

Optimal Allocation with Costly Verification¹

Elchanan Ben-Porath² Eddie Dekel³ Barton L. Lipman⁴

First Preliminary Draft
October 2011

Current Draft
March 2014

¹We thank Ricky Vohra, Benjy Weiss, numerous seminar audiences, and Andy Skrzypacz 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.

²Department of Economics and Center for Rationality, Hebrew University. Email: benporat@math.huji.ac.il.

³Economics Department, Northwestern University, and School of Economics, Tel Aviv University. Email: dekel@northwestern.edu.

⁴Department of Economics, Boston University. Email: blipman@bu.edu.

Abstract

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 *favoured-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 iff her report is confirmed. All optimal mechanisms are essentially randomizations over optimal favored-agent mechanisms.

1 Introduction

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 iff 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 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

¹Two mechanisms have the same “outcome” if the interim probabilities of checking and allocating the good are the same; see Section 2 for details.

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 which 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 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 be 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 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, these papers focus 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, 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 3 shows that all optimal mechanisms are essentially randomizations

²In a similar vein, Banerjee, Hanna, and Mullainathan (2011) 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.

³There 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 Abdulkadiroglu and Sonmez (2013) for a more recent one.

over optimal favored-agent mechanisms. In Section 4, we characterize the set of best favored-agent mechanisms. In Section 5, we give comparative statics and discuss various properties of the optimal mechanism. In Section 6, we sketch the proof of our uniqueness result, while Section 7 discusses some simple extensions. Section 8 concludes. Proofs not contained in the text are either in the Appendix or the Online Appendix.

2 Model

The set of agents is $\mathcal{I} = \{1, \dots, I\}$. 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 non-negative and the value to the principal of assigning the object to no one is zero. As we explain in Section 7, 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 interval⁴ $T_i \equiv [\underline{t}_i, \bar{t}_i]$ where $0 \leq \underline{t}_i < \bar{t}_i < \infty$. We use F_i to denote the corresponding distribution function. Let $T = \prod_i T_i$.

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 7.

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

⁴It 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 7 we allow the principal to prefer keeping the object, it is similarly straightforward to drop the assumption that the support is bounded below.

⁵Suppose we let the payoff of agent i from receiving the good be $\bar{u}_i(t_i)$ and let his utility from not receiving it be $\underline{u}_i(t_i)$ where $\bar{u}_i(t_i) > \underline{u}_i(t_i)$ for all i and all t_i . Then it is simply a renormalization to let $\bar{u}_i(t_i) = 1$ 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

$$p\bar{u}_i(t_i) + (1-p)\underline{u}_i(t_i) \geq p'\bar{u}_i(t_i) + (1-p')\underline{u}_i(t_i)$$

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 iff $p \geq p'$, exactly the incentive

reservation utility is less than or equal to his utility from not receiving the 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 A 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 $p_i(t)$ denote the total probability i is assigned the good and $q_i(t)$ the probability i is assigned the good and checked. Then these functions must satisfy $p_i : T \rightarrow [0, 1]$, $q_i : T \rightarrow [0, 1]$, $\sum_i p_i(t) \leq 1$ for all $t \in T$, and $q_i(t) \leq p_i(t)$ for all $i \in \mathcal{I}$ and all $t \in T$. Henceforth, the word “mechanism” will be used only to denote such a tuple of functions, generally denoted (p, q) for simplicity.

The principal’s objective function is

$$\mathbb{E}_t \left[\sum_i (p_i(t)t_i - q_i(t)c_i) \right].$$

The incentive compatibility constraint for i is

$$\mathbb{E}_{t_{-i}} p_i(t) \geq \mathbb{E}_{t_{-i}} [p_i(t'_i, t_{-i}) - q_i(t'_i, t_{-i})], \quad \forall t_i, t'_i \in T_i, \quad \forall i \in \mathcal{I}.$$

Given a mechanism (p, q) , let

$$\hat{p}_i(t_i) = \mathbb{E}_{t_{-i}} p_i(t)$$

constraint we have if $\bar{u}_i(t_i) = 1$ and $\underline{u}_i(t_i) = 0$.

⁶The 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 discussion and an extension to a class of verification models which does not include ours.

and

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

The $2I$ tuple of functions $(\hat{p}, \hat{q})_{i \in \mathcal{I}}$ is the *reduced form* of the mechanism (p, q) . We say that (p^1, q^1) and (p^2, q^2) are *equivalent* if $\hat{p}^1 = \hat{p}^2$ and $\hat{q}^1 = \hat{q}^2$ 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 (p^1, q^1) is an optimal incentive compatible mechanism, (p^2, q^2) must be as well. Therefore, we can only identify the optimal mechanism up to equivalence.

3 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 (p, 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 $p_{i^*}(t) = 1$ and $q_i(t) = 0$ 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 $p_i(t) = q_i(t) = 1$ and $p_k(t) = q_k(t) = 0$ 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 2, if a mechanism (p, 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

⁷If 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.

⁸Randomizing refers to the pointwise convex combination of mechanisms (p, q) .

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.*

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 6 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 $T_i = [0, 1]$ for all i . Clearly, the principal would ideally give the object to the agent with the highest t_i . Of course, this isn't 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's 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

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

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 $t_i - c_i$ 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 iff 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.

4 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

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

¹⁰In 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 iff 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.

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 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.

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 $E(t_1)$.

¹¹To 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 $t_i^* \rightarrow \infty$. 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 can have $t_i^* > \bar{t}_i$. If so, then $t_i^* = E(t_i) + c_i > \bar{t}_i$.

Recall that $t_1^* > t_1$. 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_1^* . 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 (p, q) . Hence there is a unique best mechanism in \mathcal{F}_i . ■

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.

¹²Since the first draft of this paper, Doval (2013) considered a search model where this definition does appear. In Doval, 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.

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 towards 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 C 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

$$\begin{aligned} & F_i(\hat{t}_i)F_j(t_j^*)\mathbb{E}[t_i \mid t_i \leq \hat{t}_i] + [1 - F_i(\hat{t}_i)]F_j(t_j^*)\mathbb{E}[t_i \mid t_i > \hat{t}_i] \\ & \quad + F_i(\hat{t}_i)[1 - F_j(t_j^*)] \{ \mathbb{E}[t_j \mid t_j > t_j^*] - c_j \} \\ & \geq F_i(\hat{t}_i)F_j(t_j^*)\mathbb{E}[t_j \mid t_j \leq t_j^*] + [1 - F_i(\hat{t}_i)]F_j(t_j^*) \{ \mathbb{E}[t_i \mid t_i > \hat{t}_i] - c_i \} \\ & \quad + F_i(\hat{t}_i)[1 - F_j(t_j^*)]\mathbb{E}[t_j \mid t_j > t_j^*]. \end{aligned}$$

Rewriting,

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

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

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

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

Recall that t_i^* is defined by

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

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

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

or

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

or

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

Hence

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

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

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

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 $i \in \operatorname{argmax}_j (t_j^* - c_j)$ 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 $\operatorname{argmax}_j (t_j^* - c_j)$ 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

¹³We thank Yi-Chun Chen and Siyang Xiong for showing us a proof of this result.

5 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 $[.99, 1.99]$. It is easy to see that as $\varepsilon \downarrow 0$, we have $t_2^* - c_2 \downarrow t_2 = .99 < 1 = t_1^* - c_1$. Hence for ε 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 $t_2 \sim U[.99, 1.99]$, so for ε small, 1 receives the good slightly more than 1% 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

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

¹⁴We thank an anonymous referee for raising this issue with a similar example.

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 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 4. 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 $F_i^{-1}(\hat{F}_i(\hat{t}_i))$ 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 $F_i^{-1}(\hat{F}_i(\hat{t}_i))$. 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 $F_i^{-1}(\hat{F}_i(\hat{t}_i))$ 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 $\hat{t}_i \geq F_i^{-1}(\hat{F}_i(\hat{t}_i))$, so the principal is better off.¹⁶

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 iff $t_i - c_i$ 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 or negative for the favored agent. For

¹⁵For example, if 1 is the favored agent and 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 $t_1 - c_1$.

¹⁶We thank Andy Skrzypacz for suggesting this approach.

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 D. 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.

6 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 $\hat{p}_i(t_i) = E_{t_{-i}} p_i(t_i, t_{-i})$ and $\hat{q}_i(t_i) = E_{t_{-i}} q_i(t_i, t_{-i})$. We can write the incentive compatibility constraint as

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

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 iff 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

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

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

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

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

$$\begin{aligned} \mathbb{E}_t \left[\sum_i p_i(t)t_i - \sum_i c_i q_i(t) \right] &= \sum_i \mathbb{E}_{t_i} [\hat{p}_i(t_i)t_i - c_i \hat{q}_i(t_i)] \\ &= \sum_i \mathbb{E}_{t_i} [\hat{p}_i(t_i)(t_i - c_i) + \varphi_i c_i] \end{aligned} \quad (5)$$

$$= \mathbb{E}_t \left[\sum_i [p_i(t)(t_i - c_i) + \varphi_i c_i] \right]. \quad (6)$$

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 “nonreduced” mechanism and so rely on the second expression, (6).

From this point forward, we treat the principal’s problem as choosing the φ_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 tradeoff in the model. Recall that φ_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 φ_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 φ_i ’s as given and optimizing over the p_i ’s. This partial characterization enables us to reduce the problem to choosing the φ_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 $p_i : T \rightarrow [0, 1]$ for $i \in \mathcal{I}$ subject to $\sum_i p_i(t) \leq 1$ for all t and $\mathbb{E}_{t_i} p_i(t) \geq \varphi_i$ for all t_i and all i .¹⁷ Note that since the φ_i ’s are fixed, we can take the objective function to be

$$\sum_i \mathbb{E}_{t_i} \hat{p}_i(t_i)(t_i - c_i) = \mathbb{E}_t \left[\sum_i p_i(t)(t_i - c_i) \right].$$

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

¹⁷The constraint set is nonempty iff $\sum_i \varphi_i \leq 1$. 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.

properties. First, if $t_i - c_i < v^*$, then $\hat{p}_i(t_i) = \varphi_i$. Second, for any 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 φ_i 's, a threshold mechanism is equivalent to a favored-agent mechanism.

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

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 $E_{t_{-i}} p_i(t) \geq \varphi_i$. 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 $p_i(t) = 0$ for all i if $t_i < c_i$ for all i and otherwise $p_i(t) = 1$ 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{p} is completely pinned down as a function of v^* and the φ_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 $t_i < v^*$, 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 φ_i 's. Note that Theorem 4 does *not* say that any v^* is consistent with given φ_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 φ_i , then the threshold v^* is not feasible given the φ_i 's.

It turns out to be more convenient to fix the threshold v^* and ask which φ_i 's are consistent with it. We show in Appendix A (Lemma 1) that for any v^* , the set of consistent φ_i 's is convex. It is also easy to show that the objective function of the principal, holding v^* fixed, is linear in the φ_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 φ_i ’s. Furthermore, every optimal choice of the φ_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 φ_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 φ_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 φ_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.¹⁸

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 com-

¹⁸As we discuss in Section 7, if the principal has a strictly positive value for keeping the good, then this extreme point can be optimal.

plex. 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 two types have the same probability.¹⁹ Then consider the mechanism 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{t} where $\hat{t}_i \notin \{t_i, t'_i\}$, we let $p_j^*(\hat{t}) = p_j(\hat{t})$ for every j . For any t_{-i} , we set $p_j^*(t_i, t_{-i}) = p_j(t'_i, t_{-i})$ and $p_j^*(t'_i, t_{-i}) = p_j(t_i, t_{-i})$ 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 $\hat{t}_i \in [t_i, \bar{t}_i]$, 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 $t = (t_1, \dots, t_I)$ such that $t_i - c_i > t_j - c_j > \hat{t}_j - c_j$, then any solution to the relaxed problem has $p_j(t) = 0$. To see this, suppose to the contrary that we have a solution with $p_j(t) > 0$. Then we can change

¹⁹If, 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.

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 $t = (t_1, \dots, t_I)$ such that $t_i > \hat{t}_i$ and $t_j < \hat{t}_j$, we must have $p_j(t) = 0$. 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 $t = (t_1, \dots, t_I)$ 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 $p_i(t) = 0$. However, from our third step, $t_i > \hat{t}_i$ and $t_k < \hat{t}_k$ implies $p_k(t) = 0$ 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 $p_i(t) = 1$, 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 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 $p_j(t) = 0$. Hence no agent below the threshold can receive the object. As argued in the last step, we must also have $\sum_i p_i(t) = 1$. 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 $p_j(t) = 0$. 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, 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.²⁰ We show that for a type profile $t = (t_1, t_2)$ such that $t_1 > \hat{t}_1$ and $t_2 < \hat{t}_2$, we must have $p_2(t) = 0$. 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 2 below where $\tilde{t}_1 > \hat{t}_1$ while $\tilde{t}_2 < \hat{t}_2$. So suppose that at α , $p_2(\tilde{t}) > 0$ so player 1 receives the good with probability strictly less than 1.

Then at any point directly below α but above \hat{t}_1 , such as the one labeled $\beta = (t'_1, \tilde{t}_2)$,

²⁰The case of three or more agents works similarly but the construction of the improving mechanism is more complex.

player 1 must receive the good with probability zero. This follows because if 1 did receive the good with strictly positive probability at β , we could change the mechanism by lowering this probability slightly, giving the good to 2 at β with higher probability, and increasing the probability with which 1 receives the good at α . By choosing these probabilities appropriately, we do not affect $\hat{p}_2(\tilde{t}_2)$ so this remains at φ_2 . Also, by making the reduction in p_1 small enough, $\hat{p}_1(t'_1)$ will remain above φ_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.²¹

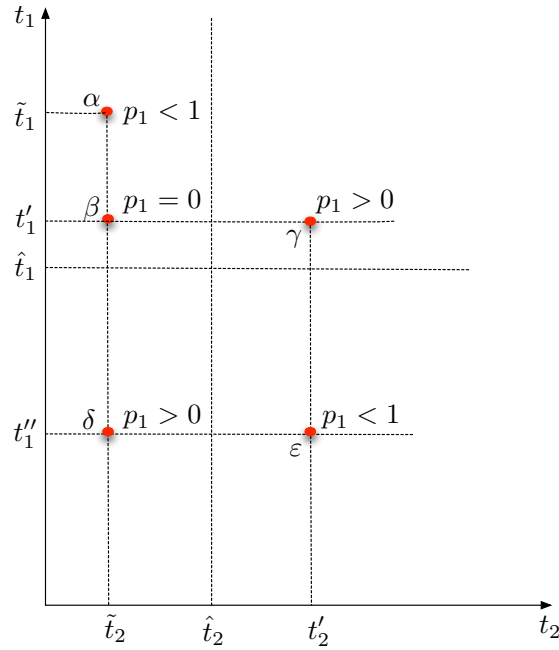


Figure 2

Similar reasoning implies that for every $t_1 \neq \tilde{t}_1$, we must have $\sum_i p_i(t_1, \tilde{t}_2) = 1$. Otherwise, the principal would be strictly better off increasing $p_2(t_1, \tilde{t}_2)$, decreasing $p_2(\tilde{t}_1, \tilde{t}_2)$, and increasing $p_1(\tilde{t}_1, \tilde{t}_2)$. Again, 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, an improvement.

Since player 1 receives the good with zero probability at β 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 γ relative to this cutoff plays no role in the argument to follow.

²¹Since $\hat{p}_2(\tilde{t}_2)$ 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.

