UNIVERSITY OF
EXETER

Economics Department
Discussion Papers Series
ISSN 1473 – 3307

Optimal Allocation without Transfer Payments

Surajeet Chakravarty and Todd R. Kaplan

Paper number 10/04

URL: http://business-school.exeter.ac.uk/economics/papers/

Optimal Allocation without Transfer Payments[1]

Surajeet Chakravarty[2]   Todd R. Kaplan[3]

February 2010

JEL Classification: C70, D44, D89.

[2]School of Business and Economics, University of Exeter, Exeter, UK; email: S.Chakravarty@exeter.ac.uk
[3]School of Business and Economics, University of Exeter, Exeter, UK and Department of Economics, University of Haifa, Haifa, Israel; email: Dr@ToddKaplan.com.

Abstract

Often an organization or government must allocate goods without collecting payment in return. This may pose a difficult problem either when agents receiving those goods have private information in regards to their values or needs or when discriminating among agents using known differences is not a viable option. In this paper, we find an optimal mechanism to allocate goods when the designer is benevolent. While the designer cannot charge agents, he can receive a costly but wasteful signal from them. We find conditions for which ignoring these costly signals by giving agents equal share (or using lotteries if the goods are indivisible) is optimal. In other cases, those that send the highest signal should receive the goods; however, we then show that there exist cases where more complicated mechanisms are superior. Finally, we show that the optimal mechanism is independent of the scarcity of the goods being allocated.

1 Introduction

One of the basic problems in economics is how to allocate scarce resources or goods. One of the fundamental difficulties afflicting such allocation is private information: knowing who desires the goods the most. While markets work well with such allocation, the market is not always a feasible or desired mechanism for allocation. In case of kidneys it may be unethical to have a market, while in case of sports or concert tickets it may be undesirable to sell the tickets to the highest bidder.[1] Finally, with the allocation of charitable goods, it is not only undesirable to collect payment in return but those needing it the most are also the least able to pay for it.[2] Hence, we often see markets being replaced with other mechanisms.

One method used is instead of goods being allocated to the person who is willing to pay the most; they are allocated to who is willing to work the hardest to get them. Sport and concert tickets are given, often using first-come first-serve mechanism, that is, whoever is willing to wait the longest before the promoters start selling, gets the right to buy tickets. Allocation of research funds by agencies like National Science Foundation in USA and Economic and Social Science Research Council in UK to various universities and individuals are done based on research proposals (where a well-crafted proposal has a higher chance of being funded). A common feature in these examples is that in order to convey their valuation, individuals must incur a socially wasteful cost. As with waiting overnight in a long line, generally at least part of this effort is socially wasted.[3]

Another mechanism that is common with charity, but, surprisingly, also prevalent elsewhere, is to allocate evenly or randomly using a lottery (among those appearing identical when classified according to public information).[4] Often baseball playoff tickets are offered

---

[1]See Roth et al. (2004) and Roth (2007) for a description of the current method used to allocate kidneys and the perceived ethical difficulties (repugnance) of moving to a market-based system for organ donation and other potentially distasteful transactions. With tickets, there is sometimes a desire for a wider audience. Indeed, the Metropolitan Opera in New York received a several million dollar grant to widen audience by selling prime orchestra tickets for $20 each, 10 percent of their usual price (*USA Today,* October 5, 2006).

[2]Che and Gale (2006) provide further examples of non-market allocation caused by wealth constrained agents.

[3]Without grant money at stake, most researchers would not start a project by writing a detailed, polished research proposal. This indicates at least some of the effort is wasted (used inefficiently).

[4]To avoid confusion, we call a lottery as a mechanism that randomly allocates an object or objects. Unless

via a lottery.[5] Likewise, NCAA College bowl tickets have a lottery amongst only the season ticket holders. Research funds are often handed evenly amongst certain groups or individuals. For example, most universities hand out fixed research grants to all new staff. Allocating goods equally (ex-ante) has the disadvantage of ignoring any private information, but has the advantage of saving the potential recipients' effort.

In this paper, we find the optimal mechanism to allocate homogeneous, not-necessarily-divisible goods when the bids made by the agents competing for the goods are socially wasted. Initially, we maximize the social surplus (ex-ante optimality) when values and costs of signalling is private information. We find the necessary and sufficient conditions for when allocating the goods randomly is optimal. In addition, we find the necessary and sufficient conditions for when distributing the objects to those who work the hardest (a contest) is optimal. We also find cases when other mechanisms can be optimal, such as using a contest but randomly allocating the objects amongst any that meet a certain threshold of effort (a contest with a bid cap). One interesting result of our paper is that an optimal mechanism does not depend upon scarcity of the goods being allocated. We also show that our results extend to where a designer may favor one type over another (other interim efficient allocations).

The intuition that drives our results are that using signals increases the probability that the good will be allocated to the person who values it the most; however, this naturally also increases the costs due to signals being wasted. Which mechanism is optimal depends upon this trade-off determined by the distribution of values and signalling costs.

There are many papers examining contests or lotteries, but as opposed to this paper, most study the case where a seller wishes to maximize revenue. Amongst these, Moldovanu and Sela (2001) study the best way to split prize money in a contest, and Gavious, Moldovanu

---

mentioned otherwise, everyone is given an equal chance. Note this usage of lottery distinguishes it from a raffle in that in a raffle chances are sold.

[5]More precisely, the price is set below the market clearing price. Since the demand exceeds supply, a lottery was used to determine who has the right to buy tickets. Among the baseball teams that have used a lottery system was 2006 New York Mets, 2007 Cleveland Indians, 2008 Chicago Cubs, and the 2009 Philadelphia Phillies (see Mucha, 2009).

and Sela (2002) analyze contests, where depending on the nature of the cost function bid caps may be more profitable or not. Also, Goeree et al. (2005) rank lotteries and contests in fund raising mechanisms and Fullerton and McAfee (1999) model research tournaments and show that it is optimal to limit the number of participants to two.

One paper that does examine allocation with a benevolent designer and thus close in spirit to our paper is Che and Gale (2009). They find that when agents have wealth constraints in a pure market, those that value goods the most cannot necessarily afford them. Hence, sometimes a random allocation can be superior.[6] Again with a benevolent designer, Hoppe, Moldovanu and Sela (2009), compare match making between two groups (such as men with women) by pairing those displaying the highest costly signal to that of random matching. Hartline and Roughgarden (2008) study the benefit of wasteful signalling in a computer science application. Finally, Condorelli (2009) and Yoon (2009) both study a similar design problem to that in this paper but under different assumptions.[7]

In the next section we discuss the allocation problem and convert it into a mechanism design problem in Section 3. In Section 4, we present the results of our analysis. Finally, in Section 5, we make our final remarks and present our conclusions.

2   Allocation Problem

The designer's problem is to allocate M homogeneous, not-necessarily-divisible goods among N agents where $M < N$. The designer is benevolent and wishes to maximize the social surplus. Each agent i has a privately known type $\theta_i \in [0, 1]$ that is drawn independently

---

[6]Another path to solving the allocation problem is by using psuedo-market systems where exogenously given points are substituted for money. These are used in the allocation of both interviews and courses for MBA students (see Brams and Taylor, 1996; Brams and Kilgour, 2001; Sönmez and Ünver, 2005). However, to work, these require more than one type of objects to be allocated (for an alternative use of points) and may be costly to implement.

[7]More specifically, Condorelli (2009) unlike this paper assumes all agents have the same signalling costs (as with Hartline and Roughgarden, 2008), but allows ex-ante asymmetry between agents. Yoon (2009) restricts design to a mechanism where the probability of an agent receiving an object depends only his rank among the order of signals, but allows a proportion of the signal to have public benefit. In our framework, in some cases the optimal mechanism is not one of those considered by Yoon.

from cumulative distribution F. Agent i has value $v(\theta_i)$ for at most one object, where $v(\theta_i)$ is strictly positive for $\theta_i > 0$, continuous and twice differentiable. (If goods are divisible, the value to agent i is $\min\{q_i, 1\} \cdot v(\theta_i)$ where $q_i \geq 0$ is the fraction of good agent i receives.) Each agent i is able to send a costly signal $x_i \in \mathbb{R}_+$ to the designer. The cost to the agent of sending signal x_i depends upon his type and equals $c(x_i) \cdot g(\theta_i) \geq 0$, where $c(x_i)$ is weakly positive, continuous and strictly increasing while $g(\theta_i)$ is strictly positive for $\theta_i < 1$, continuous and twice differentiable. The function $g(\theta_i)$ captures how the agent's type affects the cost of signalling. So if for instance $g(\theta_i) = \theta_i^{-1}$, then the higher the type of the agent, the less costly it is for him to send a high signal. Likewise, if $g(\theta_i) = \theta_i + 1$, then the higher the type of the agent, the more costly it is for him to send a high signal. When $g(\theta_i)$ depends upon θ_i, the designer is able to see the signal x_i, but does not know the agent's cost of sending the signal. For instance, if the signal is standing in line x_i hours, the designer is able to see how long the agent stands in line, but is unable to determine the (opportunity) cost to the agent.

Following Milgrom (2004, page 111), without loss of generality we can rewrite our problem using the uniform distribution on $[0, 1]$ in place of F.[8] Finally, we assume that $v(\theta_i)/g(\theta_i)$ is weakly increasing for $0 < \theta_i < 1$.[9] This can be interpreted as higher the type, the higher the maximal willingness to send a signal for the object.[10] This is equivalent to the assumption that $v'(\theta_i)/v(\theta_i) \geq g'(\theta_i)/g(\theta_i)$ for $0 < \theta_i < 1$. This condition assures that if an agent of type θ_i is willing to send signal x to have a certain chance of receiving the object then all agents with types $\widetilde{\theta}_i \geq \theta_i$ will also be willing to send that signal for the same chance (single crossing).[11] We also assume that v, g are analytic.

