *favored-agent mechanism*. There is a threshold value and a favored agent, say $i$. If each agent other than $i$ reports a value for the good below the threshold, then the good goes to the favored agent and no verification is required. If some agent other than $i$ reports a value above the threshold, then the agent who reports the highest value is checked. This agent receives the good if and only if his claims are verified and the good goes to any other agent otherwise.

In addition, we show that *every* optimal mechanism is essentially a randomization over optimal favored-agent mechanisms. In this sense, we can characterize the

* Ben-Porath: Department of Economics and Center for the Study of Rationality, Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905 (e-mail: benporat@math.huji.ac.il); Dekel: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, and School of Economics, Tel Aviv University (e-mail: dekel@northwestern.edu); Lipman: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215 (e-mail: blipman@bu.edu). We thank Ricky Vohra, Benjy Weiss, numerous seminar audiences, and three anonymous referees for helpful comments. We also thank the National Science Foundation, grants SES-0820333 and SES-1227434 (Dekel) and SES-0851590 (Lipman), and the US–Israel Binational Science Foundation (Ben-Porath and Lipman) for support for this research. Lipman also thanks Microsoft Research New England for their hospitality while the first draft was in progress and Alex Poterack for proofreading. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

† Go to http://dx.doi.org/10.1257/aer.104.12.3779 to visit the article page for additional materials and author disclosure statement(s).

full set of optimal mechanisms by focusing entirely on favored-agent mechanisms. By "essentially," we mean that any optimal mechanism has the same outcomes as such a randomization up to sets of measure zero.[1] An immediate implication is that if there is a unique optimal favored-agent mechanism, then there is essentially a unique optimal mechanism.

Finally, we give a variety of comparative statics. In particular, we show that an agent is more likely to be the favored agent the higher is the cost of verifying him, the "better" is his distribution of values in the sense of first-order stochastic dominance (FOSD), and the less risky is his distribution of values in the sense of second-order stochastic dominance (SOSD).

The standard mechanism-design approach to an allocation problem is to construct a mechanism with monetary transfers and ignore the possibility of the principal verifying the agent's information. In many cases obtaining information about the agent's type at a cost is quite realistic (see examples below). Hence we think it is important to add this option. In our exploration of this option, we take the opposite extreme position from the standard model and do not allow transfers. This obviously simplifies the problem, but we also find it reasonable to exclude transfers. Indeed, in many cases they are not used. In some situations, this may be because transfers have efficiency costs that are ignored in the standard approach. More specifically, the monetary resources each agent has might matter to the principal, so changing the allocation of these resources in order to allocate a good might be costly. In other situations, the value to the principal of giving the good to agent $i$ may differ from the value to agent $i$ of receiving the good, which reduces the usefulness of monetary transfers. For example, if the value to the principal of giving the good to $i$ and the value to agent $i$ of receiving it are independent, then, from the point of view of the principal, giving the good to the agent who values it most is the same as a random allocation. For these reasons, we adopt the opposite assumption to the standard one: we allow costly verification but do not allow for transfers.

We now discuss some examples of the environment described above. A firm may need to choose a unit to head up a new, prestigious project. A venture capital firm may need to choose which of a set of competing startups to fund. A government may need to choose in which town to locate a new hospital. A funding agency may have a grant to allocate to one of several competing researchers. A dean may have a job slot to allocate to one of several departments in the university. The head of personnel for an organization may need to choose one of several applicants for a job with a predetermined salary.

In each case, the head of the organization wishes to allocate a good to that agent who would use it in the way that best promotes the interests of the organization as a whole. Each agent, on the other hand, has his own reasons for wanting the good which may not align entirely with the incentives of the organization. For example, each town may want the hospital and put little or no value on having the hospital located in a different town. Each unit in the firm may want the resources and prestige associated with heading up the new project without regard to whether it is the unit which will best carry out the project. The venture capital firm wishes to maximize its

[1] Two mechanisms have the same "outcome" if the interim probabilities of checking and allocating the good are the same; see Section I for details.

profits, but each of the competing startups desires funding independently of which startup will yield the highest return.

It is also natural to assume that the agents have private information relevant to the choice by the head of the organization. The researchers know much more about the likelihood of a breakthrough than the funding agency. The individual towns know more about the likely level of use of a hospital than the central government. The departments know more about the characteristics of the people they would hire than does the dean.

In many of these situations, the head of the organization can, at a cost, obtain and process some or all of this information. The funding agency can investigate the research areas and progress to date of some or all of the competing researchers. The government can carry out a careful study of the towns. The firm can audit past performance of a unit and its current capabilities in detail. The head of personnel can verify some of the job applicants' claims.

Finally, monetary transfers are not practical or at least not used in many of these cases. Firms allocate budgets to individual units based on what activities they want these units to carry out—it would be self-defeating to have units bid these resources for the right to head up the new project. Similarly, it would be odd for a funding agency to ask researchers to pay in order to receive grants. Governments may ask towns to share in the cost of a hospital—but if part of the purpose of the project is to serve the poor, such transfers would undermine this goal.[2]

*Literature Review.*—Townsend (1979) initiated the literature on the principal-agent model with costly state verification. See also Gale and Hellwig (1985); Border and Sobel (1987); and Mookherjee and Png (1989). These models differ from what we consider in that they include only one agent and allow monetary transfers. In this sense, one can see our work as extending the costly state verification framework to multiple agents when monetary transfers are not possible. Our work is also related to Glazer and Rubinstein (2004, 2006), particularly the former which can be interpreted as a model of a principal and one agent with limited but costless verification and no monetary transfers. Finally, it is related to the literature on mechanism design and implementation with evidence—see Green and Laffont (1986); Bull and Watson (2007); Deneckere and Severinov (2008); Ben-Porath and Lipman (2012); Kartik and Tercieux (2012); and Sher and Vohra (2011). With the exception of Sher and Vohra (2011), these papers focus on general issues, rather than on specific mechanisms and allocation problems. Sher and Vohra do consider a specific allocation question, but it is a bargaining problem between a seller and a buyer, very different from what is considered here.

There is a somewhat less related literature on allocations without transfers but with costly signals (McAfee and McMillan 1992; Hartline and Roughgarden 2008; Yoon 2011; Condorelli 2012; and Chakravarty and Kaplan 2013).[3] In these papers,

[2] In a similar vein, Banerjee, Hanna, and Mullainathan (2013) give the example of a government that wishes to allocate free hospital beds. Their focus is the possibility that corruption may emerge in such mechanisms where it becomes impossible for the government to entirely exclude willingness to pay from playing a role in the allocation. We do not consider such possibilities here.

[3] There is also a large literature on allocations without transfers, namely the matching literature; see, e.g., Roth and Sotomayor (1990) for a classic survey and Abdulkadiroğlu and Sönmez (2013) for a more recent one.

agents can waste resources to signal their values and the principal's payoff is the value of the type receiving the good less the cost of the wasted resources. The papers differ in their assumptions about the cost, the number of goods to allocate, and so on, but the common feature is that wasting resources can be useful in allocating efficiently and that the principal may partially give up on allocative efficiency to save on these resources. See also Ambrus and Egorov (2012) who allow both monetary transfers and wasting of resources in a delegation model.

The remainder of the paper is organized as follows. In the next section, we present the model. Section II shows that all optimal mechanisms are essentially randomizations over optimal favored-agent mechanisms. In Section III, we characterize the set of best favored-agent mechanisms. In Section IV, we give comparative statics and discuss various properties of the optimal mechanism. In Section V, we sketch the proof of our uniqueness result, while Section VI discusses some simple extensions. Section VII concludes. Proofs not contained in the text are either in the Appendix or the online Appendix.

## I. Model

The set of agents is $\mathcal{I} = \{1, \ldots, I\}$. There is a single indivisible good to allocate among them. The value to the principal of assigning the object to agent $i$ depends on information which is known only to $i$. Formally, the value to the principal of allocating the good to agent $i$ is $t_i$ where $t_i$ is private information of agent $i$. We normalize so that types are always nonnegative and the value to the principal of assigning the object to no one is zero. As we explain in Section VI, the assumption that the principal always prefers allocating the object to the agents is used only to simplify some statements—the results easily extend to the case where the principal sometimes prefers to keep the object. We assume that the $t_i$s are independently distributed. The distribution of $t_i$ has a strictly positive density $f_i$ over the interval[4] $\mathcal{T}_i \equiv [\underline{t}_i, \bar{t}_i]$ where $0 \leq \underline{t}_i < \bar{t}_i < \infty$. We use $F_i$ to denote the corresponding distribution function. Let $\mathcal{T} = \prod_i \mathcal{T}_i$.

The principal can *check* the type of agent $i$ at a cost $c_i > 0$. We interpret checking as obtaining information (e.g., by requesting documentation, interviewing the agent, or hiring outside evaluators) which perfectly reveals the type of the agent being checked. The cost to the agent of providing information is assumed to be zero. We discuss these assumptions and the extent to which they can be relaxed in Section VI.

We assume that every agent strictly prefers receiving the object to not receiving it. Consequently, we can take the payoff to an agent to be the probability he receives the good. The intensity of the agents' preferences plays no role in the analysis, so these intensities may or may not be related to the types.[5] We also assume that each agent's reservation utility is less than or equal to his utility from not receiving the

[4] It is straightforward to drop the assumption of a finite upper bound for the support as long as all expectations are finite. Also, when in Section VI we allow the principal to prefer keeping the object, it is similarly straightforward to drop the assumption that the support is bounded below.

[5] Suppose we let the payoff of agent $i$ from receiving the good be $\overline{u}_i(t_i)$ and let his utility from not receiving it be $\underline{u}_i(t_i)$ where $\overline{u}_i(t_i) > \underline{u}_i(t_i)$ for all $i$ and all $t_i$. Then it is simply a renormalization to let $\overline{u}_i(t_i) = 1$ and $\underline{u}_i(t_i) = 0$ for all $t_i$. To see this, note that each of the incentive constraints will take the form

$$p\,\overline{u}_i(t_i) + (1 - p)\underline{u}_i(t_i) \geq p'\,\overline{u}_i(t_i) + \left(1 - p'\right)\underline{u}_i(t_i),$$

good. Since monetary transfers are not allowed, this is the worst payoff an agent could receive under a mechanism. Consequently, individual rationality constraints do not bind and so are disregarded throughout.

In its most general form, a mechanism can be quite complex, having multiple stages of cheap-talk statements by the agents and checking by the principal, where who can speak and which agents are checked depend on past statements and the results from past checks, finally culminating in the allocation of the good, perhaps to no one. However, it is not hard to show that one can use an argument similar to the Revelation Principle to restrict attention to a simple class of mechanisms.[6] Specifically, we show in part 1 of the online Appendix that we can consider only direct mechanisms (i.e., mechanisms which ask agents to report their types) for which truthful revelation is a Nash equilibrium and which have the following properties. First, for any vector of reported types, the mechanism selects (perhaps via randomization) at most one agent who is checked. If an agent is checked and (as will happen in equilibrium) found to have told the truth, then he receives the good. If no agent is checked, then the mechanism (again, perhaps randomly) selects which agent, if any, receives the good.

Hence we can write a mechanism as specifying for each vector of reports, two probabilities for each agent: the probability he is awarded the object without being checked and the probability he is awarded the object conditional on a successful check. Let $p_i(\mathbf{t})$ denote the total probability $i$ is assigned the good and $q_i(\mathbf{t})$ the probability $i$ is assigned the good and checked. Then these functions must satisfy $p_i : \mathcal{T} \to [0,1]$, $q_i : \mathcal{T} \to [0,1]$, $\sum_i p_i(\mathbf{t}) \leq 1$ for all $\mathbf{t} \in \mathcal{T}$, and $q_i(\mathbf{t}) \leq p_i(\mathbf{t})$ for all $i \in \mathcal{I}$ and all $\mathbf{t} \in \mathcal{T}$. Henceforth, the word "mechanism" will be used only to denote such a tuple of functions, generally denoted $(\mathbf{p}, \mathbf{q})$ for simplicity.

The principal's objective function is

$$E_{\mathbf{t}}\Big[\sum_i \big(p_i(\mathbf{t})t_i - q_i(\mathbf{t})c_i\big)\Big].$$

The incentive compatibility constraint for $i$ is

$$E_{\mathbf{t}_{-i}}\, p_i(\mathbf{t}) \geq E_{\mathbf{t}_{-i}}\big[p_i(t_i', \mathbf{t}_{-i}) - q_i(t_i', \mathbf{t}_{-i})\big], \quad \forall t_i, t_i' \in \mathcal{T}_i, \quad \forall i \in \mathcal{I}.$$

Given a mechanism $(\mathbf{p}, \mathbf{q})$, let

$$\hat{p}_i(t_i) = E_{\mathbf{t}_{-i}} p_i(t_i, \mathbf{t}_{-i})$$

and

$$\hat{q}_i(t_i) = E_{\mathbf{t}_{-i}} q_i(t_i, \mathbf{t}_{-i}).$$

where $p$ is the probability player $i$ gets the good when he is type $t_i$ and tells the truth and $p'$ is the probability he gets the good if he lies in some particular fashion. It is easy to rearrange this equation as $p[\bar{u}_i(t_i) - \underline{u}_i(t_i)] \geq p'[\bar{u}_i(t_i) - \underline{u}_i(t_i)]$. Since $\bar{u}_i(t_i) > \underline{u}_i(t_i)$, this holds if and only if $p \geq p'$, exactly the incentive constraint we have if $\bar{u}_i(t_i) = 1$ and $\underline{u}_i(t_i) = 0$.

[6] The usual version of the Revelation Principle does not apply to games with verification and hence cannot be used to obtain this conclusion. See Townsend (1988) for a discussion and an extension to a class of verification models which does not include ours.

The $2I$ tuple of functions $(\hat{\mathbf{p}}, \hat{\mathbf{q}})_{i \in \mathcal{I}}$ is the *reduced form* of the mechanism $(\mathbf{p}, \mathbf{q})$. We say that $(\mathbf{p}^1, \mathbf{q}^1)$ and $(\mathbf{p}^2, \mathbf{q}^2)$ are *equivalent* if $\hat{\mathbf{p}}^1 = \hat{\mathbf{p}}^2$ and $\hat{\mathbf{q}}^1 = \hat{\mathbf{q}}^2$ up to sets of measure zero. It is easy to see that we can write the incentive compatibility constraints and the objective function of the principal as a function only of the reduced form of the mechanism. Hence if $(\mathbf{p}^1, \mathbf{q}^1)$ is an optimal incentive compatible mechanism, $(\mathbf{p}^2, \mathbf{q}^2)$ must be as well. Therefore, we can only identify the optimal mechanism up to equivalence.