Finally, there must be a $t_1'' \neq \tilde{t}_1$ (not necessary below \hat{t}_1) corresponding to points δ and ε where p_1 is strictly positive at δ and strictly less than 1 at ε . 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 $p_1(t_1, \tilde{t}_2) = 1$. Since $\sum_i p_i(t_1, \tilde{t}_2) = 1$ for all $t_1 \neq \tilde{t}_1$, this implies that 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 $t_1 \neq \tilde{t}_1$. 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 p at the points α , γ , ε , and δ in a way that lowers the probability that 1 gets the object at γ and raises the probability he gets it at α by the same amount, while maintaining the constraints. Since 1's type is higher at α , 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.²² Let $\Delta > 0$ be a “small” positive number. All the changes in p that we now define involve increases and decreases by the same amount Δ . At γ , lower p_1 and increase p_2 . At ε , 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 Δ is small enough, $\hat{p}_1(t_1')$ remains above φ_1 . Thus the constraints are maintained. Now that we have increased p_1 at ε , we can decrease it at δ while increasing p_2 , keeping $\hat{p}_1(t_1'')$ unchanged as \tilde{t}_2 and t_2' have the same probability. Finally, since we have increased p_2 at δ , we can decrease it at α while increasing p_1 , keeping $\hat{p}_2(\tilde{t}_2)$ unchanged. Note that the overall effect of these changes is a reduction of Δ in the probability that 1 gets the object at γ and an increase of Δ in the probability that he gets the object at α , while maintaining all constraints.

7 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.

²²If they do not have the same probability, the same approach as outlined in footnote 19 can be used to complete the argument.

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 $t_i > 0$ for all i .

This change affects very little of our analysis. To see this, refer again to the proof sketch in Section 6. 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

$$\max_{p_1, \dots, p_I} E_t \left[\sum_i p_i(t) (t_i - c_i) \right]$$

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 φ_i 's, the linearity of the objective function in φ_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 φ . 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's 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.

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 $t_i^* - c_i = \max_{j=0, \dots, I} (t_j^* - c_j)$.

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’s 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 .²³

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

²³On 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.

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 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

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

so that incentive compatibility holds iff

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

Analogously to the way we characterized the optimal mechanism in Section 6, we can treat φ_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 ε , 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 φ_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

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

Substituting, this implies that

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

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

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

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.²⁴ 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.

²⁴We thank Andy Skrzypacz for useful observations exploited in the discussion which follows.

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.²⁵ 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.

²⁵As 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.

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 $i \in \operatorname{argmax}_j E(t_j)$. 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.

8 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 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 5, 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 isn't 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 7, 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 7, 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.

A Proof of Theorem 1

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

The following definition was given in the text.

Definition 1. (p, q) is a threshold mechanism if there exists $v^* \in \mathbf{R}$ such that the following holds for all t up to sets of measure zero. First, if there exists any i with $t_i - c_i > v^*$, then $p_i(t) = 1$ 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 $q_i(t) = 0$ 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 B 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 (p, q) denote any optimal mechanism. In light of Theorem 4, we know (p, q) is a threshold mechanism. Hence we can specify $\hat{p}_i(t_i)$ for each agent as a function only of v^* and φ . To see this, fix v^* and φ and consider t_i such that $t_i - c_i > v^*$. Since (p, 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 φ . More specifically, the principal's payoff is

$$\begin{aligned} \mathbf{E}_t \left[\sum_i [p_i(t)(t_i - c_i) + \varphi_i c_i] \right] &= \sum_i F_i(v^* + c_i) \mathbf{E}_t [p_i(t)(t_i - c_i) \mid t_i < v^* + c_i] \\ &+ \sum_i \int_{v^* + c_i}^{\bar{t}_i} \left[\prod_{j \neq i} F_j(t_i - c_i + c_j) \right] (t_i - c_i) f_i(t_i) dt_i + \sum_i \varphi_i c_i. \end{aligned}$$

Note that

$$\begin{aligned} \mathbf{E}_t [p_i(t)(t_i - c_i) \mid t_i < v^* + c_i] &= \mathbf{E}_{t_i} \{ \mathbf{E}_{t_{-i}} [p_i(t)(t_i - c_i)] \mid t_i < v^* + c_i \} \\ &= \mathbf{E}_{t_i} [\hat{p}_i(t_i)(t_i - c_i) \mid t_i < v^* + c_i] \\ &= \mathbf{E}_{t_i} [\varphi_i(t_i - c_i) \mid t_i < v^* + c_i] \\ &= \varphi_i (\mathbf{E}_{t_i} [t_i \mid t_i < v^* + c_i] - c_i). \end{aligned}$$

Hence we can rewrite the objective function as

$$\begin{aligned}
& \sum_i F_i(v^* + c_i) \varphi_i [\mathbf{E}_{t_i}(t_i \mid t_i < v^* + c_i) - c_i] \\
& \quad + \sum_i \int_{v^* + c_i}^{\bar{t}_i} \left[\prod_{j \neq i} F_j(t_i - c_i + c_j) \right] (t_i - c_i) f_i(t_i) dt_i + \sum_i \varphi_i c_i \\
& = \sum_i \varphi_i \{ F_i(v^* + c_i) \mathbf{E}_{t_i}(t_i \mid t_i < v^* + c_i) + [1 - F_i(v^* + c_i)] c_i \} \\
& \quad + \sum_i \int_{v^* + c_i}^{\bar{t}_i} \left[\prod_{j \neq i} F_j(t_i - c_i + c_j) \right] (t_i - c_i) f_i(t_i) dt_i.
\end{aligned}$$

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 φ fixed).

For a fixed v^* , the principal's objective function is linear in φ . Given v^* , the set of feasible φ vectors is convex. To be precise, recall that a given specification of p_i and φ_i , $i \in \mathcal{I}$, is feasible iff each $p_i : T \rightarrow [0, 1]$, each $\varphi_i \in [0, 1]$, $\sum_i p_i(t) \leq 1$ for all t , and $\mathbf{E}_{t_i} p_i(t) \geq \varphi_i$ for all $t_i \in T_i$ and all i . From Theorem 4, we know the exact value of $p_i(t)$ for all i for (almost) any 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 9, we know this holds on a set of strictly positive measure.) We say that a profile φ_i , $i \in \mathcal{I}$, is *feasible* given v^* if there exists p_i functions satisfying the properties above given v^* and these φ_i 's.

Lemma 1. *The set of φ_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 iff his is the highest type. 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

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

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

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

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

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

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 φ 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 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 .

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 φ satisfying the necessary conditions for such a v^* . Again, it is easy to see that this is feasible since it is achieved for any 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, φ_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 φ satisfying the conditions of the lemma. To see that this φ is feasible, set p as follows. For any t such that $\max_i t_i - c_i > v^*$, let $p_i(t) = 1$ for that i with $t_i - c_i > \max_{j \neq i} t_j - c_j$. For any t with $\max_i t_i - c_i < v^*$, let $p_i(t) = \varphi_i / \prod_{j \neq i} F_j(v^* + c_j)$ for every i . Since $\varphi_i \in [0, 1]$, $p_i(t)$ is non-negative for all i . Also,

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

By our third condition, this is less than 1.

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

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

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 φ is feasible given v^* . ■

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

The following lemma characterizes the extreme points. Recall that

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

Lemma 2. *If \mathcal{I}^0 is not a singleton, then φ^* is an extreme point of the set of feasible φ given v^* iff 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 φ^* is an extreme point of the set of feasible φ given v^* iff $\varphi_j^* = 0$ 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 φ 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 φ 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 φ^* 's stated in the lemma must all be extreme points. To see this, suppose that there exists a feasible φ^1 and φ^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 φ^* 's stated in the lemma. Obviously, we cannot have φ^* 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 φ^2 violates the third condition for feasibility of φ given v^* , a contradiction.

Hence we only need to show that there are no other extreme points. To show this, we show that any φ which is feasible given v^* can be written as a convex combination of these points. So fix any such φ . 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 φ^* 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 φ is not an extreme point unless it equals one of the φ^* 's. ■

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

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

Proof. First, suppose $\#\mathcal{I}^0 \geq 2$. In Lemma 1, we showed that the only feasible φ 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 $p_i^*(t) = 1$ 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 iff 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 φ 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

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

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 $\mathcal{T}^0 = \emptyset$. 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 $\varphi^*(k)$, $k = 1, \dots, I$. 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 φ_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 $t_i \geq 0$ and $c_i > 0$. Hence whenever there is an feasible φ other than $\mathbf{0}$, it *must* yield the principal a higher payoff than setting φ 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

- [1] Abdulkadiroglu, A., and T. Sonmez, “Matching Markets: Theory and Practice,” in Volume 1 of Acemoglu et. al., *Advances in Economics and Econometrics: Proceedings of the Tenth World Congress of the Econometric Society*, 2013, Cambridge University Press.
- [2] Ambrus, A., and G. Egorov, “Delegation and Nonmonetary Incentives,” Duke University working paper, July 2012.
- [3] Banerjee, A., R. Hanna, and S. Mullainathan, “Corruption,” working paper, January 2011.
- [4] Ben-Porath, E., and B. Lipman, “Implementation and Partial Provability,” *Journal of Economic Theory*, **147**, September 2012, 1689–1724.
- [5] Border, K., and J. Sobel, “Samurai Accountant: A Theory of Auditing and Plunder,” *Review of Economic Studies*, **54**, October 1987, 525–540.
- [6] Bull, J., and J. Watson, “Hard Evidence and Mechanism Design,” *Games and Economic Behavior*, **58**, January 2007, 75–93.
- [7] Chakravarty, S., and T. Kaplan, “Optimal Allocation without Transfer Payments,” *Games and Economic Behavior*, **77**, January 2013, 1–20.
- [8] Che, Y.-K., W. Dessein, and N. Kartik, “Pandering to Persuade,” *American Economic Review*, forthcoming.
- [9] Condorelli, D., “What Money Can’t Buy: Efficient Mechanism Design with Costly Signals,” *Games and Economic Behavior*, **75**, July 2012, 613–624.
- [10] Deneckere, R. and S. Severinov, “Mechanism Design with Partial State Verifiability,” *Games and Economic Behavior*, **64**, November 2008, 487–513.
- [11] Doval, L., “Whether or Not to Open Pandora’s Box,” Northwestern University working paper, November 2013.
- [12] Gale, D., and M. Hellwig, “Incentive–Compatible Debt Contracts: The One–Period Problem,” *Review of Economic Studies*, **52**, October 1985, 647–663.
- [13] Glazer, J., and A. Rubinstein, “On Optimal Rules of Persuasion,” *Econometrica*, **72**, November 2004, 1715–1736.
- [14] Glazer, J., and A. Rubinstein, “A Study in the Pragmatics of Persuasion: A Game Theoretical Approach,” *Theoretical Economics*, **1**, December 2006, 395–410.

- [15] Green, J., and J.-J. Laffont, “Partially Verifiable Information and Mechanism Design,” *Review of Economic Studies*, **53**, July 1986, 447–456.
- [16] Hartline, J. D., and T. Roughgarden, “Optimal Mechanism Design and Money Burning,” *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, 2008, 75–84.
- [17] Kartik, N., and O. Tercieux, “Implementation with Evidence,” *Theoretical Economics*, **7**, May 2012, 323–355.
- [18] McAfee, R. P., and J. McMillan, “Bidding Rings,” *American Economic Review*, **82**, June 1992, 579–599.
- [19] Mookherjee, D., and I. Png, “Optimal Auditing, Insurance, and Redistribution,” *Quarterly Journal of Economics*, **104**, May 1989, 399–415.
- [20] Roth, A. E., and M. A. O. Sotomayor, *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, 1990, Cambridge University Press.
- [21] Sher, I., and R. Vohra, “Price Discrimination through Communication,” working paper, June 2011.
- [22] Townsend, R., “Optimal Contracts and Competitive Markets with Costly State Verification,” *Journal of Economic Theory*, **21**, 1979, 265–293.
- [23] Townsend, R., “Information Constrained Insurance: The Revelation Principle Extended,” *Journal of Monetary Economics*, **21**, March–May 1988, 411–450.
- [24] Weitzman, M., “Optimal Search for the Best Alternative,” *Econometrica*, **47**, May 1979, 641–654.
- [25] Yoon, K., “Optimal Mechanism Design When Both Allocative Inefficiency and Expenditure Inefficiency Matter,” *Journal of Mathematical Economics*, **47**, December 2011, 670–676.

Online Appendix

A Reduction

In this Appendix, we show that we can reduce the principal's problem to the choice of (p, q) functions as in the text. We begin with an arbitrary mechanism which could have 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. Think of such a dynamic mechanism as a game in extensive form between the agents and the principal where the principal is committed in advance to his strategy. The principal's actions specify decisions about which agent or agents to check at various points and, ultimately, which (if any) to allocate the good to. Fix such a dynamic mechanism, deterministic or otherwise, and any equilibrium, say σ , in pure or mixed strategies. We show that the principal's payoff in this mechanism can be duplicated or improved by the appropriate choice of (p, q) functions.

There are three key steps to the reduction. The first step is a version of the Revelation Principle appropriate to this setting which shows that, without loss of generality, we can restrict attention to truth-telling equilibria in direct mechanisms. In the second step, we use our assumption of perfect verification to show that we can restrict attention to mechanisms where all checks the principal carries out are done simultaneously. Finally, the last step identifies two simple properties of the optimal allocation.

To show the first step, we take the equilibrium, σ , of the original mechanism and construct a new mechanism as follows. Each player i reports a type $t_i \in T_i$. Given a vector of reports t , the principal simulates what would happen in the original mechanism when the agents play the strategies $\sigma(t)$. That is, they play as they would have in the equilibrium if the true types were t . As the principal simulates the mechanism, he may check some agents' types. If he gets all the way through the simulation without any checks revealing that some agent's report is false, he allocates the good as in the simulation.

Suppose that some checks reveal that one or more agents have lied. If more than one agent must have lied, then the principal allocates the good arbitrarily. Since we will only be interested in truth-telling equilibria, only unilateral lies will be relevant for incentives.

So suppose that checks reveal only that agent i has lied — that is, the outcome of checks are consistent with the reports of all agents $j \neq i$. If agent i has another move after this point, the principal can no longer simulate the mechanism using the strategy

$\sigma_i(t_i)$. By definition, the information set corresponding to this later move is one that could not have been reached if i really were type t_i , so no strategy for type t_i would say anything about this information set. To continue the simulation, the principal chooses any feasible strategy for i from this point forward. Again, he completes the simulation and then allocates the good according to the result of the simulation.

It is easy to see that truth telling is an equilibrium of this game. Fix any player i of type t_i and assume that all agents $j \neq i$ report truthfully. Then i 's payoff from reporting truthfully as well is exactly the same as in the equilibrium of the original mechanism. His payoff from reporting any other type is exactly the same as his payoff to a certain deviation in the original mechanism. Hence the fact that the original strategies formed an equilibrium implies that truth telling is a best reply. Clearly, the principal's payoff in the truth telling equilibrium is the same as in the original mechanism.

Next, we show that given our assumption that verification is perfect — that is, if the principal checks agent i , he learns i 's true type — it is without loss of generality to assume that the principal carries out whatever checks he does all at once. In other words, there is no need for him to decide whether to check an agent based on the outcome of earlier checks.

To see this, again, fix any mechanism and any equilibrium. Now we construct a direct mechanism as follows. If the reported type profile is t , then the principal computes the probability distribution over which agents would be checked in the equilibrium of the original mechanism given that the true types are t . He then randomizes over the set of agents to check using this probability distribution, but carries out these checks simultaneously rather than sequentially. For example, if in the original mechanism, he would have checked agent 1, then randomized 50–50 over whether to check agent 2, the principal randomizes 50–50 over checking just agent 1 or checking both 1 and 2 simultaneously. Similarly, if the principal would have checked agent 1 and then only checked 2 if he learned 1 had some type other than t_1 , then he just checks agent 1 since this is what would happen conditional on the types being t .