---

[8]By having θ drawn from an arbitrary distribution with v and g functions of θ, we have a extra degree of freedom. By assuming a uniform distribution, we eliminate this extra degree of freedom and gain simplicity of the expressions in the paper. Also by doing so, we are able to continue to treat $v(\theta)$ and $g(\theta)$ even-handedly (not write one in terms of the other).

[9]This is fairly innocuous since we can always reorder v and g to obtain continuous functions that do satisfy this condition (and approximate them using twice differentiable functions).

[10]The maximal willingness m is such that $v(\theta) - g(\theta)m = 0$ or $m = v(\theta)/g(\theta)$. Note m is in terms of cost, that is, one would be willing to send a signal x in order to receive the object if $c(x) \leq m$. With this interpretation it is again clear one can reorder types to have v/g weakly increasing.

[11]Let us call the chance of receiving the object W (we make use of this notation later in the paper). An agent of type θ_i will be willing to send a signal x to receive the object with that chance would so if the

After the agents send costly signals, the designer receives these signals $(x_1, \ldots, x_N)$ and uses them to allocate the M goods by rule $a : \mathbb{R}_+^N \rightarrow [0,1]^N$ where $\sum_i a_i(x_1, \ldots, x_N) \leq M$ guarantees feasibility. (Note that a_i indicates the probability that agent i receives the good when the goods are indivisible and the fraction of the good received.)[12] Denote A as the set of feasible allocation rules. Given allocation rule a, the agents form a Bayes-Nash equilibrium by choosing a strategy $x_i(\theta_i, a)$ to maximize their expected surplus given the strategies of other agents. The designer's problem is to choose rule a to maximize the equilibrium social surplus of the agents given the future Bayes-Nash equilibria of the agents, that is, the designer solves

$$\max_{a \in A} \sum_i E[v(\theta_i) \cdot a_i(x_1(\theta_1, a), \ldots, x_N(\theta_N, a)) - c(x_i(\theta_i, a)) \cdot g(\theta_i)].$$

3 Mechanism Design Problem

For simplicity of analysis we will invoke the revelation principle and look at direct mechanisms where each agent i sends a costless (but not necessarily truthful) signal $\widetilde{\theta}_i$.[13] Given the set of signals $\{\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N\}$, the mechanism gives an object to agent i with probability $W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_i, \ldots, \widetilde{\theta}_N)$. (Under divisibility, this will represent the fraction good that agent i receives.)[14] Likewise, the mechanism charges agent i an amount $e_i(\widetilde{\theta}_i)$. Note that this charge depends only on $\widetilde{\theta}_i$. Feasibility requires $\sum_i W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N) \leq M$. Although the agent incurs a cost $e_i(\widetilde{\theta}_i) \cdot g(\theta_i)$, the designer does not receive any benefit from the signal $\widetilde{\theta}_i$, that is, the cost actually incurred by the agent is wasted. The mechanism is truthful if it is incentive compatible (IC) to report truthfully and individually rational (IR) to participate. Once we

---

expected payoff $W \cdot v(\theta_i) - c(x) \cdot g(\theta_i) \geq 0$. The derivative of this payoff w.r.t. θ_i is $W \cdot v'(\theta_i) - c(x) \cdot g'(\theta_i)$. From the inequality (and that v is strictly positive) we have $W \geq c(x) \cdot g(\theta_i)/v(\theta_i)$. This shows that the derivative is larger than $c(x)[(g(\theta_i)/v(\theta_i)) \cdot v'(\theta_i) - g'(\theta_i)]$ of which in turn is larger than zero.

[12] Note that there is free disposal. Also, we consider giving an agent more than measure one of the good the equivalent to giving the agent measure one and disposing of the excess.

[13] Note that there is no benefit to having a sequential mechanism such as in Crémer et al. (2009) since all parties know their private information initally.

[14] A lottery would allocate objects with probability $W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N) = M/N$. A contest with $M = 1$ would allocate objects according to $W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N)$ equals 0 if $\widetilde{\theta}_i < \max\{\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N\}$ or equals $1/\#$ if $\widetilde{\theta}_i = \max\{\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N\}$ where $\#$ is the number of j where $\widetilde{\theta}_j = \max\{\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N\}$.

solve for the optimal direct mechanism, then we can implement the solution by choosing an appropriate allocation rule, that is by setting $a_i(x) = W_i(e_1^{-1}(c(x_1)), \ldots, e_N^{-1}(c(x_N)))$, we have $c(x_i(\theta_i, a)) = e_i(\theta_i)$. Notice that this can be implemented since $e_i(\theta_i)$ does not depend upon θ_j $(j \neq i)$.

Now by limiting ourselves to symmetric mechanisms, we can denote $W(\widetilde{\theta}_i)$ as the probability of agent i receiving an object with message $\widetilde{\theta}_i$ when everyone else reports truthfully and $e(\widetilde{\theta}_i)$ as the expected cost given that everyone else reports truthfully.[15,16] For simplicity of notation, we drop the i subscript. Both $W(\widetilde{\theta})$ and $e(\widetilde{\theta})$ are assumed to be increasing in $\widetilde{\theta}$. Now an agent of type θ reporting $\widetilde{\theta}$ (with all others reporting truthfully) has payoff $\pi(\theta, \widetilde{\theta}) \stackrel{def}{=} W(\widetilde{\theta})v(\theta) - e(\widetilde{\theta}) \cdot g(\theta)$. The agent solves $\max_{\widetilde{\theta}} \pi(\theta, \widetilde{\theta})$ which in a truthful mechanism should equal $\pi(\theta, \theta)$.

The designer chooses $W(\widetilde{\theta})$ and $e(\widetilde{\theta})$ to maximize $N \cdot E[\pi(\theta, \theta)] = N \cdot E[W(\theta) \cdot v(\theta) - e(\theta) \cdot g(\theta)]$ subject to $W(\theta)$ being consistent with feasibility, IC $(\pi(\theta, \theta) \geq \pi(\theta, \widetilde{\theta}))$ and IR $(\pi(\theta, \theta) \geq 0)$. Before getting to our results, in the following lemma we simplify the designer's problem by eliminating from the problem, the IC and IR constraints as well as the choice of $e(\theta)$.

Lemma 1 *The designer's problem reduces to choosing $W(\theta)$ that is increasing and consistent with feasibility in such a way to maximize the social surplus given by*

$$N \int_0^1 W(\theta) \left(\frac{v(\theta)}{g(\theta)} \right)' \left(\int_\theta^1 g(\hat{\theta}) d\hat{\theta} \right) d\theta + N \ W(0) \frac{v(0)}{g(0)} \int_0^1 g(\theta) d\theta - N \ W(1)\ell_1, \quad (1)$$

where $\ell_1 = \lim_{\theta \to 1} \left\{ \frac{v(\theta)}{g(\theta)} \int_\theta^1 g(\hat{\theta}) d\hat{\theta} - v(\theta)(1 - \theta) \right\}$.

Proof. See the Appendix. ∎

---

[15] Note it is not optimal to restrict the number of participants as opposed to other environments where the designer values the signals such as Fullerton and McAfee (1999), indicative that our symmetric assumption is not crucial.

[16] Note that in order to be able to determine the allocation rule, we would need to decompose $W(\theta)$ into $W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N)$. This is not a problem since any mechanism described by $W(\theta)$ that is feasible with increasing $W(\theta)$ can be decomposed into a mechanism $W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N)$ that is symmetric and satisfies monotonicity: if $\widetilde{\theta}_i > \widetilde{\theta}_i'$ then $W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_i, \ldots, \widetilde{\theta}_N) \geq W_i(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_i', \ldots, \widetilde{\theta}_N)$. Likewise, any mechanism that satisfies monotonicity and symmetry has an increasing $W(\theta)$.

Notice that as one may expect increasing $v(\theta)$ to $\alpha \cdot v(\theta)$ increases surplus by factor α. Also, somewhat less obvious, increasing $g(\theta)$ to $\alpha \cdot g(\theta)$ does not alter the surplus. This is for the same reason that our cost function, $c(x)$, does not appear in the surplus. Namely, if agents have proportionately lower waiting costs, they will simply dissipate the gains by proportionately increasing the cost of effort.

We can interpret the designer's problem of Lemma (1) in terms of virtual surplus.[17] The designer wishes to maximize the expected virtual surplus of someone receiving a good, that is, the expected value of $W(\theta)$ times the virtual surplus of type θ. Actually, while this terminology is sufficient for standard mechanism design problems, we need to expand the terminology into two types of virtual surplus. The expression $s(\theta) \stackrel{def}{=} \left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right)$ is the *marginal virtual surplus* for $\theta \in (0,1)$. There are also what can be thought of as *fixed virtual surpluses* at the end points of 0 and 1. Denote these as $S_0 \stackrel{def}{=} \frac{v(0)}{g(0)} \int_0^1 g(\widetilde{\theta})d\widetilde{\theta}$ and $S_1 \stackrel{def}{=} -\ell_1$. The expression for social surplus in (1) now becomes:

$$N \left[\int_0^1 W(\theta)s(\theta)d\theta + W(0)S_0 + W(1)S_1 \right]$$

4   Results

4.1   Optimal Mechanism

We now use the simplified designer's problem to solve for an optimal mechanism in a manner similar to Myerson (1981).[18] Using the notation of the previous section, we now denote the cumulative virtual surplus as

---

[17]Virtual surplus is the surplus generated by giving the object to someone of type θ taking into account the full cost this has on incentives. Namely, the cost to designer of keeping the other (higher) types truthtelling rather than pretending to be a lower type. For use of virtual surplus in the more standard mechanism design framework, see Fudenberg and Tirole, 1991, page 266.