## II. The Sufficiency of Favored-Agent Mechanisms

Our main result in this section is that we can restrict attention to a class of mechanisms we call *favored-agent mechanisms*. To be specific, we show that every optimal mechanism is equivalent to a randomization over favored-agent mechanisms. Hence to compute the set of optimal mechanisms, we can simply optimize over the much simpler class of favored-agent mechanisms. In the next section, we use this result to characterize optimal mechanisms in more detail.

We say that $(\mathbf{p}, \mathbf{q})$ is a *favored-agent mechanism* if there exists a *favored agent* $i^* \in \mathcal{I}$ and a *threshold* $v^* \in \mathbf{R}$ such that the following holds up to sets of measure zero. First, if $t_i - c_i < v^*$ for all $i \neq i^*$, then $p_{i^*}(\mathbf{t}) = 1$ and $q_i(\mathbf{t}) = 0$ for all $i$. That is, if every agent other than the favored agent reports a "value" $t_i - c_i$ below the threshold, then the favored agent receives the object and no agent is checked. Second, if there exists $j \neq i^*$ such that $t_j - c_j > v^*$ and $t_i - c_i > \max_{k \neq i}(t_k - c_k)$, then $p_i(\mathbf{t}) = q_i(\mathbf{t}) = 1$ and $p_k(\mathbf{t}) = q_k(\mathbf{t}) = 0$ for all $k \neq i$. That is, if any agent other than the favored agent reports a value above the threshold, then the agent with the highest reported value (regardless of whether or not he is the favored agent) is checked and, if his report is verified, receives the good.[7]

Note that this is a very simple class of mechanisms. Optimizing over this set of mechanisms simply requires us to pick one of the agents to favor and a number for the threshold, as opposed to probability distributions over checking and allocation decisions as a function of the types.

Obviously, any randomization over optimal mechanisms is optimal.[8] Also, as noted in Section I, if a mechanism $(\mathbf{p}, \mathbf{q})$ is optimal, then any mechanism which is essentially the same in the sense of having the same reduced form up to sets of measure zero must also be optimal. Hence given any set of optimal mechanisms, we know that all mechanisms that are essentially equivalent to a randomization over mechanisms in this set must also be optimal.

THEOREM 1: *A mechanism is optimal if and only if it is essentially a randomization over optimal favored-agent mechanisms.*

[7] If there are several agents that maximize $t_i - c_i$, then one of them is chosen arbitrarily. This event has probability zero and so does not affect incentives or the principal's payoff.

[8] Randomizing refers to the pointwise convex combination of mechanisms $(\mathbf{p}, \mathbf{q})$.

Hence we can restrict attention to favored-agent mechanisms without loss of generality.[9] Furthermore, if there is a unique optimal favored-agent mechanism, then there is essentially a unique optimal mechanism.

Section V contains a sketch of the proof of this result.

A very incomplete intuition for this result is the following. For simplicity, suppose $c_i = c$ for all $i$ and suppose $\mathcal{T}_i = [0, 1]$ for all $i$. Clearly, the principal would ideally give the object to the agent with the highest $t_i$. Of course, this is not incentive compatible as each agent would claim to have type 1. By always checking the agent with the highest report, the principal can make this allocation of the good incentive compatible. So suppose the principal uses this mechanism.

Consider what happens when the highest reported type is below $c$. Obviously, it is better for the principal not to check in this case since it costs more to check than it could possibly be worth. Thus we can improve on this mechanism by only checking the agent with the highest report when that report is above $c$, giving the good to no one (and checking no one) when the highest report is below $c$. It is not hard to see that this mechanism is incentive compatible and, as noted, an improvement over the previous mechanism.

However, we can improve on this mechanism as well. Obviously, the principal could select any agent at random if all the reports are below $c$ and give the good to that agent. Again, this is incentive compatible. Since all the types are positive, this mechanism improves on the previous one.

The principal can do still better by further exploiting his selection of the person to give the good to when all the reports are below $c$. To see this, suppose the principal gives the good to agent 1 if all reports are below $c$. Continue to assume that if any agent reports a type above $c$, then the principal checks the highest report and gives the good to this agent if the report is true. This mechanism is clearly incentive compatible. However, the principal can also achieve incentive compatibility and the same allocation of the good while saving on checking costs: he doesn't need to check 1's report when he is the only agent to report a type above $c$. To see why this cheaper mechanism is also incentive compatible, note that if everyone else's type is below $c$, 1 gets the good no matter what he says. Hence 1 only cares what happens if at least one other agent's report is above $c$. In this case, he will be checked if he has the high report and hence cannot obtain the good by lying. Hence it is optimal for him to tell the truth.

This mechanism is the favored-agent mechanism with 1 as the favored agent and $v^* = 0$. Of course, if the principal chooses the favored agent and the threshold $v^*$ optimally, he must improve on this payoff.

This intuition does not show that some more complex mechanism cannot be superior, so it does not establish existence of an optimal mechanism in the favored-agent class, much less the uniqueness part of Theorem 1. Indeed, the proof of this theorem is rather complex.

REMARK 1: *It is worth noting that the favored-agent mechanism is ex post incentive compatible. To see this, note first that any agent i with $t_i - c_i$ above the threshold has a dominant strategy to report honestly. Second, any nonfavored agent i with*

[9] It is straightforward to show that an optimal mechanism exists, so Theorem 1 is not vacuous.

*$t_i - c_i$ below the threshold gets the good with zero probability regardless of his report. Finally, if the favored agent has $t_i - c_i$ below the threshold, he obtains the good if and only if all the other agents report values below the threshold, independently of his report.*[10] *Since ex post incentive compatibility is stricter than incentive compatibility, this implies that the favored-agent mechanism is also the optimal ex post incentive compatible mechanism.*

## III. Optimal Favored-Agent Mechanisms

We complete the specification of the optimal mechanism by characterizing the optimal threshold and the optimal favored agent. We show that conditional on the selection of the favored agent, the optimal favored-agent mechanism is unique. After characterizing the optimal threshold given the choice of the favored agent, we consider the optimal selection of the favored agent.

For each $i$, define $t_i^*$ by

$$E(t_i) = E\big(\max\{t_i, t_i^*\}\big) - c_i. \tag{1}$$

It is easy to show that $t_i^*$ is well-defined.[11]

To interpret $t_i^*$, suppose the principal is using a favored-agent mechanism with $i$ as the favored agent. Note that the principal does not take into account $i$'s report unless at least one of the other agents reports a value above the threshold. For intuition, think of the principal as not even asking $i$ for a report unless this happens. Suppose the profile of reports of the other agents is $\mathbf{t}'_{-i}$. Let $v = \max_{j \neq i}(t'_j - c_j)$. Suppose the principal is not committed to the threshold and consider his decision at this point. He can choose a threshold above $v$ or, equivalently, give the object to agent $i$ without checking him. If he does so, his expected payoff is $E(t_i)$. Alternatively, he can choose a threshold below $v$ or, equivalently, ask for a report from agent $i$ and give the object to the agent with the highest value of $t_j - c_j$ after a check. In expectation, this yields the principal $E\max\{t_i - c_i, v\}$. Hence equation (1) says that when $v = t_i^* - c_i$, the principal is indifferent between these two options.

A slight extension of this reasoning yields a proof of the following result.

[10] In other words, truth telling is an *almost dominant* strategy in the sense that it is an optimal strategy for every type of every player regardless of the strategies of his opponents. It is not dominant for an agent who is not favored and whose type is such that $t_i - c_i$ is below the threshold since his payoff is zero regardless of his report. We can make truth telling a dominant strategy by changing the mechanism off the equilibrium path. Specifically, suppose we modify our mechanism only by assuming that if an agent is checked and found to have lied, we select another agent at random, check him, and give the good to him if and only if he is found to have told the truth. It is easy to see that truth telling is a dominant strategy in this mechanism and that it generates the same allocation as the original mechanism.

[11] To see this, observe that the right-hand side of equation (1) is continuous and strictly increasing in $t_i^*$ for $t_i^* \geq \underline{t}_i$, below the left-hand side at $t_i^* = \underline{t}_i$, and above it as $t_i^* \to \infty$. Hence there is a unique solution. Note that if we allowed it, then when $c_i = 0$, we would have $t_i^* = \underline{t}_i$. (At $c_i = 0$, $t_i^*$ is not uniquely defined, but it is natural to take it to be the limit value as $c_i \downarrow 0$ which gives $t_i^* = \underline{t}_i$.) This fact together with what we show below implies the unsurprising observation that if all the costs are zero, the principal always checks the agent who receives the object and gets the same payoff as under complete information. Note also that if $c_i$ is very large, we have $t_i^* = E(t_i) + c_i > \bar{t}_i$.

THEOREM 2: *Within the set of favored-agent mechanisms with i as the favored agent, the unique optimal mechanism is obtained by setting the threshold* $v^*$ *equal to* $t_i^* - c_i$.

PROOF:

For notational convenience, let the favored agent $i$ equal 1. Contrast the principal's payoff to thresholds $t_1^* - c_1$ and $\hat{v}^* > t_1^* - c_1$. Given a profile of types for the agents other than 1, let $x = \max_{j \neq 1}(t_j - c_j)$—that is, the highest value of (and hence reported by) one of the other agents. Then the principal's payoff as a function of the threshold and $x$ is given by

| | $x < t_1^* - c_1 < \hat{v}^*$ | $t_1^* - c_1 < x < \hat{v}^*$ | $t_1^* - c_1 < \hat{v}^* < x$ |
|---|---|---|---|
| $t_1^* - c_1$ | $E(t_1)$ | $E\max\{t_1 - c_1, x\}$ | $E\max\{t_1 - c_1, x\}$ . |
| $\hat{v}^*$ | $E(t_1)$ | $E(t_1)$ | $E\max\{t_1 - c_1, x\}$ |

To see this, note that if $x < t_1^* - c_1 < \hat{v}^*$, then the principal gives the object to agent 1 without a check using either threshold. If $t_1^* - c_1 < \hat{v}^* < x$, then the principal gives the object to either 1 or the highest of the other agents with a check and so receives a payoff of either $t_1 - c_1$ or $x$, whichever is larger. Finally, if $t_1^* - c_1 < x < \hat{v}^*$, then with threshold $t_1^* - c_1$, the principal's payoff is the larger of $t_1 - c_1$ and $x$, while with threshold $\hat{v}^*$, she gives the object to agent 1 without a check and has payoff $E(t_1)$.

Recall that $t_1^* > \underline{t}_1$. Hence $t_1 < t_1^*$ with strictly positive probability. Therefore, for $x > t_1^* - c_1$, we have

$$E\max\{t_1 - c_1, x\} > E\max\{t_1 - c_1, t_1^* - c_1\}.$$

But the right-hand side is $E\max\{t_1, t_1^*\} - c_1$ which equals $E(t_1)$ by definition of $t_i^*$. Thus,

$$E\max\{t_1 - c_1, x\} > E(t_1).$$

Hence given that 1 is the favored agent, the threshold $t_1^* - c_1$ weakly dominates any larger threshold. A similar argument shows that the threshold $t_1^* - c_1$ weakly dominates any smaller threshold, establishing that it is optimal.

To see that the optimal mechanism in this class is unique, note that the comparison of threshold $t_1^* - c_1$ to a larger threshold $v^*$ is strict unless the middle column of the table above has zero probability. That is, the only situation in which the principal is indifferent between the threshold $t_1^* - c_1$ and the larger threshold $v^*$ is when the allocation of the good and checking decisions are the same with probability 1 given either threshold. That is, indifference occurs only when changes in the threshold do not change $(\mathbf{p}, \mathbf{q})$. Hence, conditional on favoring $i$, there is a unique best mechanism.

While $t_i^*$ resembles an index or reservation value of the sort often seen in the search literature, this specific definition is not standard in that literature. Indeed, it

is not straightforward to interpret it in terms of search. To see the point, compare it to the index identified in Weitzman (1979). In his model (with some minor adjustments for easier comparability), the value to learning the payoff to option $i$ is characterized by a critical value $t_i^*$ defined by

$$t_i^* = E(\max\{t_i, t_i^*\}) - c_i,$$

where $c_i$ is the cost of finding out the payoff of option $i$ and $t_i$ is the random variable giving the value of this payoff. Note that this is identical to our expression, except that $E(t_i)$ appears on the left-hand side of equation (1), not $t_i^*$. Weitzman's expression is easily interpreted. If the best alternative found so far has value $t_i^*$, then the agent is indifferent between stopping his search and choosing it versus checking option $i$ and then stopping. Our expression is not as obviously interpreted in terms of search.[12]

Now that the optimal threshold is characterized given the choice of the favored agent, it remains only to characterize the optimal favored agent.

THEOREM 3: *The optimal choice of the favored agent is any $i$ with $t_i^* - c_i = \max_j(t_j^* - c_j)$.*

A rough intuition for this result is that $t_i^* - c_i$ can be thought of as the standard that agents $j \neq i$ must satisfy to persuade the principal to not give the good to $i$ without checking anyone. Intuitively, the higher is this threshold, the more inclined the principal is to give the good to $i$ without checking anyone. It is not surprising that the agent toward whom the principal is most inclined in this sense is the principal's choice for the favored agent.

We sketch the main part of the proof here (see part 3 of the online Appendix for the omitted details). Fix any two agents, $i$ and $j$, and assume $t_i^* - c_i \geq t_j^* - c_j$. We will show that this inequality implies that given a threshold of $t_j^* - c_j$, it is weakly better for the principal to favor agent $i$ than agent $j$. By Theorem 2, favoring $i$ with a threshold of $t_i^* - c_i$ is better still, establishing the result.

If any agent $k$ other than $i$ or $j$ reports $t_k - c_k$ above the threshold, then the agent with the highest report is checked and receives the object, independent of which agent was favored. Hence we may as well condition on the event that all agents other than $i$ and $j$ report values below the threshold. Also, if *both* $i$ and $j$ report values above the threshold, again, it does not matter to the principal which agent was favored. Hence we only need to consider realizations such that at least one of these agents reports a value below the threshold.