If what the principal observes from the checks is consistent with what he would have seen in the equilibrium (that is, for every agent j he checks, he sees that j 's type is t_j), then he allocates the good exactly as he would have done in the equilibrium after these observations. If there is only a single player, say i , who is found to have type $t'_i \neq t_i$, then the allocation of the good is the same as it would have been in the original equilibrium if the type profile were (t'_i, t_{-i}) , players $j \neq i$ used their equilibrium strategies, and player i deviated to the equilibrium strategy of type t_i . Finally, the allocation is arbitrary if the principal learns that two or more players have types different from their reports.

As before, truth telling is an equilibrium of this game. For any player i , consider the best reply of type t_i to truth-telling by the other agents. Just as before, i 's payoff from

reporting truthfully is the same as in the equilibrium of the original mechanism. Just as before, his payoff to reporting any other type is the same as his payoff to a certain deviation in the original mechanism. Therefore, the fact that we began with equilibrium strategies for the original mechanism implies that a best reply for t_i is to report t_i . Clearly, the payoff for the principal is the same as before.

With imperfect verification, sequential checking procedures may be needed. However, this is just a matter of computing the statistical tools available to the principal. That is, a sequential procedure for checking gives the principal a certain probability distribution over observations and costs as a function of the true types. One can simply determine the set of such conditional distributions and treat the principal as picking among them. That is, we can translate the problem into one with *different* stochastic verification technologies, corresponding to what can be done in the sequential environment and have the principal use these simultaneously.

For the third step, we give two simple but useful properties of the optimal allocation. First, given that we focus on truth telling equilibria, all situations in which agent i 's report is checked and found to be false are off the equilibrium path. The specification of the mechanism for such a situation cannot affect the incentives of any agent $j \neq i$ since agent j will expect i 's report to be truthful. Thus the specification only affects agent i 's incentives to be truthful. Since we want i to have the strongest possible incentives to report truthfully, we may as well assume that if i 's report is checked and found to be false, then the good is given to agent i with probability 0. Hence we can further reduce the complexity of a mechanism to specify which agents are checked and which agent receives the good as a function of the reports, where the latter applies only when the checked reports are accurate.

Finally, any agent's incentive to reveal his type is unaffected by the possibility of being checked in situations where he does not receive the object regardless of the outcome of the check. That is, if an agent's report is checked even when he would not receive the object if found to have told the truth, his incentives to report honestly are not affected. Since checking is costly for the principal, this means that if the principal checks an agent, then (if he is found to have been honest), he must receive the object with probability 1.

Therefore, we can think of the mechanism as specifying 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. As in the text, we let $q_i(t)$ denote the probability i is awarded the object conditional on a successful check and let $p_i(t)$ be the total probability i is awarded the object.

B Proof of Theorem 4

The proof of Theorem 4 proceeds with a series of lemmas. Throughout we write the distribution of t_i as a measure μ_i . Recall that we have assumed this measure is absolutely continuous with respect to Lebesgue measure on the interval $T_i \subset \mathbf{R}$. We let μ be the product measure on the product Borel field of T . For any $S \subseteq T$, let

$$S(t_i) = \{t_{-i} \in T_{-i} \mid (t_i, t_{-i}) \in S\}$$

denote the t_i fiber of S . Let S_i denote the projection of S on T_i , and S_{-ij} the projection on $\prod_{k \notin \{i,j\}} T_k$.

We begin with a technical lemma.²⁶

Lemma 4. *Given any Borel measurable $S \subset \mathbf{R}^i$ with $\mu(S) > 0$, there exists $S^* \subseteq S$ with $\mu(S^*) = \mu(S)$ such that the following holds. First, for every i and every $t_i \in T_i$, the measure of every fiber is strictly positive. That is, $\mu_{-i}(S(t_i)) > 0$ for all i and all $t_i \in T_i$. Second, for all i , the projection on i of S^* , S_i^* , is measurable.*

*Moreover, given any j , there exists $\varepsilon > 0$ and $S^{**} \subseteq S$ with $\mu(S^{**}) > 0$ such that the following holds. First, for every $i \neq j$ and every $t_i \in T_i$, the measure of every fiber is strictly positive. That is, $\mu_{-i}(S^{**}(t_i)) > 0$. Second, for every $t_j \in S_j^{**}$, the fiber $S^{**}(t_j)$ has measure bounded below by ε . That is, $\mu_{-j}(S^{**}(t_j)) > \varepsilon$. Finally, for all i , S_i^{**} , the projection on i of S^{**} , is measurable.*

Proof. We first prove this for $I = 2$, and then show how to extend it to $I > 2$. So, to simplify notation for the first step, denote by x and y the two dimensions. Fix a Borel measurable S with $\mu(S) > 0$. We need to show that there is an equal measure subset of S , S^* , such that all fibers of S^* have strictly positive measure and all projections of S^* are measurable. So we need to show (1) $\mu_x(S^*(y)) > 0$ for all y , (2) $\mu_y(S^*(x)) > 0$ for all x , and (3) the projections of S^* are measurable.

First, we observe that if all the fibers have strictly positive measure, then the projections are measurable. To see this, note that the function $f : X \rightarrow \mathbf{R}$ given by $f(x) = \mu_y(S^*(x))$ is measurable by Fubini's Theorem. Hence the set $\{x \mid \mu_y(S^*(x)) > 0\}$ is measurable. But this is just the projection on the first dimension if the fiber has positive measure. An analogous argument applies to the y coordinate.

Let S^1 denote the set S after we delete all x fibers with μ_y measure zero. That is, $S^1 = S \cap [\{x \mid \mu_y(S(x)) > 0\} \times \mathbf{R}]$. We know that S^1 is measurable, has the same

²⁶We thank Benjy Weiss for suggesting the idea of the following proof.

measure as S (by Fubini, because we deleted only fibers of zero measure), all its x fibers have strictly positive y measure, and its projection on x is measurable.

We do not know that the projection of S^1 on y is measurable nor that the y fibers have strictly positive x measure. Let S^2 denote the set S^1 after we delete all y fibers with μ_x measure zero. That is, $S^2 = S^1 \cap [\{y \mid \mu_x(S^1(y)) > 0\} \times \mathbf{R}]$. We know that S^2 is measurable with the same measure as S^1 , that its projection on y is measurable, and all its y fibers have strictly positive y measure.

Again, we do not know that its projection on x is measurable nor that the x fibers have strictly positive y measure. But at this step we do know that the set of x fibers that have zero measure is contained in a set of measure zero. Put differently,

$$\mu_x \{x \mid \mu_y(S^2(x)) > 0\} = \mu_x(S^1) = \mu_x \{x \mid \mu_y(S^1(x)) > 0\}. \quad (8)$$

To see this, suppose not. Then

$$\mu_x \{x \mid \mu_y(S^2(x)) > 0\} < \mu_x \{x \mid \mu_y(S^1(x)) > 0\}$$

as

$$\{x \mid \mu_y(S^2(x)) > 0\} \subseteq \{x \mid \mu_y(S^1(x)) > 0\}.$$

Let

$$\Delta = \{x \mid \mu_y(S^1(x)) > 0\} \setminus \{x \mid \mu_y(S^2(x)) > 0\}.$$

If $\mu(\Delta) > 0$, then

$$\begin{aligned} \mu(S^1) &= \int_{\{x \mid \mu_y(S^1(x)) > 0\}} \mu_y(S^1(x)) \mu_x(dx) \\ &= \int_{\{x \mid \mu_y(S^2(x)) > 0\}} \mu_y(S^1(x)) \mu_x(dx) + \int_{\Delta} \mu_y(S^1(x)) \mu_x(dx) \\ &> \int_{\{x \mid \mu_y(S^2(x)) > 0\}} \mu_y(S^2(x)) \mu_x(dx) \\ &= \mu(S^2) \end{aligned}$$

as $S^1(x) \supseteq S^2(x)$ and $\mu(\Delta) > 0$. But this contradicts $\mu(S^2) = \mu(S^1)$. Hence equation (8) holds.

Finally, let S^3 denote S^2 after we delete all x fibers with μ_y measure zero. That is, $S^3 = S^2 \cap [\{x \mid \mu_y(S^2(x)) > 0\} \times \mathbf{R}]$. We know that S^3 is measurable with the same measure as S^2 , that its projection on x is measurable, and that all its x fibers have strictly positive y measure. But now we also know that all the y fibers have strictly positive x measure, since in going from S^2 to S^3 , we deleted a set of x 's contained in a set of zero measure. Hence each y fiber has the same measure as before.

We now extend this to $I > 2$. For brevity, we only describe the extension to $I = 3$, the more general result following the same lines. Denote the coordinates by x , y , and z . Consider a set S with $\mu(S) > 0$. We show there exists $S^* \subseteq S$ such that $\mu_{yz}(S^*(x)) > 0$ for all $x \in S_x^*$, and similarly for all $y \in S_y^*$ and all $z \in S_z^*$.

From the case of $I = 2$, we know there exists $S^1 \subseteq S$ with $\mu(S^1) = \mu(S)$ such that for all $x \in S_x^1$, we have $\mu_{yz}(S^1(x)) > 0$ and for all $(y, z) \in S_{yz}^1$, we have $\mu_x(S^1((y, z))) > 0$. Applying $I = 2$ result again to the set S_{yz}^1 , we have $G \subseteq S_{yz}^1$ with $\mu_{yz}(G) = \mu_{yz}(S_{yz}^1)$ such that for all $y \in G_y$, we have $\mu_z(G(y)) > 0$ and for all $z \in G_z$, we have $\mu_y(G(z)) > 0$. (Note that this implies that $\mu_{yz}(G) > 0$.)

Now define

$$S^2 = S^1 \cap (\mathbf{R} \times G) = \{(x, y, z) \mid (x, y, z) \in S^1 \text{ and } (y, z) \in G\}.$$

Since $G \subseteq S_{yz}^1$ and $\mu_{yz}(G) = \mu_{yz}(S_{yz}^1)$, we have $\mu(S^2) = \mu(S^1)$. Clearly, $S_y^2 = G_y$ and $S_z^2 = G_z$. Fix any $y \in S_y^2$. Since $y \in G_y$, we have $\mu_z\{z \mid (y, z) \in G\} > 0$. Since $G \subseteq S_{yz}^1$ for every $(y, z) \in G$, we have $\mu_x(S^2(y, z)) = \mu_x(S^1(y, z)) > 0$. By Fubini's Theorem, $\mu_{xz}(S^2(y)) = \int_{z \in G(y)} \mu_x(S^2(y, z)) \mu_z(dz)$ and hence $\mu_{xz}(S^2(y)) > 0$. A similar argument implies that for all $z \in S^2(z)$, we have $\mu_{xy}(S^2(z)) > 0$. However, we do not know that for every $x \in S_x^2$, we have $\mu_{y,z}(S^2(x)) > 0$. Hence we now define the set S^3 by

$$S^3 = S^2 \cap (\{x \mid \mu_{yz}(S^2(x)) > 0\} \times \mathbf{R}^2).$$

Clearly, S_x^3 is measurable and we have $\mu_{yz}(S^3(x)) > 0$ for every $x \in S_x^3$. Furthermore, $S^3 \subseteq S^2 \subseteq S^1$ and hence $S_x^3 \subseteq S_x^1$. In fact, $\mu(S^3) = \mu(S^2) = \mu(S^1)$ implies $\mu(S_x^3) = \mu(S_x^1)$. To see this, suppose not. Then $\mu_x(S_x^3) < \mu_x(S_x^1)$. Since for each $x \in S_x^1$, we have $\mu_{yz}(S^1(x)) > 0$, we obtain that $\mu(S^3) < \mu(S^1)$, a contradiction.

We claim that S^3 satisfies the properties stated in the first part of the lemma. That is, (1) S_y^3 and S_z^3 are measurable, (2) for all $y \in S_y^3$, we have $\mu_{x,z}(S^3(y)) > 0$, and (3) for all $z \in S_z^3$, we have $\mu_{x,y}(S^3(z)) > 0$. Consider an element $y \in S_y^2$. We have seen that for all $z \in G(y)$, we have $\mu_x(S^2(y, z)) > 0$. Since our construction of S^3 removes from S^2 a set of elements x in S_x^2 that is contained in a set of measure zero, we must have $\mu_x(S^3(y, z)) = \mu_x(S^2(y, z)) > 0$. Hence $S_y^3 = S_y^2$ and for every $y \in S_y^3$, we have $\mu_{xz}(S^3(y)) = \mu_{xz}(S^2(y)) > 0$. A similar argument establishes that $S_z^3 = S_z^2$ and that for $z \in S_z^3$, we have $\mu_{xy}(S^3(y)) > 0$. By defining $S^* = S^3$, we obtain a set S^* with the properties claimed in the first part of the lemma.

It remains to prove the ‘‘moreover’’ claim. This follows from a similar argument where in defining S^1 , we remove all x 's whose fibers do not have probability at least ε for an appropriately chosen ε . We provide the proof for the case $I = 2$. The proof for $I > 2$ is similar.

Note that

$$\{x \mid \mu_y(S(x)) > 0\} = \bigcup_{n=1}^{\infty} \{x \mid \mu_y(S(x)) > 1/n\}.$$

Since $\mu_x(\{x \mid \mu_y(S(x)) > 0\}) > 0$, there exists \hat{n} such that $\mu_x(\{x \mid \mu_y(S(x)) > 1/\hat{n}\}) > 0$. Define $\varepsilon = 1/\hat{n}$ and define $S^1 = S \cap (\{x \mid \mu_y(S(x)) > \varepsilon\} \times \mathbf{R})$.

The rest of the argument is essentially identical to the argument given in the proof of the first part of the lemma. Specifically, we know that S_x^1 is measurable and that for every $x \in S_x^1$, we have $\mu_y(S^1(x)) > \varepsilon$. Define

$$S^2 = S^1 \cap (\{y \mid \mu_x(S^1(y)) > 0\} \times \mathbf{R})$$

$$S^3 = S^2 \cap (\{x \mid \mu_y(S^2(x)) > \varepsilon\} \times \mathbf{R}).$$

We have $S^3 \subseteq S^2 \subseteq S^1$. Fubini's Theorem implies that $\mu(S^2) = \mu(S^1)$ which in turn implies that

$$\mu_x(\{x \mid \mu_y(S^2(x)) > \varepsilon\}) = \mu_x(\{x \mid \mu_y(S^1(x)) > \varepsilon\}).$$

To see this, suppose not. Then $S^2 \subsetneq S^1$ and the fact that $\mu_y(S^1(x)) > \varepsilon$ for all $x \in S_x^1$ implies that $\mu(S^2) < \mu(S^1)$, a contradiction.