[18]There is a unique ex-ante optimal allocation if the set $\{\theta \in [0,1]| \ d\left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right)/d\theta = 0\}$ has measure 0.

$$
z(\theta) \stackrel{def}{=} \begin{cases} 0 & \text{if } \theta = 0, \\ S_0 + \int_0^\theta s(\widetilde{\theta})d\widetilde{\theta} & \text{if } \theta < 1, \\ S_0 + \int_0^1 s(\widetilde{\theta})d\widetilde{\theta} + S_1 & \text{if } \theta = 1. \end{cases}
$$

The expression $z(\theta)/\theta$ is the average virtual surplus of all types from 0 to θ. Note that Myerson does not have the fixed virtual surplus (and of course the marginal virtual surplus is a different expression) but a similar technique does indeed work.

While this is non-standard, one can understand this better by using the analogy of a cost function $c(y)$ where $c'(y)$ is the marginal cost, $c(y)/y$ is the average cost, and $\lim_{y \searrow 0} c(y)$ the fixed cost.[19]

Let $\mathcal{C}(\theta)$ be the convex envelope of $z(\theta)$, that is, $\mathcal{C}(\theta) = Conv(z(\theta))$ where

$$
Conv(z(\theta)) = \inf_{r,\theta_1,\theta_2} rz(\theta_1) + (1-r)z(\theta_2)
$$
$$
\text{s.t. } r, \theta_1, \theta_2 \in [0,1] \text{ and } r\theta_1 + (1-r)\theta_2 = \theta.
$$

Since $\mathcal{C}$ is convex, $\mathcal{C}'' \geq 0$.

Proposition 1 *An optimal mechanism has the following allocation: the interval $[0,1]$ can be divided into regions where on each region either $\mathcal{C}''(\theta) > 0$ (a contest) or $\mathcal{C}''(\theta) = 0$ (a lottery). While the signals are sent simultaneously, goods are allocated sequentially after (truthful) signals are sent in the following steps:*

(a) If the highest type falls within a contest region, then the agent of that type receives the good.

(b) Otherwise, there is a lottery among all agents whose types fall within that region.

(c) The agent receiving the good is then removed and the process is repeated with the remaining agents until all goods have been allocated.[20,21]

---

[19] We may have $\lim_{\theta \nearrow 1} z(\theta) >$ or $< z(1)$. This has no parallel with cost functions but becomes important with monotonicity of the probability of receiving an object.

[20] Note that by removing an agent each time an object is allocated, we may move from one region to another. This occurs if there are no remaining agents with types in the original region.

[21] When $\ell_1 < 0$, we define $\theta = 1$ as a contest region.

Proof. See the Appendix. ∎

The required monotonicity of the $W(\theta)$ function causes the allocation to be in regions of contests and lotteries. If a designer needing to allocate one object between two agents receives signals of θ_1 and θ_2, where $\theta_1 > \theta_2$, then the designer can either give the object to the player with θ_1 or randomly allocate the object between both players. If the designer receives signals $\widehat{\theta}_1$ and θ_2, where $\theta_1 > \widehat{\theta}_1 > \theta_2$ and would choose a lottery between θ_1 and θ_2, then he should also choose a lottery between $\widehat{\theta}_1$ and θ_2. If the designer receives signals $\widehat{\widehat{\theta}}_1$ and θ_2, where $\widehat{\widehat{\theta}}_1 > \theta_1 > \theta_2$ and would choose θ_1 over θ_2, then he should also $\widehat{\widehat{\theta}}_1$ over θ_2. (We explain this further in the examples).

In order to see why the convex envelope $\mathcal{C}$ yields the optimal allocation, start with allocating objects according to the highest type. Take the graph of $z(\theta)$. If there is a region, $[\theta_1, \theta_2]$, that is concave (decreasing slope) on $[\theta_1, \theta_2]$, then allocating this by the highest type would not yield the highest revenue since one would more likely be giving it to someone with a lower virtual surplus (lower slope of z). Replacing this with a lottery on $[\theta_1, \theta_2]$ will yield higher surplus.[22] The average virtual surplus in that region is $(z(\theta_2) - z(\theta_1))/(\theta_2 - \theta_1)$. This is equal to the slope of the line from $(\theta_1, z(\theta_1))$ to $(\theta_2, z(\theta_2))$ on the graph of $z(\theta)$. We can thus replace the region with a new $z(\theta)$ equal to the old $z(\theta)$ except with a line between those points. Now, it is worthwhile to expand any regions by including the type at θ_2 to the lottery on $[\theta_1, \theta_2]$ if slope of $z(\theta_2)$ at θ_2 is smaller than the slope of the line from $(\theta_1, z(\theta_1))$ to $(\theta_2, z(\theta_2))$. Thus, the region should be expanded until the average virtual surplus in the lottery region stops decreasing. Graphically, this is until the line from $(\theta_1, z(\theta_1))$ to the end of the lottery stops decreasing in slope.[23] By repeating this process, we arrive at the convex

---

[22] If f_1 (the probability of receiving an object) is strictly increasing and f_2 (the virtual surplus) is strictly decreasing, then $\int_{\theta_1}^{\theta_2} f_1(\theta) f_2(\theta) d\theta < \frac{\int_{\theta_1}^{\theta_2} f_1(\theta) d\theta}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} f_2(\theta) d\theta$. (Note the LHS is the expected virtual surplus when using a contest in the region and the RHS is the expected virtual surplus when using a lottery in the region.) This inequality holds since $\mu \overset{def}{=} \frac{\int_{\theta_1}^{\theta_2} f_1(\theta) d\theta}{\theta_2 - \theta_1}$ is the average of f_1 on $[\theta_1, \theta_2]$. The inequality now becomes $\int_{\theta_1}^{\theta_2} (f_1(\theta) - \mu) f_2(\theta) d\theta < 0$. Define θ^* such that $\mu = f_1(\theta^*)$. We now have $\int_{\theta_1}^{\theta_2} (f_1(\theta) - \mu) f_2(\theta) d\theta = \int_{\theta_1}^{\theta^*} (f_1(\theta) - \mu) f_2(\theta) d\theta + \int_{\theta^*}^{\theta_2} (f_1(\theta) - \mu) f_2(\theta) d\theta < \int_{\theta_1}^{\theta^*} (f_1(\theta) - \mu) f_2(\theta^*) d\theta + \int_{\theta^*}^{\theta_2} (f_1(\theta) - \mu) f_2(\theta^*) d\theta = 0$.

[23] The beginning $\theta = 0$, can be thought of as a degenerate lottery region. If $\frac{v(0)}{g(0)} \int_0^1 g(\widetilde{\theta}) d\widetilde{\theta} > 0$, then there is the equivalent of a fixed cost and average virtual surplus would always initially be decreasing resulting in

envelope $\mathcal{C}$ which generates the optimal allocation.

Again returning to our analogy of a cost function, one can understand the above intuition by using a graph of $c(y)$ with a fixed cost (assume $c''(y) > 0$). For instance, $c(y) = 1 + y^2$. At what point on the graph does average cost in stop decreasing? The average cost is simply $c(y)/y$ and is represented by the slope of the line from the origin to the point $(y, c(y))$. It stops decreasing when this line is tangent to $c(y)$ (where average cost equals marginal cost). Beyond this point, marginal cost is above average cost and below this line marginal cost is below average cost. Basically, at each point θ, the designer is choosing between the marginal z and average z over a region leading up to θ. So the designer chooses a contest when $z'(\theta)$ is larger than the average $z(\theta)$ in any the region directly below it and if not the designer wants to choose a lottery.

Now making use of Proposition 1, we find which mechanisms will result in the optimal allocation.

4.2   When a Lottery is Optimal

We now with help of Proposition 1, we can specify necessary and sufficient conditions for a lottery to be optimal.

Proposition 2 *A lottery is an optimal mechanism if and only if for all θ, the average virtual surplus up to θ is greater than the (overall) average virtual surplus, that is, $z(\theta)/\theta \geq z(1)$ for all θ.*

Proof. The mechanism presented in Proposition 1 will be a lottery if and only if $\mathcal{C}''(\theta) = 0$ for all θ. Graphically, this means that $z(\theta)$ must be above the line from $(0, z(0))$ to $(1, z(1))$. This will cause the convex envelope to be obtained from this line, that is, $\mathcal{C}(\theta) = (1 - \theta) z(0) + \theta z(1) = \theta z(1)$ and hence $\mathcal{C}''(\theta) = 0$. Thus, a lottery is optimal if and only if $z(\theta) \geq \mathcal{C}(\theta) = \theta z(1)$. ∎

a region of a lottery starting at 0 to be part of the optimal allocation.

Example 1 *When $N = 2$, $M = 1$, which has higher surplus: (a) holding one lottery or (b) holding a lottery among types below a θ^* and a lottery among types above a θ^*?*

Before doing the calculations, let us explain how such a mechanism would work in practice. While a designer does not know the types of the agents, he can design a mechanism that allocates based upon types (by using the signals). The above scheme can be implemented by having a minimum signal. Any object is first allocated randomly to any agent sending that minimum signal or higher. Objects left over (if any) are randomly allocated to the remaining agents. For instance, Duke basketball tickets are randomly allocated by having students expend effort in waiting over several weeks (see Ariely, 2008, for a description of the mechanism). Tickets are randomly allocated among students in front of a certain point in line. Only agents with types above a certain value will be willing to send a signal above the minimal signal.

Now let us look at the specific calculation. The overall lottery yields higher surplus over the two lotteries if and only if $z(1) \geq \theta^{*2}\left(\frac{z(\theta^*)}{\theta^*}\right) + (1 - \theta^{*2})\left(\frac{z(1) - z(\theta^*)}{1 - \theta^*}\right) = (1 + \theta^*)z(1) - z(\theta^*)$, yielding $z(\theta^*) \geq \theta^* z(1)$. We see this in agreement with Proposition 2. $\square$

While it is useful to have necessary and sufficient conditions for a lottery to be optimal, we find it also useful to examine particular sufficient conditions as well.