Note that $j$'s report is above the threshold if $t_j - c_j \geq t_j^* - c_j$—i.e., $t_j \geq t_j^*$. On the other hand, $i$'s report is above the threshold if $t_i - c_i \geq t_j^* - c_j$ or $t_i \geq \hat{t}_i \equiv t_j^* - c_j + c_i$. Given this, we see that the principal is better off favoring $i$ than $j$ if

[12] Since the first draft of this paper, Doval (2013) considered a search model where this definition does appear. In Doval (2013), the searcher could choose option $i$ without checking it, yielding payoff $E(t_i)$. Thus $t_i^*$, computed according to our definition, emerges as the cutoff for the best option found so far with the property that it leaves the agent indifferent between taking the last option without checking it and checking the last option and then choosing between it and the outside option. To the best of our knowledge, this is the first time this definition has appeared in the search literature.

$$F_i(\hat{t}_i)F_j(t_j^*)E[t_i|t_i \leq \hat{t}_i] + [1 - F_i(\hat{t}_i)]F_j(t_j^*)E[t_i|t_i > \hat{t}_i]$$

$$+ F_i(\hat{t}_i)[1 - F_j(t_j^*)]\{E[t_j|t_j > t_j^*] - c_j\}$$

$$\geq F_i(\hat{t}_i)F_j(t_j^*)E[t_j|t_j \leq t_j^*] + [1 - F_i(\hat{t}_i)]F_j(t_j^*)\{E[t_i|t_i > \hat{t}_i] - c_i\}$$

$$+ F_i(\hat{t}_i)[1 - F_j(t_j^*)]E[t_j|t_j > t_j^*].$$

Rewriting,

$$F_i(\hat{t}_i)F_j(t_j^*)\big(E[t_i|t_i \leq \hat{t}_i] - E[t_j|t_j \leq t_j^*]\big)$$

$$+ \big[1 - F_i(\hat{t}_i)\big]F_j(t_j^*)c_i - F_i(\hat{t}_i)\big[1 - F_j(t_j^*)\big]c_j \geq 0.$$

This equation summarizes the change in the principal's payoff from switching from $j$ as the favored agent to $i$. The first term is the change in the expected payoff conditional on giving the object to the favored agent without a check, while the last two terms give the change in the expected costs of checking.

It is easy to show that this must hold if $F_i(\hat{t}_i) = 0$, so assume $F_i(\hat{t}_i) > 0$. It follows from the definition of $t_j^*$ that $F_j(t_j^*) > 0$, so we can rewrite this as

$$E[t_i|t_i \leq \hat{t}_i] + \frac{c_i}{F_i(\hat{t}_i)} - c_i \geq E[t_j|t_j \leq t_j^*] + \frac{c_j}{F_j(t_j^*)} - c_j. \tag{2}$$

We now show that this is implied by $t_i^* - c_i \geq t_j^* - c_j$.

Recall that $t_i^*$ is defined by

$$E(t_i) = E\max\{t_i, t_i^*\} - c_i,$$

so $t_i^* - c_i \geq t_j^* - c_j$ or, equivalently, $t_i^* \geq \hat{t}_i$ implies

$$E(t_i) = E\max\{t_i, t_i^*\} - c_i \geq E\max\{t_i, \hat{t}_i\} - c_i$$

or

$$\int_{\underline{t}_i}^{\hat{t}_i} t_i f_i(t_i)\, dt_i \geq F_i(\hat{t}_i)\hat{t}_i - c_i$$

or

$$E[t_i|t_i \leq \hat{t}_i] + \frac{c_i}{F_i(\hat{t}_i)} \geq \hat{t}_i = t_j^* - c_j + c_i.$$

Hence

$$E[t_i|t_i \leq \hat{t}_i] + \frac{c_i}{F_i(\hat{t}_i)} - c_i \geq t_j^* - c_j.$$

But the same rearranging of the definition of $t_j^*$ shows that

$$t_j^* - c_j = E[t_j | t_j \leq t_j^*] + \frac{c_j}{F_j(t_j^*)} - c_j.$$

Combining the last two inequalities yields equation (2) as asserted.

Summarizing, we see that the set of optimal favored-agent mechanisms is easily characterized. A favored-agent mechanism is optimal if and only if the favored agent $i$ satisfies $i \in \arg\max_j(t_j^* - c_j)$ and the threshold $v^*$ satisfies $v^* = \max_j(t_j^* - c_j)$. Thus the set of optimal mechanisms is equivalent to picking a favored-agent mechanism with threshold $v^* = \max_j(t_j^* - c_j)$ and randomizing over which of the agents in $\arg\max_j(t_j^* - c_j)$ to favor. Clearly for generic checking costs, there will be a unique $j$ maximizing $t_j^* - c_j$ and hence a unique optimal mechanism. Moreover, fixing $c_i$ and $c_j$, the set of $(F_i, F_j)$ such that $t_i^* - c_i = t_j^* - c_j$ is nowhere dense in the product weak* topology. Hence in either sense, such ties are non-generic.[13]

## IV. Properties of Optimal Mechanisms

Given that optimal mechanisms are favored-agent mechanisms, it is easy to compare outcomes under the optimal mechanism to the first best. For simplicity, suppose there are two agents and that the optimal mechanism favors agent 1 and has threshold $v^*$. Then there are exactly three ways the outcome can be inferior for the principal to the first best. First, if $t_2 - c_2 < v^*$ but $t_2 > t_1$, then agent 1 receives the object, even though the principal would be better off giving it to agent 2. Second, if $t_2 - c_2 > v^*$, then the principal ends up checking one of the agents, a cost he would avoid in the first best. Finally, if $t_2 - c_2 > v^*$, the good could still go to the "wrong" agent relative to the first best. In particular, if $t_1 > t_2$ but $t_2 - c_2 > t_1 - c_1$, then agent 2 will receive the good even though the principal prefers agent 1 to have it and similarly if we reverse the roles of 1 and 2.

Also, our characterization of the optimal favored agent and threshold makes it easy to compute optimal mechanisms and analyze their properties. Consider the following example. There are two agents. Agent 1's cost of being checked is large in the sense that $c_1 > \bar{t}_1 - E(t_1)$. As discussed in footnote 11, this implies $t_1^* = E(t_1) + c_1$. For concreteness, assume $E(t_1) = 1$. Suppose $c_2 = \varepsilon$ where $\varepsilon > 0$ but very small. Finally, assume $t_2$ is uniformly distributed over $[0.99, 1.99]$. It is easy to see that as $\varepsilon \downarrow 0$, we have $t_2^* - c_2 \downarrow \underline{t}_2 = 0.99 < 1 = t_1^* - c_1$. Hence for $\varepsilon$ sufficiently small, 1 will be the favored agent and the threshold $v^*$ will equal 1. However, suppose 2 reports $t_2 - c_2 > v^* = 1$. Note that $t_1 - c_1 < \bar{t}_1 - c_1 < E(t_1) = 1$. Thus if 2 is above the threshold, he receives the good for sure. Therefore, 1, even though he is favored, only receives the good if $t_2 - c_2 < 1$, or $t_2 < 1 + \varepsilon$. Recall $t_2 \sim U[0.99, 1.99]$, so for $\varepsilon$ small, 1 receives the good slightly more than 1 percent of the time, even though he is favored. Note that we have assumed very little about the distribution of $t_1$, so it could well be true that 1 is very likely to have the higher type.[14]

[13] We thank Yi-Chun Chen and Siyang Xiong for showing us a proof of this result.
[14] We thank an anonymous referee for raising this issue with a similar example.

This example highlights the fact that the favored agent is not necessarily the agent with the highest probability of receiving the good, even conditional on his type. So in what sense is the favored agent "favored?"

Compare a favored-agent mechanism with favored agent $i$ and threshold $v_i^*$ to a favored-agent mechanism with $j \neq i$ favored and threshold $v_j^*$. Then for *any* $v_i^*$, $v_j^*$, and $j \neq i$, agent $i$ (at least weakly) prefers the former mechanism to the latter. That is, it is *always* better to be favored than not. To see this, simply note that in the mechanism where $j$ is favored, a *necessary* condition for $i$ to receive the good is that $t_i - c_i \geq t_k - c_k$ for all $k$. This is true because $i$ can only receive the good if he or some other agent is above the threshold and $i$'s value is the highest. However, in the mechanism where $i$ is the favored agent, a *sufficient* condition for $i$ to receive the good is that $t_i - c_i \geq t_k - c_k$ for all $k$. Thus $i$ must weakly prefer being favored and typically will strictly prefer it.

For comparative statics, it is useful to give an equivalent definition of $t_i^*$. Our original definition (see equation (1)) can be rewritten as

$$\int_{\underline{t}_i}^{t_i^*} t_i f_i(t_i)\, dt_i = t_i^* F_i(t_i^*) - c_i$$

or

$$c_i = t_i^* F_i(t_i^*) - \int_{\underline{t}_i}^{t_i^*} t_i f_i(t_i)\, dt_i = \int_{\underline{t}_i}^{t_i^*} F_i(\tau)\, d\tau.$$

So an equivalent definition of $t_i^*$ is

$$\int_{\underline{t}_i}^{t_i^*} F_i(\tau)\, d\tau = c_i. \tag{3}$$

From (3), it is easy to see that an increase in $c_i$ increases $t_i^*$. Also, from our first definition of $t_i^*$, note that $t_i^* - c_i$ is that value of $v_i^*$ solving $E(t_i) = E\max\{t_i - c_i, v_i^*\}$. Obviously for fixed $v_i^*$, the right-hand side is decreasing in $c_i$, so $t_i^* - c_i$ must be increasing in $c_i$. Hence, all else equal, the higher is $c_i$, the more likely $i$ is to be selected as the favored agent. To see the intuition, note that if $c_i$ is larger, then the principal is less willing to check agent $i$'s report. Since the agent who is favored is the one the principal checks least often, this makes it more desirable to favor $i$.

It is also easy to see that a first-order or second-order stochastic dominance shift upward in $F_i$ reduces the left-hand side of equation (3) for fixed $t_i^*$, so to maintain the equality, $t_i^*$ must increase. Therefore, such a shift makes it more likely that $i$ is the favored agent and increases the threshold in this case. Hence both "better" (FOSD) and "less risky" (SOSD) agents are more likely to be favored.

The intuition for the effect of a first-order stochastic dominance increase in $t_i$ is clear. If agent $i$ is more likely to have high type, he is a better choice to be the favored agent. The intuition for why less risky agents are favored is that there is less benefit from checking $i$ if there is less uncertainty about his type.

Finally, equation (3) shows that if we change $F_i$ only at values of $t_i$ larger than $t_i^*$, then $t_i^*$ is unaffected. In this sense, the optimal favored agent and the optimal threshold are independent of the upper tails of the distributions of the $t_i$s. Intuitively, this is because the choice of threshold conditional on favoring agent $i$ is based on comparing $E(t_i)$ versus $E(\max\{t_i - c_i, v\})$ where $v = \max_{j \neq i}(t_j - c_j)$, for the reasons explained in Section III. Changing the probabilities for very high values of $t_i$ affects both parts of this comparison symmetrically and hence are irrelevant.

Now that we have shown how changes in the parameters affect the optimal mechanism, we turn to how these changes affect the payoffs of the principal and agents. First, consider changes in the realized type vector. Obviously, an increase in $t_i$ increases agent $i$'s probability of receiving the good and thus his ex post payoff. Therefore, his ex ante payoff increases with an FOSD shift upward in $F_i$. Similarly, the ex post payoffs of other agents are decreasing in $t_i$, so their ex ante payoffs decrease with an FOSD shift upward in $F_i$. However, the principal's ex post payoff does not necessarily increase as an agent's type increases: if at some profile, the favored agent is receiving the good without being checked, an increase in another agent's type might result in the same allocation but with costly verification.[15]

Nevertheless, an FOSD increase in any $F_i$ does increase the principal's ex ante payoff. To see this, suppose $t_i$ has distribution function $F_i$ and $\hat{t}_i$ has distribution function $\hat{F}_i$ where $\hat{F}_i$ dominates $F_i$ in the sense of FOSD. It is not hard to see that $F_i^{-1}\big(\hat{F}_i(\hat{t}_i)\big)$ has the same distribution as $t_i$. So suppose after changing the distribution of $i$'s type from $F_i$ to $\hat{F}_i$, the principal uses the same mechanism as he used before the shift, but converts $i$'s report to keep the distribution of $i$'s reports unchanged. That is, if $i$ reports $\hat{t}_i$, the principal treats this as a report of $F_i^{-1}\big(\hat{F}_i(\hat{t}_i)\big)$. Given this, we see that the principal's payoff is affected by the shift in the distribution only if he gives the good to agent $i$. In this case, he would have received $F_i^{-1}\big(\hat{F}_i(\hat{t}_i)\big)$ under the original distribution, but receives $\hat{t}_i$ instead. Since $F_i(\hat{t}_i) \geq \hat{F}_i(\hat{t}_i)$ by FOSD, we see that $\hat{t}_i \geq F_i^{-1}\big(\hat{F}_i(\hat{t}_i)\big)$, so the principal is better off.

Turning to the effect of changes in $c_i$, it is obvious that a decrease in $c_i$ makes the principal better off as he could use the same mechanism and save on costs. It is also easy to see that if agent $i$ is not favored, then increases in $c_i$ make him worse off and make all other agents better off, as long as the increase in $c_i$ does not change the identity of the favored agent. This is true simply because $i$ receives the good if and only if $t_i - c_i$ is large enough, so a higher $c_i$ makes $i$ less likely to receive the good and other agents more likely to do so.

On the other hand, changes in the cost of checking the favored agent have ambiguous effects in general. This is true because $t_i^* - c_i$ is *increasing* in $c_i$. Hence if the cost of the favored agent increases, all other agents are less likely to be above the threshold. This effect makes the favored agent better off and the other agents worse off. However, it is also true that if $t_j - c_j$ is above the threshold, then it is the comparison of $t_j - c_j$ to $t_i - c_i$ that matters. Clearly, an increase in $c_i$ makes this comparison worse for the favored agent $i$ and better for $j$. The total effect can be positive

[15] For example, if 1 is the favored agent and $\mathbf{t}$ satisfies $t_1 > t_1^*$ and $t_i - c_i < t_1^* - c_1$ for all $i \neq 1$, the payoff to the principal is $t_1$. If $t_2$, say, increases to $t_2'$ such that $t_1^* - c_1 < t_2' - c_2 < t_1 - c_1$, then the principal's payoff falls to $t_1 - c_1$.

or negative for the favored agent. For example, if $I = 2$ and $F_1 = F_2 = F$, then the favored agent benefits from an increase in his cost of being checked if the density $f$ is increasing and conversely if it is decreasing; see online Appendix 4. In short, every agent has an incentive to increase his cost of being checked if this can make him the favored agent and, depending on the density, the favored agent may have incentives to increase his cost of being checked even beyond that point. Clearly, such an incentive is potentially costly for the principal.

