Optimal Allocation with Costly Verification[1]

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

Preliminary Draft
January 2012

[1]We thank Ricky Vohra and numerous seminar audiences for helpful comments. We also thank the National Science Foundation, grants SES–0820333 (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.

[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@nwu.edu.

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

1 Introduction

Consider the problem of the head of an organization — say, a dean — who has an indivisible resource (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. The problem faced by the dean is made more complex by the fact that each department has much more information regarding the availability of promising candidates and the likelihood that these candidates will produce valuable research, teach well, and more generally be of value to the university.

The standard mechanism design approach to this situation would construct a mechanism whereby each department would report its type to the dean. Then the slot would be allocated and various monetary transfers made as a function of these reports. The drawback of this approach is that the monetary transfers between the dean and the departments are assumed to have no efficiency consequences. In reality, the monetary resources each department has is presumably chosen by the dean in order to ensure that the department can achieve certain goals the dean sees as important. To take back such funds as part of an allocation of a job slot would undermine the appropriate allocation of these resources. In other words, such monetary transfers are part of the overall allocation of all resources within the university and hence do have important efficiency consequences. We focus on the admittedly extreme case where no monetary transfers are possible at all.

Of course, without some means to ensure incentive compatibility, the dean cannot extract any information from the departments. In many situations, it is natural to assume that the head of the organization can demand to see documentation which proves that the division or department's claims are correct. Processing such information is costly, though, to the dean and departments and so it is optimal to restrict such information requests to the minimum possible.

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 the opposite group. 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 instead.[1]

[1]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.

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.

We characterize optimal mechanisms for such settings, considering both Bayesian and ex post incentive compatibility. We construct a mechanism which is optimal under either approach and which has a particularly simple structure which we call a *favored–agent mechanism*. There is a threshold value and a favored agent, say i. If all agents other than i report a value for the resource below the threshold, then the resource goes to the favored agent and no documentation is required. If some agent other than i reports a value above the threshold, then the agent which reports the highest value is required to document its claims. This agent receives the resource if his claims are verified and the resource goes to any other agent otherwise. We also give a variety of comparative statics. In particular, we show that an agent is more likely to be the favored agent the higher is his cost of being verified, the "better" is his distribution of values, and the less risk is his distribution of values. We also show that the mechanism is, in a sense, *almost* a dominant strategy mechanism.

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 there is only one agent and monetary transfers are allowed. 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 (2011), Kartik and Tercieux (2011), 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.

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 Bayesian incentive compatible mechanisms, showing that the favored–agent mechanism is an optimal mechanism. We show that the favored–agent mechanism also satisfies ex post incentive compatibility and hence is also an optimal ex post incentive compatible mechanism. Section 3 also contains comparative statics and discusses other properties of the mechanism.

In Section 4, we sketch the proof of optimality of our mechanism and discuss several other issues. Section 5 concludes. Proofs not contained in the text are in the Appendix.

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 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. 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 $0 \le \underline{t}_i < \bar{t}_i < \infty$. (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 requiring documentation by agent i to demonstrate what his type is. 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 providing this documentation. The costs to the agent of providing documentation is zero. To understand this, think of the agent's resources as used for 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 be used in either case. 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.[2] In Section 4, we show one way to generalize our model to allow agents to bear some costs of providing documentation 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.[3] We also assume that each agent's reservation utility is 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

[2] 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 threats to require documentation as a way of "fining" agents and thus to help achieve incentive compatibility. This both complicates the analysis and indirectly introduces a form of the transfers we wish to exclude.

[3] 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 $\underline{u}_i(t_i) = 0$.