Since

$$\{x \mid \mu_y(S^2(x)) > \varepsilon\} = \{x \mid \mu_y(S^3(x)) > \varepsilon\} = S_x^3,$$

we see that $\mu_x(S_x^1) = \mu_x(S_x^3)$. Hence in moving from S^2 to S^3 , the set of x 's that is deleted is contained in a set of measure zero. Since for all $y \in S_y^2$, we have $\mu_x(S^2(y)) > 0$, we see that $S_y^3 = S_y^2$ and that $\mu_x(S^3(y)) > 0$ for all $y \in S_y^3$. Thus the set S^3 satisfies all the properties stated in the second paragraph of the lemma. ■

For the remaining lemmas, fix p and φ that maximize

$$\mathbf{E}_t \left[\sum_i [p_i(t)(t_i - c_i) + \varphi_i c_i] \right] = \sum_i \{ \mathbf{E}_{t_i} [\hat{p}_i(t_i)(t_i - c_i)] + \varphi_i c_i \}$$

subject to $\sum_i p_i(t) \leq 1$ for all t and $\hat{p}_i(t_i) \geq \varphi_i \geq 0$ for all i and t_i where $\hat{p}_i(t_i) = \mathbf{E}_{t-i} p_i(t)$. As explained in Section 6, the optimal q will then be any feasible q satisfying $\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i$ for all i and t_i where $\hat{q}_i(t_i) = \mathbf{E}_{t-i} q_i(t)$.

Lemma 5. *There is a set $T' \subseteq T$ with $\mu(T') = 1$ such that the following hold:*

1. For each i , if $t_i < c_i$ and $t_i \in T'_i$, then $\hat{p}_i(t_i) = \varphi_i$.
2. For each $t \in T'$, if $t_i > c_i$ for some i , then $\sum_j p_j(t) = 1$.
3. For any $t \in T'$, if $\hat{p}_i(t_i) > \varphi_i$ for some i , then $\sum_j p_j(t) = 1$.

Proof. Proof of 1. If $t_i < c_i$, then the objective function is strictly decreasing in $\hat{p}_i(t_i)$. Obviously, reducing $\hat{p}_i(t_i)$ makes the other constraints easier to satisfy. Since we improve the objective function and relax the constraints by reducing $\hat{p}_i(t_i)$, we must have $\hat{p}_i(t_i) = \varphi_i$ at the optimum. This completes the proof of part 1. Since we only characterize optimal mechanisms up to sets of measure zero, we abuse notation, and redefine T to equal a measure 1 subset of T on which property 1 is satisfied, and whose projections are measurable (which exists by Lemma 4).

Proof of 2. Suppose not. Then there exists an agent i and a set \hat{T} with positive measure such that for every $t \in \hat{T}$, we have $t_i > c_i$ and yet $\sum_j p_j(t) < 1$. Define an allocation function p^* by

$$p_j^*(t) = \begin{cases} p_j(t), & \text{if } j \neq i \text{ or } t \notin \hat{T} \\ 1 - \sum_{j \neq i} p_j(t), & \text{otherwise.} \end{cases}$$

It is easy to see that p^* satisfies all the constraints and improves the objective function, a contradiction.

Proof of 3. Suppose to the contrary that we have a positive measure set of t such that $\sum_j p_j(t) < 1$ but for each t , there exists some i with $\hat{p}_i(t_i) > \varphi_i$. Then there exists i and a positive measure set of t such that for each t , we have $\sum_j p_j(t) < 1$ and $\hat{p}_i(t_i) > \varphi_i$. From part 1, we know that for all $t_i \in T_i$ with $\hat{p}_i(t_i) > \varphi_i$ we have $t_i > c_i$. Hence from part 2, the mechanism is not optimal, a contradiction. ■

Abusing notation, redefine T to equal a measure 1 subset of $T \setminus T'$ whose projections are measurable (which exists by Lemma 4) on which all the properties of Lemma 5 are satisfied everywhere.

Lemma 6. *There is a set $T' \subseteq T$ with $\mu(T') = 1$ such that for any $t \in T'$, if $t_i - c_i > t_j - c_j$ and $\hat{p}_j(t_j) > \varphi_j$, then $p_j(t) = 0$.*

Proof. Suppose not. Then we have a positive measure set S such that for all $t \in S$, $t_i - c_i > t_j - c_j$, $\hat{p}_j(t_j) > \varphi_j$, and $p_j(t) > 0$. Hence there exists $\alpha > 0$ and $\varepsilon > 0$ such that $\mu(\hat{S}) > 0$ where

$$\hat{S} = \{t \in T \mid t_i - c_i - (t_j - c_j) \geq \alpha, \hat{p}_j(t_j) \geq \varphi_j + \varepsilon, \text{ and } p_j(t) \geq \varepsilon\}.$$

Define p^* by

$$p_j^*(t) = \begin{cases} p_k(t), & \text{for } k \neq i, j \text{ or } t \notin \hat{S} \\ p_j(t) - \varepsilon, & \text{for } k = j \text{ and } t \in \hat{S} \\ p_i(t) + \varepsilon, & \text{for } k = i \text{ and } t \in \hat{S}. \end{cases}$$

Since $p_j(t) \geq \varepsilon$ for all $t \in \hat{S}$, we have $p_k^*(t) \geq 0$ for all k and t . Obviously, $\sum_k p_k^*(t) = \sum_k p_k(t)$, so the constraint that the p_k 's sum to less than one must be satisfied.

Turning to the lower bound constraint on the \hat{p}_k 's, obviously, for $k \neq j$, we have $\hat{p}_k^*(t_k) \geq \hat{p}_k(t_k) \geq \varphi_k$, so the constraint is satisfied for all $k \neq j$ and all t_k . For any t_j , either $\hat{p}_j^*(t_j) = \hat{p}_j(t_j)$ or

$$\hat{p}_j^*(t_j) = \hat{p}_j(t_j) - \varepsilon \mu_{-j}(\hat{S}(t_j)) \geq \hat{p}_j(t_j) - \varepsilon.$$

But for each t_j for which $\hat{p}_j^*(t_j) \neq \hat{p}_j(t_j)$, we have $\hat{p}_j(t_j) \geq \varphi_j + \varepsilon$, so

$$\hat{p}_j^*(t_j) \geq \varphi_j + \varepsilon - \varepsilon = \varphi_j.$$

Hence for every k and every t_k , we have $\hat{p}_k^*(t_k) \geq \varphi_k$. Therefore, p^* is feasible given φ .

Finally, the change in the principal's payoff in moving from p to p^* is

$$\mu(\hat{S})\varepsilon \left[\mathbb{E}(t_i - c_i \mid t \in \hat{S}) - \mathbb{E}(t_j - c_j \mid t \in \hat{S}) \right] \geq \mu(\hat{S})\varepsilon\alpha > 0.$$

Hence p could not have been optimal, a contradiction. ■

Thus the set $S^0 = \{t \in T \mid t_i - c_i - (t_j - c_j) > 0, \hat{p}_j(t_j) > \varphi_j, \text{ and } p_j(t) > 0\}$ has measure zero. Abusing notation, redefine T to equal a measure 1 subset of $T \setminus S^0$ whose projections are measurable (which exists by Lemma 4).

Lemma 7. *There is a set of measure one $T' \subseteq T$ such that for all $t', t'' \in T'$ such that $t'_j = t''_j$, $p_j(t') > 0$, $\hat{p}_i(t'_i) > \varphi_i$, $t''_i < t'_i$, and $\hat{p}_i(t''_i) > \varphi_i$, we have $p_i(t'') = 0$.*

The idea that underlies the proof is simple. Consider two profiles t' and t'' that have the properties stated in the lemma. That is, $t'_j = t''_j$, $p_j(t') > 0$, $\hat{p}_i(t'_i) > \varphi_i$, $t''_i < t'_i$, and $\hat{p}_i(t''_i) > \varphi_i$. Suppose the claim is false, so that $p_i(t'') > 0$. Clearly, there is some $\varepsilon > 0$ such that $p_j(t') > \varepsilon$, $\hat{p}_i(t'_i) > \varphi_i + \varepsilon$, and $\hat{p}_i(t''_i) > \varphi_i + \varepsilon$, and $p_i(t'') > \varepsilon$. For simplicity, assume $\mu(t') = \mu(t'') = \delta > 0$. (The formal proof will extend the argument to the case that μ is a general atomless probability measure.) Consider the following transfer of allocation probabilities between agents i and j . For the profile t' , increase $p_i(t')$ by ε and decrease $p_j(t')$ by ε . For the profile t'' , decrease $p_i(t'')$ by ε and increase $p_j(t'')$ by ε . Let p^* denote the resulting probability function. It is easy to see that p^* satisfies all the constraints. Also, it increases the value of the objective function because the net effect of the transfers is to move a probability $\varepsilon\delta$ of allocating the object from type t''_i to type t'_i where $t'_i > t''_i$. This argument is not sufficient for the general proof, of course, since μ is atomless, implying that we must change p on a positive measure set of types to have an effect.

Proof. Given any rational number α and any $t'_j \in T_j$, let

$$\hat{A}_{-j}(\alpha, t'_j) = \{t'_{-j} \in T_{-j} \mid t'_i > \alpha, \hat{p}_i(t'_i) > \varphi_i, \text{ and } p_j(t') > 0\}$$

$$\begin{aligned}\hat{B}_{-j}(\alpha, t'_j) &= \{t'_{-j} \in T_{-j} \mid t'_i < \alpha, \hat{p}_i(t'_i) > \varphi_i, \text{ and } p_i(t') > 0\}. \\ \hat{C}_j(\alpha) &= \{t'_j \in T_j \mid \mu_{-j}(\hat{A}_{-j}(\alpha, t'_j)) > 0, \mu_{-j}(\hat{B}_{-j}(\alpha, t'_j)) > 0\}.\end{aligned}$$

Also let

$$\begin{aligned}\tilde{A}_{-j}(\alpha, t'_j, \varepsilon, \delta) &= \{t'_{-j} \in T_{-j} \mid t'_i > \alpha + \delta, \hat{p}_i(t'_i) > \varphi_i + \varepsilon, \text{ and } p_j(t') > \varepsilon\} \\ \tilde{B}_{-j}(\alpha, t'_j, \varepsilon, \delta) &= \{t'_{-j} \in T_{-j} \mid t'_i < \alpha - \delta, \hat{p}_i(t'_i) > \varphi_i + \varepsilon, \text{ and } p_i(t') > \varepsilon\}.\end{aligned}$$

and

$$\begin{aligned}\tilde{A}(\alpha, \varepsilon, \delta) &= \{t \in T \mid t_i > \alpha + \delta, \hat{p}_i(t_i) > \varphi_i + \varepsilon, \text{ and } p_j(t) > \varepsilon\} \\ &= \bigcup_{t'_j \in T_j} \{t'_j\} \times \tilde{A}_{-j}(\alpha, t'_j, \varepsilon, \delta)\end{aligned}$$

$$\begin{aligned}\tilde{B}(\alpha, \varepsilon, \delta) &= \{t \in T \mid t_i < \alpha - \delta, \hat{p}_i(t_i) > \varphi_i + \varepsilon, \text{ and } p_i(t) > \varepsilon\} \\ &= \bigcup_{t'_j \in T_j} \{t'_j\} \times \tilde{B}_{-j}(\alpha, t'_j, \varepsilon, \delta)\end{aligned}$$

$$\tilde{C}_j(\alpha, \varepsilon, \delta) = \{t'_j \in T_j \mid \mu_{-j}(\tilde{A}_{-j}(\alpha, t'_j, \varepsilon, \delta)) > 0, \mu_{-j}(\tilde{B}_{-j}(\alpha, t'_j, \varepsilon, \delta)) > 0\}.$$

Finally let

$$\begin{aligned}\bar{A}(\alpha, \varepsilon, \delta) &= \bigcup_{t_j \in \tilde{C}_j(\alpha, \varepsilon, \delta)} \{t_j\} \times \tilde{A}_{-j}(\alpha, t_j, \varepsilon, \delta) = \left(\tilde{C}_j(\alpha, \varepsilon, \delta) \times T_{-j}\right) \cap \tilde{A}(\alpha, \varepsilon, \delta) \\ \bar{B}(\alpha, \varepsilon, \delta) &= \bigcup_{t_j \in \tilde{C}_j(\alpha, \varepsilon, \delta)} \{t_j\} \times \tilde{B}_{-j}(\alpha, t_j, \varepsilon, \delta) = \left(\tilde{C}_j(\alpha, \varepsilon, \delta) \times T_{-j}\right) \cap \tilde{B}(\alpha, \varepsilon, \delta)\end{aligned}$$

Measurability of all the sets defined above follows from standard arguments.

We now show that for every rational number α , we have $\mu_j(\hat{C}_j(\alpha)) = 0$. So suppose not. Fix the rational α for which it fails. Then there must be $\varepsilon > 0$ and $\delta > 0$ such that $\mu_j(\tilde{C}_j(\alpha, \varepsilon, \delta)) > 0$. For notational simplicity, we drop the arguments $\alpha, \varepsilon, \delta$ in the next step of the argument as they are fixed in this step.

Define p^* as follows. For $k \neq i, j$ and any t , $p_k^*(t) = p_k(t)$. Also, for any $t \notin \bar{A} \cup \bar{B}$ and all k , $p_k^*(t) = p_k(t)$. For $t \in \bar{A}$,

$$p_j^*(t) = p_j(t) - \varepsilon \mu_{-j}(\tilde{B}_{-j}(t_j)) \quad \text{and} \quad p_i^*(t) = p_i(t) + \varepsilon \mu_{-j}(\tilde{B}_{-j}(t_j)).$$

For $t \in \bar{B}$,

$$p_j^*(t) = p_j(t) + \varepsilon \mu_{-j}(\tilde{A}_{-j}(t_j)) \quad \text{and} \quad p_i^*(t) = p_i(t) - \varepsilon \mu_{-j}(\tilde{A}_{-j}(t_j)).$$

For $t \in \bar{A}$, we have $p_j(t) \geq \varepsilon$, while for $t \in \bar{B}$, $p_i(t) \geq \varepsilon$. Hence p^* satisfies non-negativity. Clearly, for any t , $\sum_k p_k^*(t) = \sum_k p_k(t)$, so p^* satisfies the constraint that the sum of the p 's is less than 1.

Obviously, for $k \neq i, j$, we have $\hat{p}_k^*(t_k) = \hat{p}_k(t_k) \geq \varphi_k$. So the lower bound constraint on $\hat{p}_k(t_k)$ holds for all t_k for all $k \neq i, j$. Clearly, for any t_j such that $p_j^*(t) \geq p_j(t)$ for all t_{-j} , we have $\hat{p}_j(t_j) \geq \varphi_j$. Otherwise, we have

$$\hat{p}_j^*(t_j) = \hat{p}_j(t_j) - \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \mu_{-j}(\tilde{A}(t_j)) + \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right).$$

So $\hat{p}_j^*(t_j) = \hat{p}_j(t_j)$. Hence $\hat{p}_j^*(t_j) \geq \varphi_j$ for all $t_j \in T_j$.

For any t_i such that $p_i^*(t) \geq p_i(t)$ for all t_{-i} , we have $\hat{p}_i(t_i) \geq \varphi_i$. So consider t_i such that $p_i^*(t) < p_i(t)$ for some t_{-i} . Then there must be t_{-i} such that $t \in \bar{B}$. Hence $\hat{p}_i(t_i) \geq \varphi_i + \varepsilon$. So

$$\begin{aligned} \hat{p}_i^*(t_i) &= \hat{p}_i(t_i) + \varepsilon \mu_{-i} \left(\{t_{-i} \mid (t_i, t_{-i}) \in \bar{A}\} \right) \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \\ &\quad - \varepsilon \mu_{-i} \left(\{t_{-i} \mid (t_i, t_{-i}) \in \bar{B}\} \right) \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \\ &\geq \hat{p}_i(t_i) - \varepsilon \\ &\geq \varphi_i + \varepsilon - \varepsilon = \varphi_i. \end{aligned}$$

Hence the lower bound constraint for \hat{p}_i also holds everywhere.