## V. Proof Sketch for Theorem 1

Mechanism design problems without transfers are very difficult to solve in general since the usual integral characterization of feasible allocation rules is unavailable. As we explain in this section, in our case, the structure of the problem provides some useful simplifications which allow for a complete characterization of optimal mechanisms.

The first step is to rewrite the optimization problem. Recall that $\hat{p}_i(t_i) = E_{\mathbf{t}_{-i}} p_i(t_i, \mathbf{t}_{-i})$ and $\hat{q}_i(t_i) = E_{\mathbf{t}_{-i}} q_i(t_i, \mathbf{t}_{-i})$. We can write the incentive compatibility constraint as

$$\hat{p}_i(t_i') \geq \hat{p}_i(t_i) - \hat{q}_i(t_i), \quad \forall t_i, t_i' \in \mathcal{T}_i.$$

That is, the payoff to type $t_i'$ from telling the truth exceeds the payoff to claiming to be type $t_i$. The unusual property of our incentive compatibility constraint is that the payoff to falsely claiming to be type $t_i$ does not depend on the true type $t_i'$. We have this structure because any lie is caught if and only if the agent is checked and because we can normalize so that an agent's payoffs to receiving or not receiving the good do not depend on his type.

Normally, one rearranges the incentive constraints to say that for each type, telling the truth is better than the best possible lie. Because the payoff to lying does not depend on the truth, we can rearrange the incentive constraint to say that the worst truth is better than any lie. In other words, a mechanism is incentive compatible if and only if it satisfies

$$\inf_{t_i' \in \mathcal{T}_i} \hat{p}_i(t_i') \geq \hat{p}_i(t_i) - \hat{q}_i(t_i), \quad \forall t_i \in \mathcal{T}_i.$$

Letting $\varphi_i = \inf_{t_i' \in \mathcal{T}_i} \hat{p}_i(t_i')$, we can rewrite the incentive compatibility constraint as

$$\hat{q}_i(t_i) \geq \hat{p}_i(t_i) - \varphi_i, \quad \forall t_i \in \mathcal{T}_i.$$

Because the objective function is strictly decreasing in $\hat{q}_i(t_i)$, this constraint must bind, so

$$\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i. \tag{4}$$

We can substitute this result into the objective function and rewrite it as

$$E_{\mathbf{t}}\Big[\sum_i p_i(\mathbf{t})t_i - \sum_i c_i q_i(\mathbf{t})\Big] = \sum_i E_{t_i}[\hat{p}_i(t_i)t_i - c_i\hat{q}_i(t_i)] \tag{5}$$

$$= \sum_i E_{t_i}[\hat{p}_i(t_i)(t_i - c_i) + \varphi_i c_i]$$

$$= E_{\mathbf{t}}\Big[\sum_i [p_i(\mathbf{t})(t_i - c_i) + \varphi_i c_i]\Big]. \tag{6}$$

Some of the arguments below will use the reduced form probabilities and hence rely on the first expression, (5), for the payoff function, while others focus on the "non-reduced" mechanism and so rely on the second expression, (6).

From this point forward, we treat the principal's problem as choosing the $\varphi_i$s and $p_i$s subject to the constraints that the $p_i$s be well-defined probabilities and the constraints implied by $\varphi_i = \inf_{t_i} \hat{p}_i(t_i)$. Given this, equation (6) expresses the key trade-off in the model. Recall that $\varphi_i$ is the minimum probability for any type of $i$ to receive the good. Since the principal only needs to check $t_i$ often enough that a fake claim of $t_i$ succeeds with probability equal to this minimum, a higher value of $\varphi_i$ means that the principal does not have to check agent $i$ as often, as shown by equation (4). Thus, as the payoff function (6) shows, this helps the principal in proportion to the cost $c_i$ that he saves. However, this puts a more severe constraint on $\hat{p}_i$, thus tending to force a less efficient allocation of the good.

Next, we give a partial characterization of the optimal mechanism taking the $\varphi_i$s as given and optimizing over the $p_i$s. This partial characterization enables us to reduce the problem to choosing the $\varphi_i$s and one other variable which will turn out to be the threshold.

Thus in what follows, we fix $\varphi_i \in [0, 1]$ for each $i$ and characterize the solution to what we will call the *relaxed problem* of maximizing (6) by the choice of functions $p_i : \mathcal{T} \to [0, 1]$ for $i \in \mathcal{I}$ subject to $\sum_i p_i(\mathbf{t}) \leq 1$ for all $\mathbf{t}$ and $E_{\mathbf{t}_{-i}} p_i(\mathbf{t}) \geq \varphi_i$ for all $t_i$ and all $i$.[16] Note that since the $\varphi_i$s are fixed, we can take the objective function to be

$$\sum_i E_{t_i}\hat{p}_i(t_i)(t_i - c_i) = E_{\mathbf{t}}\Big[\sum_i p_i(\mathbf{t})(t_i - c_i)\Big].$$

We show below that every optimal solution to the relaxed problem is what we call a *threshold mechanism*. Specifically, every solution has a threshold $v^*$ with the following properties. First, if $t_i - c_i < v^*$, then $\hat{p}_i(t_i) = \varphi_i$. Second, for any $\mathbf{t}$ such that some agent $i$ has $t_i - c_i > v^*$, the agent with the highest $t_i - c_i$ receives the good with probability 1. As we show below, for an appropriate selection of the $\varphi_i$s, a threshold mechanism is equivalent to a favored-agent mechanism.

THEOREM 4: *Every solution to the relaxed problem is a threshold mechanism.*

[16] The constraint set is nonempty if and only if $\sum_i \varphi_i \leq 1$. It is not hard to show that the solution to the overall problem will necessarily satisfy this and so we assume it in what follows.

To see the intuition for this, first note that the solution to the relaxed problem would be trivial if we did not have the constraint that $E_{\mathbf{t}_{-i}} p_i(\mathbf{t}) \geq \varphi_i$. Without this constraint, the relaxed problem is equivalent to allocating the good where the principal receives $t_i - c_i$ if he allocates the good to $i$. Hence the solution would be $p_i(\mathbf{t}) = 0$ for all $i$ if $t_i < c_i$ for all $i$ and otherwise $p_i(\mathbf{t}) = 1$ for that $i$ such that $t_i - c_i = \max_k(t_k - c_k)$. Typically, this solution will violate the lower bound constraint on $\hat{p}_i(t_i)$. For example, if $t_i < c_i$, this solution would have $\hat{p}_i(t_i) = 0$, so the constraint could only be satisfied if $\varphi_i = 0$.

Thus the constraint forces the principal to sometimes allocate the good to an agent with $t_i < c_i$ or to an agent who does not have the highest value of $t_i - c_i$. When should the principal deviate from giving the good to the agent with the highest value of $t_i - c_i$ in order to satisfy the constraint? Intuitively, it is obvious that the principal should do this only when $\max_i(t_i - c_i)$ is relatively small. That is, it is natural that there should be a threshold for $\max_i(t_i - c_i)$ such that the principal allocates the good to the agent with the highest $t_i - c_i$ when this threshold is met. It is also natural that if the principal does not want to allocate the good to a particular type of a particular agent because $t_i - c_i$ is relatively small, then he gives it to that agent with the smallest probability allowed by the constraints. It is not as immediately intuitive that the same threshold should apply to every $i$ for this second statement or that this threshold should also apply to the first statement, but this is the content of Theorem 4.

We sketch the proof of Theorem 4 below, but first explain how this result enables us to complete the proof of Theorem 1.

Theorem 4 implies that $\hat{\mathbf{p}}$ is completely pinned down as a function of $v^*$ and the $\varphi_i$s. Specifically, if $t_i - c_i > v^*$, then $\hat{p}_i(t_i)$ must be the probability that $t_i - c_i > \max_{j \neq i}(t_j - c_j)$. If $t_i < t^*$, then $\hat{p}_i(t_i) = \varphi_i$. By substituting into equation (5), we can write the principal's payoff as a function only of $v^*$ and the $\varphi_i$s. Note that Theorem 4 does *not* say that any $v^*$ is consistent with given $\varphi_i$s. To see why, note that the theorem pins down $\hat{p}_i(t_i)$ as a function of $v^*$ for any $t_i > v^* + c_i$. But if this implied value of $\hat{p}_i(t_i)$ is below $\varphi_i$, then the threshold $v^*$ is not feasible given the $\varphi_i$s.

It turns out to be more convenient to fix the threshold $v^*$ and ask which $\varphi_i$s are consistent with it. We show in Appendix A (Lemma 1) that for any $v^*$, the set of consistent $\varphi_i$s is convex. It is also easy to show that the objective function of the principal, holding $v^*$ fixed, is linear in the $\varphi_i$s. Hence given $v^*$, there must be a solution to the principal's optimization problem at an extreme point of the set of consistent $\varphi_i$s. Furthermore, every optimal choice of the $\varphi_i$s is a randomization over optimal extreme points.

The last step is to show that the optimal extreme points correspond to favored-agent mechanisms. First, we show that extreme points take on one of two forms. In one type of extreme point, all but one of the $\varphi_i$s is set to zero and the remaining one is "as large as possible." For convenience, consider the extreme point where $\varphi_j = 0$ for all $j \neq 1$ and $\varphi_1$ is set as high as possible. In this case, we have a favored-agent mechanism where 1 is the favored agent and $v^*$ is the threshold. To see this, first observe that the allocation probabilities match this. If every agent other than the favored agent has $t_j - c_j < v^*$, then each of these agents receives the good with probability $\varphi_j = 0$. It is not hard to show that making $\varphi_1$ "as large as possible" entails giving the good to agent 1 in this situation, whether or not $t_1 - c_1$ is above

$v^*$. If some agent $j \neq 1$ has $t_j - c_j > v^*$, then he receives the good if and only if he has the highest value of $t_j - c_j$, just as required by the favored-agent mechanism. Equation (4) can be used to show that the checking probabilities, up to equivalence, are as required for a favored-agent mechanism.

The other type of extreme point is where $\varphi_i = 0$ for all $i$. One way this can come about is if there are two agents, say $i$ and $j$, with $t_i - c_i > v^*$ and $t_j - c_j > v^*$ with probability 1. In this case, we have the same mechanism as a favored-agent mechanism with threshold $v^*$ and *any* agent selected as the favored agent. Since we always have more than one agent above the threshold, the identity of the favored agent is irrelevant as the mechanism always picks the agent $i$ with the largest $t_i - c_i$, checks him, and gives him the good after a successful check.

There is one other way to have an extreme point with $\varphi_i = 0$ for all $i$. This occurs when every agent has a strictly positive probability of being below the threshold. Recall that Theorem 4 implies that if $t_i - c_i < v^*$, then $\hat{p}_i(t_i) = \varphi_i$. Hence this mechanism has the property that for any $t$ such that all agents have $t_i - c_i < v^*$, no agent receives the good. It is easy to use our assumption that the value of the good to the principal is zero and that types are positive with probability 1 to show that this extreme point cannot be optimal.[17]

We now sketch the proof of Theorem 4. The proof of this theorem has some technical complications because of our use of a continuum of types. The continuum has the significant advantage that it makes the statement of the optimal mechanism much cleaner, while we would have messy specifications of the mechanism at certain "boundary" types if we assumed finitely many types instead. On the other hand, the continuum of types assumption introduces measurability considerations and makes the argument more complex. We could retain the simplicity of the statement of the mechanism but simplify the proof if we only wished to characterize *an* optimal mechanism rather than *all* optimal mechanisms. To characterize an optimal mechanism, we could approximate the continuum type space with a finite partition, characterize the optimal mechanism measurable with respect to this partition (and hence focus on an effectively finite model), and take limits. It is not hard to show that the limiting mechanism would have to be optimal in the continuum model, but we know of no way to show that all optimal mechanisms in the continuum model would be characterized this way. Thus we are led to work with the continuum model in spite of its complexities.

For the purposes of this sketch, we will simplify by considering the case of finite type spaces and will disregard certain boundary issues. This enables us to convey the basic ideas of the construction without getting bogged down in measurability issues, allowing a much simpler approach. We now use this simplification to sketch the argument that any solution to the relaxed problem is a threshold mechanism.

The proof sketch has four steps. First, we observe that every solution to the relaxed problem is monotonic in the sense that higher types are more likely to receive the object. That is, for all $i$, $t_i > t_i'$ implies $\hat{p}_i(t_i) \geq \hat{p}_i(t_i')$. To see why this holds, suppose we have a solution which violates this monotonicity property so that we have types $t_i$ and $t_i'$ such that $\hat{p}_i(t_i) < \hat{p}_i(t_i')$ even though $t_i > t_i'$. For simplicity, suppose that these

[17] As we discuss in Section VI, if the principal has a strictly positive value for keeping the good, then this extreme point can be optimal.

two types have the same probability.[18] Then consider the mechanism $\mathbf{p}^*$ which is the same as this one *except* we flip the roles of $t_i$ and $t_i'$. That is, for any type profile $\hat{\mathbf{t}}$ where $\hat{t}_i \notin \{t_i, t_i'\}$, we let $p_j^*(\hat{\mathbf{t}}) = p_j(\hat{\mathbf{t}})$ for every $j$. For any $\mathbf{t}_{-i}$, we set $p_j^*(t_i, \mathbf{t}_{-i}) = p_j(t_i', \mathbf{t}_{-i})$ and $p_j^*(t_i', \mathbf{t}_{-i}) = p_j(t_i, \mathbf{t}_{-i})$ for every $j$. Since the probabilities of $t_i$ and $t_i'$ are the same, our independence assumption implies that for every $j \neq i$, agent $j$ is unaffected by the change in the sense that $\hat{p}_j^*(t_j) = \hat{p}_j(t_j)$ for all $t_j$. Obviously, $\hat{p}_i(t_i') = \hat{p}_i^*(t_i) > \hat{p}_i^*(t_i') = \hat{p}_i(t_i)$. Since the original mechanism satisfied the constraint that $\hat{p}_j(t_j) \geq \varphi_j$ for all $t_j$ and all $j$, the new mechanism satisfies this constraint as well. It is easy to see from equation (5) that this change improves the objective function, so the original mechanism could not have been optimal.

This monotonicity property implies that any solution to the relaxed problem has the property that there is a cutoff type, say $\hat{t}_i \in [\underline{t}_i, \bar{t}_i]$, such that $\hat{p}_i(t_i) = \varphi_i$ for $t_i < \hat{t}_i$ and $\hat{p}_i(t_i) > \varphi_i$ for $t_i > \hat{t}_i$.

