Optimal Allocation with Costly Verification[1]

Elchanan Ben-Porath[2] Eddie Dekel[3] Barton L. Lipman[4]

First Preliminary Draft
October 2011

Current Draft
March 2013

[1]We thank Ricky Vohra, Benjy Weiss, and numerous seminar audiences 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 this draft was in progress and Alex Poterack for proofreading.

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

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

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

Abstract

A principal (dean) has an object (job slot) to allocate to one of I agents (departments). Each agent has a strictly positive value for receiving the object. Each agent also has private information which determines the value to the principal of giving the object to him. There are no monetary transfers but the principal can check the value associated with any individual at a cost which may vary across individuals. We characterize the class of optimal Bayesian mechanisms, that is, mechanisms which maximize the expected value to the principal from his assignment of the good minus the costs of checking values. One particularly simple mechanism in this class, which we call the favored–agent mechanism, specifies a threshold value v^* and a favored 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 with probability 1 and receives the good iff her report is confirmed. We show that *all* optimal mechanisms are essentially randomizations over optimal favored–agent mechanisms.

1 Introduction

Consider the following environment: a principal has 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). We also show that the mechanism is, in a sense, *almost* a dominant strategy mechanism and consequently is ex post incentive compatible.[2]

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; this is the main goal of this paper. In our exploration of this option, we

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

[2] More precisely, truthful reporting is an optimal strategy for agent i given any profile of reporting strategies for the other agents.

take the opposite extreme position from the standard model and do not allow transfers. This obviously simplifies the problem, but we also find it reasonable to exclude transfers. Indeed, in many cases they are not used. In some situations, this may be because transfers have efficiency costs that are ignored in the standard approach. More specifically the monetary resources each agent has might matter to the principal, so changing the allocation of monetary resources in order to allocate a good might be costly. In other situations, the value to the principal of agent i getting the good may differ from the value to agent i of getting 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 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. Consider the problem of the head of an organization — say, a dean — who has an indivisible resource or good (say, a job slot) that can be allocated to any of several divisions (departments) within the organization (university). Naturally, the dean wishes to allocate this slot to that department which would fill the position in the way which best promotes the interests of the university as a whole. Each department, on the other hand, would like to hire in its own department and puts less, perhaps no, value on hires in other departments. In addition, the department may well value candidates differently than the dean. The department, of course, has much more information regarding the candidates' qualities and hence the value of the candidate to the dean. In such a situation, it is natural to assume that the head of the organization can obtain information that can confirm whether the department's claims are correct. Obtaining and/or processing such information is costly to the dean so this will also be taken into account.

Similar problems arise in areas other than organizational economics. For example, governments allocate various goods or subsidies which are intended not for those willing and able to pay the most but for those most in need. Hence allocation mechanisms based on auctions or similar approaches cannot achieve the government's goal, often leading to the use of mechanisms which rely instead on some form of verification.[3]

As another example, consider the problem of choosing which of a set of job applicants to hire for a job with a predetermined salary. Each applicant wants the job and presents claims about his qualifications for the job. The person in charge of hiring can verify these claims but doing so is costly.

[3]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. The matching literature also studies allocation problems where transfers are considered inappropriate and hence are not used.

Literature review. Townsend (1979) initiated the literature on the principal–agent model with costly state verification. 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. See also Gale and Hellwig (1985), Border and Sobel (1987), and Mookherjee and Png (1989). 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 more on general issues of mechanism design and implementation in these environments 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)).[4] 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 contains the characterization of the class of optimal mechanisms, showing all optimal mechanisms are essentially randomizations over optimal favored–agent mechanisms. Since these results show that we can restrict attention to favored–agent mechanisms, we turn in Section 4 to characterizing the set of best mechanisms in this class. In Section 5, we give comparative statics and some examples. In Section 6, we sketch the proof of our uniqueness result and discuss several other issues. Section 7 concludes. Proofs not contained in the text are in the Appendix.

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

2 Model

The set of agents is $\mathcal{I} = \{1, \ldots, I\}$. There is a single indivisible good to allocate among them. The value to the principal of assigning the object to agent i depends on information which is known only to i. Formally, the value to the principal of allocating the good to agent i is t_i where t_i is private information of agent i. The value to the principal of assigning the object to no one is normalized to zero.[5] We assume that the t_i's are independently distributed. The distribution of t_i has a strictly positive density f_i over the interval $T_i \equiv [\underline{t}_i, \bar{t}_i]$ where $0 \leq \underline{t}_i < \bar{t}_i < \infty$. (All results extend to allowing the support to be unbounded above.) 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). If the principal checks some agent, she learns that agent's type. The cost c_i is interpreted as the direct cost to the principal of reviewing the information provided plus the costs to the principal associated with the resource cost to the agent of being checked. The cost to the agent of providing information is assumed to be zero. To understand this, think of the agent's resources as allocated to activities which are either directly productive for the principal or which provide information for checking claims. The agent is indifferent over how these resources are used since they will all be used regardless. Thus by directing the agent to spend resources on providing information, the principal loses some output the agent would have produced with the resources otherwise while the agent's utility is unaffected.[6] In Section 6.3, we show one way to generalize our model to allow agents to bear some costs of providing evidence which does not change our results qualitatively.

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.[7] We also assume that each agent's 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

[5] The case where the value to the principal is $V > 0$ is equivalent to the case where his value is zero but there is an agent $I + 1$ with $t_{I+1} = V$ with probability 1. See Section 6.3 for more detail.

[6] One reason this assumption is a convenient simplification is that dropping it allows 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, just as in the literature on allocations without transfers but with costly signals discussed in the introduction. This both complicates the analysis and indirectly introduces a form of the transfers we wish to exclude.

[7] In other words, suppose we let the payoff of 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.

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 past outcome of checking, finally culminating in the allocation of the good, perhaps to no one. Without loss of generality, we can restrict attention to truth telling equilibria of mechanisms where each agent sends a report of his type to the principal who is committed to (1) a probability distribution over which agents (if any) are checked as a function of the reports and (2) a probability distribution over which agent (if any) receives the good as a function of the reports and the outcome of checking. While this does not follow directly from the usual Revelation Principle,[8] the argument is similar. Fix a dynamic mechanism, deterministic or otherwise, and any equilibrium, in pure or mixed strategies.

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 determines 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. If what he 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.

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 to reporting any other type is exactly the same as his payoff to deviating to that type's strategy 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.

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

[8] The usual version of the Revelation Principle does not apply to games with verification. See Townsend (1988) for discussion and an extension to a class of verification models.

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, it is not hard to see that 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. 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. So a mechanism is a $2I$ tuple of functions, $(p_i, q_i)_{i\in\mathcal{I}}$ where $p_i : T \to [0,1]$, $q_i : T \to [0,1]$, $\sum_i p_i(t) \leq 1$ 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

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

The incentive compatibility constraint for i is then

$$\mathrm{E}_{t_{-i}} p_i(t) \geq \mathrm{E}_{t_{-i}} \left[p_i(\hat{t}_i, t_{-i}) - q_i(\hat{t}_i, t_{-i}) \right], \quad \forall \hat{t}_i, t_i \in T_i, \quad \forall i \in \mathcal{I}.$$

Given a mechanism (p, q), let

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

and

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

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 more specific, first we show that there is always a favored–agent mechanism which is an optimal mechanism. Second, we show that every Bayesian 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.