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, allowing the principal to decide which agents to check as a function of the outcome of previous checks and cheap talk statements for multiple stages before finally allocating the good or deciding to not allocate it at all. 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. To see this, fix a dynamic mechanism and any equilibrium. The equilibrium defines a function from type profiles into probability distributions over outcomes. More specifically, an outcome is a sequence of checks and an allocation of the good (perhaps to no one). Replace this mechanism with a direct mechanism where agents report types and the outcome (or distribution over outcomes) given a vector of type reports is what would happen in the equilibrium if this report were true. Clearly, just as in the usual Revelation Principle, truth telling is an equilibrium of this mechanism and this equilibrium yields the same outcome as the original equilibrium in the dynamic mechanism. We can replace any outcome which is a *sequence* of checks with an outcome where exactly these checks are done *simultaneously*. All agents and the principal are indifferent between these two outcomes. Hence the altered form of the mechanism where we change outcomes in this way also has a truth telling equilibrium and yields an outcome which is just as good for the principal as the original equilibrium of the dynamic 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 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 any agent $j \neq i$ with probability 1. 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 the principal either checks the agent she is giving the object to or no agent.

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 overall probability i is assigned the good and $q_i(t)$ the probability i is assigned the good and checked. In light of this, we define a mechanism to be 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, with the exception of some discussion in Section 4, the word "mechanism" will be used only to denote such a tuple of functions.

The incentive compatibility constraint for i is then

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

The principal's objective function is

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

We will also characterize the optimal ex post incentive compatible mechanism. It is easy to see that the reduction arguments above apply equally well to the ex post case. For ex post incentive compatibility, then, we can write the mechanism the same way. In this case, the objective function is the same, but the incentive constraints become

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

Of course, the ex post incentive constraints are stricter than the Bayesian incentive constraints.

3 Results

In this section, we show two main results. First, every optimal mechanism takes the form of what we call a *threshold mechanism*. Second, there is always an optimal mechanism with the more specific form of a *favored–agent mechanism*. We also give a simple characterization of the optimal favored–agent mechanism.

More specifically, we say that $\mathcal{M}$ is a *threshold mechanism* if there exists a *threshold* $v^* \in \mathbf{R}_+$ such that the following holds 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 any t, if $t_i - c_i < v^*$, then $q_i(t) = 0$ for all i and $\hat{p}_i(t_i) = \min_{t'_i \in T_i} \hat{p}_i(t_i)$. In other words, no agent with a "value" — that is, $t_i - c_i$ — below the threshold can get the object with more than his lowest possible interim probability of receiving it. Such an agent is not checked. On the other hand, if any agent reports a "value" — that is, $t_i - c_i$ — above the threshold, then the agent with the highest reported value receives the object.

Theorem 1. *Every Bayesian optimal mechanism is a threshold mechanism.*

Section 4 contains a sketch of the proof of this result.

We say that $\mathcal{M}$ 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. 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 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 $p_i(t) = q_i(t) = 1$. 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.

To see that a favored–agent mechanism is a special case of a threshold mechanism, consider a threshold mechanism with the property that $\min_{t'_i \in T_i} \hat{p}_i(t_i) = 0$ for all $i \neq i^*$ and $\min_{t'_{i^*} \in T_{i^*}} \hat{p}_{i^*}(t'_{i^*}) = \prod_{i \neq i^*} F_i(v^* + c_i)$. In this case, if any agent $i \neq i^*$ has $t_i - c_j < v^*$, he receives the object with probability 0, just as in the favored–agent mechanism. On the other hand, the favored agent receives the object with probability at least equal to the probability that all other agents have values below the threshold. Expand?

Theorem 2. *There always exists a Bayesian optimal mechanism which is a favored–agent mechanism.*

We complete the specification of the optimal mechanism by characterizing the optimal threshold and the optimal favored agent. When the type distributions and verification costs are the same for all i, of course, the principal is indifferent over which agent is favored. In this case, the principal may also be indifferent between a favored–agent mechanism and a randomization over favored–agent mechanisms with different favored agents. Loosely speaking, this is "nongeneric" in the sense that it requires a very specific relationship between the type distributions and verification costs. "Generically," there is a unique optimal favored agent.

For each i, define t_i^* by

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

It is easy to show that t_i^* is well–defined. To see this, let

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

Clearly, $\psi_i(\underline{t}_i) = \mathrm{E}(t_i) - c_i < \mathrm{E}(t_i)$. For $t^* \geq \underline{t}_i$, ψ_i is strictly increasing in t^* and goes to infinity as $t^* \to \infty$. Hence there is a unique $t_i^* > \underline{t}_i$.[4]

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 definition of t_i^* is

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

We say that i is *not isolated* if

$$t_i^* - c_i \in \operatorname{int}\left(\bigcup_{j \neq i} [\underline{t}_j - c_j, \bar{t}_j - c_j]\right).$$

Intuitively, this property says only that there are values of $t_j - c_j$, $j \neq i$, which have positive probability and are "just above" $t_i^* - c_i$ and similarly some which are "just below."