The second step shows that if we have a type profile $\mathbf{t} = (t_1, \ldots, t_I)$ such that $t_i - c_i > t_j - c_j > \hat{t}_j - c_j$, then any solution to the relaxed problem has $p_j(\mathbf{t}) = 0$. To see this, suppose to the contrary that we have a solution with $p_j(\mathbf{t}) > 0$. Then we can change the mechanism by lowering this probability slightly and raising the probability of giving the good to $i$. Since $t_j > \hat{t}_j$, we have $\hat{p}_j(t_j) > \varphi_j$ before the change, so if the change is small enough, we still satisfy this constraint. Since $t_i - c_i > t_j - c_j$, the value of the objective function in equation (6) increases, so the original mechanism could not have been optimal.

The third step is to show that for a type profile $\mathbf{t} = (t_1, \ldots, t_I)$ such that $t_i > \hat{t}_i$ and $t_j < \hat{t}_j$, we must have $p_j(\mathbf{t}) = 0$. Because this step is more involved, we postpone explaining it till the end of the argument. So we continue the proof sketch taking this step as given.

The fourth step is to show that $\hat{t}_i - c_i = \hat{t}_j - c_j$ for all $i$ and $j$. To see this, suppose to the contrary that $\hat{t}_j - c_j > \hat{t}_i - c_i$. Then consider a type profile $\mathbf{t} = (t_1, \ldots, t_I)$ such that $\hat{t}_j - c_j > t_j - c_j > t_i - c_i > \hat{t}_i - c_i$ and $t_k < \hat{t}_k$ for all $k \neq i, j$. From our second step, the fact that $t_j - c_j > t_i - c_i > \hat{t}_i - c_i$ implies $p_i(\mathbf{t}) = 0$. However, from our third step, $t_i > \hat{t}_i$ and $t_k < \hat{t}_k$ implies $p_k(\mathbf{t}) = 0$ for all $k \neq i$. But $t_i > \hat{t}_i$ implies $\hat{p}_i(t_i) > \varphi_i$. The only way this could be optimal is if $t_i > c_i$. But then we can improve the objective function at $t$ by setting $p_i(\mathbf{t}) = 1$, a contradiction. Hence $\hat{t}_i - c_i = \hat{t}_j - c_j$ for all $i$ and $j$.

Let $v^*$ equal the common value of $\hat{t}_i - c_i$. To sum up the four steps, we see that any solution to the relaxed problem is characterized by a threshold $v^*$. If $t_i - c_i < v^*$, then since $v^* = \hat{t}_i - c_i$, we have $t_i < \hat{t}_i$ and hence $\hat{p}_i(t_i) = \varphi_i$. Also, suppose we have a type profile $\mathbf{t}$ such that $t_i - c_i > v^*$ for some $i$. From our third step, if $t_i - c_i > v^* > t_j - c_j$, then $p_j(\mathbf{t}) = 0$. Hence no agent below the threshold can receive the object. As argued in the last step, we must also have $\sum_i p_i(\mathbf{t}) = 1$. Hence if only one agent is above the threshold, he receives the object. From our second step, if we have $t_i - c_i > t_j - c_j > v^*$, then $p_j(\mathbf{t}) = 0$. Hence if two or more agents are above the threshold, the agent with the highest $t_i - c_i$ receives the object. Either way, then,

[18] If, say, $t_i$ has higher probability than $t_i'$, we can simply create two "versions" of $t_i$, one of which has the same probability as $t_i'$. Then apply this argument to $t_i'$ and the version of $t_i$ with the same probability, leaving the other version unchanged.

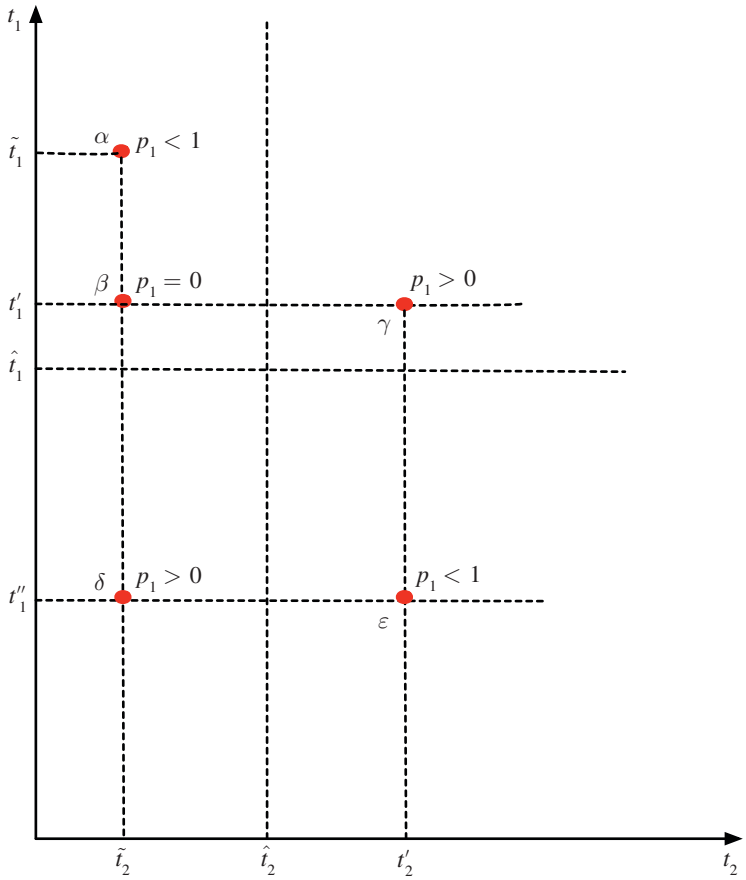


FIGURE 1

if there is at least one agent $i$ with $t_i - c_i$ above the threshold, the agent with the highest $t_i - c_i$ receives the object. Hence we have a threshold mechanism.

This concludes the proof sketch, except for proving step 3 to which we now turn. For added simplicity, we sketch this argument only for the case of two agents.[19] We show that for a type profile $\mathbf{t} = (t_1, t_2)$ such that $t_1 > \hat{t}_1$ and $t_2 < \hat{t}_2$, we must have $p_2(\mathbf{t}) = 0$. To see this, suppose we have a solution to the relaxed problem such that this property is violated at the point labeled $\alpha = (\tilde{t}_1, \tilde{t}_2)$ in Figure 1, where $\tilde{t}_1 > \hat{t}_1$ while $\tilde{t}_2 < \hat{t}_2$. So suppose that at $\alpha$, $p_2(\tilde{\mathbf{t}}) > 0$ so player 1 receives the good with probability strictly less than 1.

Then at any point directly below $\alpha$ but above $\hat{t}_1$, such as the one labeled $\beta = (t'_1, \tilde{t}_2)$, player 1 must receive the good with probability zero. This follows because if 1 did receive the good with strictly positive probability at $\beta$, we could change the mechanism by lowering this probability slightly, giving the good to 2 at $\beta$ with higher probability, and increasing the probability with which 1 receives the good at $\alpha$. By choosing these probabilities appropriately, we do not affect $\hat{p}_2(\tilde{t}_2)$ so this remains at $\varphi_2$. Also, by making the reduction in $p_1$ small enough,

[19] The case of three or more agents works similarly but the construction of the improving mechanism is more complex.

$\hat{p}_1(t_1')$ will remain above $\varphi_1$. Hence this new mechanism would satisfy the constraints for the relaxed problem. Since it would switch probability from one type of player 1 to a higher type, the new mechanism would be better than the old one, implying the original one was not a solution to the relaxed problem.[20]

Similar reasoning implies that for every $t_1$, we must have $\sum_i p_i(t_1, \tilde{t}_2) = 1$. Otherwise, as we show below, we could increase $\hat{p}_1(\tilde{t}_1)$ without affecting any other $\hat{p}_i$. This increase would improve the objective function as $\tilde{t}_1 > \hat{t}_1$ implies $\tilde{t}_1 > c_1$, a contradiction. For $t_1 = \tilde{t}_1$, it is obvious that we could increase $\hat{p}_1(\tilde{t}_1)$ since we could increase $p_1(\tilde{t}_1, \tilde{t}_2)$. For $t_1 \neq \tilde{t}_1$, this follows because otherwise, the principal could increase $p_2(t_1, \tilde{t}_2)$, decrease $p_2(\tilde{t}_1, \tilde{t}_2)$, and increase $p_1(\tilde{t}_1, \tilde{t}_2)$. If we choose the sizes of these changes appropriately, $\hat{p}_2(\tilde{t}_2)$ is unchanged but $\hat{p}_1(\tilde{t}_1)$ is increased.

Since player 1 receives the good with zero probability at $\beta$ but type $t_1'$ does have a strictly positive probability overall of receiving the good (as $t_1' > \hat{t}_1$), there must be some point like the one labeled $\gamma = (t_1', t_2')$ where 1 receives the good with strictly positive probability. We do not know whether $t_2'$ is above or below $\hat{t}_2$; the position of $\gamma$ relative to this cutoff plays no role in the argument to follow.

Finally, there must be a $t_1'' \neq \tilde{t}_1$ (not necessary below $\hat{t}_1$) corresponding to points $\delta$ and $\varepsilon$ where $p_1$ is strictly positive at $\delta$ and strictly less than 1 at $\varepsilon$. To see that such a $t_1''$ must exist, suppose not. Then for all $t_1 \neq \tilde{t}_1$, either $p_1(t_1, \tilde{t}_2) = 0$ or $p_1(t_1, t_2') = 1$. Since $\sum_i p_i(t_1, \tilde{t}_2) = 1$ for all $t_1 \neq \tilde{t}_1$, either $p_2(t_1, \tilde{t}_2) = 1$ or $p_2(t_1, \tilde{t}_2') = 0$. Either way, $p_2(t_1, \tilde{t}_2) \geq p_2(t_1, t_2')$ for all $t_1 \neq \tilde{t}_1$.

We now show that $p_2(\tilde{t}_1, t_2') = 0$, implying $p_2(\tilde{t}_1, \tilde{t}_2) \geq p_2(\tilde{t}_1, t_2')$ as well. Suppose by contradiction that $p_2(\tilde{t}_1, t_2') > 0$, so $p_1(\tilde{t}_1, t_2') < 1$. Then we could slightly lower $p_1$ and raise $p_2$ at the point $\gamma = (t_1', t_2')$ and raise $p_1$ and lower $p_2$ at the point $(\tilde{t}_1, t_2')$. By choosing these changes appropriately, $\hat{p}_2(t_2')$ is unaffected and $\hat{p}_1(t_1')$ remains above $\varphi_1$. Since $\tilde{t}_1 > t_1'$, this improves the payoff, a contradiction.

Hence, $p_2(t_1, \tilde{t}_2) \geq p_2(t_1, t_2')$ for every $t_1$. But we also have $p_2(t_1', \tilde{t}_2) = 1 > 1 - p_1(t_1', t_2') \geq p_2(t_1', t_2')$. So $\hat{p}_2(\tilde{t}_2) > \hat{p}_2(t_2')$. But $\hat{p}_2(\tilde{t}_2) = \varphi_2$, so this implies $\hat{p}_2(t_2') < \varphi_2$, which violates the constraints of the relaxed problem.

Now we use $p_1(t_1'', \tilde{t}_2) > 0$ and $p_1(t_1'', t_2') < 1$ to derive a contradiction to the optimality of the mechanism. Specifically, we change the specification of $\mathbf{p}$ at the points $\alpha$, $\gamma$, $\varepsilon$, and $\delta$ in a way that lowers the probability that 1 gets the object at $\gamma$ and raises the probability he gets it at $\alpha$ by the same amount, while maintaining the constraints. Since 1's type is higher at $\alpha$, this is an improvement, implying that the original mechanism does not solve the relaxed problem. For simplicity, assume $\tilde{t}_1$, $t_1'$, and $t_1''$ all have the same probability and that $\tilde{t}_2$ and $t_2'$ have the same probability.[21] Let $\Delta > 0$ be a "small" positive number. All the changes in $\mathbf{p}$ that we now define involve increases and decreases by the same amount $\Delta$. At $\gamma$, lower $p_1$ and increase $p_2$. At $\varepsilon$, do the opposite—i.e., raise $p_1$ and lower $p_2$. Because the probabilities of $t_1'$ and $t_1''$ are the same, $\hat{p}_2(t_2')$ is unchanged. Also, if $\Delta$ is small enough, $\hat{p}_1(t_1')$ remains above $\varphi_1$. Thus the constraints are maintained. Now that we have increased $p_1$ at $\varepsilon$, we can decrease it at $\delta$ while increasing $p_2$, keeping $\hat{p}_1(t_1'')$ unchanged as $\tilde{t}_2$ and

[20] Since $\hat{p}_2(\tilde{t}_2)$ is unchanged, the ex ante probability of type $\tilde{t}_1$ getting the good goes up by the same amount that the ex ante probability of the lower type $t_1'$ getting it goes down.

[21] If they do not have the same probability, the same approach as outlined in footnote 18 can be used to complete the argument.

$t_2'$ have the same probability. Finally, since we have increased $p_2$ at $\delta$, we can decrease it at $\alpha$ while increasing $p_1$, keeping $\hat{p}_2(\tilde{t_2})$ unchanged. Note that the overall effect of these changes is a reduction of $\Delta$ in the probability that 1 gets the object at $\gamma$ and an increase of $\Delta$ in the probability that he gets the object at $\alpha$, while maintaining all constraints.

## VI. Reservation Values, Small Fines/Costly Disclosure, Weaker Verification

In this section, we discuss three issues. First, we show that our analysis extends to the case where the principal has a strictly positive payoff to keeping the object. Second, under an additional simplifying assumption, we characterize the optimal mechanism when it is costly to an agent to have his report verified. A byproduct of this analysis is a characterization of the optimal mechanism when the principal can impose limited fines on dishonest agents. Finally, we discuss the robustness of our analysis to weakening our assumptions on verification.

Turning to the first point, let $R > 0$ denote the principal's payoff from retaining the object. To analyze this case, note first that we can renormalize by subtracting $R$ from every type of every player and changing the principal's payoff from keeping the object to 0. This effectively subtracts $R$ from the principal's payoff to every action and so cannot change the optimal mechanism. Thus we can consider this case simply by relaxing our assumption that $\underline{t}_i > 0$ for all $i$.

This change affects very little of our analysis. To see this, refer again to the proof sketch in Section V. It is easy to see that our analysis of incentive compatibility is unaffected by allowing some $t_i$s to be negative since the $t_i$s play no role in incentive compatibility. Hence the rewriting of the objective function in the form of equations (5) and (6) is unaffected.