Proposition 3 *(i) If $\left(\frac{v(\theta)}{g(\theta)}\right)'' \leq 0$, then a lottery is optimal.*
(ii) If $v(1)$ is bounded and $g(1) > 0$, then there exists an $\alpha > 0$ where for the same g and a different $\widehat{v}(\theta)$, where $\widehat{v}(\theta) \overset{def}{=} v(\theta) + \alpha g(\theta)$, a lottery will be the optimal mechanism.[24]

Proof. First, we prove (i). If $\left(\frac{v(\theta)}{g(\theta)}\right)'\left(\int_\theta^1 g(\hat\theta)d\hat\theta\right)$ is weakly decreasing and $\ell_1 \geq 0$, then a lottery is an optimal mechanism. Begin by looking at the first part of the expression in Lemma 1. Since the part inside the integral (not considering $W(\theta)$) is decreasing, the integral is maximized with a lottery (flat $W(\theta)$). Now notice that the second part of the expression in Lemma 1 is maximized by maximizing $W(0)$, this happens also with a lottery

---

[24] It is also possible to show that if $(1 - \theta)v'(\theta)$ is decreasing in θ and $g'(\theta) \geq 0$, then a lottery is socially optimal.

(since W must be weakly increasing). Finally, since $\ell_1 \geq 0$, the third part is maximized by minimizing $W(1)$, which also happens with a lottery.

Now notice that

$$\frac{d\left[\left(\frac{v(\theta)}{g(\theta)}\right)'\left(\int_\theta^1 g(\hat\theta)d\hat\theta\right)\right]}{d\theta} = \left(\frac{v(\theta)}{g(\theta)}\right)''\left(\int_\theta^1 g(\hat\theta)d\hat\theta\right) - \left(\frac{v(\theta)}{g(\theta)}\right)'g(\theta).$$

By the mean value theorem, we have

$$\left(\frac{v(\theta)}{g(\theta)}\right)''(1-\theta)\,g(\theta^*(\theta)) - \left(\frac{v(\theta)}{g(\theta)}\right)'g(\theta)$$

where $\theta^*(\theta)$ is between θ and 1. If $\left(\frac{v(\theta)}{g(\theta)}\right)'' \leq 0$, then this expression is negative. If $\left(\frac{v(\theta)}{g(\theta)}\right)'' \leq 0$, then $\frac{v(1)}{g(1)}$ is bounded. If $v(1)$ is also bounded, then $\ell_1 = 0$. If $v(1)$ is unbounded, then for $\frac{v(1)}{g(1)}$ to be bounded we must have $g(\theta)$ increasing as θ approaches 1. If so, by the mean value theorem, $\frac{v(\theta)}{g(\theta)}\int_\theta^1 g(\hat\theta)d\hat\theta - v(\theta)(1-\theta) \geq 0$ and hence $\ell_1 \geq 0$.

The proof of part (ii) is in the Appendix. ∎

Note that part (i) that a lottery is optimal if v/g is weakly concave. From the surplus, slope of v/g does not affect the decision upon which mechanism to employ. However, if this slope is increasing, it becomes increasingly expensive in terms of value for the high types to signal. In this case, the messages made by the higher-value agents are too costly for the society to waste and therefore it is better to run a lottery, that is, allocate the good randomly.[25] While the proof of part (ii) is not particularly straightforward, the intuition of the result is clear. Since $\hat v(\theta)/g(\theta) = v(\theta)/g(\theta) + \alpha$, a larger α makes the different types relatively close in terms of $v(\theta)/g(\theta)$. Hence, when the differences between types become relatively smaller, that is, a larger part of the value is common, a lottery would be the optimal mechanism. We now see that the sufficient conditions can be useful in finding examples where

---

[25]With possibility of costless messages, a lottery is only the unique optimal mechanism among monotonic mechanisms (as described in footnote 16). For instance, when $N = 2$, $M = 1$, one can ask players to send a costless signal s_i of their types. If $s_i > s_j$ then only if $s_i - s_j < 0.5$, the mechanism would allocate it to i. If players truth tell, the expected probability of winning will be the same. Monotonicity is violated since if $s_2 = 0.8$, player 1 wins with a signal of $s_1 = 0.1$, but loses with a signal of $s_1 = 0.4$.

a lottery is optimal.

Example 2 $N = 2$, $M = 1$, $v(\theta) = \theta$, and $g(\theta) = 1$.

In this case, $\left(\frac{v(\theta)}{g(\theta)}\right)'' = 0$. Thus, the social planner does better by running a lottery, with surplus $\int_0^1 (1 - \theta)\, d\theta = \frac{1}{2}$, than a contest, with surplus $2\int_0^1 (1 - \theta)\theta\, d\theta = \frac{1}{3}$, (or any other mechanism). $\square$

This contrasts with a mechanism that maximizes revenue when payments can be collected. Namely, types $\theta < \frac{1}{2}$ are excluded and the agent with the highest θ $(\geq \frac{1}{2})$ receives the object.

Example 3 $v(\theta) = (\gamma \cdot \theta + c)^\alpha$ and $g(\theta) = (\gamma \cdot \theta + c)^\beta$ where $\alpha, c, \gamma > 0$ and $0 < \alpha - \beta \leq 1$.

In this case, $\left(\frac{v(\theta)}{g(\theta)}\right)'' = \gamma^2(\alpha - \beta)(\alpha - \beta - 1)(\gamma \cdot \theta + c)^{\alpha - \beta - 2} \leq 0$. Hence, a lottery is socially optimal. $\square$

This is a generalization of the previous example.[26] Notice that a lottery can be optimal when $g'(\theta) < 0$ or $g'(\theta) > 0$.

4.3   When a Contest is Optimal

We can also use Proposition 1 to specify the necessary and sufficient conditions for a contest to be optimal.[27]

Proposition 4 *A contest is an optimal mechanism if and only if $s(\theta)$ is weakly increasing, $S_0 = 0$ and $S_1 \geq 0$.*

---

[26] When $c = 0$, a lottery is still socially optimal even though our assumption that $g(0) > 0$ is violated. This is because $\lim_{\theta \searrow 0} \frac{v(\theta)}{g(\theta)} = 0$.

[27] Such conditions would require that the designer would allocate to the objects to the agent with the highest type. In the standard auction framework, first-price, second-price and all-pay auctions would all achieve this with the same expected payments for each agent of a particular type. However, a first-price or second-price auction would first require a signal to be sent and then payment to be paid. In our model, this would require both a costless initial signal and a wasteful payment to be made after the allocation decision has been decided. It seems to us that a contest is a more viable mechanism both politically and for credible commitment reasons. For instance, if we have agents write down how long they would be willing to wait in line for an object and then choose the agent willing to wait the longest to receive the object. There would be pressure to not actually have the agent waste his time waiting for the specified time. A contest (or regions with contests) solves this issue.

Proof. This follows from Proposition 1. A contest is optimal only if $\mathcal{C}''(\theta) > 0$ for all θ. This will happen only if $\mathcal{C}(\theta) = z(\theta)$ for all θ. If $\mathcal{C}(\theta) = z(\theta)$ for all θ, then a contest is optimal (since in the regions where $\mathcal{C}''(\theta) = 0$ and $\mathcal{C}(\theta) = z(\theta)$, the designer would be indifferent to a contest and a lottery). We have $\mathcal{C}(\theta) = z(\theta)$ if and only if $S_0 = 0$, $z''(\theta) \geq 0$, and $S_1 \geq 0$. Well, $z''(\theta) \geq 0$ if and only if $s(\theta)$ is weakly increasing. ∎

The conditions in the Proposition ensure that replacing a region with a lottery (as discussed in section 4.1) is never beneficial.

Corollary 1 *For a contest to be optimal, we must have*

(i) $\left(\frac{v(\theta)}{g(\theta)}\right)'' \geq 0$ *for all* θ,

(ii) $v(0)/g(0) = 0$,

(iii) and either $v(1)$ *unbounded or* $g(1) = 0$.

Proof. From Proposition 4, $s(\theta)$ must be weakly increasing and $S_0 = 0$ for a contest to be optimal (we do not make use of the condition $S_1 \geq 0$). We prove the Corollary by showing: (i) is necessary for $s(\theta)$ to be weakly increasing, (ii) is necessary to have $S_0 = 0$, and (iii) is necessary for $s(\theta)$ to be increasing near 1.

From the proof of Proposition 2 (i), we have

$$s'(\theta) = \frac{d\left[\left(\frac{v(\theta)}{g(\theta)}\right)'\left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right)\right]}{d\theta} = \left(\frac{v(\theta)}{g(\theta)}\right)''(1-\theta)\,g(\theta^*(\theta)) - \left(\frac{v(\theta)}{g(\theta)}\right)'g(\theta).$$

This is positive only if $\left(\frac{v(\theta)}{g(\theta)}\right)'' \geq 0$ everywhere. Since $S_0\left(= \frac{v(0)}{g(0)}\int_0^1 g(\theta)d\theta\right)$ equals 0 only if $v(0)/g(0) = 0$ (since $\int_0^1 g(\theta)d\theta > 0$ as $g(\theta)$ is strictly positive for $\theta < 1$).

From the proof of Proposition 2 (ii), if $v(1)$ is bounded and $g(1) > 0$, then $\lim_{\theta \to 1}\left(\frac{v(\theta)}{g(\theta)}\right)'\left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right) = 0$. This shows that $\left(\frac{v(\theta)}{g(\theta)}\right)'\left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right)$ cannot be weakly increasing. ∎

Note that the conditions in Corollary 1 are somewhat difficult to be satisfied, thus (pure) contests may be used somewhat rarely; even so, we now provide two numerical examples where a contest is the best way to allocate a good. In the first case, v is unbounded. In the second case, $g(1) = 0$.

Example 4 $N = 2$, $M = 1$, $v(\theta) = \frac{2}{(1-\theta)^{0.5}} - 2$, and $g(\theta) = 1$.