Finally, the change in the principal's payoff from switching to p^* from p is

$$\begin{aligned} &\int_{t \in \bar{A}} \left\{ \left[(t_i - c_i) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] - \left[(t_j - c_j) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] \right\} \mu(dt) \\ &\quad + \int_{t \in \bar{B}} \left\{ \left[-(t_i - c_i) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] + \left[(t_j - c_j) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] \right\} \mu(dt) \\ &= \int_{\tilde{C}_j} \left(\int_{\tilde{A}_{-j}(t_j)} \left\{ \left[(t_i - c_i) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] - \left[(t_j - c_j) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] \right\} \mu_{-j}(dt_{-j}) \right. \\ &\quad \left. + \int_{\tilde{B}_{-j}(t_j)} \left\{ \left[-(t_i - c_i) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] + \left[(t_j - c_j) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] \right\} \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j). \end{aligned}$$

Note that $(t_j - c_j) \varepsilon \mu_{-j}(\tilde{A}_{-j}(t_j))$ and $(t_j - c_j) \varepsilon \mu_{-j}(\tilde{B}_{-j}(t_j))$ are functions only of t_j , not

t_{-j} . Hence we can rewrite the above as

$$\begin{aligned}
& \int_{\tilde{C}_j} \left(- \left[(t_j - c_j) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] \int_{\tilde{A}_{-j}(t_j)} \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j) \\
& + \int_{\tilde{C}_j} \left(\left[(t_j - c_j) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] \int_{\tilde{B}_{-j}(t_j)} \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j) \\
& + \int_{\tilde{C}_j} \left(\int_{\tilde{A}_{-j}(t_j)} \left[(t_i - c_i) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j) \\
& - \int_{\tilde{C}_j} \left(\int_{\tilde{B}_{-j}(t_j)} \left[(t_i - c_i) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j)
\end{aligned}$$

The first two lines sum to zero. For the last two lines, recall that $t \in \bar{A}$ implies $t_i \geq \alpha + \delta$, while $t \in \bar{B}$ implies $t_i \leq \alpha - \delta$. Hence the last two lines sum to at least

$$\begin{aligned}
& \int_{\tilde{C}_j} \left(\int_{\tilde{A}_{-j}(t_j)} \left[(\alpha + \delta) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j) \\
& - \int_{\tilde{C}_j} \left(\int_{\tilde{B}_{-j}(t_j)} \left[(\alpha - \delta) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] \mu_{-j}(dt_{-j}) \right) \mu_j(dt_j) \\
& = \int_{\tilde{C}_j} \left[(\alpha + \delta) \varepsilon \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \right] \mu_j(dt_j) \\
& - \int_{\tilde{C}_j} \left[(\alpha - \delta) \varepsilon \mu_{-j} \left(\tilde{A}_{-j}(t_j) \right) \mu_{-j} \left(\tilde{B}_{-j}(t_j) \right) \right] \mu_j(dt_j) \\
& > 0.
\end{aligned}$$

Hence the payoff difference for the principal between p^* and p is strictly positive. Hence p could not have been optimal, a contradiction.

This establishes that for every rational α , $\mu_j(\hat{C}_j(\alpha)) = 0$.

To complete the proof, let

$$\hat{A}_j(\alpha) = \{t_j \in T_j \mid \mu_{-j}(\hat{A}_{-j}(\alpha, t_j)) = 0\}$$

and

$$\hat{B}_j(\alpha) = \{t_j \in T_j \mid \mu_{-j}(\hat{B}_{-j}(\alpha, t_j)) = 0\}.$$

It is easy to see that for any α , $\hat{A}_j(\alpha) \cup \hat{B}_j(\alpha) \cup \hat{C}_j(\alpha) = T_j$. Let

$$\begin{aligned}
A(\alpha) &= \bigcup_{t_j \in \hat{A}_j(\alpha)} \{t_j\} \times \hat{A}_{-j}(\alpha, t_j) \\
&= \{t \in T \mid t_i > \alpha, \hat{p}_i(t_i) > \varphi_i, \text{ and } p_j(t) > 0\} \cap \left[\hat{A}_j(\alpha) \times T_{-j} \right]
\end{aligned}$$

$$\begin{aligned}
B(\alpha) &= \bigcup_{t_j \in \hat{B}_j(\alpha)} \{t_j\} \times \hat{B}_{-j}(\alpha, t_j) \\
&= \{t \in T \mid t_i < \alpha, \hat{p}_i(t_i) > \varphi_i, \text{ and } p_i(t) > 0\} \cap \left[\hat{B}_j(\alpha) \times T_{-j} \right]
\end{aligned}$$

$$C(\alpha) = \bigcup_{t_j \in \hat{C}_j(\alpha)} \{t_j\} \times T_{-j}$$

and

$$D(\alpha) = A(\alpha) \cup B(\alpha) \cup C(\alpha).$$

Once again measurability of the sets just defined is straightforward.

Note that $\mu(A(\alpha)) = 0$, since

$$\begin{aligned}
\mu(A(\alpha)) &= \int_{\hat{A}_j(\alpha)} \mu_{-j}(A_{-j}(\alpha, t_j)) \mu_j(dt_j) \\
&= \int_{\hat{A}_j(\alpha)} \mu_{-j}(\hat{A}_{-j}(\alpha, t_j)) \mu_j(dt_j) \\
&= 0,
\end{aligned}$$

where the last equality follows from $\mu_{-j}(\hat{A}_{-j}(\alpha, t_j)) = 0$ for all $t_j \in \hat{A}_j(\alpha)$. Similarly, $\mu(B(\alpha)) = 0$. Also, $\mu(C(\alpha)) = \mu_j(\hat{C}_j(\alpha)) \mu_{-j}(T_{-j})$ which is 0 by the first step. Hence $\mu(D(\alpha)) = 0$.

Let $S = \bigcup_{\alpha \in \mathbb{Q}} D(\alpha)$ where \mathbb{Q} denotes the rationals. Clearly $\mu(S) = 0$.

To complete the proof, suppose that, contrary to our claim, there exists $t', t'' \in T \setminus S$ such that $p_j(t') > 0$, $\hat{p}_i(t'_i) > \varphi_i$, $t''_i < t'_i$, and $\hat{p}_i(t''_i) > \varphi_i$, but $p_i(t'_j, t''_i, t''_{-ij}) > 0$. Obviously, there exists a rational α such that $t''_i < \alpha < t'_i$. Hence $(t'_i, t'_{-ij}) \in \hat{A}_{-j}(\alpha, t'_j)$ and $(t''_i, t''_{-ij}) \in \hat{B}_{-j}(\alpha, t'_j)$. Since t' is not in S , we know that $t' \notin A(\alpha)$, implying that $t'_j \notin \hat{A}_j(\alpha)$. Similarly, since t'' is not in S , we have $t'' \notin B(\alpha)$, so $t'_j \notin \hat{B}_j(\alpha)$. Similarly, $t' \notin C(\alpha)$, implying $t'_j \notin \hat{C}_j(\alpha)$. But $\hat{A}_j(\alpha) \cup \hat{B}_j(\alpha) \cup \hat{C}_j(\alpha) = T_j$, a contradiction. ■

Abusing notation, define T to be a measure one subset of T' whose projections are measurable and such that for all $t', t'' \in T$ for which $t'_j = t''_j$, $p_j(t') > 0$, $\hat{p}_i(t'_i) > \varphi_i$, $t''_i < t'_i$, and $\hat{p}_i(t''_i) > \varphi_i$, we have $p_i(t'') = 0$.

Lemma 8. *There is a set of measure one T' such that if $\hat{p}_j(t_j) = \varphi_j$, $\hat{p}_i(t_i) > \varphi_i$, and*

$$\mu_i(\{t'_i \in T_i \mid t'_i < t_i \text{ and } \hat{p}_i(t'_i) > \varphi_i\}) > 0,$$

then $p_j(t) = 0$.

Proof. Let

$$T_i^* = \{t_i \in T_i \mid \hat{p}_i(t_i) > \varphi_i \text{ and } \mu_i(\{t'_i \mid \hat{p}_i(t'_i) > \varphi_i \text{ and } t'_i < t_i\}) > 0\}.$$

To see that T_i^* is measurable. note that

$$T_i^* = \tilde{T}_i^* \cap \{t_i \in T_i \mid \hat{p}_i(t_i) > \varphi_i\}$$

where

$$\tilde{T}_i^* = \{t_i \in T_i \mid \mu_i(\{t'_i \mid \hat{p}_i(t'_i) > \varphi_i \text{ and } t'_i < t_i\}) > 0\}.$$

Since \tilde{T}_i^* is an interval (i.e., $\hat{t}_i \in \tilde{T}_i^*$ and $t''_i > \hat{t}_i$ implies $t''_i \in \tilde{T}_i^*$), it is measurable. Hence T_i^* is the intersection of two measurable sets and so is measurable.

Suppose the claim of the lemma is not true. Then there exists $\varepsilon > 0$ such that $\mu(S) > 0$ and such that S has measurable projections where

$$S = \{t \in T \mid t_i \in T_i^*, \hat{p}_j(t_j) = \varphi_j, \hat{p}_i(t_i) > \varphi_i + \varepsilon, \text{ and } p_j(t) \geq \varepsilon\},$$

and where we use Lemma 4 and take an equal measure subset if necessary.

Since $\mu(S) > 0$, we must have $\mu_i(S_i) > 0$ and hence $\mu_i(T_i^*) > 0$ since $S_i \subseteq T_i^*$. Choose measurable sets $L_i, M_i, U_i \subset S_i$ such that the following hold. First, all three sets have strictly positive measure. Second, $\sup L_i < \inf M_i$ and $\sup M_i < \inf U_i$. (Think of U, M , and L as standing for “upper,” “middle,” and “lower” respectively.) Third, there is an $\varepsilon' > 0$ such that $\mu(\hat{S}) > 0$ where \hat{S} is defined as follows. Let

$$S'' = \bigcup_{t_i \in U_i} \{t_i\} \times \{t_{-i} \in T_{-i} \mid (t_i, t_{-i}) \in S\} = \{t \in T \mid t_i \in U_i, \hat{p}_j(t_j) = \varphi_j, \text{ and } p_j(t) \geq \varepsilon\}.$$

Clearly $\mu(S'') > 0$. By Lemma 4, there exists a positive measure set $\hat{S} \subset S''$ and a number $\varepsilon' > 0$ satisfying the following. First, \hat{S} has strictly positive measure fibers. That is, for all i and all t_i , $\mu_{-i}(\hat{S}(t_i)) > 0$. Second, the j fibers of \hat{S} have measure bounded below by ε' . That is, $\mu_{-j}(\hat{S}(t_j)) > \varepsilon'$.

Let $E = \{t \in T \mid p_i(t) > \varphi_i, t_i \in L_i\}$. Since $\hat{p}_i(t_i) > \varphi_i$ for all $t_i \in L_i \subset T_i^*$, E has strictly positive measure. By taking a subset if necessary, we know that for all k , the projections E_k on T_k have strictly positive measure, as do the projections on $-i$ and on $-\{i, j\}$. ($E_{-i}(t_i)$ denotes, as usual, the t_i fiber of E .)

Let $A = M_i \times E_{-i}$. Since $\mu_i(M_i) > 0$ and $\mu_{-i}(E_{-i}) > 0$, we see that $\mu(A) > 0$. Taking subsets if necessary, and using Lemma 4, we know that we can find an equal measure subset (also, abusing notation, denoted A) all of whose fibers have strictly positive measure and whose projections are measurable. We now show that $p_i(t) = 1$ for almost all $t \in A$.

To see this, suppose not. Then we have a positive measure set of such $t \in A$ with $p_i(t) < 1$. For all $t \in A$, we have $\hat{p}_i(t_i) > \varphi_i$. In light of Lemma 5, this implies $\sum_k p_k(t) = 1$. Therefore, there exists $k \neq i$ and a positive measure set $\hat{A} \subseteq A$ such that $p_k(t) > 0$ for all $t \in \hat{A}$.

But fix any $t' \in \hat{A}$. By construction, $t'_i \in M_i$ and $t'_{-i} \in E_{-i}(t''_i)$ for some $t''_i \in L_i$. Since $t'_i \in M_i$ and $t''_i \in L_i$, we have $t'_i > t''_i$, $\hat{p}_i(t_i) > \varphi_i$, and $\hat{p}_i(t''_i) > \varphi_i$. By definition of $E_{-i}(t''_i)$, we have $p_i(t'_i, t'_{-i}) > 0$. Finally, we have $p_k(t') > 0$. Letting $t'' = (t''_i, t'_{-i})$, we see that this is impossible given that we removed the set \bar{S} defined in Lemma 7 from T . Hence $p_i(t) = 1$ for all $t \in A$.

Let $B = M_i \times \hat{S}_j \times E_{-ij}$. Recall that $\mu_i(M_i) > 0$. Also, $\mu(\hat{S}) > 0$ implies $\mu_j(\hat{S}_j) > 0$. Finally, $\mu_{-ij}(E_{-ij}) > 0$. Hence $\mu(B) > 0$. Again, taking subsets if necessary, and using Lemma 4, we know that we can find an equal measure subset (also, abusing notation, denoted B) all of whose fibers have strictly positive measure and whose projections are measurable. We now show that for all $t \in B$, we have $p_j(t) = 1$.

To see this, suppose not. Just as before, Lemma 5 then implies that there exists $k \neq j$ and $\hat{B} \subseteq B$ such that for all $t \in \hat{B}$, $p_k(t) > 0$. First, we show that $k \neq i$. To see this, suppose to the contrary that $k = i$. Fix any $t'' \in \hat{B}$. By assumption, $p_i(t'') > 0$. By definition of B , $t''_i \in M_i$, so $\hat{p}_i(t''_i) > \varphi_i$. Also by definition of B , $t''_j \in \hat{S}_j$. So fix $t' \in \hat{S}$ such that $t'_j = t''_j$. By definition of \hat{S} , $t'_i \in U_i$, implying both $\hat{p}_i(t'_i) > \varphi_i$ and $t'_i > t''_i$ (as $t''_i \in M_i$). The definition of \hat{S} also implies $p_j(t') > 0$. Just as before, this contradicts the removal of \bar{S} from T . Hence $k \neq i$.

So fix any $t' \in \hat{B}$. By assumption, $p_k(t') > 0$. By definition of B , $t'_i \in M_i$, so $\hat{p}_i(t'_i) > \varphi_i$. Also, the definition of B implies that $t'_{-ij} \in E_{-ij}(t''_i)$ for some $t''_i \in L_i$. Hence there exists t''_j such that $(t''_j, t'_{-ij}) \in E_{-i}(t''_i)$. Let $t'' = (t''_i, t''_j, t'_{-ij})$. By construction, $t'_k = t''_k$. Also, since $t''_i \in L_i$, we have $\hat{p}_i(t''_i) > \varphi_i$ and $t'_i > t''_i$ (as $t'_i \in M_i$). Finally, by definition of $E_{-i}(t''_i)$, we have $p_i(t'') > 0$. Again, this contradicts the removal of \bar{S} from T . Hence for all $t \in B$, $p_j(t) = 1$.

Summarizing, for every $t \in A$, we have $p_i(t) = 1$ (and hence $p_j(t) = 0$) and $\hat{p}_i(t_i) \geq \varphi_i + \varepsilon$, while for almost every $t \in B$, we have $p_j(t) = 1$ and $\hat{p}_i(t_i) \geq \varphi_i + \varepsilon$.