Theorem 3. *If i is the favored agent, then $t_i^* - c_i$ is an optimal threshold. If i is not isolated, then $t_i^* - c_i$ is the uniquely optimal threshold. When there are multiple optimal thresholds, all optimal thresholds give the same allocation and checking probabilities.*

Proof. For notational convenience, let 1 be the favored agent. Contrast the principal's payoff to thresholds $t_1^* - c_1$ and $\hat{v}^* > t_1^* - c_1$. Let t denote the profile of reports and let x be the highest $t_j - c_j$ reported by one of the other agents. Then the principal's payoff as a function of the threshold and x is given by

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

[4] Note that if we allowed $c_i = 0$, we would have $t^* = \underline{t}_i$. 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.

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 give 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 $\mathrm{E}(t_1)$. Note that $x > t_1^* - c_1$ implies

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

Hence given that 1 is the favored agent, the threshold $t_1^* - c_1$ is better than any larger threshold. This comparison is strict for every $v^* > t_1^* - c_1$ if for every such v^*, there is a strictly positive probability that there is some j with $t_1^* - c_1 < t_j - c_j < \hat{v}^*$. A similar argument shows the threshold $t_1^* - c_1$ is better than any smaller threshold, strictly so if for every $v^* < t_1^* - c_1$, there is a strictly positive probability that there is some j with $v^* < t_j - c_j < t_1^* - c_1 < t_j - c_j$. Thus $t_1^* - c_1$ is always optimal and is uniquely optimal under the condition stated.

Finally, note that the only time that $t_i^* - c_i$ is not strictly better for the principal than v^* is when the middle column of the table above has zero probability and hence the allocation of the good and checking probabilities are the same whether $t_i^* - c_i$ or v^* is the threshold. ▌

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

Proof. 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

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

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

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

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

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

From equation (2), the right–hand side equation (4) is $t_2^* - c_2$. Hence we need to show

$$\mathrm{E}[t_1 \mid t_1 \leq \hat{t}_1] + \frac{c_1}{F_1(\hat{t}_1)} - c_1 \geq t_2^* - c_2. \tag{4}$$

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

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

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

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

implying equation (4). 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, $\hat{t}_1 = t_1^*$. Since $\underline{t}_1 < t_1^*$, this implies $F_1(\hat{t}_1) > 0$. Hence 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 $t_1^* - c_1 = t_2^* - c_2$.

Given this, we may as well consider only 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

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

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 (4) 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 must strictly prefer 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 $\underline{t}_j - c_j \geq t_2^* - c_2$. 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 $j \neq 1, 2$ with $(t_2^* - c_2, t_1^* - c_1) \cap (\underline{t}_j - c_j, \bar{t}_j - c_j) \neq \emptyset$. From the above, we know there is a $j \neq 1, 2$ with $\underline{t}_j - c_j \geq t_2^* - c_2$. Also, $\underline{t}_j - c_j < t_j^* - c_j \leq t_1^* - c_1$. Hence $t_2^* - c_2 < \underline{t}_j - c_j < t_1^* - c_1$, completing the proof. ▌

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

$$\int_{\underline{t}_i}^{t_i^*} F_i(\tau)\, d\tau = c_i. \tag{5}$$

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 (5) for fixed t_i^*, so to maintain the equality, t_i^* must increase. Therefore, such a shift makes it more likely than 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 value, he is a better choice to be the favored agent.

The intuition for why less risky agents are favored is not as immediate. One way to see the idea is to suppose that there is one agent whose type is completely riskless — i.e., is known to the principal. Obviously, there is no reason for the principal to check this agent since his type is known. Thus setting him as the favored agent — the least likely agent to be checked — seems natural.