Notice that $\left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat\theta)d\hat\theta\right) = \frac{1}{(1-\theta)^{0.5}}$ is increasing on $[0, 1]$ and $\ell_1 = 0$ (since g is constant). Surplus from a pure contest is $2\int_0^1 \frac{1}{(1-\theta)^{0.5}}\theta d\theta = 2\frac{2}{3}$ and is more than that from a lottery $\int_0^1 \frac{1}{(1-\theta)^{0.5}}d\theta = 2$. Here, the contest does best. $\square$

Example 5 $N = 2$, $M = 1$, $v(\theta) = \theta$, and $g(\theta) = (1 - \theta)^2$.

Here $\left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat\theta)d\hat\theta\right) = (1 + \theta)\frac{1}{3}$ which is increasing in θ, $v(0) = 0$ and $\ell_1 = 0$. $\square$

Remark 1 *The optimal mechanism depends upon v and g, not just v/g.*

We see this by means of the following example which has the same v/g ratio as the previous example but instead of a contest being optimal, a lottery is optimal.

Example 6 $N = 2$, $M = 1$, $v(\theta) = \theta(1 - \theta)$, and $g(\theta) = (1 - \theta)^3$.

Here, $\int_0^\theta \left(\frac{v(\widetilde\theta)}{g(\widetilde\theta)}\right)' \left(\int_{\widetilde\theta}^1 g(\hat\theta)d\hat\theta\right) d\widetilde\theta = \int_0^\theta \left(1 - \widetilde\theta^2\right)\frac{1}{4}d\widetilde\theta = \frac{\theta}{4} - \frac{\theta^3}{12}$. Since $v(0) = 0$, $\frac{v(0)}{g(0)}\int_0^1 g(\hat\theta)d\hat\theta = 0$. The expected surplus of running a lottery is $\int_0^1 \theta(1 - \theta)d\theta = \frac{1}{6}$. Since $\frac{\theta}{4} - \frac{\theta^3}{12} \geq \frac{1}{6}\theta$ for all $\theta \in [0, 1]$, a lottery is optimal. $\square$

4.4   When More Complicated Mechanisms Are Optimal

There are still possibilities under Proposition 1 when neither a lottery nor contest is optimal.

Remark 2 *Under certain conditions, a more complicated mechanism can be optimal such as*

(i)   a contest with a bid cap,

(ii) a lottery among low types but a contest for high types,

(iii) several regions where each region is of contest or lottery.

We illustrate the variety of possible mechanisms by means of examples. In example 7, we show that a contest with bid caps can be optimal, while in example 8 we indicate that a

lottery, when all values are low, followed by a contest for higher values is optimal. Finally, in example 9, we show that a contest followed by a lottery and then again followed by a contest is the optimal mechanism. Note that in all the examples the main intuition for the use of a particular mechanism, is if the virtual surplus increases or falls. If the virtual surplus increases then it means that efficiency increases with a non-cooperative bidding mechanism while if the surplus falls a random allocation is better.

Example 7 $N = 2, M = 1, v(\theta) = \frac{\theta^2}{2}$ and $g(\theta) = 1$.

Notice here that $\left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat\theta)d\hat\theta\right) = \theta(1-\theta)$ increases and then decreases in θ. Consider the following allocation where θ_1 and θ_2 denote the types of the agents. If $\theta_1, \theta_2 \geq \theta^*$, then the good is allocated randomly, otherwise whoever has the higher θ gets the good. Under such an allocation rule, the probability of winning with $\theta \geq \theta^*$ is $\overline{P} \stackrel{def}{=} (1+\theta^*)/2$. (There is a θ^* chance the other agent has a value below θ^*. In this case, one wins. There is a $1-\theta^*$ chance the other agent has a value above θ^*. In this case, one wins half the time.) Under such an allocation, the social surplus is $N \int_0^{\theta^*} W(\theta)\left[\theta(1-\theta)\right]d\theta + N \int_{\theta^*}^1 \left[(1-\theta)\theta\overline{P}\right]d\theta$. Surplus is then $2\int_0^{\theta^*}(1-\theta)\theta^2 d\theta + 2\int_{\theta^*}^1(1-\theta)\theta\frac{(1+\theta^*)}{2}d\theta = \frac{1}{6}\theta^* - \frac{1}{2}\theta^{*2} + \frac{1}{2}\theta^{*3} - \frac{1}{6}\theta^{*4} + \frac{1}{6}$ which achieves its maximum of 0.18425 at $\theta^* = 1/4$.[28] Such an allocation results from running a contest with an appropriate bid cap that induces all agents with values above θ^* to bid at the bid cap (resulting in a lottery among those agents).

Examine the thin line in Figure 1. This represents the virtual surplus of giving the object to an agent of type θ. For all θ, it is also worthwhile to give the object than not to give the object. Notice that for points to the right of the graph, such as $\theta = 0.9$ and $\theta = 0.8$, one would prefer to give the object to the agent with lower θ. However, with the restriction that the probability must be increasing, the designer can at most keep the probability of receiving the object the same (holding a lottery). While the surplus reaches the peak at $\theta = 0.5$, we would still want to hold a lottery between someone with $\theta = 0.5$ and someone with $\theta = 0.4$, since under the restriction that the probability of receiving the object must be

---

[28] The equivalent bid cap of $\overline{x}$ sets $c(\overline{x}) = 0.00716$.

increasing in signal, we must choose between either holding a lottery amongst someone with $\theta = 0.4$ and all those with $\theta \geq 0.5$ or always giving the object to all those with $\theta \geq 0.5$ over someone with $\theta = 0.4$.[29] This leads us to the thick line in the graph. This line represents the average virtual surplus of all $\theta' \geq \theta$. The optimal mechanism will weigh this average against the virtual surplus of θ. When the surplus of θ is higher, then θ will be added to the lottery. When the average above θ is higher, then higher θ will be preferred as in a contest. The average surplus above θ is $\left[\int_\theta^1 \widehat{\theta}(1 - \widehat{\theta})d\widehat{\theta}\right] / (1 - \theta) = \frac{1}{6}(1 - \theta)(1 + 2\theta)$. This equals the virtual surplus when $\frac{1}{6}(1 - \theta^*)(1 + 2\theta^*) = \theta^*(1 - \theta^*)$ or when $\theta^* = 1/4$, confirming the above.

Thus far, we have presented the mechanism in terms of θ. Now, we demonstrate how this would work in terms of bids x. To do so, we must first calculate the expected effort an agent of type θ must expend in the optimal mechanism. The probability of getting an object with type $\theta < 1/4$ is just the probability that the other agent has a lower type θ. For $\theta \geq 1/4$ it is the probability that the other agent has a type less than $1/4$ plus half the chance otherwise, totalling $5/8$. Thus,

$$W_i(\theta) = \begin{cases} \theta & \text{if } \theta < 1/4, \\ \frac{5}{8} & \text{otherwise.} \end{cases}$$

We can now substitute this into equation (4), yielding[30]

$$e_i(\theta) = \int_0^\theta W_i'(\hat{\theta})\frac{\hat{\theta}^2}{2}d\hat{\theta} = \begin{cases} \frac{\theta^3}{6} & \text{if } \theta < 1/4, \\ \frac{\left(\frac{1}{4}\right)^3}{6} + \frac{3}{8} \cdot \frac{\left(\frac{1}{4}\right)^2}{2} = \frac{11}{768} & \text{otherwise.} \end{cases} \tag{2}$$

Finally, $c(x(\theta)) = e_i(\theta)$. This implies that the bid cap x^* will be set such that $c(x^*) = e_i(\theta^*) = \frac{11}{768}$. If c is linear, this is the bid cap. Notice that this is incentive compatible since the highest bid before the cap, $lim_{\theta \nearrow \frac{1}{4}} x(\theta) = lim_{\theta \nearrow \frac{1}{4}} e_i(\theta) = \frac{\left(\frac{1}{4}\right)^3}{6} = \frac{1}{384}$. The expected

---

[29] This is necessary to be consistent with monotonicity as describe in footnote 16. For instance, if we choose someone with $\theta = 0.5$ over someone with $\theta = 0.4$ while choosing someone with $\theta = 0.4$ over someone with $\theta = 0.7$, then $W_1(0.5, 0.4) > W_1(0.7, 0.4)$.

[30] Note that over the discrete jump in $W_i(\theta)$, we can use of a Dirac delta to obtain our results (the increase in payments is the probability jump multiplied by $\hat{\theta}^2/2$ evaluated at the jump. Alternatively, one can integrate by parts $\int_0^\theta W_i'(\hat{\theta})\frac{\hat{\theta}^2}{2}d\hat{\theta} = W_i(\theta)\frac{\theta^2}{2} - \int_0^\theta W_i(\hat{\theta})\hat{\theta}d\hat{\theta}$. For $\theta < 1/4$, this is $\theta^3/6$. For $\theta \geq 1/4$, $W_i(\theta)\frac{\theta^2}{2} - \int_0^\theta W_i(\hat{\theta})\hat{\theta}d\hat{\theta} = \frac{5}{8} \cdot \frac{\theta^2}{2} - \frac{\left(\frac{1}{4}\right)^3}{3} - \int_{\frac{1}{4}}^\theta \frac{5}{8}\hat{\theta}d\hat{\theta} = \frac{11}{768}$.