For any $t'_j \in A_j$ and $t''_j \in B_j$, let

$$F_{-j}(t'_j, t''_j) = \{t_{-j} \in T_{-j} \mid p_j(t'_j, t_{-j}) > p_j(t''_j, t_{-j})\}.$$

Obviously, for every t_j and hence every $t_j \in A_j$, we have $\hat{p}_j(t_j) \geq \varphi_j$. For every $t_j \in B_j$, we have $t_j \in \hat{S}_j \subseteq S_j$, so $\hat{p}_j(t_j) = \varphi_j$. Hence for every $t'_j \in A_j$ and $t''_j \in B_j$, we have $\hat{p}_j(t'_j) \geq \hat{p}_j(t''_j)$ even though $p_j(t') = 0$ and $p_j(t'') = 1$ for all $t' \in A$, $t'' \in B$. Moreover, $B_{-j} = A_{-j} = M_i \times E_{-ij}$. Hence for every $t'_j \in A_j$ and $t''_j \in B_j$, we must have

$$\mu_{-j}(F_{-j}(t'_j, t''_j)) > 0.$$

By Lemma 5, the fact that $p_j(t'') = 1$ for $t'' \in B$ implies that $t'_j > c_j$ for all $t'' \in B$. Hence, by Lemma 5, for every (t'_j, t_{-j}) with $t'_j \in B_j$, we have $\sum_k p_k(t'_j, t_{-j}) = 1$. Thus for every $t_{-j} \in F_{-j}(t'_j, t''_j)$, there exists $k \neq j$ such that $p_k(t'_j, t_{-j}) > 0$.

Let

$$G = \left\{ (t'_j, t''_j, t_{-j}) \in T_j^{(1)} \times T_j^{(2)} \times T_{-j} \mid t'_j \in A_j, t''_j \in B_j, \text{ and } t_{-j} \in F_{-j}(t'_j, t''_j) \right\}$$

where we use the superscripts on T_j to distinguish the order of components. The argument above implies that according to the product measure $\mu = \mu_j \times \mu_j \times \mu_{-j}$, G is non-null, i.e., $\mu(G) > 0$. (Specifically, $\mu(G) = \int_{A_j} \int_{B_j} \mu_{-j}(F_{-j}(t'_j, t''_j)) \mu_j(dt'_j) \mu_j(dt''_j)$ which is strictly positive since for each (t'_j, t''_j) in the domain of integration $\mu_{-j}(F_{-j}(t'_j, t''_j)) > 0$ and the domains of integration has positive μ_j measure.) The argument above also showed that for every $(t'_j, t''_j, t_{-j}) \in G$, there exists k such that $p_k(t''_j, t_{-j}) > 0$. Therefore there exists k such that $\mu(G^k) > 0$ where

$$G^k = \left\{ (t'_j, t''_j, t_{-j}) \in A_j \times B_j \times T_{-j} \mid t_{-j} \in F_{-j}(t'_j, t''_j), \text{ and } p_k(t''_j, t_{-j}) > 0 \right\}.$$

So we can find $\hat{G}^k \subset G^k$ such that $\mu(\hat{G}^k) > 0$ and for all $(t'_j, t''_j, t_{-j}) \in \hat{G}^k$, we have (1) $p_j(t'_j, t_{-j}) > p_j(t''_j, t_{-j}) + \varepsilon''$, and (2) $p_k(t''_j, t_{-j}) > \varepsilon''$. Taking subsets if necessary, and using Lemma 4, we know that we can find an equal measure subset (also, abusing notation, denoted G^k) all of whose fibers have strictly positive measure and whose projections are measurable.

Now we define

$$\begin{aligned} \hat{C} &= \text{proj}_{T_j^{(2)} \times T_{-j}} \hat{G}^k \\ \hat{D} &= \text{proj}_{T_j^{(1)} \times T_{-j}} \hat{G}^k \\ \hat{A} &= A \cap \left[\text{proj}_{T_j^{(1)}} \hat{G}^k \times T_{-j} \right] = \left\{ t \in A \mid t_j \in \text{proj}_{T_j^{(1)}} \hat{G}^k \right\} \\ \hat{B} &= B \cap \left[\text{proj}_{T_j^{(2)}} \hat{G}^k \times T_{-j} \right] = \left\{ t \in B \mid t_j \in \text{proj}_{T_j^{(2)}} \hat{G}^k \right\} \\ \tilde{S} &= \hat{S} \cap \left[\text{proj}_{T_j^{(2)}} \hat{G}^k \times T_{-j} \right] = \left\{ t \in \hat{S} \mid t_j \in \text{proj}_{T_j^{(2)}} \hat{G}^k \right\} \end{aligned}$$

All the above defined sets are measurable with strictly positive measure.²⁷

²⁷For example, \hat{A} has strictly positive measure because we defined it to have fibers with strictly positive measure. Moreover, $\text{proj}_{T_j^{(1)}} \hat{G}^k$ is a subset of A_j with strictly positive measure. So the measure of \hat{A} is the integral over a strictly positive measure set of t_j 's (those in $\text{proj}_{T_j^{(1)}} \hat{G}^k$) of the measure of the j -fibers of A , which have strictly positive measure. The same argument applies to \hat{B} and to \tilde{S} (the latter since $\hat{S}_j = B_j$).

The following is a summary of the key facts about these sets. For every $t \in \hat{A}$, we have $p_i(t) = 1$ and $\hat{p}_i(t_i) \geq \varphi_i + \varepsilon$. For every $t \in \tilde{S}$, we have $p_j(t) \geq \varepsilon$. For every $t \in \hat{C}$, we have $p_k(t) \geq \varepsilon''$. For every $t \in \hat{D}$, we have $p_j(t) \geq \varepsilon$. Finally, $\hat{A}_j = \hat{D}_j$, $\tilde{S}_j = \hat{C}_j$, and $\hat{C}_k = \hat{D}_k$. (Also $\mu_{-j}(\hat{C}_{-j}) = \mu_{-j}(\hat{D}_{-j}) > 0$, $\mu(\hat{C}) > 0$, and $\mu(\hat{D}) > 0$. To see that $\hat{A}_j = \hat{D}_j$, note that $\hat{G}_{j(1)}^k \subset A_j$. Similarly, to see that $\tilde{S}_j = \hat{C}_j$, note that $\hat{G}_{j(2)}^k \subset \hat{S}_j = B_j$.)

For each $E \in \{\hat{A}, \tilde{S}, \hat{C}, \hat{D}\}$, define a function $z_E : T \rightarrow [0, 1]$ such the following holds (where, for notational simplicity, the subscripts of Z do not include the hats and tildes):

$$z_E = 0 \text{ iff } t \notin E \quad (9)$$

$$\forall t_j \in \hat{A}_j = \hat{D}_j, \quad \mathbb{E}_{t_{-j}}[z_A(t_j, t_{-j})] = \mathbb{E}_{t_{-j}}[z_D(t_j, t_{-j})] \quad (10)$$

$$\forall t_k \in \hat{C}_k = \hat{D}_k, \quad \mathbb{E}_{t_{-k}}[z_C(t_k, t_{-k})] = \mathbb{E}_{t_{-k}}[z_D(t_k, t_{-k})] \quad (11)$$

$$\forall t_j \in \tilde{S}_j = \hat{C}_j, \quad \mathbb{E}_{t_{-j}}[z_S(t_j, t_{-j})] = \mathbb{E}_{t_{-j}}[z_C(t_j, t_{-j})] \quad (12)$$

We show below that such functions exist. Note the following useful implication of the definitions. If we multiply both sides of the first equation by $\mu_j(t_j)$ and integrate over t_j , we obtain

$$\mathbb{E}_t[z_A(t)] = \mathbb{E}_t[z_D(t)].$$

Similarly,

$$\mathbb{E}_t[z_S(t)] = \mathbb{E}_t[z_C(t)].$$

$$\mathbb{E}_t[z_C(t)] = \mathbb{E}_t[z_D(t)].$$

Hence

$$\mathbb{E}_t[z_A(t)] = \mathbb{E}_t[z_S(t)].$$

We now use this fact to construct a mechanism that improves on p .

Define p^* as follows. For any $t \notin \hat{A} \cup \tilde{S} \cup \hat{C} \cup \hat{D}$, $p^*(t) = p(t)$. Similarly, for any $\ell \notin \{i, j, k\}$, we have $p_\ell^*(t) = p_\ell(t)$ for all t . Also,

$$\forall t \in \hat{A}, \quad p_i^*(t) = p_i(t) - \varepsilon z_A(t), \quad p_j^*(t) = p_j(t) + \varepsilon z_A(t), \quad \text{and} \quad p_k^*(t) = p_k(t)$$

$$\forall t \in \tilde{S}, \quad p_i^*(t) = p_i(t) + \varepsilon z_S(t), \quad p_j^*(t) = p_j(t) - \varepsilon z_S(t), \quad \text{and} \quad p_k^*(t) = p_k(t)$$

$$\forall t \in \hat{C}, \quad p_i^*(t) = p_i(t), \quad p_j^*(t) = p_j(t) + \varepsilon z_C(t), \quad \text{and} \quad p_k^*(t) = p_k(t) - \varepsilon z_C(t)$$

$$\forall t \in \hat{D}, \quad p_i^*(t) = p_i(t), \quad p_j^*(t) = p_j(t) - \varepsilon z_D(t), \quad \text{and} \quad p_k^*(t) = p_k(t) + \varepsilon z_D(t).$$

The key facts summarized above are easily seen to imply that $p_\ell^*(t) \geq 0$ for all ℓ and all t . Also, $\sum_\ell p_\ell^*(t) = \sum_\ell p_\ell(t)$, so the constraint that p^* sum to less than 1 is satisfied.

It is easy to see that the way we defined the z functions implies that $\hat{p}_j^*(t_j) = \hat{p}_j(t_j)$ for all t_j and $\hat{p}_k^*(t_k) = \hat{p}_k(t_k)$ for all t_k . Finally, note that $p_i^*(t) < p_i(t)$ only for $t_i \in \hat{A}_i$

and that such t_i have $\hat{p}_i(t_i) \geq \varphi_i + \varepsilon$. Hence for those t_i 's with $p_i^*(t_i, t'_{-i}) < p_i(t_i, t_{-i})$ for some t_{-i} , we have

$$\hat{p}_i^*(t_i) \geq \hat{p}_i(t_i) - \varepsilon \mathbf{E}_{t_{-i}}[z_A(t_{-i}, t_i)].$$

But the fact that $z_A(t) < 1$ for all t implies that the right-hand side is at least

$$\hat{p}_i(t_i) - \varepsilon \geq \varphi_i + \varepsilon - \varepsilon = \varphi_i.$$

Hence the constraint that $p_\ell^*(t_\ell) \geq \varphi_\ell$ holds for all t_ℓ and all ℓ . Therefore, p^* is feasible given φ .

Finally, note that the principal's payoff from p^* minus his payoff from p is

$$\begin{aligned} \mathbf{E}_{t_i} [(\hat{p}_i^*(t_i) - \hat{p}_i(t_i))(t_i - c_i)] &= \varepsilon \int_{\tilde{S}} z_S(t)(t_i - c_i) \mu(dt) - \varepsilon \int_{\hat{A}} z_A(t)(t_i - c_i) \mu(dt) \\ &> \varepsilon(\inf U_i - c_i) \mathbf{E}[z_S(t)] - \varepsilon(\sup M_i - c_i) \mathbf{E}[z_A(t)] \\ &= \varepsilon \mathbf{E}[z_S(t)](\inf U_i - \sup M_i), \end{aligned}$$

where the first inequality follows from the fact that $t_i \in \tilde{S}_i$ implies $t_i \in U_i$ and $t_i \in \hat{A}_i$ implies $t_i \in M_i$ and the last equality from $\mathbf{E}[z_S(t)] = \mathbf{E}[z_A(t)]$. Recall that $\inf U_i > \sup M_i$, so the expression above is strictly positive. Hence if such z functions exist, p could not have been optimal.

To conclude, we show that for each $E \in \{\hat{A}, \tilde{S}, \hat{C}, \hat{D}\}$, z_E functions exist that satisfy equations (9), (10), (11), and (12).

Fix $\delta < 1$ and define functions as follows:

$$\begin{aligned} g(t_j) &= \delta \mu_{-j}(\hat{A}_{-j}(t_j)) \\ z_A(t_j) &= \delta \int_{\hat{D}_{-j}(t_j)} [\mu_{-k}(\hat{C}_{-k}(t_k))] \mu_{-j}(dt_{-j}) \\ z_C(t_k) &= \int_{\hat{D}_{-k}(t_k)} g(t_j) \mu_{-k}(dt_{-k}) \\ z_D(t_k, t_j) &= g(t_j) \mu_{-k}(\hat{C}_{-k}(t_k)) \\ z_S(t_j) &= \frac{\int_{\hat{C}_{-j}(t_j)} z_C(t_k) \mu_{-j}(dt_{-j})}{\mu_{-j}(\tilde{S}_{-j}(t_j))} \end{aligned}$$

where we recall that for any event S , we let $S_{-\ell}(t_\ell) = \{t_{-\ell} \in T_{-\ell} \mid (t_\ell, t_{-\ell}) \in S\}$, the t_ℓ -fiber of S . For any $\delta < 1$, it is obvious that z_A , z_C , and z_D take values in $[0, 1)$. Regarding z_S , if $\mu(\tilde{S}(t_j))$ is bounded away from above zero, then for $\delta \leq \inf_{t_j \in \tilde{S}_j} \mu_{-j}(\tilde{S}_{-j}(t_j))$, we have $z_S \in [0, 1)$. As discussed above, $\inf_{t_j \in \tilde{S}_j} \mu_{-j}(\tilde{S}_{-j}(t_j)) > \varepsilon$ so we can find such a δ .

We now verify equations (10), (11), and (12). First, consider equation (10). Note that

$$\begin{aligned} \mathbb{E}_{t-j}[z_A(t_j, t-j)] &= \int_{\hat{A}_{-j}(t_j)} z_A(t_j) \mu_{-j}(dt_{-j}) \\ &= z_A(t_j) \mu_{-j}(\hat{A}_{-j}(t_j)) \\ &= \delta \mu_{-j}(\hat{A}_{-j}(t_j)) \int_{\hat{D}_{-j}(t_j)} \mu_{-k}(\hat{C}_{-k}(t_k)) \mu_{-j}(dt_{-j}) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}_{t-j}[z_D(t_j, t-j)] &= \int_{\hat{D}_{-j}(t_j)} g(t_j) \mu_{-k}(\hat{C}_{-k}(t_k)) \mu_{-j}(dt_{-j}) \\ &= \delta \mu_{-j}(\hat{A}_{-j}(t_j)) \int_{\hat{D}_{-j}(t_j)} \mu_{-k}(\hat{C}_{-k}(t_k)) \mu_{-j}(dt_{-j}), \end{aligned}$$

where in both sets of equalities the main step is taking terms outside the integral when they do not depend on the variable of integration. Thus (10) holds.

Second, consider equation (11). Note that

$$\begin{aligned} \mathbb{E}_{t-k}[z_C(t_k, t-k)] &= z_C(t_k) \int_{\hat{C}_{-k}(t_k)} \mu_{-k}(dt_{-k}) \\ &= \left[\int_{\hat{D}_{-k}(t_k)} g(t_j) \mu_{-k}(dt_{-k}) \right] \left[\mu_{-k}(\hat{C}_{-k}(t_k)) \right] \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}_{t-k}[z_D(t_k, t-k)] &= \int_{\hat{D}_{-k}(t_k)} z_D(t_k, t-k) \mu_{-k}(dt_{-k}) \\ &= \int_{\hat{D}_{-k}(t_k)} g(t_j) \mu_{-k}(\hat{C}_{-k}(t_k)) \mu_{-k}(dt_{-k}) \\ &= \mu_{-k}(\hat{C}_{-k}(t_k)) \int_{\hat{D}_{-k}(t_k)} g(t_j) \mu_{-k}(dt_{-k}). \end{aligned}$$

Thus (11) holds.