Similarly, the threshold mechanism result, Theorem 4, is unaffected. This result characterizes the solution of

$$\max_{p_1,\ldots,p_I} E_{\mathbf{t}}\Big[\sum_i p_i(\mathbf{t})(t_i - c_i)\Big]$$

subject to feasibility and $\hat{p}_i(t_i) \geq \varphi_i$ for all $i$ and $t_i$. This characterization again has nothing to do with the sign of the $t_i$s. The sign of $t_i - c_i$ is certainly relevant, but the assumptions we used to prove Theorem 4 did not impose any particular sign on this expression. Hence this result still holds even when some of the $t_i$s can be negative.

Once we have Theorem 4, the convexity of the set of $\varphi_i$s, the linearity of the objective function in $\varphi_i$s, the conclusion that the optimum is at an extreme point, and the characterization of the extreme points all have nothing to do with the signs of the $t_i$s.

The only time the assumption that the types are positive is used in the proof of Theorem 1 is when we rule out one of the extreme points for $\varphi$. Specifically, as discussed in the proof sketch, we can have an extreme point of the form $\varphi_i = 0$ for all $i$. If every agent has a strictly positive probability of being below the threshold, then this mechanism has a strictly positive probability that no agent receives the good. When types are positive with probability 1, this cannot be optimal. But when types can be negative, it may be. Thus the only point that changes when we allow the possibility that it is better for the principal to keep the good than to allocate it to any of the agents is that this particular mechanism may now be optimal.

So what is this mechanism? If $\varphi_i = 0$ for all $i$, this says that if every agent has $t_i - c_i < v^*$, then no one gets the good. In other words, this is a "reserve-price mechanism" with reserve $v^*$.

This mechanism is easily reinterpreted as a form of a favored-agent mechanism. Simply introduce an agent 0 whose type $t_0 = 0$ with probability 1 and for whom the checking cost $c_0$ is zero. Then the principal keeping the good is the same as allocating the good to this agent. Then the reserve-price mechanism is the same as a favored-agent mechanism where agent 0 is the favored agent and $v^*$ is the threshold.

Our characterization of optimal thresholds and the optimal favored agent carries over to the inclusion of "agent 0." More specifically, it is easy to see that $t_0^* = 0$ solves the equation $E(t_0) = E\max\{t_0, t_0^*\} - c_0$ (though negative values of $t_0^*$ also solve the equation). One can show that if "agent 0" is favored, then the optimal threshold to use is $t_0^* - c_0 = 0$ and that the optimal agent (including agent 0 in this statement) to favor is any agent $i$ with $t_i^* - c_i = \max_{j=0,\ldots,I}(t_j^* - c_j)$.

The simplest way to state how the principal's value affects the optimal mechanism is to undo the renormalization and go back to a model where types are positive and the principal's value is $R > 0$. It is easy to see that undoing the renormalization makes $t_0^* - c_0 = R$. So we can restate the results above as follows. If $R$ is sufficiently large, then agent 0 will be favored. That is, if every agent $i \geq 1$ reports $t_i - c_i < R$, then the principal retains the object. If any agent $i \geq 1$ reports $t_i - c_i > R$, then the agent with the highest such report is checked and, if found not to have lied, receives the object. If $R$ is small enough that 0 is not the favored agent, then the optimal mechanism is unaffected by the principal's value. In this case, the principal may end up allocating the good to the favored agent even when the favored agent's type is below $R$.[22]

We can treat in parallel two other natural extensions. First, we can consider the case when the process of verifying an agent's claim is also costly for that agent. This complicates the analysis since such costs create a "back door" for transfers. If agents bear costs of providing documentation, then the principal can use these costs to provide incentives for truth telling. Intuitively, the agents may now trade off the value of obtaining the object with the costs of verification. An agent who values the object more highly would, of course, be willing to incur a higher expected verification cost to increase his probability of receiving it. This both complicates the analysis and indirectly introduces a form of the transfers we wish to exclude.

Second, as this intuition suggests, this case is similar to the case where the principal can impose limited fines on the agent. As above, agents who value the object more are willing to take a bigger risk of receiving such fines.

In both cases, the simplification we obtain where we can treat the agent's payoff as equal to the probability he receives the object no longer holds. While a general analysis of either of these extensions is beyond the scope of this paper, we can easily extend our analysis at the cost of adding an assumption under which this

[22] On the other hand, from the definition of $t_i^*$, we have

$$E(t_i) = E(\max\{t_i, t_i^*\}) - c_i \geq t_i^* - c_i.$$

Hence if $i \neq 0$ is the favored agent, we must have $E(t_i) \geq t_i^* - c_i > R$. That is, while the favored agent might receive the good even though his type is below $R$, his type cannot be below $R$ in expectation.

simplification continues to hold. To be specific, we assume that the value to the agent of receiving the object is 1 and the value of not receiving it is 0, regardless of his type. If we make this assumption, the extension to fines or verification costs for the agents is straightforward. To see this, first consider the case where verification imposes costs on the agent. Let $\hat{c}_i^T \geq 0$ be the cost incurred by agent $i$ from being verified by the principal if he reported his type truthfully and let $\hat{c}_i^F \geq 0$ be his cost if he lied. We assume $\hat{c}_i^T < 1$ to ensure that individual rationality always holds. This formulation nests the case where the principal can impose limited fines on the agent. To see this, simply note that the optimal mechanism in such a case is for the principal to impose the largest possible fine when the agent is found to have lied and nothing otherwise. Thus we obtain a model appropriate for this case by setting $\hat{c}_i^T = 0$ and $\hat{c}_i^F$ equal to this maximum penalty.

In this model, the incentive compatibility condition becomes

$$\hat{p}_i(t_i') - \hat{c}_i^T \hat{q}_i(t_i') \geq \hat{p}_i(t_i) - \hat{c}_i^F \hat{q}_i(t_i) - \hat{q}_i(t_i), \quad \forall t_i, t_i', \quad \forall i.$$

Let

$$\varphi_i = \inf_{t_i'} \left[ \hat{p}_i(t_i') - \hat{c}_i^T \hat{q}_i(t_i') \right],$$

so that incentive compatibility holds if and only if

$$\varphi_i \geq \hat{p}_i(t_i) - \hat{c}_i^F \hat{q}_i(t_i) - \hat{q}_i(t_i), \quad \forall t_i, \quad \forall i.$$

Analogously to the way we characterized the optimal mechanism in Section V, we can treat $\varphi_i$ as a separate choice variable for the principal where we add the constraint that $\hat{p}_i(t_i') - \hat{c}_i^T \hat{q}_i(t_i') \geq \varphi_i$ for all $t_i'$.

Given this, $\hat{q}_i(t_i)$ must be chosen so that the incentive constraint holds with equality for all $t_i$. To see this, suppose to the contrary that we have an optimal mechanism where the constraint holds with strict inequality for some $t_i$ (more precisely, some positive measure set of $t_i$s). If we lower $\hat{q}_i(t_i)$ by $\varepsilon$, the incentive constraint will still hold. Since this increases $\hat{p}_i(t_i') - \hat{c}_i^T \hat{q}_i(t_i')$, the constraint that this quantity is greater than $\varphi_i$ will still hold. Since auditing is costly for the principal, his payoff will increase, implying the original mechanism could not have been optimal, a contradiction.

Since the incentive constraint holds with equality for all $t_i$, we have

$$\hat{q}_i(t_i) = \frac{\hat{p}_i(t_i) - \varphi_i}{1 + \hat{c}_i^F}. \tag{7}$$

Substituting, this implies that

$$\varphi_i = \inf_{t_i'} \left[ \hat{p}_i(t_i') - \frac{\hat{c}_i^T}{1 + \hat{c}_i^F} \left[ \hat{p}_i(t_i') - \varphi_i \right] \right]$$

or

$$\varphi_i = \inf_{t_i'} \left[ \left\{ 1 - \frac{\hat{c}_i^T}{1 + \hat{c}_i^F} \right\} \hat{p}_i(t_i') + \frac{\hat{c}_i^T}{1 + \hat{c}_i^F} \varphi_i \right].$$

By assumption, the coefficient multiplying $\hat{p}_i(t_i')$ is strictly positive, so this is equivalent to

$$\left\{1 - \frac{\hat{c}_i^T}{1 + \hat{c}_i^F}\right\}\varphi_i = \left\{1 - \frac{\hat{c}_i^T}{1 + \hat{c}_i^F}\right\} \inf_{t_i'} \hat{p}_i(t_i'),$$

so $\varphi_i = \inf_{t_i'} \hat{p}_i(t_i')$, exactly as in our original formulation.

The principal's objective function is

$$\begin{aligned} E_{\mathbf{t}} \sum_i [p_i(\mathbf{t})t_i - c_i q_i(\mathbf{t})] &= \sum_i E_{t_i}[\hat{p}_i(t_i)t_i - c_i \hat{q}_i(t_i)] \\ &= \sum_i E_{t_i}\left[\hat{p}_i(t_i)t_i - \frac{c_i}{1 + \hat{c}_i^F}[\hat{p}_i(t_i) - \varphi_i]\right] \\ &= \sum_i E_{t_i}[\hat{p}_i(t_i)(t_i - \tilde{c}_i) + \varphi_i \tilde{c}_i], \end{aligned}$$

where $\tilde{c}_i = c_i/(1 + \hat{c}_i^F)$. This is the same as the principal's objective function in our original formulation but with $\tilde{c}_i$ replacing $c_i$.

Thus the solution changes as follows. The allocation probabilities $p_i$ are exactly the same as what we characterized but with $\tilde{c}_i$ replacing $c_i$. The checking probabilities, however, are the earlier ones divided by $1 + \hat{c}_i^F$ (see equation (7)). Intuitively, since verification or fines impose costs on the agent in this model, the threat of checking the agent is more severe than in the previous model, so the principal doesn't need to check as often.

That is, the new optimal mechanism is still a favored-agent mechanism but where the checking which had probability 1 before now has probability $1/(1 + \hat{c}_i^F)$. The optimal choice of the favored agent and the optimal threshold is exactly as before with $\tilde{c}_i$ replacing $c_i$. Note that agents with low values of $\hat{c}_i^F$ have higher values of $\tilde{c}_i$ and hence are more likely to be favored. That is, agents who find it easy to undergo an audit after lying are more likely to be favored. Note also that $\hat{c}_i^T$ has no effect on the optimal mechanism.

Finally, we discuss the robustness of our results to weakening our assumptions on the verification process. To motivate the kinds of robustness that are of interest, we first clarify how we view the model. Think of each agent as knowing a set of facts about how he would make use of the object if the principal were to give it to him. Suppose the principal has no private information and that the facts known to the agent are sufficient to determine the value the principal would receive if he gives the object to the agent. Thus verifying the agent's type is equivalent to verifying these facts known to the agent. Finally, suppose that if the principal checks the agent, he learns all the facts known to the agent. Since the agent knows these facts, he knows exactly what the principal would learn if he were to verify. Thus we can identify the agent's knowledge with the value of the object to the principal and with what the principal would learn if he verifies and call all of these things $t_i$.

There are several assumptions in this story that are worth generalizing. First, one could assume that the facts known to agent $i$ are not sufficient to determine the value

the principal would receive if he gives the object to $i$. If this is because the principal has private information which is also relevant to determining this value, then the results would depend greatly on the nature of this information.[23] So suppose the principal does not have private information. In this case, we can simply interpret $t_i$ as the expectation of the value to the principal conditional on the information of the agent and our results go through unaltered.

Second, one could assume that not all of the agent's information is verifiable. We can think of agent $i$'s type as consisting of a pair $(u_i, v_i)$ where $u_i$ is the unverifiable part of the agent's information, $v_i$ is verifiable, and the value to the principal of giving the object to $i$ is determined by the pair. More precisely, if the principal checks agent $i$, he learns $v_i$ but learns nothing about $u_i$ (except to the extent that these components are correlated). We can rewrite this variation on our model in a way which makes it identical to the original version. To see this, note that the principal cannot verify $u_i$ and hence cannot induce the agent to reveal it unless he makes the agent indifferent over all possible $u_i$ reports. That is, $i$'s $u_i$ report cannot affect the probability he receives the object. Hence it is irrelevant to the principal since it cannot affect the objective function so the principal may as well disregard the $u_i$ reports. So we can assume the principal only solicits information about $v_i$ and treat the principal's value given $v_i$ as the expectation of his value conditional on this parameter. Then, just as in the preceding paragraph, our results are unchanged.

Finally, we could assume that the verification process itself is subject to noise. That is, even if all the agent's information is verifiable and his information is sufficient to determine the value to the principal, the verification process may not perfectly "measure" the agent's information. For example, it may be that the agent's information is contained in a large number of documents and the principal checks the agent by reading a randomly chosen subset of these documents.

A thorough analysis of such noise is well beyond the scope of this paper. However, we can show that the favored-agent mechanism is robust to relaxing this assumption by allowing small amounts of noise. More specifically, for approximately perfect verification, a mechanism which is approximately a favored-agent mechanism is approximately optimal. Also, it is easy to see that if verification is pure noise, the principal has no means to enforce incentive compatibility of a nonconstant mechanism and hence will simply allocate the good to some agent $i \in \arg\max_j E(t_j)$. Note that this is also a favored-agent mechanism with a threshold that no type of any agent can meet. Thus we see that our results continue to hold, at least approximately, for very large or very small amounts of noise. The range in between is more complex.

## VII. Conclusion

A natural question to ask is whether we see mechanisms like the favored-agent mechanism in practice and, if not, why not. As argued in the introduction, the setting we consider appears to be a natural benchmark description of many real world

[23] As a very simple example, suppose the principal has a type $s$ and the value to the principal of giving the object to agent $i$ is a function of $(s_i, t_i)$. Suppose that for some values of $s$, this function is increasing in $t_i$, while for others, it is decreasing. Then it is possible that even without verification, the agent reports $t_i$ truthfully because the agent thinks he is equally likely to get the object with a high report or a low one. Such a formulation would yield very different results from ours.

situations involving allocation of desirable resources, particularly within an organization or by a government.

Mechanisms similar to the favored-agent mechanism do seem to be used in some of these settings. At its essence, the favored-agent mechanism is simply a setting of appropriate defaults. In other words, we can describe the mechanism as setting a default outcome in case no agent makes a strong enough claim for "special consideration." If such a strong claim is made, it is checked before being granted. Otherwise, the principal goes with the default outcome.

The use of defaults along this line is very common and can be seen in any of the examples from the introduction. For instance, when selecting a unit to head a new project, it seems natural that the manager has a default unit in mind if no other unit makes a strong claim of special qualifications. While in practice the principal's prior and the checking costs will surely enter in the determination of the threshold and the default agent, how these are determined may not exactly match our calculations. For example, the default unit may be the one with the most experience, or, when checking costs are low, every unit may be checked. There are a variety of implicit or explicit assumptions of the model that may be violated in some situations, suggesting why the mechanism used in practice does not exactly match our favored-agent mechanism.