profit of an agent with $\theta = 1/4$ bidding $\frac{1}{384}$ is $\frac{1}{4}\frac{\left(\frac{1}{4}\right)^2}{2} - \frac{1}{384} = \frac{1}{192}$, while bidding at the cap of $\frac{11}{768}$ yields $\frac{5}{8}\frac{\left(\frac{1}{4}\right)^2}{2} - \frac{11}{768} = \frac{1}{192}$. An agent with a higher θ, gains more from winning, while an agent with a lower θ gains less from winning. Since bidding at the bid cap wins with higher probability, all agents with $\theta < 1/4$ prefer bidding beneath the bid cap and agents with $\theta \geq 1/4$ prefer bidding at the bid cap. $\square$
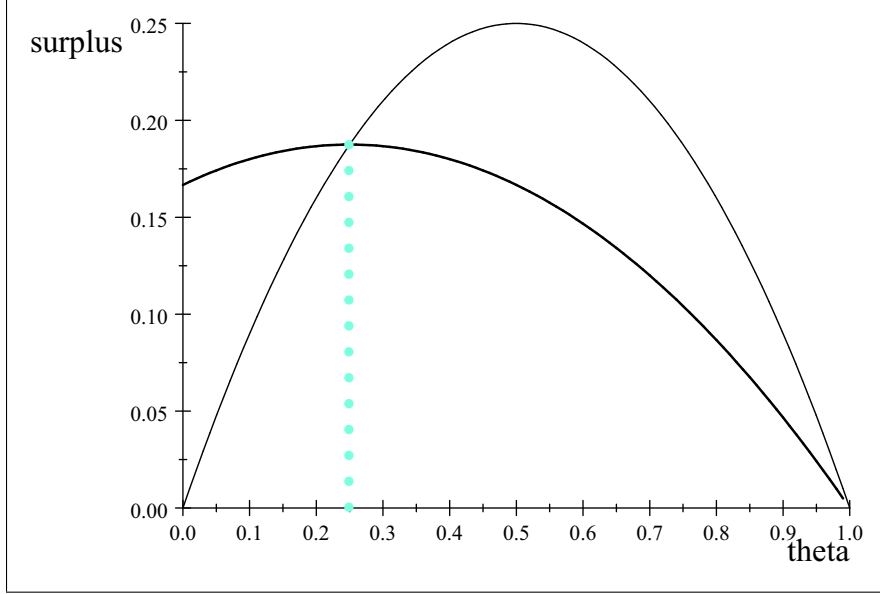
Figure 1: In example 7, the winner is by highest signal and after a threshold $\theta^* = 1/4$, by lottery. Thin line is virtual surplus and thick line is average virtual surplus above θ.

Example 8 $N = 2$, $M = 1$, $v(\theta) = \frac{1}{2\sqrt{(1-\theta)}} + \theta - \frac{1}{2}$ and $g(\theta) = 1$.

Here $\left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right) = (1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 1\right)$. It first decreases until $\theta = 0.75$, and then increases until $\theta = 1$. Consider the following allocation where θ_1 and θ_2 denote the types of the agents. If $\theta_1, \theta_2 \leq \theta^*$, then the good is allocated randomly, otherwise whichever θ is higher gets the good. Such an allocation results from running a contest with a minimum bid and allocating the good randomly if no one meets the minimum bid. From this, the social surplus is $\int_0^{\theta^*} \left[(1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 1\right)\theta^*\right]d\theta + N\int_{\theta^*}^1 \left[(1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 1\right)\theta\right]d\theta$. We will now see that this is the optimal mechanism.

In Figure 2, as before, the thin line in the graph represents the virtual surplus of giving the object to an agent of type θ. Again, for all θ, it is also worthwhile to give the object than

not to give the object. Notice that now for points to the left of the graph, such as $\theta = 0.2$ and $\theta = 0.1$, a designer prefers to give the object to the agent with lower θ. Hence, under the restriction that the probability of receiving the object must be increasing in signals, the designer would choose a lottery for those points. While the surplus reaches the minimum at $\theta = 0.75$, we would still want to hold a lottery beyond this point, for example between someone with $\theta = 0.75$ and $\theta = 0.8$. This is for similar reasons to those in example 7. Namely, since under the restriction that the probability must be increasing in signals, we need to make the choice between holding a lottery amongst someone with $\theta = 0.8$ and all those with $\theta \leq 0.75$ or always giving the object to someone with $\theta = 0.8$ over all those with $\theta \leq 0.75$.[31]

This leads us to the thick line in the graph that represents the average virtual surplus of all $\theta' \leq \theta$. The optimal mechanism will weigh this average against the virtual surplus of θ. If the surplus of θ is higher, then θ will be preferred. When the average below θ is higher, then θ will be added to the lottery. This crossing point is at $\theta^* = 0.9117$. As in example 7, we can use equation (4) to determine the minimum signal to be included in the contest. Since in the lottery phase, when $\theta < \theta^*, W'(\theta) = 0$ we have $e(\theta) = \int_{\underline{\theta}}^{\theta} W'(\hat{\theta}) \frac{v(\hat{\theta})}{g(\hat{\theta})} d\hat{\theta} = 0$. Going from the lottery to the contest, the jump in probability is $\theta^*/2$. Thus, the jump in effort must be $lim_{\theta \searrow \theta^*} e(\theta) = \frac{\theta^*}{2} \frac{v(\theta^*)}{g(\theta^*)} = \frac{\theta^*}{2} \left(\frac{1}{2\sqrt{(1-\theta^*)}} + \theta^* \right) = 1.1826$. This should equal the cost of the signal for the agent with type θ^*, hence $c(x(\theta^*)) = 1.1826$. If $c(x) = x^2$, then the minimum bid should be $\sqrt{1.1826} = 1.0875$. $\square$

---

[31] Otherwise, monotonicity from footnote 16 is broken. For instance, if we choose someone with $\theta = 0.2$ over someone with $\theta = 0.8$ while choosing someone with $\theta = 0.8$ over someone with $\theta = 0.75$, then $W_1(0.2, 0.8) > W_1(0.75, 0.8)$.
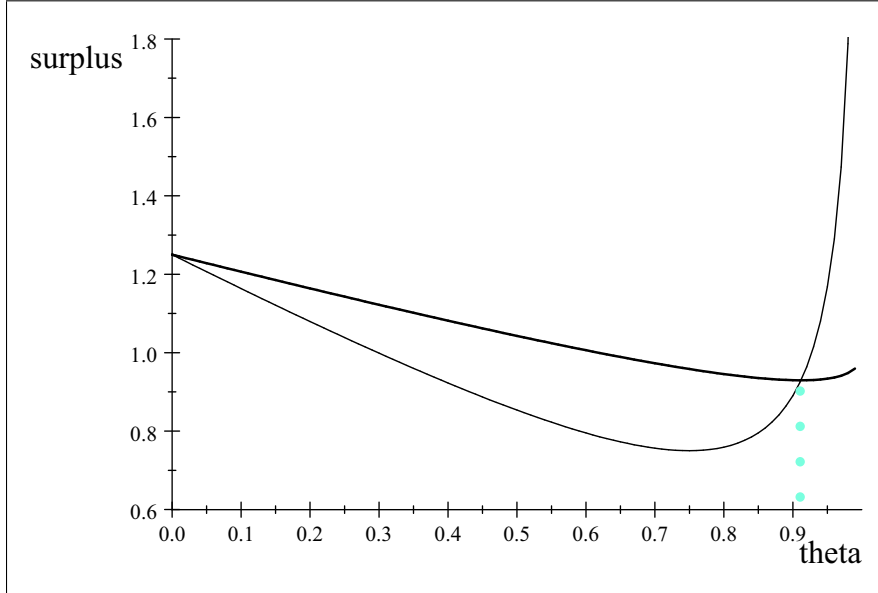
Figure 2: In example 8, the winner is by lottery, and then by highest signal. Thin line is virtual surplus and thick line is average virtual surplus above θ.

Example 9 $N = 2$, $M = 1$, $v(\theta) = \frac{1}{2(1-\theta)^{0.5}} + \frac{3}{2}\theta^2 - \frac{1}{2}$ *and* $g(\theta) = 1$.

Here $\left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right) = (1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 3\theta\right)$. This first increases, then decreases, and then again increases till $\theta = 1$. In this case, the following mechanism is optimal, where the social planner first runs a contest then a lottery and then runs a contest for the high-value agents. This yields the following allocation: θ_1 and θ_2 denote the types of the agents. If θ_1, θ_2 are in $[0.45, 0.91]$, then the good is allocated randomly. Otherwise, it is allocated to the one with the highest θ. Note that from $\theta = 0.45$ to $\theta = 0.91$ the social planner will run a lottery and in the rest of the range a contest will be used. Therefore, the social surplus is

$$2\int_0^{0.45}\left[(1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 3\theta\right)\theta\right]d\theta + \int_{0.45}^{0.91}\left[(1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 3\theta\right)(0.91 - 0.45)\right]d\theta$$
$$+2\int_{0.91}^1\left[(1-\theta)\left(\frac{0.25}{(1-\theta)^{1.5}} + 3\theta\right)\theta\right]d\theta.$$

This is a combination of our two previous examples with the lottery range in the middle. Denote the lottery range from θ_a to θ_b. We must compare the average virtual surplus of those in the range to those out of the range. We would prefer a θ in $[\theta_a, \theta_b]$ to those below θ_a

20

if the average surplus is higher than the surplus of all those below and prefer those above θ_b if the average surplus is lower than the surplus of all these values. Since the virtual surplus is increasing (and continuous) outside of this range, this can only happen if the average virtual surplus is equal to the virtual surplus on both ends: $\int_{\theta_a}^{\theta_b} s(\theta)d\theta = s(\theta_a) = s(\theta_b)$. We see this in Figure 3. Again, the thin line is the virtual surplus. Here, the thick line helps to demonstrate the range of types where a lottery should be used. With this line, both endpoints have the same virtual surplus. This virtual surplus should also equal the average virtual surplus in the range of the thick line. In order for this to happen, the area above it and below the thin line and below it and above the thin line should cancel (be equal).