To be more precise, 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 he is the favored agent or not) is checked and, if his report is verified, receives the good.

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.

Theorem 1. *There always exists a Bayesian optimal mechanism which is a favored–agent mechanism.*

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 this is a feasible mechanism.

Suppose the highest reported type is below c. Obviously, it's better for the principal to 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 selecting the "best" person to give the good to when all the reports are below c. To think more about this, suppose the principal gives the good to agent 1 if all reports are below c. Continue to assume that if any agent reports a type above c, then the principal checks the highest report and gives the good to this agent if the report is true. This mechanism is clearly incentive compatible. However, the principal can also achieve incentive compatibility and the same allocation of the good while saving on checking costs: he doesn't need to check 1's report when he is the only agent to report a type above c. To see why this cheaper mechanism is also incentive compatible, note that if everyone else's type is below c, 1 gets the good no matter what he says. Hence he 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 is far from a proof. Indeed, we prove this result as a corollary to the next theorem, a result whose proof is rather complex.

Let $\mathcal{F}$ denote the set of favored–agent mechanisms and let $\mathcal{F}^*$ denote the set of optimal favored–agent mechanisms. By Theorem 1, if a favored–agent mechanism is better for the principal than every other favored–agent mechanism, then it must be better for the principal than every other incentive compatible mechanism, whether in the favored–agent class or not. Hence every mechanism in $\mathcal{F}^*$ is an optimal mechanism even without the restriction to favored–agent mechanisms.

Given two mechanisms, (p^1, q^1) and (p^2, q^2) and a number $\lambda \in (0, 1)$, we can construct a new mechanism, say (p^λ, q^λ), by $(p^\lambda, q^\lambda) = \lambda(p^1, q^1) + (1 - \lambda)(p^2, q^2)$, where the right–hand side refers to the pointwise convex combination of these functions. The mechanism

(p^λ, q^λ) is naturally interpreted as a random choice by the principal between the mechanisms (p^1, q^1) and (p^2, q^2). It is easy to see that if (p^k, q^k) is incentive compatible for $k = 1, 2$, then (p^λ, q^λ) is incentive compatible. Also, the principal's payoff is linear in (p, q), so if both (p^1, q^1) and (p^2, q^2) are optimal for the principal, it must be true that (p^λ, q^λ) is optimal for the principal. It is easy to see that this implies that any mechanism in the convex hull of $\mathcal{F}^*$, denoted $\text{conv}(\mathcal{F}^*)$, is optimal.

Finally, as noted in Section 2, if a mechanism (p, q) is optimal, then any mechanism equivalent to it in the sense of having the same reduced form up to sets of measure zero must also be optimal. Hence Theorem 1 implies that any mechanism equivalent to a mechanism in $\text{conv}(\mathcal{F}^*)$ must be optimal. The following theorem shows the stronger result that this is precisely the set of optimal mechanisms.

Theorem 2. *A mechanism is optimal if and only if it is equivalent to some mechanism in* $\text{conv}(\mathcal{F}^*)$.

Section 6 contains a sketch of the proof of this result.

Theorem 2 says that all optimal mechanisms are, essentially, favored–agent mechanisms or randomization over such mechanisms. Hence we can restrict attention to favored–agent mechanisms without loss of generality. This result also implies that if there is a unique optimal favored–agent mechanism, then $\text{conv}(\mathcal{F}^*)$ is a singleton so that there is essentially a unique optimal 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

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

It is easy to show that t_i^* is well–defined.[9]

[9] To see this, note that the right–hand side of equation (1) is continuous and strictly increasing in t_i^* for $t_i^* \geq \underline{t}_i$, below the left–hand side at $t_i^* = \underline{t}_i$, and above it as $t_i^* \to \infty$. Hence there is a unique solution. Note also that if we allowed $c_i = 0$, we would have $t_i^* = \underline{t}_i$. 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.

It will prove useful to give two alternative definitions of t_i^*. Note that we can rearrange the definition above as

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

or

$$t_i^* = \mathrm{E}[t_i \mid t_i \leq t_i^*] + \frac{c_i}{F_i(t_i^*)}. \tag{2}$$

Finally, note that we could rearrange the next–to–last equation as

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

So a final equivalent definition of t_i^* is

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

Given any i, let $\mathcal{F}_i$ denote the set of favored–agent mechanisms with i as the favored agent.

Theorem 3. *The unique best mechanism in $\mathcal{F}_i$ 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

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

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 $\mathrm{E}(t_1)$.

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

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

But the right–hand side is $\mathrm{E}\max\{t_1, t_1^*\} - c_1$ which equals $\mathrm{E}(t_1)$ by our first definition of t_i^*. 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 $\mathcal{F}_i$. ▌

Given that the best mechanism in each $\mathcal{F}_i$ is unique, it remains only to characterize the optimal choice of i.

Theorem 4. *The optimal choice of the favored agent is any i with $t_i^* - c_i = \max_j(t_j^* - c_j)$.*

The proof is in Appendix C. Here we sketch the proof for the special case of two agents with equal verification costs, c.

From equation (2), if $t_i^* > t_j^*$, then

$$\frac{c}{F_i(t_i^*)} + \mathrm{E}[t_i \mid t_i \leq t_i^*] > \frac{c}{F_j(t_j^*)} + \mathrm{E}[t_j \mid t_j \leq t_j^*]$$

or

$$F_j(t_j^*)c + F_i(t_i^*)F_j(t_j^*)\mathrm{E}[t_i \mid t_i \leq t_i^*] > F_i(t_i^*)c + F_i(t_i^*)F_j(t_j^*)\mathrm{E}[t_j \mid t_j \leq t_j^*]$$

One can actually show the stronger result that for $t^* \in \{t_j^*, t_i^*\}$,

$$F_j(t^*)c + F_i(t^*)F_j(t^*)\mathrm{E}[t_i \mid t_i \leq t^*] > F_i(t^*)c + F_i(t^*)F_j(t^*)\mathrm{E}[t_j \mid t_j \leq t^*].$$

This in turn is equivalent to

$$c(1 - F_i(t^*))F_j(t^*) - c(1 - F_j(t^*))F_i(t^*) \tag{4}$$

$$+F_i(t^*)F_j(t^*)(\mathrm{E}[t_i \mid t_i \leq t^*] - \mathrm{E}[t_j \mid t_j \leq t^*]) \quad > \quad 0 \tag{5}$$

The first line is the savings in checking costs when using i versus j as the favored agent with threshold t^*. To see this, note that if $t_i > t^*$, $t_j < t^*$, and j is the favored agent, then i is checked. If $t_i < t^*$, $t_j > t^*$, and i is the default individual, then j is checked. Otherwise there is no difference in the checking policy between the case when

i is favored versus when j is favored. So switching from j to i saves the first term and costs the second, giving (4).

The second line compares the benefit when the good is given without checking to i instead of j. To see this, consider (5). $F_i(t^*)F_j(t^*)$ is the probability that both i and j are below the threshold, so the good is given without checking. $\mathrm{E}[t_i \mid t_i \leq t^*] - \mathrm{E}[t_j \mid t_j \leq t^*]$ is the difference in the expected values conditional on being below the threshold.