Finally, consider equation (12). We have

$$\mathbb{E}_{t-j}[z_C(t_j, t-j)] = \int_{\hat{C}_{-j}(t_j)} z_C(t_k) \mu_{-j}(dt_{-j})$$

and

$$\begin{aligned}
\mathbb{E}_{t_{-j}}[z_S(t_j, t_{-j})] &= \int_{\tilde{S}_{-j}(t_j)} z_S(t_j) \mu_{-j}(dt_{-j}) = z_S(t_j) \int_{\tilde{S}_{-j}(t_j)} \mu_{-j}(dt_{-j}) \\
&= \frac{\int_{\hat{C}_{-j}(t_j)} z_C(t_k) \mu_{-j}(dt_{-j})}{\mu_{-j}(\tilde{S}_{-j}(t_j))} \int_{\tilde{S}_{-j}(t_j)} \mu_{-j}(dt_{-j}) \\
&= \int_{\hat{C}_{-j}(t_j)} z_C(t_k) \mu_{-j}(dt_{-j}).
\end{aligned}$$

Thus (12) holds. \blacksquare

Lemma 9. For any i ,

$$\mu_i(\{t_i \in T_i \mid \hat{p}_i(t_i) = \varphi_i\}) > 0.$$

Proof. Clearly if $\varphi_i = 1$, the result holds, so assume $\varphi_i < 1$.

Suppose the claim is false. Recall that the principal's objective function is

$$\sum_i \{\mathbb{E}_{t_i}[\hat{p}_i(t_i)(t_i - c_i)] + \varphi_i c_i\}$$

and that at the optimal solution $\varphi_i = \inf_{t_i} \hat{p}_i(t_i)$.

If $\mu_i(\{t_i \mid \hat{p}_i(t_i) = \varphi_i\}) = 0$, then for any $\delta > 0$, there is an $\varepsilon > 0$ such that

$$\mu_i(\{t_i \mid \hat{p}_i(t_i) < \varphi_i + \varepsilon\}) < \delta.$$

To see this, fix a sequence ε_n converging to 0 and define

$$A_n = \{t_i \mid \hat{p}_i(t_i) < \varphi_i + \varepsilon_n\},$$

$$A_0 = \{t_i \mid \hat{p}_i(t_i) = \varphi_i\},$$

and let $\delta_n = \mu_i(A_n)$. Then $A_n \downarrow A_0$ and $\mu_i(A_0) = 0$ by assumption, so $\delta_n \downarrow 0$. Hence for any $\delta > 0$, find n such that $\delta_n < \delta$ and choose $\varepsilon = \varepsilon_n$ to get the desired property.

So given any $\delta \in (0, 1)$ and the corresponding ε , let $A_i^{\delta, \varepsilon} = \{t_i \mid \hat{p}_i(t_i) < \varphi_i + \varepsilon\}$. Choose δ small enough so that $\varphi_i + \varepsilon < 1 - I\sqrt{\delta}$. (This is possible since $\varphi_i < 1$.) So for each $t_i \in A_i^{\delta, \varepsilon}$, we have

$$\int_{T_{-i}} p_i(t_i, t_{-i}) \mu_{-i}(dt_{-i}) < 1 - I\sqrt{\delta}.$$

By hypothesis, $\hat{p}_i(t_i) > \varphi_i$ with probability 1. Hence by Lemma 5, we have $\sum_k p_k(t) = 1$ with probability 1. Therefore, for each t_i with $\hat{p}_i(t_i) < \varphi_i + \varepsilon$, there exists $k = k^{t_i, \delta, \varepsilon} \neq i$

and $V_k^{t_i, \delta, \varepsilon} \subseteq T_{-i}$ with $p_k(t_i, t_{-i}) \geq \sqrt{\delta}$ for all $t_{-i} \in V_k^{t_i, \delta, \varepsilon}$ and $\mu_{-i}(V_k^{t_i, \delta, \varepsilon}) \geq \sqrt{\delta}$. Choose a subset of $V_k^{t_i, \delta, \varepsilon}$ with measure $\sqrt{\delta}$ and for simplicity denote it by $V_k^{t_i, \delta, \varepsilon}$.

Let $\eta = \min\{\sqrt{\delta}, \varepsilon\}$. Increase φ_i by $\eta\sqrt{\delta}$. This change increases the value of the objective function by $c_i\eta\sqrt{\delta}$. However, this may violate the constraint that $\hat{p}_i(t_i) \geq \varphi_i$ for all t_i . Clearly, this can only occur for t_i such that $\hat{p}_i(t_i) < \varphi_i + \eta\sqrt{\delta}$. By our choice of η , such t_i satisfy $\hat{p}_i(t_i) < \varphi_i + \varepsilon\sqrt{\delta} < \varphi_i + \varepsilon$ as $\delta < 1$. So for all t_i such that $\hat{p}_i(t_i) < \varphi_i + \varepsilon$ and all $t_{-i} \in V_k^{t_i, \delta, \varepsilon}$, increase $p_i(t_i, t_{-i})$ by η and decrease $p_k(t_i, t_{-i})$ by η . Since $\mu_{-i}(V_k^{t_i, \delta, \varepsilon}) = \sqrt{\delta}$, this change increases $\hat{p}_i(t_i)$ by $\eta\sqrt{\delta}$. Hence we again have $\hat{p}_i(t_i) \geq \varphi_i$ for all t_i after the change.

However, the reduction in p_k may have violated the constraint $\hat{p}_k(t_k) \geq \varphi_k$ for all t_k . Hence we increase φ_k by $\eta\delta$. To see that this will ensure the constraint is satisfied, note that p_k was reduced only for t_i such that $\hat{p}_i(t_i) < \varphi_i + \varepsilon$, a set with probability less than δ . Hence for any t_k , the reduction in $\hat{p}_k(t_k)$ must be less than $\eta\delta$. After this change, the resulting p and φ tuples satisfy feasibility.

To see that the objective function has increased as a result, recall that the gain from the increase in φ_i is $c_i\eta\sqrt{\delta}$. Similar reasoning shows that the loss from decreasing φ_k is $c_k\eta\delta$. Finally, the reduction in p_k and the corresponding increase in p_i generates a loss of no more than $\eta\delta\sqrt{\delta}[(\bar{t}_k - c_k) - (\underline{t}_i - c_i)]$ since the measure of the set of t 's for which we make this change is less than $\delta\sqrt{\delta}$. Hence the objective function increases if

$$c_i\eta\sqrt{\delta} > c_k\eta\delta + \eta\delta\sqrt{\delta}[(\bar{t}_k - c_k) - (\underline{t}_i - c_i)],$$

which must hold for δ sufficiently small. ■

Recall that

$$T_i^* = \{t_i \in T_i \setminus \bar{S}_i \mid \hat{p}_i(t_i) > \varphi_i \text{ and } \mu_i(\{t'_i \mid \hat{p}_i(t'_i) > \varphi_i \text{ and } t'_i < t_i\}) > 0\}.$$

Lemma 10. *There exists v^* such that for all i ,*

$$T_i^* = \{t_i \in T_i \mid t_i - c_i > v^*\}$$

up to sets of measure zero.

Proof. First, we show that for every i and j , we have $\mu_{ij}(E_{ij}) = 0$ where

$$E_{ij} = \{(t_i, t_j) \mid t_i - c_i > t_j - c_j, \hat{p}_i(t_i) = \varphi_i, \text{ and } t_j \in T_j^*\}.$$

To see this, suppose to the contrary that $\mu_{ij}(E_{-ij}) > 0$. Clearly, this implies $T_j^* \neq \emptyset$. Let

$$F_{-ij} = \prod_{k \neq i, j} \{t_k \in T_k \mid \hat{p}_k(t_k) = \varphi_k\}$$

and let $S = E_{ij} \times F_{-ij}$. Then $\mu(S) > 0$ by Lemma 9.

By Lemma 8, the fact that $t_j \in T_j^*$ and that $\hat{p}_k(t_k) = \varphi_k$ for all $k \neq j$ implies that up to sets of measure zero, we must have $p_k(t) = 0$ for all $k \neq j$. However, by Lemma 6, the fact that $t_i - c_i > t_j - c_j$ and $\hat{p}_j(t_j) > \varphi_j$ implies that up to sets of measure zero, we have $p_j(t) = 0$. So $\sum_k p_k(t) = 0$ for almost all $t \in E \times F$, contradicting Lemma 5.

We now show that this implies that for all i and j such that $T_i^* \neq \emptyset$ and $T_j^* \neq \emptyset$, we have

$$\inf T_i^* - c_i = \inf T_j^* - c_j.$$

Without loss of generality, assume $\inf T_i^* - c_i \geq \inf T_j^* - c_j$. Suppose that there is a positive measure set of $t_i \in T_i$ such that $t_i > \inf T_i^*$ but $t_i \notin T_i^*$. Hence for each such t_i , we must have $\hat{p}_i(t_i) = \varphi_i$. By definition of the infimum, for every $r > \inf T_j^*$, there exists $t_j \in T_j^*$ such that $r > t_j \geq \inf T_j^*$. By definition of T_j^* , the measure of such t_j 's must be strictly positive since $t_j \in T_j^*$ implies that there is a positive measure set of $t'_j < t_j$ with $t'_j \in T_j^*$. But then $\mu_{ij}(E_{ij}) > 0$, a contradiction. Hence, up to sets of measure zero, $t_i > \inf T_i^*$ implies $\hat{p}_i(t_i) > \varphi_i$.

By Lemma 9, then, we must have $\inf T_i^* > \underline{t}_i$. So suppose, contrary to our claim, that $\inf T_i^* - c_i > \inf T_j^* - c_j$. Then the set of t_i such that $\inf T_i^* - c_i > t_i - c_i > \inf T_j^* - c_j$ and $\hat{p}_i(t_i) = \varphi_i$ has strictly positive probability. The same reasoning as in the previous paragraph shows that $\mu_{ij}(E_{ij}) > 0$, a contradiction.

In light of this, we can specify v^* such that the claim of the lemma holds. First, if $T_i^* = \emptyset$ for all i , then set $v^* \geq \max_i(\bar{t}_i - c_i)$. Obviously, the lemma holds in this case.

Otherwise, let $v^* = \inf T_i^* - c_i$ for any i such that $T_i^* \neq \emptyset$. From the above, we see that v^* is well-defined. Let \mathcal{I}^N denote the set of i with $T_i^* \neq \emptyset$ and \mathcal{I}^E the set of i with $T_i^* = \emptyset$. By assumption, $\mathcal{I}^N \neq \emptyset$.

First, we show that for this specification of v^* , the claim of the lemma holds for all $i \in \mathcal{I}^E$. To see this, suppose to the contrary that for some $i \in \mathcal{I}^E$, we have $\bar{t}_i - c_i > v^*$. Then there is a positive measure set of t such that $t_j \in T_j^*$ for all $j \in \mathcal{I}^N$ and $t_i - c_i > t_j - c_j$ for all $j \in \mathcal{I}^N$ and some $i \in \mathcal{I}^E$. Then Lemma 6 implies $p_j = 0$ for all $j \in \mathcal{I}^N$, Lemma 8 implies $p_i = 0$ for all $i \in \mathcal{I}^E$, and Lemma 5 implies $\sum_i p_i(t) = 1$, a contradiction. Hence for all $i \in \mathcal{I}^E$, we have $v^* \geq \bar{t}_i - c_i$.

To complete the proof, we show that the claim holds for all $i \in \mathcal{I}^N$. Fix any $i \in \mathcal{I}^N$. Obviously, up to sets of measure zero, $t_i \in T_i^*$ implies $t_i - c_i > \inf T_i^* - c_i$, so

$$T_i^* \subseteq \{t_i \in T_i \mid t_i - c_i > v^*\}.$$

To prove the converse, suppose to the contrary that there is a positive measure set of t_i such that $t_i - c_i > v^*$ and $t_i \notin T_i^*$. Hence there must be a positive measure set of t_i such

that $t_i > \inf T_i^*$ and $\hat{p}_i(t_i) = \varphi_i$. To see why, recall that $v^* = \inf T_i^* - c_i$, so $t_i - c_i > v^*$ is equivalent to $t_i > \inf T_i^*$. Also, T_i^* is the set of points that have $\hat{p}_i(t_i) > \varphi_i$ and a positive measure of smaller points also satisfying this. So if $t_i \notin T_i^*$ but does have $\hat{p}_i(t_i) > \varphi_i$, it must be that the set of smaller points satisfying this has zero measure. Hence there is a zero measure of such t_i . Hence if there's a positive measure set of points outside T_i^* , a positive measure of them have $\hat{p}_i(t_i) = \varphi_i$. Let \hat{T}_i denote this set.

If there is some $j \neq i$ with $T_j^* \neq \emptyset$, the same argument as above implies that $\mu(E_{ij}) > 0$, a contradiction. Hence we must have $T_j^* = \emptyset$ for all $j \neq i$. Hence $\hat{p}_j(t_j) = \varphi_j$ with probability 1 for all $j \neq i$. Hence Lemma 8 implies that for all $t_i \in T_i^*$, we have $p_j(t) = 0$ for $j \neq i$ for almost all t_{-i} . By Lemma 5, then $p_i(t) = 1$ for all $t_i \in T_i^*$ and almost all t_{-i} .

By definition, for $t_i \in \hat{T}_i$, we have $\hat{p}_i(t_i) = \varphi_i < 1$.²⁸ Note that $t_i \in \hat{T}_i$ implies that t_i is larger than some $t'_i \in T_i^*$. Since $t'_i \in T_i^*$ implies $\hat{p}_i(t'_i) > \varphi_i$, Lemma 5 implies $t'_i > c_i$ and hence $t_i > c_i$. Hence Lemma 5 implies that for almost every $t_i \in \hat{T}_i$ and almost every t_{-i} , we have $\sum_j p_j(t) = 1$.

This implies that for every $t_i \in \hat{T}_i$, there exists $\hat{T}_{-i}(t_i) \subseteq T_{-i}$ and $j \neq i$, such that $p_j(t_i, t_{-i}) \geq (1 - \varphi_i)/(I - 1)$ for all $t_{-i} \in \hat{T}_{-i}(t_i)$. To see this, suppose not. Then there is some $t_i \in \hat{T}_i$ such that for every t_{-i} we have $p_j(t_i, t_{-i}) < (1 - \varphi_i)/(I - 1)$. But then $\sum_{j \neq i} p_j(t_i, t_{-i}) < 1 - \varphi_i$. Recall that $\sum_j p_j(t) = 1$ for all $t_i \in \hat{T}_i$ and all t_{-i} . Hence $p_i(t_i, t_{-i}) > \varphi_i$ for all t_{-i} , so $\hat{p}_i(t_i) > \varphi_i$, contradicting $t_i \in \hat{T}_i$. Since I is finite, this implies that there exists $j \neq i$, a positive measure subset of \hat{T}_i , say \hat{T}'_i , and a positive measure subset of T_{-i} , say \hat{T}'_{-i} , such that for every $t \in \hat{T}'_i \times \hat{T}'_{-i}$, we have $p_j(t) \geq (1 - \varphi_i)/(I - 1)$.