As with examples 7 and 8, we can determine the mechanism in regards to an auction. The mechanism would combine a bid cap with a minimum bid but ironically with the bid cap first. The cap will start with $\underline{x}$ and go to $\overline{x}$. It is a cap in the sense that any bid between $\underline{x}$ and $\overline{x}$ will count only as much as $\underline{x}$. These can be determined in a similar way as in examples 7 and 8. □
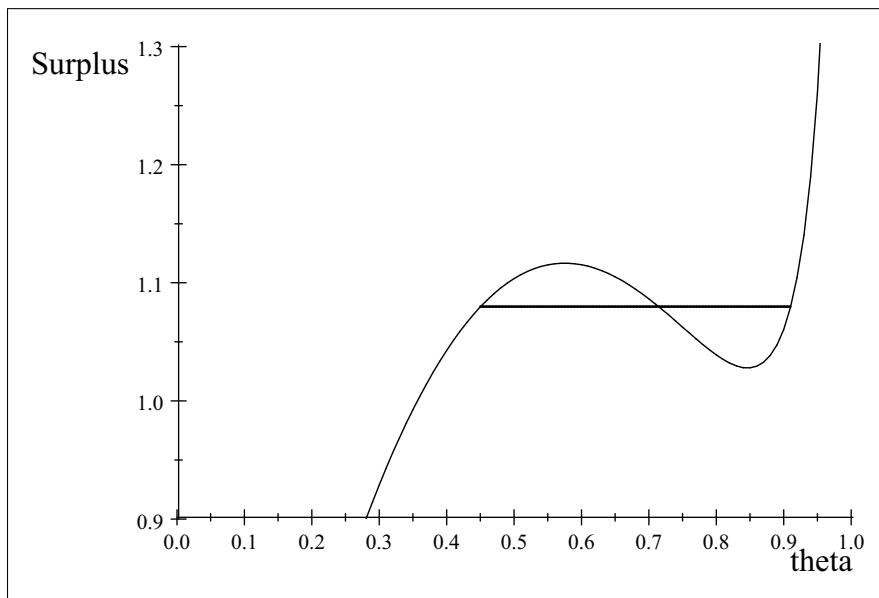
Figure 3: In example 9, the winner is by highest signal except for the interval $\theta \in [0.45, 0.91]$ Thin line is virtual surplus and thick line is the interval $[0.45, 0.91]$. The area below the thin line and above the thick line is equal to the area above the thin line and below the thick line.

We observe in numerous instances where goods are allocated by one of the more complicated methods of examples 7 to 9, that is, a method beyond a straight lottery or contest. For instance, the way example 7 could work in practice, is to allocate tickets for an event by having a lottery for anyone that waits x hours for tickets and if there are tickets left after that, allocate the tickets through lottery. Another illustration of this is the ticket distribution by All England Tennis and Croquet Club for the Wimbledon tennis tournament. The club first holds a lottery to allocate the tickets almost six months before the Wimbledon tournament and then gives them away in first come first serve basis or person willing to stand longest in the queue. (We presume that buying tickets six-months prior takes more effort.) We see a system like example 8 with the distribution of entries in the New York marathon.[32] Those that put in greater effort can qualify automatically (by completing a number of sanctioned races or making a qualifying time), remaining entries are distributed by a lottery system. Finally, in example 9, we see where it is optimal to run a contest first, and then allocate the good randomly and for the higher values again run a contest. One possible example of this is admissions to top US universities among students with high test scores. Writing an essay is part of the application. As most lecturers know, most essays are indistinguishable in level. A few good ones stand out as well as a few bad ones. It is possible that an admissions officer would first admit the good essays and then randomly select among the middle pile. If there are still slots left, the officer may start to offer admissions to the top of the lower group. Similarly, there are more students that apply to Oxford or Cambridge Universities with the highest score on the admissions tests (three As in the A-level exams) than places. To select students, interviews are held. We can interpret the interviews (where preparation can help) as the socially wasteful but necessary to signal the type of the students.

---

[32]Some may be surprised to discover that the right to run in the major marathons needs to be rationed. There are logistic limits to supply and demand for 26 miles of punishment is high. We also note that transfer of numbered bibs is currently prohibited on both medical and fairness grounds (see Blecher, 2006 for details).

4.5 Scarcity

An important question is whether scarcity of goods (the number of goods relative to the number of agents) favors one mechanism over another. To answer this question, we must first find a way to compare mechanisms across environments. We do so by the following definition.

Definition 1 *Two mechanisms are said to be equivalent if their W functions have the same regions where they are strictly increasing, that is, a mechanism characterized by W_a is equivalent to one characterized by W_b if for all $\theta_2 > \theta_1$, we have $W_a(\theta_2) > W_a(\theta_1) \iff W_b(\theta_2) > W_b(\theta_1)$.*

Intuitively, this definition says that two mechanisms are equivalent if for any set of types both mechanisms allocate the good in a similar way. For example, if in mechanism a a lottery is used among all types between $1/3$ and $1/2$, then if mechanism b is equivalent to mechanism a, then in mechanism b a lottery must also be run among those types. We can now use the above definition to show that an optimal mechanism does not depend upon the scarcity of the good.

Proposition 5 *For any mechanism that is optimal in environment (v, g, M, N) (and described by Proposition 1) there is an equivalent mechanism that is an optimal mechanism in environment $(v, g, \widehat{M}, \widehat{N})$. However, which of two equivalent non-optimal mechanisms yield higher social surplus may depend upon M or N.*

Proof. Denote $\mathcal{C}_a$ as the $\mathcal{C}$ function used to generate the mechanism (via Proposition 1) for environment (v, g, M, N) resulting in the W function, W_a. Notice that $W_a(\theta_2) > W_a(\theta_1)$ if and only if $\mathcal{C}_a(\theta_2) > \mathcal{C}_a(\theta_1)$. Now notice that $\mathcal{C}$ functions are independent of M or N (they depend only on v and g). Hence, for environment $(v, g, \widehat{M}, \widehat{N})$ the $\mathcal{C}$ function would equal $\mathcal{C}_a$ and hence the W function, W_b, would have the property that $W_a(\theta_2) > W_a(\theta_1) \iff W_b(\theta_2) > W_b(\theta_1)$.

We will now show the second part of the proposition. We do so by means of an example showing that in comparing two non-optimal mechanisms which is higher changes in either M or N. More specifically, a lottery is optimal when $M = 1$ and $N = 3$; however, a contest is optimal either when $M = 2$ and $N = 3$ or $M = 1$ and $N = 2$. The details are as follows: $g(\theta) = 1$ and $v(\theta) = \theta^\alpha$, yielding virtual surplus of $\theta^\alpha(1 - \theta)$. The social surplus with a contest with one object, $M = 1$, is $\frac{N}{(N+\alpha)(N+\alpha+1)}$. The surplus with a contest with two objects, $M = 2$, is $\frac{(3N+\alpha-3)N}{(N+\alpha)(N+\alpha+1)(N+\alpha-1)}$. The surplus with a lottery is $\frac{M}{(1+\alpha)(2+\alpha)}$ (independent of N). Take $\alpha = 5/4$. When we have $N = 2$ and $M = 1$, the surplus for the contest, 0.1447, is greater than that of a lottery, 0.13675, however when $N = 3$ and $M = 1$, this is reversed, with surplus from contest, 0.1344, which is now smaller than that of a lottery. Finally, when we move to $N = 3$ and $M = 2$, the contest is again better than a lottery, with contest surplus, 0.299, and lottery surplus, 0.273. ∎

Taylor et al. (2003) and Koh et al. (2006) analyze allocation through lotteries and queues, and find that lotteries are more efficient in comparison to waiting line auctions if the time valuation are less varied and objects are scarce. This may be surprising compared to the first part of the proposition, however the second part shows how our paper is consistent. Namely, we show that the optimal mechanism is independent of scarcity, while earlier papers compared only two different mechanisms.[33]

4.6 Other Interim-Efficient Allocations

Up to this point, we have studied ex-ante Pareto-optimal allocations. In such a manner, the designer treats every type with the same ex-ante weight. It is conceivable that the designer may want to favor some types over others. For instance, a designer may wish to count higher types more (they could be future donors) or lower types more (increasing future interest). To include such a possibility, we need a set of weights $\lambda(\theta)$ such that $\int_0^1 \lambda(\theta) d\theta = 1$ for which the designer would maximize $N \cdot E[\lambda(\theta) \pi(\theta, \theta)]$. This leads to the set interim-efficient

---

[33]Other notable earlier papers on waiting-line auctions include Barzel (1974), Holt and Sherman (1982) and Suen (1989).

allocations (see Ledyard and Palfrey, 2007). Observe that for equal weights, $\lambda(\theta) = 1$, we have our original case of an ex-ante efficient allocation.

Proposition 6 *A $\lambda(\theta)$ interim-efficient allocation for environment $v(\theta)$ and $g(\theta)$ is given by Proposition 1 where $\mathcal{C}(\theta)$ is defined using $\widehat{v}(\theta) \stackrel{def}{=} \lambda(\theta) v(\theta)$ and $\widehat{g}(\theta) \stackrel{def}{=} \lambda(\theta) g(\theta)$ instead of $v(\theta)$ and $g(\theta)$.*

Proof. In order for this proposition to hold, we need to show two problems yield the same solution: (A) choosing $W(\cdot)$ and $e(\cdot)$ to maximize $N \cdot E[\lambda(\theta) \pi(\theta, \theta)]$ subject to $W(\theta)$ being consistent with feasibility, IC and IR constraints and (B) choosing $W(\cdot)$ and $e(\cdot)$ to maximize $N \cdot E[\widehat{\pi}(\theta, \theta)]$ (where $\widehat{\pi}(\theta, \theta) \stackrel{def}{=} W(\theta) \cdot \widehat{v}(\theta) - e(\theta) \cdot \widehat{g}(\theta)$) is subject to $W(\theta)$ being consistent with feasibility, IC and IR constraints (with $\widehat{v}(\theta)$ and $\widehat{g}(\theta)$).