We illustrate with two examples. First, suppose we have two agents where $t_1 \sim U[0,1]$, $t_2 \sim U[0,2]$ and $c_1 = c_2 = c$. It is easy to calculate t_i^*. From equation (1), we have

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

For $i = 1$, if $t_1^* < 1$, it must solve

$$\frac{1}{2} = \int_0^{t_1^*} t_1^* \, ds + \int_{t_1^*} 1 s \, ds - c$$

or

$$\frac{1}{2} = (t_1^*)^2 + \frac{1}{2} - \frac{(t_1^*)^2}{2} - c$$

so

$$t_1^* = \sqrt{2c}.$$

This holds only if $c \leq 1/2$ so that $t_1^* \leq 1$. Otherwise, $\mathrm{E}\max\{t_1, t_1^*\} = t_1^*$, so $t_1^* = (1/2) + c$. So

$$t_1^* = \begin{cases} \sqrt{2c}, & \text{if } c \leq 1/2 \\ (1/2) + c, & \text{otherwise.} \end{cases}$$

A similar calculation shows that

$$t_2^* = \begin{cases} 2\sqrt{c}, & \text{if } c \leq 1 \\ 1 + c, & \text{otherwise.} \end{cases}$$

It is easy to see that $t_2^* > t_1^*$ for all $c > 0$, so 2 is the favored agent. The optimal threshold value is

$$t_2^* - c = \begin{cases} 2\sqrt{c} - c, & \text{if } c \leq 1 \\ 1, & \text{otherwise.} \end{cases}$$

Note that if $2\sqrt{c} \geq 1$, i.e., $c \geq 1/4$, then the threshold value $v^* \geq 2\sqrt{c} - c \geq 1 - c$ is set so high that it is impossible to have $t_1 - c > v^*$ since $t_1 \leq 1$ with probability 1. In this case, the favored agent mechanism corresponds to simply giving the good to agent 2 independently of the reports. If $c \in (0, 1/4)$, then there are type profiles for which agent 1 receives the good, specifically those with $t_1 > 2\sqrt{c}$ and $t_1 > t_2$.

For a second example, suppose again we have two agents, but now $t_i \sim U[0,1]$ for $i = 1, 2$. Assume $c_2 > c_1 > 0$. In this case, calculations similar to those above show that

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

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 $c_1 < 1/2$, we must have $t_2^* - c_2 > t_1^* - c_1$, so that 2 is the favored agent. If $c_1 \geq 1/2$, 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. It is not hard to show that when $c_1 < 1/2$ so that 2 is the favored agent, 2's payoff is higher than 1's. 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.

As noted earlier, 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 any lie will be caught and result in i not receiving the good. Clearly, then, truth telling weakly dominates any lie for t_i.

Continuing to assume i is not favored, consider any t_i such that $t_i - c_i < v^*$.[5] 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.

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. Hence the favored–agent mechanism which is Bayesian optimal is also ex post optimal since it maximizes the objective function and satisfies the tighter constraints imposed by ex post incentive compatibility.

While the almost–dominance property implies a certain robustness of the mechanism, the complete indifference for types below the threshold is troubling. Fortunately, there are simple modifications of the mechanism which do not change its equilibrium properties but do 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 $\bar{t}_i - c_i > v^*$.[6] 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. 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.

[5] Since this is a set of measure zero, the optimality of the mechanism does not depend on how we treat reports with $t_i - c_i = v^*$. If we treat such reports the same way we treat reports with $t_i - c_i < v^*$ or the same way we treat reports with $t_i - c_i > v^*$, the same dominance arguments apply to these reports as well.

[6] 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 since it has no effect on the outcome. Hence we may as well disregard such agents.

4 Discussion

4.1 Proof Sketch

In this section, we sketch the proofs of Theorems 1 and 2. For simplicity, the proof sketches consider the case where $c_i = c$ for all i. In this case, the threshold value v^* can be thought of as a threshold type t^* to which we compare the t_i reports.

First, it is useful to rewrite the optimization problem as follows. Let $\hat{p}_i(t_i) = \mathrm{E}_{t_{-i}} p_i(t_i, t_{-i})$ and $\hat{q}_i(t_i) = \mathrm{E}_{t_{-i}} q_i(t_i, t_{-i})$ denote the interim probabilities. With this notation, we can write the incentive compatibility constraint as

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

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 $\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i$. Hence we can rewrite the objective function as

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

Both of the last two expressions for the objective function will be useful.