Fix any $t'_i \in \hat{T}'_i$ such that $\mu_i(\{t_i \in \hat{T}'_i \mid t_i > t'_i\}) > 0$. It is easy to see that such t'_i must exist. Since $t'_i > \inf T_i^*$, it must also be true that $\mu_i(\{t_i \in T_i^* \mid t_i < t'_i\}) > 0$. Given this, for any sufficiently small $\varepsilon > 0$, we have

$$\begin{aligned} \mu_i(\{t_i \in \hat{T}'_i \mid t_i \geq t'_i + \varepsilon\}) &> 0 \\ \mu_i(\{t_i \in T_i^* \mid t_i \leq t'_i - \varepsilon\}) &> 0. \end{aligned}$$

Choose any such $\varepsilon \in (0, (1 - \varphi_i)/(I - 1))$.

Taking subsets if necessary, then, we obtain two sets, $S^1 \subseteq \hat{T}'_i$ and $S^2 \subseteq T_i^*$ satisfying the following. First, $\mu_i(S^1) = \mu_i(S^2) > 0$. Second, $t_i \in S^1$ implies $t_i \geq t'_i + \varepsilon$ and $t_i \in S^2$ implies $t_i \leq t'_i - \varepsilon$.

Define p^* as follows. For any $t \notin (S^1 \cup S^2) \times \hat{T}'_{-i}$, $p^*(t) = p(t)$. For any $k \neq i, j$, $p_k^*(t) = p_k(t)$ for all t . For $t \in S^1 \times \hat{T}'_{-i}$,

$$p_j^*(t) = p_j(t) - \varepsilon \quad \text{and} \quad p_i^*(t) = p_i(t) + \varepsilon.$$

²⁸If $\varphi_i = 1$, then $T_i^* = \emptyset$ which contradicts our assumption.

For $t \in S^2 \times \hat{T}'_{-i}$,

$$p_j^*(t) = \varepsilon \quad \text{and} \quad p_i^*(t) = 1 - \varepsilon.$$

Recall that $S^2 \subseteq T_i^*$ and that $p_i(t) = 1$ for almost all $t_i \in T_i^*$ and $t_{-i} \in T_{-i}$. Hence this is equivalent to $p_j^*(t) = p_j(t) + \varepsilon$ and $p_i^*(t) = p_i(t) - \varepsilon$. Recall that $\varepsilon < (1 - \varphi_i)/(I - 1) \leq p_j(t)$ for all $t \in S^1 \times \hat{T}'_{-i}$ and that $\varepsilon < 1$, so we have $p_k^*(t) \geq 0$ for all k and t . Also, $\sum_k p_k^*(t) = \sum_k p_k(t)$, so the constraint that the p_k 's sum to less than one is satisfied. For any $k \neq i, j$, we have $\hat{p}_k^*(t_k) = \hat{p}_k(t_k)$ for all k and t_k so for such k , the constraint that $\hat{p}_k^*(t_k) \geq \varphi_k$ obviously holds.

For any t_j , either $\hat{p}_j^*(t_j) = \hat{p}_j(t_j)$ or

$$\hat{p}_j^*(t_j) = \hat{p}_j(t_j) - \varepsilon \mu_{-j}(S^1 \times \hat{T}'_{-ij}) + \varepsilon \mu_{-j}(S^2 \times \hat{T}'_{-ij}),$$

where \hat{T}'_{-ij} is the projection of \hat{T}'_{-i} on T_{-ij} . But $\mu_i(S^1) = \mu_i(S^2)$, implying $\hat{p}_j^*(t_j) = \hat{p}_j(t_j) \geq \varphi_j$ for all t_j .

For any t_i , either $\hat{p}_i^*(t_i) \geq \hat{p}_i(t_i)$ or

$$\hat{p}_i^*(t_i) = 1 - \varepsilon \mu_{-i}(\hat{T}'_{-i}) > 1 - \varepsilon.$$

By construction, $\varepsilon < (1 - \varphi_i)/(I - 1) \leq 1 - \varphi_i$. Hence $1 - \varepsilon > \varphi_i$. Hence we have $\hat{p}_i^*(t_i) \geq \varphi_i$ for all t_i . So p^* is feasible given φ .

Finally, the change in the principal's payoff in moving from p to p^* is

$$\mu(S^1)\varepsilon [\mathbb{E}(t_i - c_i \mid t_i \in S^1) - \mathbb{E}(t_i - c_i \mid t_i \in S^2)] \geq 2\mu(S^1)\varepsilon^2 > 0.$$

Hence p was not optimal, a contradiction. ■

To see that this proves Theorem 4, let v^* be the threshold. By Lemma 10, if some i has $t_i - c_i > v^*$, then that i satisfies $\hat{p}_i(t_i) > \varphi_i$. By Lemma 6, if there is more than one such i , then only the i with the largest value (i.e., $t_i - c_i$) has a positive probability of getting the good. By Lemma 8, no j with $t_j - c_j < v^*$ has any probability of getting the good. Since $\hat{p}_i(t_i) > \varphi_i$, Lemma 5 implies that we must have $\sum_j p_j(t) = 1$. Hence if some i has $t_i - c_i > v^*$, the i with the largest such value gets the good with probability 1. If any i has $t_i - c_i < v^*$, then Lemma 10 implies that $\hat{p}_i(t_i) = \varphi_i$. Thus we have a threshold mechanism. ■

C Proof of Theorem 3

For this proof, it is useful to give an alternative definition of t_i^* . Note that we can rearrange the definition in equation (1) as

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

or

$$t_i^* = E[t_i | t_i \leq t_i^*] + \frac{c_i}{F_i(t_i^*)}. \quad (13)$$

For notational convenience, number the agents so that 1 is any i with $t_i^* - c_i = \max_j(t_j^* - c_j)$ and let 2 denote any other agent so $t_1^* - c_1 \geq t_2^* - c_2$. First, we show that the principal must weakly prefer having 1 as the favored agent at a threshold of $t_2^* - c_2$ to having 2 as the favored agent at this threshold. If $t_1^* - c_1 = t_2^* - c_2$, this argument implies that the principal is indifferent between having 1 and 2 as the favored agents, so we then turn to the case where $t_1^* - c_1 > t_2^* - c_2$ and show that it must always be the case that the principal strictly prefers having 1 as the favored agent at threshold $t_1^* - c_1$ to favoring 2 with threshold $t_2^* - c_2$, establishing the claim.

So first let us show that it is weakly better to favor 1 at threshold $t_2^* - c_2$ than to favor 2 at the same threshold. First, note that if any agent other than 1 or 2 reports a value above $t_2^* - c_2$, the designation of the favored agent is irrelevant since the good will be assigned to the agent with the highest reported value and this report will be checked. Hence we may as well condition on the event that all agents other than 1 and 2 report values below $t_2^* - c_2$. If this event has zero probability, we are done, so we may as well assume this probability is strictly positive. Similarly, if both agents 1 and 2 report values above $t_2^* - c_2$, the object will go to whichever reports a higher value and the report will be checked, so again the designation of the favored agent is irrelevant. Hence we can focus on situations where at most one of these two agents reports a value above $t_2^* - c_2$ and, again, we may as well assume the probability of this event is strictly positive.

If both agents 1 and 2 report values below $t_2^* - c_2$, then no one is checked under either mechanism. In this case, the good goes to the agent who is favored under the mechanism. So suppose 1's reported value is above $t_2^* - c_2$ and 2's is below. If 1 is the favored agent, he gets the good without being checked, while he receives the good with a check if 2 were favored. The case where 2's reported value is above $t_2^* - c_2$ and 1's is below is symmetric. For brevity, let $\hat{t}_1 = t_2^* - c_2 + c_1$. Note that 1's report is below the threshold iff $t_1 - c_1 < t_2^* - c_2$ or, equivalently, $t_1 < \hat{t}_1$. Given the reasoning above, we see

that under threshold $t_2^* - c_2$, it is weakly better to have 1 as the favored agent if

$$\begin{aligned}
& F_1(\hat{t}_1)F_2(t_2^*)\mathbb{E}[t_1 \mid t_1 \leq \hat{t}_1] + [1 - F_1(\hat{t}_1)]F_2(t_2^*)\mathbb{E}[t_1 \mid t_1 > \hat{t}_1] \\
& \quad + F_1(\hat{t}_1)[1 - F_2(t_2^*)] \{ \mathbb{E}[t_2 \mid t_2 > t_2^*] - c_2 \} \\
& \geq F_1(\hat{t}_1)F_2(t_2^*)\mathbb{E}[t_2 \mid t_2 \leq t_2^*] + [1 - F_1(\hat{t}_1)]F_2(t_2^*) \{ \mathbb{E}[t_1 \mid t_1 > \hat{t}_1] - c_1 \} \\
& \quad + F_1(\hat{t}_1)[1 - F_2(t_2^*)]\mathbb{E}[t_2 \mid t_2 > t_2^*].
\end{aligned} \tag{14}$$

If $F_1(\hat{t}_1) = 0$, then this equation reduces to

$$F_2(t_2^*)\mathbb{E}[t_1 \mid t_1 > \hat{t}_1] \geq F_2(t_2^*) \{ \mathbb{E}[t_1 \mid t_1 > \hat{t}_1] - c_1 \},$$

which must hold. If $F_1(\hat{t}_1) > 0$, then we can rewrite the equation as

$$\mathbb{E}[t_1 \mid t_1 \leq \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} - c_1 \geq \mathbb{E}[t_2 \mid t_2 \leq t_2^*] + \frac{c_2}{F_2(t_2^*)} - c_2. \tag{15}$$

From equation (13), the right-hand side of equation (15) is $t_2^* - c_2$. Hence we need to show

$$\mathbb{E}[t_1 \mid t_1 \leq \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} - c_1 \geq t_2^* - c_2. \tag{16}$$

Recall that $t_2^* - c_2 \leq t_1^* - c_1$ or, equivalently, $\hat{t}_1 \leq t_1^*$. Hence from equation (1), we have

$$\mathbb{E}(t_1) \geq \mathbb{E}[\max\{t_1, \hat{t}_1\}] - c_1.$$

A similar rearrangement to our derivation of equation (13) yields

$$\mathbb{E}[t_1 \mid t_1 \leq \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} \geq \hat{t}_1.$$

Hence

$$\mathbb{E}[t_1 \mid t_1 \leq \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} - c_1 \geq \hat{t}_1 - c_1 = t_2^* - c_2 + c_1 - c_1 = t_2^* - c_2,$$

implying equation (15). Hence as asserted, it is weakly better to have 1 as the favored agent with threshold $t_2^* - c_2$ than to have 2 as the favored agent with this threshold.

Suppose that $t_1^* - c_1 = t_2^* - c_2$. In this case, an argument symmetric to the one above shows that the principal weakly prefers favoring 2 at threshold $t_1^* - c_1$ to favoring 1 at the same threshold. Hence the principal must be indifferent between favoring 1 or 2 at threshold $t_1^* - c_1 = t_2^* - c_2$.

We now turn to the case where $t_1^* - c_1 > t_2^* - c_2$. The argument above is easily adapted to show that favoring 1 at threshold $t_2^* - c_2$ is strictly better than favoring 2 at this threshold if the event that $t_j - c_j < t_2^* - c_2$ for every $j \neq 1, 2$ has strictly positive probability. To see this, note that if this event has strictly positive probability, then the

claim follows iff equation (14) holds with a strict inequality. If $F_1(\hat{t}_1) = 0$, this holds iff $F_2(t_2^*)c_1 > 0$. By assumption, $c_i > 0$ for all i . Also, $\underline{t}_2 < t_2^*$, so $F_2(t_2^*) > 0$. Hence this must hold if $F_1(\hat{t}_1) = 0$. If $F_1(\hat{t}_1) > 0$, then this holds if equation (16) holds strictly. It is easy to use the argument above and $t_1^* - c_1 > t_2^* - c_2$ to show that this holds.

So if the event that $t_j - c_j < t_2^* - c_2$ for every $j \neq 1, 2$ has strictly positive probability, the principal strictly prefers having 1 as the favored agent to having 2. Suppose, then, that this event has zero probability. That is, there is some $j \neq 1, 2$ such that $t_j - c_j \geq t_2^* - c_2$ with probability 1. In this case, the principal is indifferent between having 1 as the favored agent at threshold $t_2^* - c_2$ versus favoring 2 at this threshold. However, we now show that the principal must strictly prefer favoring 1 with threshold $t_1^* - c_1$ to either option and thus strictly prefers having 1 as the favored agent.

To see this, recall from the proof of Theorem 2 that the principal strictly prefers favoring 1 at threshold $t_1^* - c_1$ to favoring him at a lower threshold v^* if there is a positive probability that $v^* < t_j - c_j < t_1^* - c_1$ for some $j \neq 1$. Thus, in particular, the principal strictly prefers favoring 1 at threshold $t_1^* - c_1$ to favoring him at $t_2^* - c_2$ if there is a $j \neq 1, 2$ such that the event $t_2^* - c_2 < t_j - c_j < t_1^* - c_1$ has strictly positive probability. By hypothesis, there is a $j \neq 1, 2$ such that $t_2^* - c_2 < t_j - c_j$ with probability 1, so we only have to establish that for this j , we have a positive probability of $t_j - c_j < t_1^* - c_1$. Recall that $\underline{t}_j - c_j < t_j^* - c_j$ by definition of t_j^* . By hypothesis, $t_j^* - c_j < t_1^* - c_1$. Hence we have $t_j - c_j < t_1^* - c_1$ with strictly positive probability, completing the proof. ■

D Comparative Statics Proof

In this appendix, we show the claim in the text regarding the effect of changes in the cost of checking the favored agent when $I = 2$ and $F_1 = F_2 = F$. For notational ease, let 1 be the favored agent. Then the probability 1 gets the good is

$$F(t_1^*)F(t_1^* - c_1 + c_2) + \int_{t_1^*}^{\bar{t}} F(t_1 - c_1 + c_2)f(t_1) dt_1.$$

Differentiating with respect to c_1 gives

$$\begin{aligned} f(t_1^*)F(t_1^* - c_1 + c_2) \frac{\partial t_1^*}{\partial c_1} + F(t_1^*)f(t_1^* - c_1 + c_2) \left[\frac{\partial t_1^*}{\partial c_1} - 1 \right] \\ - F(t_1^* - c_1 + c_2)f(t_1^*) \frac{\partial t_1^*}{\partial c_1} - \int_{t_1^*}^{\bar{t}} f(t_1 - c_1 + c_2)f(t_1) dt_1 \end{aligned}$$

or

$$F(t_1^*)f(t_1^* - c_1 + c_2) \left[\frac{\partial t_1^*}{\partial c_1} - 1 \right] - \int_{t_1^*}^{\bar{t}} f(t_1 - c_1 + c_2)f(t_1) dt_1.$$

Recall that t_1^* is defined by

$$\int_{\underline{t}}^{t_1^*} F(s) ds = c_1.$$

Using this, it's easy to see that

$$\frac{\partial t_1^*}{\partial c_1} = \frac{1}{F(t_1^*)}.$$

Substituting, the derivative is

$$\begin{aligned} f(t_1^* - c_1 + c_2)[1 - F(t_1^*)] - \int_{t_1^*}^{\bar{t}} f(t_1 - c_1 + c_2)f(t_1) dt_1 \\ = \int_{t_1^*}^{\bar{t}} [f(t_1^* - c_1 + c_2) - f(t_1 - c_1 + c_2)]f(t_1) dt_1. \end{aligned}$$

Hence if f is increasing throughout the relevant range, this is negative, implying that the probability 1 gets the good is decreasing in c_1 . If f is decreasing throughout the relevant range, this is positive, so 1's probability of getting the good increases in c_1 . If the types have a uniform distribution, the derivative is zero.