Since these two changes are the only effects of changing the identity of the favored agent from j to i, we see that the change strictly increases the principal's payoff. Hence i is the optimal favored agent if $t_i^* > t_j^*$.

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 $t_i^* - c_i = \max_j t_j^* - c_j$ and the threshold v^* satisfies $v^* = \max_j t_j^* - c_j$. Thus the set of optimal mechanisms is equivalent to picking a favored–agent mechanism with threshold $v^* = \max_j t_j^* - c_j$ and randomizing over which of the agents i with $t_i^* - c_i$ equal to this threshold to favor. Clearly for generic checking costs, there will be a unique i with $t_i^* - c_i = \max_j t_j^* - c_j$ 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.[10]

5 Comparative Statics and Examples

Our characterization of the optimal favored agent and threshold makes it easy to give comparative statics. Recall our third expression for t_i^* which is

$$\int_{\underline{t}_i}^{t_i^*} F_i(\tau)\, d\tau = c_i. \tag{6}$$

Hence 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 $\mathrm{E}(t_i) = \mathrm{E}\max\{t_i - c_i, v_i^*\}$. 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 favored agent is the one the principal checks least often, this makes it more desirable to make i the favored agent.

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 (6) for fixed t_i^*, so to maintain the

[10] We thank Yi-Chun Chen and Siyang Xiyong for showing us a proof of this result.

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.

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.[11] Nevertheless, an FOSD increase in any F_i does increase the principal's ex ante payoff. See Appendix D for proof.

Turning to the effect of changes in c_i, it is obvious that a decrease in c_i makes the principal better off as she 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 the report by i is $t_i - c_i$, so a higher c_i means that i's reports are all "worse." Hence he is less likely to receive the good and other agents are more likely to do so.

On the other hand, changes in the cost of 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 agent j's $t_j - c_j$ 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 example, if $I = 2$ and $F_1 = F_2 = F$, then the favored agent benefits from an increase in c_i if the density f is increasing and conversely if it is decreasing; see Appendix D. In the uniform example presented below, the two effects cancel out.

[11] For 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 is $t_1 - c_1$.

Finally, note that an increase in c_i increases $t_i^* - c_i$, so this can cause i to go from not being favored to being favored. Such a change necessarily increases i's payoff. It is not hard to see that the payoff to i increases discontinuously at the c_i which makes him the favored agent.

To illustrate the effects of changes in c_i further, continue to assume $I = 2$ and now suppose that $t_1, t_2 \sim U[0,1]$. It is easy to calculate t_i^*. From equation (1), we have

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

so

$$\frac{1}{2} = \int_0^{t_i^*} t_i^* \, ds + \int_{t_i^*}^1 s \, ds - c_i$$

or

$$\frac{1}{2} = (t_i^*)^2 + \frac{1}{2} - \frac{(t_i^*)^2}{2} - c_i$$

so

$$t_i^* = \sqrt{2c_i}.$$

This holds only if $c_i \leq 1/2$ so that $t_i^* \leq 1$. Otherwise, $\mathrm{E}\max\{t_i, t_i^*\} = t_i^*$, so $t_i^* = (1/2) + c_i$. Hence

$$t_i^* = \begin{cases} \sqrt{2c_i}, & \text{if } c_i \leq 1/2 \\ (1/2) + c_i, & \text{otherwise} \end{cases}$$

so

$$t_i^* - c_i = \begin{cases} \sqrt{2c_i} - c_i, & \text{if } c_i \leq 1/2 \\ 1/2, & \text{otherwise.} \end{cases}$$

It is easy to see that $\sqrt{2c_i} - c_i$ is an increasing function for $c_i \in (0, 1/2)$. Thus if $c_1 < 1/2$ and $c_1 < c_2$, we must have $t_2^* - c_2 > t_1^* - c_1$, so that 2 is the favored agent. If $c_2 \geq c_1 \geq 1/2$, then $t_1^* - c_1 = t_2^* - c_2 = 1/2$, so the principal is indifferent over which agent should be favored. Note that in this case, the cost of checking is so high that the principal never checks, so that the favored agent simply receives the good independent of the reports. Since the distributions of t_1 and t_2 are the same, it is not surprising that the principal is indifferent over who should be favored in this case.

In Figure 1 below, we show agent 2's expected payoff as a function of his cost c_2 given a fixed value of $c_1 < 1/2$. Note that when $c_2 > c_1$, so that 2 is the favored agent, 2's payoff is higher than when his cost is below c_1 where 1 is favored. That is, it is advantageous to be favored. Note that this implies that agents may have incentives to increase the cost of being checked in order to become favored, an incentive which is costly for the principal.

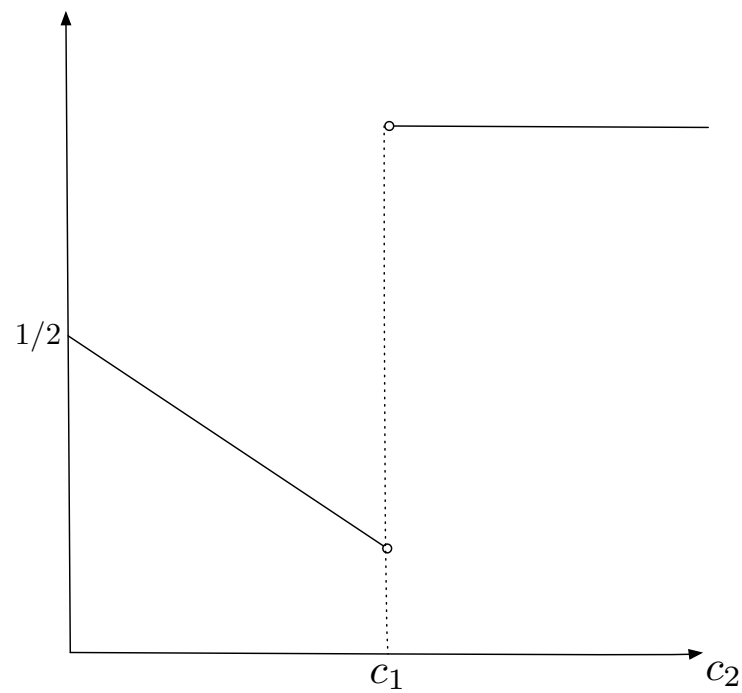

Figure 1

Note that, as discussed above, the effect of changes in the cost of checking the favored agent on the payoffs of the agents is ambiguous in general, so the fact that this payoff is constant for the uniform distribution is special. As noted above, the fact that the payoff to the agent who is not favored is decreasing in his cost of being checked (and hence the payoff to the other agent is increasing in this cost) plus the fact that i's payoff increases discontinuously at the cost that makes him favored are general properties.

6 Discussion

6.1 Proof Sketch

In this section, we sketch the proof of Theorem 2. It is easy to see that Theorem 1 is a corollary.