Hence 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.

Our proof of Theorem 1 works with finite approximations to the continuous type space, so the remainder of the argument we sketch here focuses on finite type spaces.

We prove the result in a series of four steps. The first step is to show 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 for this result, 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. 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 not giving it to anyone). 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 $p_1(t) = 1$. To see this, consider the point labeled $\alpha = (t_1^*, t_2^*)$ in Figure 1 below and note that $t_1^* > \hat{t}_1$ while $t_2^* < \hat{t}_2$. Suppose that at α, player 1 receives the good with probability strictly less than 1. It is not hard to see that at any point, such as the one labeled $\beta = (t_1', t_2^*)$ directly below α but above $\hat{t}_1$, player 1 must receive the good with probability zero. To see this, simply note that if 1 did receive the good with strictly positive probability here, we could change the mechanism by lowering this probability slightly and increasing the probability 1 receives the good at α. By choosing these probabilities appropriately, we do not affect $\hat{p}_2(t_2^*)$ 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.

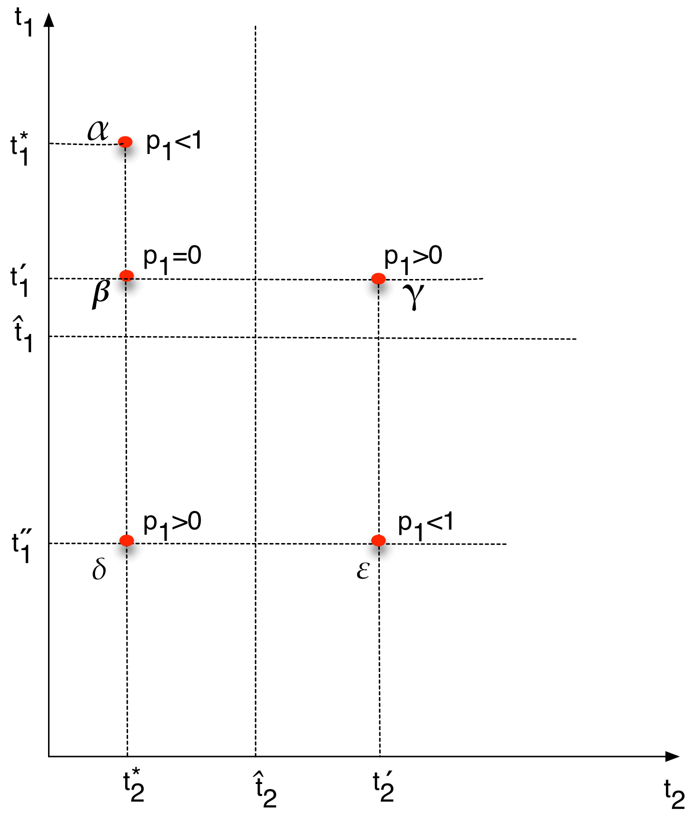

Figure 1

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

Finally, there must be a 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, note that $\hat{p}_2(t_2^*) = \varphi_2 \leq \hat{p}_2(t_2')$. On the other hand, $p_2(t_1', t_2^*) = 1 < p_2(t_1', t_2')$. So there must be some t_1''

where $p_2(t_1'', t_2^*) > p_2(t_1'', t_2')$. Hence p_2 must be strictly positive at δ, implying $p_1 < 1$ there. Similarly, $p_2 < 1$ at ε, implying $p_1 > 0$ there.

From this, we can derive a contradiction to the optimality of the mechanism. Lower p_1 at γ and raise it at ε in such a way that $\hat{p}_2(t_2')$ is unchanged. In doing so, keep the reduction of p_1 at γ small enough that $\hat{p}_2(t_1')$ remains above φ_1. This is clearly feasible. Now that we have increased p_1 at ε, we can lower it at δ in such a way that $\hat{p}_1(t_1'')$ remains unchanged. Finally, since we have lowered p_1 at δ, we can increase it at α in such a way that $\hat{p}_2(t_2^*)$ is unchanged.