First, as noted at the end of Section IV, the favored-agent mechanism gives every agent an incentive to increase his costs of being checked in order to become the favored agent, cost increases which make the principal worse off. Hence in settings where agents can affect these costs, it may be optimal for the principal to modify the favored-agent structure to avoid such incentives.

Second, in some settings, there may be concerns that the principal is biased in favor of one of the agents. For example, some divisions of an organization may suspect that the head of the organization is not just maximizing profits, but is biased toward one division in the sense that his payoff is skewed toward them. Such beliefs could have counterproductive effects as these divisions engage in strategic behavior to overcome this bias or protect themselves from it. Similar issues arise with government allocation of public goods. If a government has to locate a hospital, for example, and uses a mechanism which favors one town, this could create political pressure by the other towns. Thus it may be important for the mechanism designer to signal a lack of bias by choosing a mechanism which treats all agents symmetrically, even (or especially) if it is costly to do so.

Third, as noted in Section VI, if each agent bears costs associated with being checked by the principal and his value of receiving the good is increasing in his type, then the favored-agent mechanism may not be optimal. In some situations, the assumption that checking does not impose costs on the agent is natural, but in others, it is more natural to assume that this involves time-consuming and hence costly paperwork for the agent.

Fourth, as discussed in Section VI, it may be of interest to consider models where either the principal has private information relevant to determining his value and/or there is an intermediate level of noise in the verification technology.

# Appendix

## A. *Proof of Theorem 1*

In the Appendix, we work with the restated version of the optimization problem for the principal derived in Section V (in particular, equations (5) and (6) and the constraints that follow).

The following definition was given in the text.

DEFINITION 1: $(\mathbf{p}, \mathbf{q})$ *is a* threshold mechanism *if there exists* $v^* \in \mathbf{R}$ *such that the following holds for all* $\mathbf{t}$ *up to sets of measure zero. First, if there exists any* $i$ *with* $t_i - c_i > v^*$, *then* $p_i(\mathbf{t}) = 1$ *for that* $i$ *such that* $t_i - c_i > \max_{j \neq i}(t_j - c_j)$. *Second, for all* $i$, *if* $t_i - c_i < v^*$, *then* $q_i(\mathbf{t}) = 0$ *and* $\hat{p}_i(t_i) = \min_{t_i' \in T_i} \hat{p}_i(t_i')$.

As discussed in the text, Theorem 4, proved in part 2 of the online Appendix, states that every solution to the relaxed problem is a threshold mechanism. We now use this result to prove Theorem 1.

Let $(\mathbf{p}, \mathbf{q})$ denote any optimal mechanism. In light of Theorem 4, we know $(\mathbf{p}, \mathbf{q})$ is a threshold mechanism. Hence we can specify $\hat{p}_i(t_i)$ for each agent as a function only of $v^*$ and $\varphi$. To see this, fix $v^*$ and the $\varphi_i$s and consider $t_i$ such that $t_i - c_i > v^*$. Since $(\mathbf{p}, \mathbf{q})$ is a threshold mechanism, $t_i$ receives the object with probability 1 if $t_i - c_i > \max_{j \neq i} t_j - c_j$ and with probability 0 if $\max_{j \neq i} t_j - c_j > t_i - c_i$. Hence $\hat{p}_i(t_i) = \prod_{j \neq i} F_j(t_i - c_i + c_j)$. For any $t_i$ such that $t_i < v^*$, the definition of a threshold mechanism requires $\hat{p}_i(t_i) = \varphi_i$. Since we can write the principal's payoff as a function of the $\hat{p}_i$s, this means we can write his payoff as a function only of $v^*$ and the $\varphi_i$s. More specifically, the principal's payoff is

$$E_{\mathbf{t}}\Big[\sum_i [p_i(\mathbf{t})(t_i - c_i) + \varphi_i c_i]\Big] = \sum_i F_i(v^* + c_i) E_{\mathbf{t}}[p_i(\mathbf{t})(t_i - c_i) \,|\, t_i < v^* + c_i]$$

$$+ \sum_i \int_{v^*+c_i}^{\bar{t}_i} \Big[\prod_{j \neq i} F_j(t_i - c_i + c_j)\Big](t_i - c_i) f_i(t_i)\, dt_i + \sum_i \varphi_i c_i.$$

Note that

$$\begin{aligned} E_{\mathbf{t}}[p_i(\mathbf{t})(t_i - c_i) \,|\, t_i < v^* + c_i] &= E_{t_i}\big\{E_{\mathbf{t}_{-i}}[p_i(\mathbf{t})(t_i - c_i)] \,|\, t_i < v^* + c_i\big\} \\ &= E_{t_i}[\hat{p}_i(t_i)(t_i - c_i) \,|\, t_i < v^* + c_i] \\ &= E_{t_i}[\varphi_i(t_i - c_i) \,|\, t_i < v^* + c_i] \\ &= \varphi_i(E_{t_i}[t_i \,|\, t_i < v^* + c_i] - c_i). \end{aligned}$$

Hence we can rewrite the objective function as

$$\sum_i F_i(v^* + c_i)\varphi_i[E_{t_i}(t_i \mid t_i < v^* + c_i) - c_i]$$

$$+ \sum_i \int_{v^*+c_i}^{\bar{t}_i} \left[\prod_{j\neq i} F_j(t_i - c_i + c_j)\right](t_i - c_i)f_i(t_i)\, dt_i + \sum_i \varphi_i c_i$$

$$= \sum_i \varphi_i \{F_i(v^* + c_i)E_{t_i}(t_i \mid t_i < v^* + c_i) + [1 - F_i(v^* + c_i)]c_i\}$$

$$+ \sum_i \int_{v^*+c_i}^{\bar{t}_i} \left[\prod_{j\neq i} F_j(t_i - c_i + c_j)\right](t_i - c_i)f_i(t_i)\, dt_i.$$

Without loss of generality, we can restrict attention to $v^* \geq \max_i \underline{t}_i - c_i$ and $v^* \leq \max_i \bar{t}_i - c_i$. To see this, note that if $v^* < \max_i \underline{t}_i - c_i$, then with probability 1, there will be some $i$ with $t_i - c_i > v^*$. Hence the principal's payoff is the same if $v^* = \max_i \underline{t}_i - c_i$ as it would be at any lower $v^*$. Similarly, if $v^* > \max_i \bar{t}_i - c_i$, then with probability 1, every $i$ will have $t_i - c_i < v^*$. Hence, again, the principal's payoff at $v^* = \max_i \bar{t}_i - c_i$ is the same as it would be at any higher $v^*$ (holding $\varphi$ fixed).

For a fixed $v^*$, the principal's objective function is linear in the vector $\varphi$. Given $v^*$, the set of feasible $\varphi$ vectors is convex. To be precise, recall that a given specification of $p_i$ and $\varphi_i$, $i \in \mathcal{I}$, is feasible if and only if each $p_i : \mathcal{T} \to [0, 1]$, each $\varphi_i \in [0, 1]$, $\sum_i p_i(\mathbf{t}) \leq 1$ for all $\mathbf{t}$, and $E_{\mathbf{t}_{-i}} p_i(\mathbf{t}) \geq \varphi_i$ for all $t_i \in \mathcal{T}_i$ and all $i$. From Theorem 4, we know the exact value of $p_i(\mathbf{t})$ for all $i$ for (almost) any $\mathbf{t}$ such that $t_i - c_i > v^*$ for some $i$. Finally, Theorem 4 also tells us that $\hat{p}_i(t_i) = \varphi_i$ for (almost) all $t_i < v^* + c_i$ for all $i$. (From Lemma 6 in the online Appendix, we know this holds on a set of strictly positive measure.) We say that a profile $\varphi_i$, $i \in \mathcal{I}$, is *feasible* given $v^*$ if there exists $p_i$ functions satisfying the properties above given $v^*$ and these $\varphi_i$s.

LEMMA 1: *The set of $\varphi_i$, $i \in \mathcal{I}$, that is feasible given $v^*$ is the set satisfying $\varphi_i \in [0, 1]$ for all $i$,*

$$\varphi_i = \prod_{j\neq i} F_j(v^* + c_j), \quad \forall i \text{ such that } F_i(v^* + c_i) = 0,$$

and

$$\sum_i \varphi_i F_i(v^* + c_i) \leq \prod_i F_i(v^* + c_i).$$

PROOF:

Since $\varphi_i = \hat{p}_i(t_i)$ on a set of strictly positive measure, it is obviously necessary to have $\varphi_i \in [0, 1]$. To see the necessity of the second condition, consider some $i$ with $F_i(v^* + c_i) = 0$ or, equivalently, $v^* \leq \underline{t}_i - c_i$. Since we must have $v^* \geq \max_j \underline{t}_j - c_j$, this implies $v^* = \underline{t}_i - c_i$. For any $t_i \in (\underline{t}_i, \bar{t}_i)$, then, we have $t_i - c_i > v^*$, so type $t_i$ receives the good if and only if his is the highest value. That is, $\hat{p}_i(t_i) = \prod_{j\neq i} F_j(t_i - c_i + c_j)$. Thus,

$$\varphi_i = \inf_{t_i} \hat{p}_i(t_i) = \lim_{t_i \downarrow \underline{t}_i} \hat{p}_i(t_i) = \prod_{j\neq i} F_j(v^* + c_j),$$

implying that the second condition is necessary.

For necessity of the third condition, note that

$$\begin{aligned}\sum_i \varphi_i F_i(v^* + c_i) &= \sum_i \int_{\underline{t}_i}^{v^*+c_i} \hat{p}_i(t_i) f_i(t_i)\, dt_i \\ &= \sum_i \int_{\underline{t}_i}^{v^*+c_i} E_{\mathbf{t}_{-i}} p_i(t_i, t_{-i}) f_i(t_i)\, dt_i \\ &= \sum_i \int_{\underline{t}_i}^{v^*+c_i} \int_{\mathbf{t}_{-i}} p_i(\mathbf{t}) f_i(t_i) f_{-i}(\mathbf{t}_{-i})\, d\mathbf{t}_{-i}\, dt_i.\end{aligned}$$

But for any $\mathbf{t}_{-i}$ such that $t_j - c_j > v^*$ for some $j \neq i$, we must have $p_i(t_i, \mathbf{t}_{-i}) = 0$. Hence

$$\begin{aligned}\sum_i \int_{\underline{t}_i}^{v^*+c_i} \int_{\mathbf{t}_{-i}} p_i(\mathbf{t}) f_i(t_i) f_{-i}(\mathbf{t}_{-i})\, d\mathbf{t}_{-i}\, dt_i &= \sum_i \int_{\mathbf{t}|t_j<v^*+c_j,\ \forall j} p_i(\mathbf{t}) f(\mathbf{t})\, d\mathbf{t} \\ &= \int_{\mathbf{t}|t_j<v^*+c_j,\ \forall j} \left[\sum_i p_i(\mathbf{t})\right] f(\mathbf{t})\, d\mathbf{t} \\ &\leq \int_{\mathbf{t}|t_j<v^*+c_j,\ \forall j} f(\mathbf{t})\, d\mathbf{t} \\ &= \prod_j F_j(v^* + c_j).\end{aligned}$$

Hence the third condition is necessary.

Note for use below that the third condition and $\varphi_i \geq 0$ implies

$$\varphi_i F_i(v^* + c_i) \leq \prod_j F_j(v^* + c_j).$$

If $F_i(v^* + c_i) \neq 0$, this implies $\varphi_i \leq \prod_{j\neq i} F_j(v^* + c_j)$. As the second condition shows, if $F_i(v^* + c_i) = 0$, we still require this condition, though with equality.

To see that these conditions are sufficient, we consider three cases. Let

$$\mathcal{I}^0 = \{i \in \mathcal{I} | F_i(v^* + c_i) = 0\} = \{i \in \mathcal{I} | v^* = \underline{t}_i - c_i\}.$$

The first case is where $\#\mathcal{I}^0 \geq 2$ (where # denotes cardinality). In this case, we have $\prod_{j\neq i} F_j(v^* + c_j) = 0$ for all $i$. Hence the third condition implies $\varphi_i = 0$ for all $i$ such that $F_i(v^* + c_i) \neq 0$. If $F_i(v^* + c_i) = 0$, then the second condition applies to $i$, so, again, $\varphi_i = 0$. Hence the only $\varphi$ satisfying the necessary conditions for such a $v^*$ is the zero vector. It is easy to see that this is feasible since it is achieved for any $\mathbf{p}$ satisfying $p_i(\mathbf{t}) = 1$ for that $i$ with $t_i - c_i > \max_{j\neq i} t_j - c_j$ for every $\mathbf{t}$.

The second case is where $\#\mathcal{I}^0 = 1$. Let $k$ denote the unique element of $\mathcal{I}^0$. Then the third condition implies that $\varphi_i = 0$ for all $i \neq k$. The second condition implies $\varphi_k = \prod_{j \neq k} F_j(v^* + c_j)$. Hence, again, there is a unique $\varphi$ satisfying the necessary conditions for such a $v^*$. Again, it is easy to see that this is feasible since it is achieved for any $\mathbf{p}$ satisfying $p_i(t) = 1$ for that $i$ with $t_i - c_i > \max_{j \neq i} t_j - c_j$ for every $t$. To see this, note that $k \in \mathcal{I}^0$ implies $t_k - c_k > v^*$ with probability 1, so the threshold mechanism must always allocate the good to the agent with the highest value. If every other agent has value below $v^*$, $k$ must get the good, regardless of his value; $\varphi_k$ is the probability this occurs.

Finally, suppose $\mathcal{I}^0 = \emptyset$. In this case, $\prod_{j \neq i} F_j(v^* + c_j) > 0$ for all $i$. Fix any $\varphi$ satisfying the conditions of the lemma. To see that this $\varphi$ is feasible, set $\mathbf{p}$ as follows. For any $\mathbf{t}$ such that $\max_i t_i - c_i > v^*$, let $p_i(\mathbf{t}) = 1$ for that $i$ with $t_i - c_i > \max_{j \neq i} t_j - c_j$. For any $\mathbf{t}$ with $\max_i t_i - c_i < v^*$, let $p_i(\mathbf{t}) = \varphi_i / \prod_{j \neq i} F_j(v^* + c_j)$ for every $i$. Since $\varphi_i \in [0, 1]$, $p_i(\mathbf{t})$ is non-negative for all $i$. Also,

$$\sum_i p_i(\mathbf{t}) = \sum_i \frac{\varphi_i F_i(v^* + c_i)}{\prod_j F_j(v^* + c_j)} = \frac{\sum_i \varphi_i F_i(v^* + c_i)}{\prod_j F_j(v^* + c_j)}.$$

By our third condition, this is less than 1.