First, it is useful to rewrite the optimization problem as follows. Recall that $\hat{p}_i(t_i) = \mathrm{E}_{t_{-i}} p_i(t_i, t_{-i})$ and $\hat{q}_i(t_i) = \mathrm{E}_{t_{-i}} q_i(t_i, t_{-i})$. We can write the incentive compatibility con-

straint as

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

Clearly, this holds if and only if

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

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

$$\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i. \tag{7}$$

Hence we can rewrite the objective function as

$$\begin{aligned}
\mathrm{E}_t \left[\sum_i p_i(t) t_i - c_i \sum_i q_i(t) \right] &= \sum_i \mathrm{E}_{t_i} \left[\hat{p}_i(t_i) t_i - c_i \hat{q}_i(t_i) \right] \\
&= \sum_i \mathrm{E}_{t_i} \left[\hat{p}_i(t_i)(t_i - c_i) + \varphi_i c_i \right] \qquad (8) \\
&= \mathrm{E}_t \left[\sum_i [p_i(t)(t_i - c_i) + \varphi_i c_i] \right]. \qquad (9)
\end{aligned}$$

Some of the arguments below will use the reduced form probabilities and hence rely on the first expression, (8), for the payoff function, while others focus on the "nonreduced" mechanism and so rely on the second expression, (9).

Summarizing, we can replace the choice of p_i and q_i functions for each i with the choice of a number $\varphi_i \in [0, 1]$ for each i and a function $p_i : T \to [0, 1]$ satisfying $\sum_i p_i(t) \leq 1$ and $\mathrm{E}_{t_{-i}} p_i(t) \geq \varphi_i \geq 0$. Note that this last constraint implies $\mathrm{E}_t p_i(t) \geq \varphi_i$, so

$$\sum_i \varphi_i \leq \sum_i \mathrm{E}_t p_i(t) = \mathrm{E}_t \sum_i p_i(t) \leq 1.$$

Hence the constraint that $\varphi_i \leq 1$ cannot bind and so can be ignored.

The remainder of the proof sketch is more complex and so we introduce several simplifications. First, the proof sketch assumes $I = 2$ and $c_i = c$ for all i. The equal costs assumption implies that the threshold value v^* can be thought of as defining a threshold type t^* to which we compare the t_i reports. Second, we assume that each F_i is uniform which simplifies the calculations.

Third, we will consider the case of finite type spaces and disregard certain boundary issues. The reason for this is that the statement of Theorem 2 is made cleaner by our

use of a continuum of types. Without this, we would have "boundary" types where there is some arbitrariness to the optimal mechanism, making a statement of uniqueness more complex. On the other hand, the continuum of types adds significant technical complications to the proof of our characterization of optimal mechanisms. In this proof sketch, we explain how the proof would work if we focused on finite type spaces, ignoring what happens at boundaries. The proof in Appendix A can be seen as a generalization of these ideas to continuous type spaces.

The proof sketch has five steps. First, we observe that every optimal mechanism 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 the intuition, suppose we have an optimal mechanism 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 $t_i > t'_i$. To simplify further, suppose that these two types have the same probability. 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 $p^*_i(\hat{t}) = p_i(\hat{t})$. For any type profile of the form (t_i, t_{-i}) we assign the p's the original mechanism assigned to (t'_i, t_{-i}) and conversely. Since the probabilities of these types are the same, our independence assumption implies that for every $j \neq i$, agent j is unaffected by the change in the sense that $\hat{p}^*_j = \hat{p}_j$. Obviously, $\hat{p}^*_i(t_i) \geq \hat{p}^*_i(t'_i) = \hat{p}_i(t_i)$. Since the original mechanism was feasible, we must have $\hat{p}_i(t_i) \geq \varphi_i$, so this mechanism must be feasible. It is easy to see that this change improves the objective function, so the original mechanism could not have been optimal.

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

The second step shows that if we have a type profile $t = (t_1, t_2)$ such that $t_2 > t_1 > \hat{t}_1$, then the optimal mechanism has $p_2(t) = 1$. To see this, suppose to the contrary that $p_2(t) < 1$. Then we can change the mechanism by increasing this probability slightly and lowering the probability of giving the good to 1 (or, if the probability of giving it to 1 was 0, lowering the probability that the good is not given to either agent). Since $t_1 > \hat{t}_1$, we have $\hat{p}_1(t_1) > \varphi_1$ before the change, so if the change is small enough, we still satisfy this constraint. Since $t_2 > t_1$, the value of the objective function increases, so the original mechanism could not have been optimal.

The third step is to 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 $p_1(t) = 1$. Because this step is more involved, we postpone explaining it till the end of the rest of the argument. So we continue the proof sketch taking this step as given.

The fourth step is to show that $\hat{t}_1 = \hat{t}_2$. To see this, suppose to the contrary that $\hat{t}_2 > \hat{t}_1$. Then consider a type profile $t = (t_1, t_2)$ such that $\hat{t}_2 > t_2 > t_1 > \hat{t}_1$. From our

second step, the fact that $t_2 > t_1 > \hat{t}_1$ implies $p_2(t) = 1$. However, from our third step, $t_1 > \hat{t}_1$ and $t_2 < \hat{t}_2$ implies $p_1(t) = 1$, a contradiction. Hence there cannot be any such profile of types, implying $\hat{t}_2 \leq \hat{t}_1$. Reversing the roles of the players then implies $\hat{t}_1 = \hat{t}_2$.

Let $t^* = \hat{t}_1 = \hat{t}_2$. This common value of these individual "thresholds" will yield the threshold of our favored–agent mechanism as we will see shortly.

To sum up the first four steps, we can characterize any optimal mechanism by specifying t^*, φ_1, and φ_2. From our second step, if we have $t_2 > t_1 > t^*$, then $p_2(t) = 1$. That is, if both agents are above the threshold, the higher type agent receives the object. From our third step, if $t_1 > t^* > t_2$, then $p_1(t) = 1$. That is, if only one agent is above the threshold, this agent receives the object. Either way, then, if there is at least one agent whose type is above the threshold, the agent with the highest type receives the object. Also, by definition, if $t_i < t^*$, then $\hat{p}_i(t_i) = \varphi_i = \inf_{t_i'} \hat{p}_i(t_i')$. Recall that we showed $\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i$, so $\hat{q}_i(t_i) = 0$ whenever $t_i < t^*$. That is, if an agent is below the threshold, he receives the good with the lowest possible probability and is not checked.

This implies that $\hat{p}$ is completely pinned down as a function of t^*, φ_1, and φ_2. If $t_i > t^*$, then $\hat{p}_i(t_i)$ must be the probability $t_i > t_j$. If $t_i < t^*$, then $\hat{p}_i(t_i) = \varphi_i$. By equation (7), we know that $\hat{q}$ is pinned down by $\hat{p}$ and the φ_i's, so the reduced form is a function only of t^* and the φ_i's. Since we can write the principal's payoff as a function only of the reduced form, the principal's payoff is completely pinned down once we specify t^* and the φ_i's. The fact that the principal's payoff is linear in the φ_i's and the set of feasible φ vectors is convex implies that, given v^*, there must be a solution to the principal's problem at an extreme point of the set of feasible (φ_1, φ_2). Furthermore, every optimal choice of the φ's is a randomization over optimal extreme points.