Note the overall effect: $\hat{p}_1$ is unaffected at t_1'' and lowered in a way which retains feasibility at t_1'. $\hat{p}_2$ is unchanged at t_2^* and at t_2'. Hence the resulting p is feasible. But we have shifted some of the probability that 1 gets the object from γ to α. Since 1's type is higher at α, this is an improvement, implying that the original mechanism was not optimal.

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" is then the threshold of the threshold mechanism. To see that this establishes that the optimal mechanism is a threshold mechanism, recall the definition of such a mechanism. Restating the definition for the two agent case where $c_i = c$ for all i, a threshold mechanism is one where there exists t^* such that the following holds up to sets of measure zero. First, if there exists any i with $t_i > t^*$, then $p_i(t) = 1$ for that i such that $t_i > \max_{j \neq i} t_j$. Second, for any profile t, if $t_i < t^*$, then $q_i(t) = 0$ and $\hat{p}_i(t_i) = \min_{t_i' \in T_i} \hat{p}_i(t_i)$.

It is easy to see that our second and third steps above imply the first of these properties. 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.

It is also not hard to see that the second property of threshold mechanisms must be satisfied as well. By definition, if $t_i < t^*$, then $\hat{p}_i(t_i) = \varphi_i = \inf_{t_i'} \hat{p}_i(t_i')$. To see that this implies that i is not checked, recall that we showed $\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i$. Since $\hat{p}_i(t_i) = \varphi_i$ for $t_i < t^*$, we have $\hat{q}_i(t_i) = 0$ for such t_i. Obviously, $\mathrm{E}_{t_{-i}} q_i(t_i, t_{-i}) = 0$ if and only if $q_i(t_i, t_{-i}) = 0$ for all t_{-i}. Hence, as asserted, i is not checked.

In short, any optimal mechanism must be a threshold mechanism.

Turning to Theorem 2, note that Theorem 1 established that there is a threshold t^* with the property that if any t_i exceeds t^*, then the agent with the highest type receives the object. Also, for every i with $t_i < t^*$, $\hat{p}_i(t_i) = \varphi_i$. Since we can write the principal's payoff can be written as a function only of the $\hat{p}_i$'s — the interim probabilities — this implies that the principal's payoff is completely pinned down once we specify the φ_i's. It is not hard to see that the principal's payoff is linear in the φ_i's. Because of this and the fact that the set of feasible φ vectors is convex, there must be a solution to the principal's problem at an extreme point of the set of feasible $(\varphi_1, \ldots, \varphi_I)$.

Such extreme points correspond to identifying a favored agent. It is not hard to see that an extreme point is where all but one of the φ_i's is set to zero and the remaining one is set to the highest feasible value.[7] For notational convenience, consider the extreme point where $\varphi_i = 0$ for all $i \neq 1$ and φ_1 is set as high as possible.

As we now explain, this specification does not uniquely identify the mechanism, but identifies all but some of the probabilities of checking agent 1. In particular, we can resolve the remaining flexibility in such a way as to create a favored agent mechanism with 1 as the favored agent.

To see this, first observe that since $\varphi_i = 0$ for all $i \neq 1$, no agent other than 1 can receive the object if his type is below t^*, just as in the favored agent mechanism where 1 is favored. If all agents but 1 report types below the threshold and 1's type is above, then the properties of a threshold mechanism already ensure that 1 receives the object, just as in the favored agent mechanism. The only point left to identify as far as the allocation of the good is concerned is what happens when all agents are below the threshold. It is not hard to show that the statement "φ_1 is as high as possible" implies that 1 must receive the object with probability 1 in this situation. Thus as far as the allocation of the object is concerned, the mechanism is the favored agent mechanism with 1 as the favored agent.