Also, for any $i$ and any $t_i < v^* + c_i$, we have

$$\begin{aligned}
\hat{p}_i(t_i) &= \left[\prod_{j \neq i} F_j(v^* + c_j)\right] E\big(p_i(\mathbf{t}) \,|\, t_j \leq v^* + c_j, \quad \forall j \neq i\big) \\
&\quad + \left[1 - \prod_{j \neq i} F_j(v^* + c_j)\right] E\big(p_i(\mathbf{t}) \,|\, t_j > v^* + c_j, \quad \textit{for some } j \neq i\big) \\
&= \left[\prod_{j \neq i} F_j(v^* + c_j)\right] \left[\frac{\varphi_i}{\prod_{j \neq i} F_j(v^* + c_j)}\right] + \left[1 - \prod_{j \neq i} F_j(v^* + c_j)\right](0) \\
&= \varphi_i.
\end{aligned}$$

If $F_i(v^* + c_i) = 1$, this implies $\inf_{t_i} \hat{p}_i(t_i) = \varphi_i$. Otherwise, for $t_i > v^* + c_i$, we have

$$\hat{p}_i(t_i) = \prod_{j \neq i} F_j(t_i - c_i + c_j) \geq \prod_{j \neq i} F_j(v^* + c_j) \geq \varphi_i,$$

where the last inequality follows from the necessary conditions. Hence, again, $\inf_{t_i} \hat{p}_i(t_i) = \varphi_i$, so $\varphi$ is feasible given $v^*$.

Given Lemma 1, we see that the set of feasible $\varphi$ given $v^*$ is the set satisfying a finite system of linear inequalities and hence is convex. Since the objective function is linear in $\varphi$ and the feasible set is convex, we see that given $v^*$, there is an optimal $\varphi$ which is an extreme point. Furthermore, the set of optimal $\varphi$ is the convex hull of the set of optimal extreme points.

The following lemma characterizes the extreme points. Recall that

$$\mathcal{I}^0 = \{i \in \mathcal{I} \,|\, F_i(v^* + c_i) = 0\} = \{i \in \mathcal{I} \,|\, v^* = \underline{t}_i - c_i\}.$$

LEMMA 2: *If $\mathcal{I}^0$ is not a singleton, then $\varphi^*$ is an extreme point of the set of feasible $\varphi$ given $v^*$ if and only if either $\varphi^* = 0$ or there exists $i$ such that $\varphi_j^* = 0$ for all $j \neq i$ and $\varphi_i^* = \prod_{j \neq i} F_j(v^* + c_j)$. If $\mathcal{I}^0 = \{i\}$, then $\varphi^*$ is an extreme point of the set of feasible $\varphi$ given $v^*$ if and only if $\varphi_j^* = 0$ for all $j \neq i$ and $\varphi_i^* = \prod_{j \neq i} F_j(v^* + c_j)$.*

PROOF:

If $\#\mathcal{I}^0 \geq 2$, then, as shown in the proof of Lemma 1, the only feasible $\varphi$ is the 0 vector. Note, though, that $\prod_{j \neq i} F_j(v^* + c_j) = 0$ for all $i$, so the description in the statement of the lemma applies. If $\mathcal{I}^0 = \{i\}$, then the proof of Lemma 1 shows that the only feasible $\varphi$ is the one stated as the extreme point in the lemma, so again the lemma follows.

So for the rest of this proof, assume $\mathcal{I}^0 = \emptyset$. That is, $F_i(v^* + c_i) > 0$ for all $i$. It is easy to see that the $\varphi^*$s stated in the lemma must all be extreme points. To see this, suppose that there exist a feasible $\varphi^1$ and $\varphi^2$ such that $\varphi^1 \neq \varphi^2$ and there exists $\lambda \in (0, 1)$ such that $\lambda\varphi^1 + (1 - \lambda)\varphi^2 = \varphi^*$ for one of the $\varphi^*$s stated in the lemma. Obviously, we cannot have $\varphi^*$ equal to the zero vector since $\varphi_i^k \geq 0$ for all $i$ and $k$ would then imply $\varphi^1 = \varphi^2 = 0$, a contradiction. So suppose there is some $i$ with $\varphi_j^* = 0$ for all $j \neq i$ and $\varphi_i^* = \prod_{j \neq i} F_j(v^* + c_j)$. Again, we must have $\varphi_j^1 = \varphi_j^2 = 0$ for all $j \neq i$. Since we cannot have $\varphi^1 = \varphi^2$, without loss of generality, we must have $\varphi_i^1 < \prod_{j \neq i} F_j(v^* + c_j) < \varphi_i^2$. But then $\varphi^2$ violates the third condition for feasibility of $\varphi$ given $v^*$, a contradiction.

Hence we only need to show that there are no other extreme points. To show this, we show that any $\varphi$ which is feasible given $v^*$ can be written as a convex combination of these points. So fix any such $\varphi$. Define $r_i = \varphi_i / \prod_{j \neq i} F_j(v^* + c_j)$. By the necessary conditions stated in Lemma 1, $r_i \geq 0$. Also,

$$\sum_{i=1}^{I} r_i = \sum_{i=1}^{I} \frac{\varphi_i F_i(v^* + c_i)}{\prod_j F_j(v^* + c_j)} = \frac{\sum_{i=1}^{I} \varphi_i F_i(v^* + c_i)}{\prod_j F_j(v^* + c_j)}.$$

By the third necessary condition, then, $\sum_{i=1}^{I} r_i \leq 1$. Finally, let $r_0 = 1 - \sum_{i=1}^{I} r_i$ Hence $\sum_{i=0}^{I} r_i = 1$. Let $\varphi^*(i)$ denote the $\varphi^*$ of the lemma which has $\varphi_j^* = 0$ for all $j \neq i$ and $\varphi_i^* = \prod_{j \neq i} F_j(v^* + c_j)$. It is easy to see that

$$\varphi = \sum_{i=1}^{I} r_i \varphi^*(i) + r_0(\mathbf{0}),$$

where $\mathbf{0}$ denotes the 0 vector. Hence $\varphi$ is not an extreme point unless it equals one of the $\varphi^*$s.

Summarizing, any optimal mechanism has its reduced form completely specified by a choice of $v^*$ and a vector $\varphi$. Given any $v^*$, the set of optimal $\varphi$s is the convex hull of the set of optimal extreme $\varphi$s, characterized in Lemma 2. We now show that for any $v^*$ and any optimal extreme $\varphi$, there is a favored-agent mechanism with the same reduced form as that determined by $v^*$ and $\varphi$.

LEMMA 3: *Given any $v^*$ and any optimal extreme $\varphi$, let $(\hat{\mathbf{p}}^*, \hat{\mathbf{q}}^*)$ be the reduced form specified by $v^*$ and $\varphi$. Then there is a favored agent mechanism $(\mathbf{p}, \mathbf{q})$ with $\hat{\mathbf{p}} = \hat{\mathbf{p}}^*$ and $\hat{\mathbf{q}} = \hat{\mathbf{q}}^*$.*

PROOF:

First, suppose $\#\mathcal{I}^0 \geq 2$. In Lemma 1, we showed that the only feasible $\varphi$ in this case is the zero vector. Because $\#\mathcal{I}^0 \geq 2$, we have at least two agents $i$ with $t_i - c_i > v^*$ with probability 1. Hence $p_i^*(\mathbf{t}) = 1$ for that $i$ such that $t_i - c_i > \max_{j \neq i} t_j - c_j$. Thus for all $i$ and all $t_i$, $\hat{p}_i^*(t_i) = \prod_{j \neq i} F_j(t_i - c_i + c_j)$. Since $\varphi_i = 0$ for all $i$, we have $\hat{q}_i^*(t_i) = \hat{p}_i^*(t_i)$. We generate the same reduced form from the favored-agent mechanism with threshold $v^*$ for any selection of the favored agent. Since there are always at least two agents with values above the threshold, the selection of the favored agent is irrelevant—any agent receives the good if and only if he has the highest value and is checked in this case.

Next, suppose $\mathcal{I}^0 = \{k\}$. In the proof of Lemma 1, we showed that the only feasible $\varphi$ in this case is $\varphi^*(k)$ defined by

$$\varphi_i^*(k) = \begin{cases} 0, & \text{if } i \neq k \\ \prod_{i \neq k} F_i(v^* + c_i), & \text{if } i = k. \end{cases}$$

The reduced form generated by this extreme point is as follows. First, consider any $j \neq k$. Since $\varphi_j = 0$, we know that $\hat{q}_j^*(t_j) = \hat{p}_j^*(t_j)$. By Theorem 4, if $t_j - c_j < v^*$, then $\hat{p}_j^*(t_j) = \varphi_j = 0$. For $t_j - c_j > v^*$, $p_j^*(t_j) = \prod_{i \neq j} F_i(t_j - c_j + c_i)$. Also, for every $t_k$, $\hat{p}_k^*(t_k) = \prod_{j \neq k} F_j(t_k - c_k + c_j)$ and

$$\begin{aligned} \hat{q}_k^*(t_k) &= \hat{p}_k^*(t_k) - \prod_{j \neq k} F_j(v^* + c_j) \\ &= \Pr\left[t_k - c_k > t_j - c_j, \quad \forall j \neq k\right] - \Pr[t_j - c_j < v^*, \quad \forall j \neq k] \\ &= \Pr\left[v^* < \max_{j \neq k} t_j - c_j < t_k - c_k\right]. \end{aligned}$$

It is easy to see that a favored-agent mechanism with $k$ as the favored agent and threshold $v^*$ generates the same reduced form.

Finally, suppose $\mathcal{I}^0 = \emptyset$. We showed in the proof of Lemma 1 that the set of extreme points consists of the zero vector $\mathbf{0}$ and the collection of vectors $\varphi^*(k)$, $k = 1, \ldots, I$. The same argument as for the previous case shows that any of the extreme points $\varphi^*(k)$ generates the same reduced form as the favored-agent mechanism with $k$ as the favored agent and $v^*$ as the threshold.

We now complete the proof by showing that $\mathbf{0}$ cannot be the optimal extreme point. To see this, simply note that the term multiplying $\varphi_i$ in the principal's objective function is

$$F_i(v^* + c_i) E_{t_i}(t_i \mid t_i \leq v^* + c_i) + [1 - F_i(v^* + c_i)] c_i.$$

It is easy to see that this term must be strictly positive since $\underline{t}_i \geq 0$ and $c_i > 0$. Hence whenever there is a feasible $\varphi$ other than $\mathbf{0}$, it *must* yield the principal a higher payoff than setting $\varphi$ to the zero vector.

Hence the set of optimal mechanisms given a particular $v^*$ is equivalent to the convex hull of the set of optimal favored-agent mechanisms with $v^*$ as the threshold. Therefore, the set of optimal mechanisms is equivalent to the convex hull of the set of optimal favored-agent mechanisms where we optimize over $v^*$ as well as the identity of the favored-agent.

## REFERENCES

**Abdulkadiroğlu, Atila, and Tayfun Sönmez.** 2013. "Matching Markets: Theory and Practice." In *Advances in Economics and Econometrics: Tenth World Congress*, Vol. 1, edited by Daron Acemoglu, Manuel Arellano, and Eddie Dekel, 3–47. New York: Cambridge University Press.

**Ambrus, Attila, and Georgy Egorov.** 2012. "Delegation and Nonmonetary Incentives." Unpublished.

**Banerjee, Abhijit, Rema Hanna, and Sendhil Mullainathan.** 2013. "Corruption." In *Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 1109–47. Princeton: Princeton University Press.

**Ben-Porath, Elchanan, and Barton L. Lipman.** 2012. "Implementation with Partial Provability." *Journal of Economic Theory* 147 (5): 1689–1724.

**Border, Kim C., and Joel Sobel.** 1987. "Samurai Accountant: A Theory of Auditing and Plunder." *Review of Economic Studies* 54 (4): 525–40.

**Bull, Jesse, and Joel Watson.** 2007. "Hard Evidence and Mechanism Design." *Games and Economic Behavior* 58 (1): 75–93.

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

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

**Deneckere, Raymond, and Sergei Severinov.** 2008. "Mechanism Design with Partial State Verifiability." *Games and Economic Behavior* 64 (2): 487–513.

**Doval, Laura.** 2013. "Whether or Not to Open Pandora's Box." Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science Discussion Paper 1574.

**Gale, Douglas, and Martin Hellwig.** 1985. "Incentive-Compatible Debt Contracts: The One-Period Problem." *Review of Economic Studies* 52 (4): 647–63.

**Glazer, Jacob, and Ariel Rubinstein.** 2004. "On Optimal Rules of Persuasion." *Econometrica* 72 (6): 1715–36.

**Glazer, Jacob, and Ariel Rubinstein.** 2006. "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach." *Theoretical Economics* 1 (4): 395–410.

**Green, Jerry R., and Jean-Jacques Laffont.** 1986. "Partially Verifiable Information and Mechanism Design." *Review of Economic Studies* 53 (3): 447–56.

**Hartline, Jason D., and Tim Roughgarden.** 2008. "Optimal Mechanism Design and Money Burning." In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, 75–84. New York: ACM.

**Kartik, Navin, and Olivier Tercieux.** 2012. "Implementation with Evidence." *Theoretical Economics* 7 (2): 323–55.

**McAfee, R. Preston, and John McMillan.** 1992. "Bidding Rings." *American Economic Review* 82 (3): 579–99.

**Mookherjee, Dilip, and Ivan Png.** 1989. "Optimal Auditing, Insurance, and Redistribution." *Quarterly Journal of Economics* 104 (2): 399–415.

**Roth, Alvin E., and Marilda A. Oliveira Sotomayor.** 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. New York: Cambridge University Press.

**Sher, Itai, and Rakesh Vohra.** 2011. "Price Discrimination through Communication." Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science Discussion Paper 1536.

**Townsend, Robert M.** 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (2): 265–93.

**Townsend, Robert M.** 1988. "Information Constrained Insurance: The Revelation Principle Extended." *Journal of Monetary Economics* 21 (2-3): 411–50.

**Weitzman, Martin L.** 1979. "Optimal Search for the Best Alternative." *Econometrica* 47 (3): 641–54.

**Yoon, Kiho.** 2011. "Optimal Mechanism Design when Both Allocative Inefficiency and Expenditure Inefficiency Matter." *Journal of Mathematical Economics* 47 (6): 670–76.