The last step is to show that such extreme points correspond to favored–agent mechanisms. It is not hard to see that at an extreme point, one of the φ_i's is set to zero and the other is "as large as possible."[12] For notational convenience, consider the extreme point where $\varphi_2 = 0$ and φ_1 is set as high as possible. Recall that φ_1 is the probability 1 gets the good conditional on $t_1 < t^*$. Recall also that 2 gets the good whenever $t_1 < t^*$ and $t_2 > t_2^*$. Hence $\varphi_1 \leq F_2(t^*)$. With the favored agent mechanism where 1 is favored, $\varphi_1 = F_2(t^*)$, so this value of φ_1 is feasible. Hence it must be the largest φ_1 can be. As noted, this corresponds to a favored agent mechanism.

This concludes the proof sketch, except for proving step 3 to which we now turn. 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 $p_1(t) = 1$. To see this, first consider the point labeled $\alpha = (\tilde{t}_1, \tilde{t}_2)$ in Figure 2 below where $\tilde{t}_1 > \hat{t}_1$ while $\tilde{t}_2 < \hat{t}_2$. Suppose that at α, player 1 receives the good with

[12] Under some conditions, $\varphi_i = 0$ for all i is also an extreme point. See Lemma 9 of Appendix B for details.

probability strictly less than 1. Then at any point directly below α but above $\hat{t}_1$, such as the one labeled $\beta = (t_1', \tilde{t}_2)$, player 1 must receive the good with probability zero. This follows because if 1 did receive the good with strictly positive probability here, 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 be feasible. 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 optimal.[13]

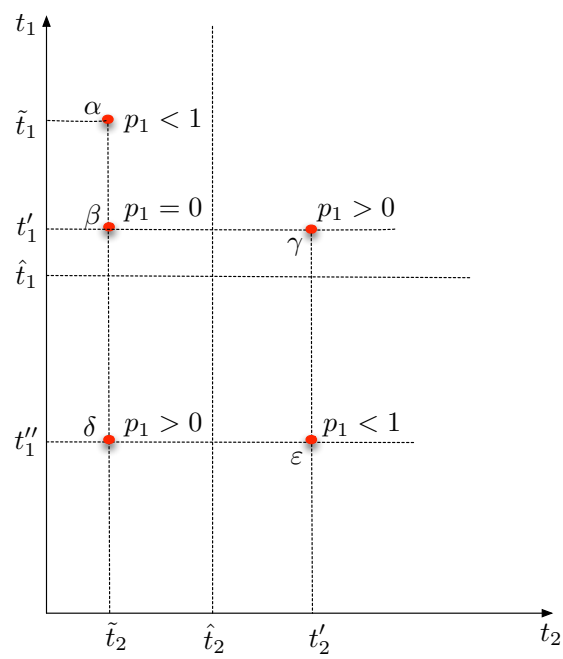

Figure 2

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

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

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.

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 $p_1(t_1, t_2') = 1$. 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 $p_2(t_1, t_2') = 0$. Either way, $p_2(t_1, \tilde{t}_2) \geq p_2(t_1, t_2')$ for all $t_1 \neq t_1'$. 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 on our optimization 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 was not optimal. 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 F_1 is uniform, $\hat{p}_2(t_2')$ 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 by the uniformity of F_2. 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 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 he gets the object at α, while maintaining all constraints.

This completes the sketch of the proof of Theorem 2. The proof itself is divided to two parts. The main part, Theorem 5, establishes certain properties of every optimal mechanism. The proof of this result corresponds to steps (1) through (4) in the sketch above and is contained in Appendix A. The proof of Theorem 2 from Theorem 5, step (5) in the proof sketch, is given in Appendix B.

6.2 Almost Dominance and Ex Post Incentive Compatibility

One appealing property of the favored–agent mechanism is that it is *almost* a dominant strategy mechanism. That is, for every agent, truth telling is a best response to *any* strategies by the opponents. It is not always a dominant strategy, however, as the agent may be completely indifferent between truth telling and lies.

To see this, consider any agent i who is not favored and a type t_i such that $t_i - c_i > v^*$.

If t_i reports his type truthfully, then i receives the object with strictly positive probability under a wide range of strategy profiles for the opponents. Specifically, any strategy profile for the opponents with the property that $t_i - c_i$ is the highest report for some type profiles has this property. On the other hand, if t_i lies, then i receives the object with zero probability given any strategy profile for the opponents. This follows because i is not favored and so cannot receive the object without being checked. Hence for such a type, truth telling weakly dominates any lie for t_i.

Continuing to assume i is not favored, consider any t_i such that $t_i - c_i < v^*$. For *any* profile of strategies by the opponents, t_i's probability of receiving the object is zero regardless of his report. To see this, simply note that if i reports truthfully, he cannot receive the good (since it will either go to another nonfavored agent if one has the highest $t_j - c_j$ and reports honestly or to the favored agent). Similarly, if i lies, he cannot receive the object since he will be caught lying when checked. Hence truth telling is *an* optimal strategy for t_i, though it is not weakly dominant since the agent is indifferent over all reports given any strategies by the other agents.

A similar argument applies to the favored agent. Again, if his type satisfies $t_i - c_i > v^*$, truth telling is dominant, while if $t_i - c_i < v^*$, he is completely indifferent over all strategies. Either way, truth telling is an optimal strategy regardless of the strategies of the opponents.

Because of this property, the favored–agent mechanism is ex post incentive compatible. Formally, (p, q) is ex post incentive compatible if

$$p_i(t) \geq p_i(\hat{t}_i, t_{-i}) - q_i(\hat{t}_i, t_{-i}), \quad \forall \hat{t}_i, t_i \in T_i, \quad \forall t_{-i} \in T_{-i}, \quad \forall i \in \mathcal{I}.$$

That is, t_i prefers reporting honestly to lying even conditional on knowing the types of the other agents. It is easy to see that the favored–agent mechanism's almost–dominance property implies this. Of course, the ex post incentive constraints are stricter than the Bayesian incentive constraints, so this implies that that favored–agent mechanism is ex post optimal.

While the almost–dominance property implies a certain robustness of the mechanism, the complete indifference for types below the threshold is troubling. There are simple modifications of the mechanism which do not change its equilibrium properties but make truth telling weakly dominant rather than just almost dominant. For example, suppose there are at least three agents and that every agent i satisfies $\bar{t}_i - c_i > v^*$.[14] Suppose we modify the favored– agent mechanism as follows. If an agent is checked and found to have lied, then one of the other agents is chosen at random and his report is checked.

[14]Note that if $\bar{t}_i - c_i < v^*$, then the favored agent mechanism *never* gives the object to i, so i's report is entirely irrelevant to the mechanism. Thus we cannot make truth telling dominant for such an agent, but the report of such an agent is irrelevant anyway.

If it is truthful, he receives the object. Otherwise, no agent receives it. It is easy to see that truth telling is still an optimal strategy and that the outcome is unchanged if all agents report honestly. It is also still weakly dominant for an agent to report the truth if $t_i - c_i > v^*$. Now it is also weakly dominant for an agent to report the truth even if $t_i - c_i < v^*$. To see this, consider such a type and assume i is not favored. Then if t_i lies, it is impossible for him to receive the good regardless of the strategies of the other agents. However, if he reports truthfully, there is a profile of strategies for the opponents where he has a strictly positive probability of receiving the good — namely, where one of the nonfavored agents lies and has the highest report. Hence truth telling weakly dominates any lie. A similar argument applies to the favored agent.