Hence we have only the checking probabilities left to determine. Recall that $\hat{q}_i(t_i) = \hat{p}_i(t_i) - \varphi_i$. For $i \neq 1$, $\varphi_i = 0$, so $\hat{q}_i(t_i) = \hat{p}_i(t_i)$. Recall that $\hat{p}_i(t_i)$ is the expected probability that t_i is assigned the good, while $\hat{q}_i(t_i)$ is the expected probability that t_i is assigned the good and is checked. Thus $\hat{q}_i(t_i) = \hat{p}_i(t_i)$ says that the expected probability that t_i is assigned the good and *not* checked is zero. Of course, the only way this can be true is if for all t_{-i}, the probability that t_i is assigned the good and not checked is zero. Hence $q_i(t) = p_i(t)$ for all $i \neq 1$. That is, any agent other than the favored agent is checked if and only if he is assigned the good. Therefore, for agents other than the favored agent, the probability of being checked is uniquely identified and is as specified by the favored agent mechanism.

[7] $\varphi_i = 0$ for all i is also an extreme point, but is easily shown to be inferior for the principal.

For the favored agent, we have some flexibility. Again, we have $\hat{q}_1(t_1) = \hat{p}_1(t_1) - \varphi_1$, but now $\varphi_1 \neq 0$. Hence φ_1 is the expected probability that t_1 receives the good without being checked and $\hat{q}_1(t_1)$ is the expected probability t_1 receives the good with a check. These numbers are uniquely pinned down, but the way that $q_1(t_1, t_{-1})$ depends on t_{-1} is not unique in general. For $t_1 < t^*$, we know that $\hat{p}_1(t_1) = \varphi_1$, so $\hat{q}_1(t_1) = 0$ in this range. The only way this can be true is if $\hat{q}_1(t_1, t_{-1}) = 0$ for all t_{-1}. So, just as specified for the favored agent mechanism, if 1's type is below the threshold, he is never checked and only receives the good if all other agents are below the threshold.

However, consider some $t_1 > t^*$. The favored agent mechanism specifies that this type receives the good iff all other agents report types below t_1, a property that we have established. The favored agent mechanism also specifies that this type must be checked in this event. However, all we can pin down is this type's *expected* probability of getting the good with a check. More specifically, we know that $\hat{p}_1(t_1) = \mathrm{E}_{t_{-1}} p_1(t) = \Pr[t_j < t_1, \ \forall j \neq 1]$, that $\hat{q}_1(t_1) = \hat{p}_1(t_1) - \varphi_1$, and that $\varphi_1 = \Pr[t_j < t^*, \ \forall j \neq 1]$. One specification consistent with this is that of the favored agent mechanism where t_1 is checked if and only if $t_j < t_1$ for all $j \neq 1$ but $t_j > t^*$ for some $j \neq 1$. On the other hand, another specification consistent with this would be that t_1 receives the good with a check with probability $\hat{p}_1(t_1) - \varphi_1$ for *every* t_{-1}.

4.2 Extension: When Verification is Costly for Agent

A 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 a stronger 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. For example, in the example where the principal is a dean and the agents are academic departments, this assumption holds if each department wants the job slot independently of the value they would produce for the dean. 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 $\hat{c}_i^T$ be the cost incurred by agent i from being verified by the principal if he report his type truthfully and let $\hat{c}_i^F$ be his cost if he lied. We assume $1 + \hat{c}_i^F > \hat{c}_i^T \geq 0$. (Otherwise, verification costs hurt honest types more than dishonest ones.) Then 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 earlier, 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 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

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

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

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

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

In short, 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 $1 + \hat{c}_i^F$. Intuitively, since verification costs the agent, the threat of verification is more severe, so the principal doesn't need to check as often. In short, 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.

5 Conclusion

There are many natural extensions to consider. For example, in the previous subsection, we discussed the extension to where 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.

A second natural extension would be to allow costly monetary transfers. We argued in the introduction that within organizations, monetary transfers are difficult 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.

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.

References

[1] Banerjee, A., R. Hanna, and S. Mullainathan, "Corruption," working paper, January 2011.

[2] Border, K., and J. Sobel, "Samurai Accountant: A Theory of Auditing and Plunder," *Review of Economic Studies*, **54**, October 1987, 525–540.

[3] Bull, J., and J. Watson, "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, **58**, January 2007, 75–93.

[4] Deneckere, R. and S. Severinov, "Mechanism Design with Partial State Verifiability," *Games and Economic Behavior*, **64**, November 2008, 487–513.

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

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

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

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

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

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

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