6.3 Extensions

In this subsection, we discuss some simple extensions. First, it is straightforward to generalize to allow the principal to have a strictly positive value, say $R > 0$, to retaining the object. To see this, simply introduce an agent $I+1$ whose type is R with probability 1. Allocating the object to this agent is the same as keeping it. It is easy to see that our analysis then applies to this modified model directly. If R is sufficiently large, then agent $I+1$ will be favored. That is, if every agent $i \leq I$ reports $t_i - c_i < R$, then the principal retains the object. If any agent $i \leq I$ 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. This mechanism is the analog of a reserve price mechanism with R as the reserve price. If R is small enough that $I+1$ is not the favored agent, then the optimal mechanism is unaffected by the principal's value.

Another natural extension to consider is when the process of verifying an agent's claim is also costly for that agent. In our example where the principal is a dean and the agents are departments, it seems natural to say that departments bear a cost associated with providing documentation to the dean.

The main complication associated with this extension is that 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. Thus the simplification we obtain where we can treat the agent's payoff as simply equal to the probability he receives the object no longer holds.

On the other hand, we can retain this simplification at the cost of adding an assumption. To be specific, we can simply 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 verification costs for the agents is straightforward. We can

also allow the cost to the agent of being verified to differ depending on whether the agent lied or not. To see this, 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. The incentive compatibility condition becomes

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

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

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

Analogously to the way we characterized the optimal mechanism in Section 6.1, 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

$$\hat{q}_i(t_i) = \frac{\hat{p}_i(t_i) - \varphi_i}{1 + \hat{c}_i^F}. \tag{10}$$

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

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

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 (10)). Intuitively, since verification imposes costs on the agent in this model, the threat of verification 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 optimal mechanism.

As a third extension, suppose that the principal can impose (limited) fines on the agents. Assume, as in the second extension, that the value to an agent of receiving the object is 1 and the payoff to not receiving it is zero. It is easy to see that the optimal mechanism for the principal is impose the largest possible fine when the agent is found to have lied and nothing otherwise. The analysis of this model is then identical to that of our second extension where we set $\hat{c}_i^T = 0$ and $\hat{c}_i^F$ equal to this maximum penalty.

7 Conclusion

There are many natural extensions to consider. For example, in the previous subsection, we discussed the extension in which the agents bear some costs associated with verification, but under the restriction that the value to the agent of receiving the object is independent of his type. A natural extension of interest would be to drop this restriction.[15]

[15] See Ambrus and Egorov (2012) for an analysis of how imposing costs on agents can be useful for the principal in a setting without verification.

A second natural extension would be to allow costly monetary transfers. We argued in the introduction that within organizations, monetary transfers are costly to use and hence have excluded them from the model. It would be natural to model these costs explicitly and determine to what extent the principal allows inefficient use of some resources to obtain a better allocation of other resources.

Another direction to consider is to generalize the nature of the principal's allocation problem. For example, what is the optimal mechanism if the principal has to allocate some *tasks*, as well as some resources? In this case, the agents may prefer to *not* receive certain "goods." Alternatively, there may be some common value elements to the allocation in addition to the private values aspects considered here.

In the context of the example of the dean allocating a job slot to one of several departments, a natural extension would be to allow each department to have multiple candidates, one of which it can put forward. In this case, the department faces a tradeoff between those the dean would prefer and those the department wants most. This tradeoff implies that the value to the department of receiving the slot matters, unlike in the analysis here. Thus this model would look more like a multiple sender version of Che, Dessein, and Kartik (forthcoming) with costly verification.

Another natural direction to consider is alternative specifications of the information structure and verification technology. Here each agent knows exactly what value he can create for the principal with the object. Alternatively, the principal may have private information which determines how he interprets an agent's information. Also, it is natural to consider the possibility that the principal partially verifies an agent's report, choosing how much detail to go into. For example, an employer dealing with a job applicant can decide how much checking to do of the applicant's resume and references.

Appendix

A Proof of Theorem 5

In this section, we state and prove a theorem which is used in Appendix B to prove Theorem 2. Specifically, in Theorem 5 below, we show that every optimal mechanism is what we call a *threshold mechanism*. In Appendix B, we show that optimal threshold mechanisms are convex combinations of favored–agent mechanisms. In this and subsequent sections, we work with the restated version of the optimization problem for the principal derived in Section 6.1 (in particular, equations (8) and (9) and the constraints that follow).

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* $\hat{p}_i(t_i) = \min_{t_i' \in T_i} \hat{p}_i(t_i)$*.*

Theorem 5. *Every optimal mechanism is a threshold mechanism.*

The proof of Theorem 5 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.[16]

Lemma 1. *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.*

[16] We thank Benjy Weiss for suggesting the idea of of the following proof.

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 $f : X \to \mathbf{R}$ 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 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,

$$\mu_x \left\{x \mid \mu_y \left(S^2(x)\right) > 0\right\} = \mu_x \left(S^1_x\right) = \mu_x \left\{x \mid \mu_y \left(S^1(x)\right) > 0\right\}. \tag{11}$$

To see this, suppose not. Then

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

as

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

Let

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

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

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

as $S^1(x) \supseteq S^2(x)$ and $\mu(\Delta) > 0$. But this contradicts $\mu(S^2) = \mu(S^1)$. Hence equation (11) 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 $x \in S^1_x$, we have $\mu_{yz}(S^1(x)) > 0$ and for all $(y, z) \in S^1_{yz}$, we have $\mu_x(S^1((y, z))) > 0$. Applying $I = 2$ result again to the set S^1_{yz}, we have $G \subseteq S^1_{yz}$ with $\mu_{yz}(G) = \mu_{yz}(S^1_{yz})$ 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

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

Since $G \subseteq S^1_{yz}$ and $\mu_{yz}(G) = \mu_{yz}(S^1_{yz})$, we have $\mu(S^2) = \mu(S^1)$. Clearly, $S^2_y = G_y$ and $S^2_z = G_z$. Fix any $y \in S^2_y$. Since $y \in G_y$, we have $\mu_z\{z \mid (y, z) \in G\} > 0$. Since $G \subseteq S^1_{yz}$ 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 $x \in S^2_x$, we have $\mu_{y,z}(S^2(x)) > 0$. Hence we now define the set S^3 by

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

Clearly, S^3_x is measurable and we have $\mu_{yz}(S^3(x)) > 0$ for every $x \in S^3_x$. Furthermore, $S^3 \subseteq S^2 \subseteq S^1$ and hence $S^3_x \subseteq S^1_x$. In fact, $\mu(S^3) = \mu(S^2) = \mu(S^1)$ implies $\mu(S^3_x) = \mu(S^1_x)$. To see this, suppose not. Then $\mu_x(S^3_x) < \mu_x(S^1_x)$. Since for each $x \in S^1_x$, 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^3_y and S^3_z are measurable, (2) for all $y \in S^3_y$, we have $\mu_{x,z}(S^3(y)) > 0$, and (3) for all $z \in S^3_z$, we have $\mu_{x,y}(S^3(z)) > 0$. Consider an element $y \in S^2_y$. 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^2_x 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 $S^3_y = S^2_y$ and for every $y \in S^3_y$, we have $\mu_{xz}(S^3(y)) = \mu_{xz}(S^2(y)) > 0$. A similar argument establishes that $S^3_z = S^2_z$ and that for $z \in S^3_z$, 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^1_x is measurable and that for every $x \in S^1_x$, we have $\mu_y(S^1(x)) > \varepsilon$. Define

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

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

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 $S^2 \subseteq S^1$ and the fact that $\mu_y(S^1(x)) > \varepsilon$ for all $x \in S^1_x$ implies that $\mu(S^2) < \mu(S^1)$, a contradiction.

Since

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

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

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

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 $\hat{p}_i(t_i) = \mathrm{E}_{t_{-i}} p_i(t)$. As explained in Section 6.1, 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 $\hat{q}_i(t_i) = \mathrm{E}_{t_{-i}} q_i(t)$.

Lemma 2. *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 1).

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 1) on which all the properties of Lemma 2 are satisfied everywhere.

Lemma 3. *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

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

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 1).

Lemma 4. *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

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

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

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

Also let

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

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

and

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

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

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

Finally let

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

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

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 $t \in \bar{A}$,

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

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

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

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

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

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

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

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 $S = \cup_{\alpha \in \mathbb{Q}} D(\alpha)$ 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 $t'_j \notin C_j(\alpha)$. 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 5. *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*

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

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

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

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 1 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 1, 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 1, 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 2, 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 4 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 1, 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 2 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 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, $B_{-j} = A_{-j} = M_i \times E_{-i,j}$. 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 2, 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 2, 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

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

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, $\mu(G) = \int_{A_j} \int_{B_j} \mu_{-j}\left(F_{-j}\left(t'_j, t''_j\right)\right) \mu_j\left(dt'_j\right) \mu_j\left(dt''_j\right)$ which is strictly positive since for each $\left(t'_j, t''_j\right)$ in the domain of integration $\mu_{-j}\left(F_{-j}\left(t'_j, t''_j\right)\right) > 0$ and the domains of integration has positive μ_j measure.) The argument above also showed that for every $\left(t'_j, t''_j, t_{-j}\right) \in G$, there exists k such that $p_k(t''_j, t_{-k}) > 0$. Therefore there exists k such that $\mu(G^k) > 0$ where

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

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 1, 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.[17]

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 $\hat{G}^k_{j^{(1)}} \subset A_j$. Similarly, to see that $\tilde{S}_j = \hat{C}_j$, note that $\hat{G}^k_{j^{(2)}} \subset \hat{S}_j = B_j$.)

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

$$z_E = 0 \text{ iff } t \notin E \tag{12}$$

$$\forall t_j \in \hat{A}_j = \hat{D}_j, \quad \mathrm{E}_{t_{-j}}[z_A(t_j, t_{-j})] = \mathrm{E}_{t_{-j}}[z_D(t_j, t_{-j})] \tag{13}$$

$$\forall t_k \in \hat{C}_k = \hat{D}_k, \quad \mathrm{E}_{t_{-k}}[z_C(t_k, t_{-k})] = \mathrm{E}_{t_{-k}}[z_D(t_k, t_{-k})] \tag{14}$$

$$\forall t_j \in \tilde{S}_j = \hat{C}_j, \quad \mathrm{E}_{t_{-j}}[z_S(t_j, t_{-j})] = \mathrm{E}_{t_{-j}}[z_C(t_j, t_{-j})] \tag{15}$$

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

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

[17] For 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 $\hat{S}_j = B_j$).

Similarly,

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

Hence

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

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

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

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

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

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 $\mathrm{E}[z_S(t)] = \mathrm{E}[z_A(t)]$. 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 (12), (13), (14), and (15).

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

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

where we recall that for any event S, we let $S_{-\ell}(t_\ell) = \{t_{-\ell} \in T_{-\ell} \mid (t_\ell, t_{-\ell}) \in E\}$, the t_ℓ-fiber of E. 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 (13), (14), and (15). First, consider equation (13). Note that

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

and

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

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 (13) holds.

Second, consider equation (14). Note that

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

and

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

Thus (14) holds.

Finally, consider equation (15). We have

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

and

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

Thus (15) holds. ▮

Lemma 6. *For any i,*

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

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

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

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

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

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 2, 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

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

which must hold for δ sufficiently small. ▌

Recall that

$$T_i^* = \left\{t_i \in T_i \setminus \bar{S}_i \mid \hat{p}_i(t_i) > \varphi_i \text{ and } \mu_i(\{t_i' \mid \hat{p}_i(t_i') > \varphi_i \text{ and } t_i' < t_i\}) > 0\right\}.$$

Lemma 7. *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

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

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 6.

By Lemma 5, 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 3, 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 2.

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 6, 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 3 implies $p_j = 0$ for all $j \in \mathcal{I}^N$, Lemma 5 implies $p_i = 0$ for all $i \in \mathcal{I}^E$, and Lemma 2 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 $\hat{p}_i(\hat{t}_i) > \varphi_i$, 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 5 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 2, 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 $\hat{p}_i(t_i) = \varphi_i < 1$.[18] 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 2 implies $t_i' > c_i$ and hence $t_i > c_i$. Hence Lemma 2 implies that for almost every $t_i \in \hat{T}_i$ and almost every t_{-i}, we have $\sum_j p_j(t) = 1$.

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

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

$$\mu_i\left(\{t_i \in \hat{T}'_i \mid t_i \geq t'_i + \varepsilon\}\right) > 0$$

$$\mu_i\left(\{t_i \in T^*_i \mid t_i \leq t'_i - \varepsilon\}\right) > 0.$$

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 $t \in S^1 \times \hat{T}'_{-i}$,

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

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

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

Hence p was not optimal, a contradiction. ▌

To see that this proves Theorem 5, let v^* be the threshold. By Lemma 7, if some i has $t_i - c_i > v^*$, then that i satisfies $\hat{p}_i(t_i) > \varphi_i$. By Lemma 3, 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 5, no j with $t_j - c_j < v^*$ has any probability of getting the good. Since $\hat{p}_i(t_i) > \varphi_i$, Lemma 2 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 7 implies that $\hat{p}_i(t_i) = \varphi_i$. Thus we have a threshold mechanism. ▌

B Proof of Theorem 2

In this section, we use Theorem 5 to complete the proof of Theorem 2.

Let (p, q) denote any optimal mechanism. In light of Theorem 5, 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

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

Note that

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

Hence we can rewrite the objective function as

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

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 $p_i : T \to [0,1]$, each $\varphi_i \in [0,1]$, $\sum_i p_i(t) \leq 1$ for all t, and $\mathrm{E}_{t_{-i}} p_i(t) \geq \varphi_i$ for all $t_i \in T_i$ and all i. From Theorem 5, 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 5 also tells us that $\hat{p}_i(t_i) = \varphi_i$ for (almost) all $t_i < v^* + c_i$ for all i. (From Lemma 6, 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 8. *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 $t_i \in (\underline{t}_i, \overline{t}_i)$, 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

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

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

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

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 8, 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 9. *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 8, 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 8 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 $\varphi^k_i \geq 0$ 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 $\varphi^1_j = \varphi^2_j = 0$ for all $j \neq i$. Since we cannot have $\varphi^1 = \varphi^2$, without loss of generality, we must have $\varphi^1_i < \prod_{j \neq i} F_j(v^* + c_j) < \varphi^2_i$. 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 8, $r_i \geq 0$. 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 9. 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 10. *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 8, 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 8, 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 5, 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 $\mathcal{I}^0 = \emptyset$. We showed in the proof of Lemma 8 that the set of extreme points consists of the zero vector $\mathbf{0}$ and the collection of vectors $\varphi^*(k)$, $k = 1, \ldots, I$. The same argument as for the previous case shows that any of the extreme points $\varphi^*(k)$ generates the same reduced form as the favored–agent mechanism with k as the favored agent and v^* as the threshold.

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

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

It is easy to see that this term must be strictly positive since $\underline{t}_i \geq 0$ and $c_i > 0$. Hence whenever there is 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. ▌

C Proof of Theorem 4

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

$$\begin{aligned} F_1(\hat{t}_1)F_2(t_2^*)\mathrm{E}[t_1 \mid t_1 \le \hat{t}_1] &+ [1 - F_1(\hat{t}_1)]F_2(t_2^*)\mathrm{E}[t_1 \mid t_1 > \hat{t}_1] \\ &+ F_1(\hat{t}_1)[1 - F_2(t_2^*)]\left\{\mathrm{E}[t_2 \mid t_2 > t_2^*] - c_2\right\} \\ &\ge F_1(\hat{t}_1)F_2(t_2^*)\mathrm{E}[t_2 \mid t_2 \le t_2^*] + [1 - F_1(\hat{t}_1)]F_2(t_2^*)\left\{\mathrm{E}[t_1 \mid t_1 > \hat{t}_1] - c_1\right\} \\ &+ F_1(\hat{t}_1)[1 - F_2(t_2^*)]\mathrm{E}[t_2 \mid t_2 > t_2^*]. \end{aligned} \tag{16}$$

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

$$F_2(t_2^*)\mathrm{E}[t_1 \mid t_1 > \hat{t}_1] \ge F_2(t_2^*)\left\{\mathrm{E}[t_1 \mid t_1 > \hat{t}_1] - c_1\right\},$$

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

$$\mathrm{E}[t_1 \mid t_1 \le \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} - c_1 \ge \mathrm{E}[t_2 \mid t_2 \le t_2^*] + \frac{c_2}{F_2(t_2^*)} - c_2. \tag{17}$$

From equation (2), the right–hand side of equation (17) is $t_2^* - c_2$. Hence we need to show

$$\mathrm{E}[t_1 \mid t_1 \le \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} - c_1 \ge t_2^* - c_2. \tag{18}$$

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

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

A similar rearrangement to our derivation of equation (2) yields

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

Hence

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

implying equation (17). 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 (16) 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 (18) 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 3 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 Proofs

First, we show that an FOSD shift upward in any F_i increases the the principal's ex ante payoff. Let $W(t_1, \ldots, t_I)$ denote the principal's payoff under the optimal mechanism given that the realized types are $(t_1, \ldots, t_I)$. We show that $\mathrm{E}_{t_{-i}} W(t_i, t_{-i})$ is increasing in t_i for all i, implying the claim. It is easy to see that $W(t)$ is increasing in t_i if i is favored, so we only need to consider the case where i is *not* the favored agent.

For notational simplicity, assume 1 is the favored agent. Then if $t_i - c_i < t_1^* - c_1$, we see that $\mathrm{E}_{t_{-i}} W(t_i, t_{-i})$ is independent of t_i and so is trivially (weakly) increasing. So consider any t_i such that $t_i - c_i > t_1^* - c_1$. In this range,

$$\mathrm{E}_{t_{-i}} W(t_i, t_{-i}) = \mathrm{E}_{t_{-i}} \max_j (t_j - c_j)$$

which is obviously increasing in t_i. We complete the argument by showing that $\mathrm{E}_{t_{-i}} W(t_i, t_{-i})$ is continuous in t_i at $t_i = t_1^* - c_1 + c_i$. To see this, note that

$$\begin{aligned}
\lim_{t_i \downarrow t_1^* - c_1 + c_i} \mathrm{E}_{t_{-i}} W(t_i, t_{-i}) &= \mathrm{E}_{t_{-i}} \max\{t_1^* - c_1, \max_{j \neq i}(t_j - c_j)\} \\
&= \Pr\left[\max_{j \neq 1,i}(t_j - c_j) < t_1^* - c_1\right] \mathrm{E}\left(\max\{t_1 - c_1, t_1^* - c_1\}\right) \\
&\quad + \Pr\left[\max_{j \neq 1,i}(t_j - c_j) > t_1^* - c_1\right] \mathrm{E}\left(\max_{j \neq 1,i}(t_j - c_j) \mid \max_{j \neq 1,i}(t_j - c_j) > t_1^* - c_1\right) \\
&= \Pr\left[\max_{j \neq 1,i}(t_j - c_j) < t_1^* - c_1\right] \mathrm{E}(t_1) \\
&\quad + \Pr\left[\max_{j \neq 1,i}(t_j - c_j) > t_1^* - c_1\right] \mathrm{E}\left(\max_{j \neq 1,i}(t_j - c_j) \mid \max_{j \neq 1,i}(t_j - c_j) > t_1^* - c_1\right)
\end{aligned}$$

where the last equality follows from the definition of t_1^*. It is not hard to see that this last expression is equal to $\mathrm{E}_{t_{-i}} W(t_i, t_{-i})$ for any $t_i < t_1^* - c_1 + c_i$. Hence $\mathrm{E}_{t_i} W(t_i, t_{-i})$ is increasing in t_i for all i, so an FOSD shift upward in the distribution of t_i must increase $\mathrm{E}_t W(t)$.

Next, 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

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

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

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

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.

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] Gale, D., and M. Hellwig, "Incentive–Compatible Debt Contracts: The One–Period Problem," *Review of Economic Studies*, **52**, October 1985, 647–663.

[12] Glazer, J., and A. Rubinstein, "On Optimal Rules of Persuasion," *Econometrica*, **72**, November 2004, 1715–1736.

[13] Glazer, J., and A. Rubinstein, "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, **1**, December 2006, 395–410.

[14] Green, J., and J.-J. Laffont, "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, **53**, July 1986, 447–456.

[15] 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.

[16] Kartik, N., and O. Tercieux, "Implementation with Evidence," *Theoretical Economics*, **7**, May 2012, 323–355.

[17] McAfee, R. P., and J. McMillan, "Bidding Rings," *American Economic Review*, **82**, June 1992, 579–599.

[18] Mookherjee, D., and I. Png, "Optimal Auditing, Insurance, and Redistribution," *Quarterly Journal of Economics*, **104**, May 1989, 399–415.

[19] Roth, A. E., and M. A. O. Sotomayor, *Two–Sided Matching: A Study in Game–Theoretic Modeling and Analysis*, 1990, Cambridge University Press.

[20] Sher, I., and R. Vohra, "Price Discrimination through Communication," working paper, June 2011.

[21] Townsend, R., "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, **21**, 1979, 265–293.

[22] Townsend, R., "Information Constrained Insurance: The Revelation Principle Extended," *Journal of Monetary Economics*, **21**, March–May 1988, 411–450.

[23] Yoon, K., "Optimal Mechanism Design When Both Allocative Inefficiency and Expenditure Inefficiency Matter," *Journal of Mathematical Economics*, **47**, December 2011, 670–676.