Notice that $N \cdot E[\lambda(\theta) \pi(\theta, \theta)] = N \cdot E[W(\theta) \cdot \lambda(\theta) v(\theta) - e(\theta) \cdot \lambda(\theta) g(\theta)] = N \cdot E[\widehat{\pi}(\theta, \theta)]$. Thus, the objective function is the same. We now show that any allocation that satisfies IC and IR constraints for $\widehat{v}(\theta)$ and $\widehat{g}(\theta)$ will also satisfy IC and IR constraints for $v(\theta)$ and $g(\theta)$ and vice-versa.

Define $\widehat{\pi}(\theta, \widetilde{\theta}) \stackrel{def}{=} W(\widetilde{\theta}) \cdot \widehat{v}(\theta) - e(\widetilde{\theta}) \cdot \widehat{g}(\theta)$. Now $\widehat{\pi}(\theta, \theta) \geq \widehat{\pi}(\theta, \widetilde{\theta}) \iff \pi(\theta, \theta) \geq \pi(\theta, \widetilde{\theta})$ because $\widehat{\pi}(\theta, \widetilde{\theta}) = W(\widetilde{\theta}) \cdot \lambda(\theta) v(\theta) - e(\widetilde{\theta}) \cdot \lambda(\theta) g(\theta) = \lambda(\theta) \pi(\theta, \widetilde{\theta})$. Likewise, $\widehat{\pi}(\theta, \theta) \geq 0 \iff \pi(\theta, \theta) \geq 0$. Feasibility is also the same. Since the objective and constraints are the same, the solution must be the same. (Since $\widehat{v}(\theta)/\widehat{g}(\theta) = v(\theta)/g(\theta)$, our assumption that v/g is increasing is still sufficient.) ∎

As we see from the following Corollary, the Proposition expands our results.

Corollary 2 *The results of sections 4.1 to 4.5 hold for $\widehat{v}(\theta)$ and $\widehat{g}(\theta)$ in place of $v(\theta)$ and $g(\theta)$.*

One result that may seem surprising is that if $\left(\frac{v(\theta)}{g(\theta)}\right)'' \leq 0$, then a lottery is optimal for any weighting of types. Intuitively, one may expect when higher types are weighted more, then a lottery may not be optimal. However, $\left(\frac{v(\theta)}{g(\theta)}\right)'' \leq 0$ is enough to ensure that the virtual surplus is decreasing for any $\lambda(\cdot)$, since $\frac{d\left[\left(\frac{v(\theta)}{g(\theta)}\right)'\left(\int_\theta^1 \lambda(\hat{\theta}) g(\hat{\theta}) d\hat{\theta}\right)\right]}{d\theta} = \left(\frac{v(\theta)}{g(\theta)}\right)'' \left(\int_\theta^1 \lambda(\hat{\theta}) g(\hat{\theta}) d\hat{\theta}\right) - \left(\frac{v(\theta)}{g(\theta)}\right)' \lambda(\theta) g(\theta)$.

5 Concluding remarks

This paper makes a contribution in the allocation of goods when signalling one's desire for the good is a wasteful activity. Under such conditions, there is a trade-off between efficient allocation and wasted resources. A mechanism such as a contest, which allocates by the highest signal, will allocate goods to the people who value them the most but the act of signalling will be wasteful. Allocating an equal share to everyone (or a random allocation by a lottery) saves any waste from signalling, but leads to an ex-post inefficient allocation.

This analysis has many additional applications. Contests are also used to grant the Olympic games, where the individual cities submit bids, and part of the bids are often socially wasted. In the UK, governmental research funds are distributed through two main channels: research councils or quality-related (QR) funding. Research councils allocate funds to institutions by gathering private signals through research grant applications, for which the paper work can be considered socially wasteful. QR funding allocates funds through publicly available information such as publications, which presumably is less costly to gather (this is done through the Research Assessment Exercise, RAE). Our analysis can help design an appropriate mechanism. If the cost incurred by the institutions to make the case for grants is too high, the government should favour QR funding. Policy research into which system is best is an important area in which our paper contributes.

Our results also has implications for bidding rings (cartels) in auctions. In this literature, McAfee and McMillan (1992) find an optimal mechanism for collusion that agrees with our results. Namely, if the hazard rate is decreasing, bidders should participate in the auction non-cooperatively; however, if the hazard rate is increasing, then bidders should bid the reserve price whenever they value the object more than the reserve. In this application, the cartel's objective is congruent to that of our designer while the bids are analogous to our wasteful signal. Hence, our results indicate that the McAfee and McMillan results apply more generally.[34] Moreover, a connection would show that in many cases the optimal collusive

---

[34]There is still the discrepancy of the all-pay nature of our signals versus the first-price payments in an auction.

policy would be something more complicated such as an increasing bid function that reaches a peak or bidding the reserve price for low values and then jump to bidding higher values. In another direction, our idea of having private opportunity costs (in addition to values) can be applied to other applications where such costs have so far been assumed to be identical (such as in scheduling of appointments, McAfee and Miller, 2009). Another area where our analysis of non-market mechanisms may apply is club membership with limited vacancies (see Hillman and Swan, 1983).

There are many possible changes to our basic model for which similar analysis could be used. One change is to relax our key assumption of a wasteful signal to a partial wasteful signals. For instance, an agent of type θ sending a signal of x may have private value of sending the signal of $\lambda c(x)g(\theta)$. This means the true waste is only $(1 - \lambda)c(x)g(\theta)$; however, as we saw that the cost function $c(x)$ has no affect on the optimal mechanism, this will also have no affect. If instead there were a public benefit to such a signal, then the higher the benefit, the more likely the optimal mechanism would favour contests.[35] We can also partially relax the assumption on payments to the designer. The necessary element for our analysis to apply is that there is a maximum price that the designer can charge and at this price, there is an excess demand (as the case with playoff tickets). The timing of the signals in our mechanism can also be changed while keeping the same nature of our results. For instance, a war of attrition can be used to allocate goods in place of a contest. A war of attrition with a time limit can be used in lieu of a contest with a bid cap.

While in this paper, we examined only the case where each agent has one of two possible allocations: with an object or without an object, we can apply this in any case with two possible outcomes. Think of the case where a course is offered twice and students have to decide which time they want to be scheduled for (with all students being assigned to one of the two slots). If there is an excess demand for one of the time slots, then one can use our analysis to determine how to allocate the slots to the students demanding the popular time slot. (All students demanding the less popular slot will get it.)

---

[35] Yoon (2009) did such analysis under a more restrictive set of mechanisms (see footnote 7).

A natural extension for our paper would be to consider several types of goods. Doing so would make it possible to explore a link to papers on pseudo-markets (markets using points rather than money), except we will optimize the method for obtaining points as a function of effort (better grades yield more points in course markets). An exogenous allocation of points is similar to a lottery while points solely as a function of effort is like a contest.

Fairness issues may also be considered in determining a mechanism. A lottery is ex-ante (and interim) fair in that everyone has an equal chance of receiving the good, but ex-post, those not allocated the good envy those allocated the good. The losers in a contest are worse off than those not given the good if a lottery were run in its place; however, the losers in a contest envy the winners less than in a lottery (since in a contest the winners pay more than the losers). Hence, which mechanism is deemed fairer could depend upon the circumstances.[36]

References

Ariely, D., Predictably irrational: The hidden forces that shape our decisions, 2008, Harper Collins, London.

Barzel, Y., "A theory of rationing by waiting," *Journal of Law and Economics*, April 1974, 17 (1), 73 − 95.

BBC News, "Waiting Over for Would-be Buyers," May 31, 2007. http://news.bbc.co.uk/1/hi/england/devon/6706825.stm

BBC News, "Debate Over School Lottery System," June 12, 2007. http://news.bbc.co.uk/1/hi/england/sussex/6745069.stm

Blecher, J., "Marathon: Paying Top Dollar For Punishment, 26 Miles' Worth," *New York*

---

[36] In explaining why they selected a first-come first-serve mechanism for selling 23 flats in a popular complex in Exeter, manager Sally Henley said "we've had more than 1,000 people interested in the properties, so we felt this was the fairest way to go about it" (*BBC News*, May 31, 2007). In explaining why the Brighton Council is using a lottery to allocate school spaces, education director David Hawker said "the council remains confident the scheme is the fairest solution" (*BBC News*, June 12, 2007).

Times, September 16, 2006.

Brams, S. J. and Kilgour, M.D., "Competitive Fair Division," *Journal of Political Economy*, 2001, 109 (2), 418 – 443.

Brams, S. J. and Taylor, Alan D., "Fair Division: From Cake-Cutting to Dispute Resolution," *Cambridge University Press*, 1996.

Che, Y.-K. and Gale, I., "Market versus Non-Market Assignment of Ownership," Working Paper, Dept. of Economics, Columbia University, 2009.

Condorelli, D., "What Money Can't Buy: Allocations with Priority Lists, Lotteries and Queues," Working Paper, 2009.

Crémer, J., Spiegel, J. and Zheng, C. Z., "Auctions with Costly Information Acquisition," *Economic Theory*, 2009, 38 (1), 41 – 72.

Fudenberg, D. and Tirole, J., Game Theory, 1991, MIT Press, Cambridge.

Fullerton, R. and McAfee, P., "Auctioning Entry into Tournaments," *Journal of Political Economy,* 1999, 107 (3), 573 – 605.

Gardner, E., "Metropolitan Doors Open Even Wider," *USA Today,* October 5, 2006.
`http://www.usatoday.com/travel/destinations/2006-10-05-metropolitan-opera_x.htm`

Gavious, A., Moldovanu, B. and Sela, A., "Bid Costs and Endogenous Bid Caps," *RAND Journal of Economics*, 2002, 33 (4), 709 – 722.

Goeree, J. K., Maasland, E., Onderstal, S. and Turner, J. L.,"How Not to (Raise) Money," *Journal of Political Economy*, August 2005, 113 (4), 897 – 918.

Hartline, J.D. and Roughgarden, T., "Optimal Mechansim Design and Money Burning,"
`http://www.citebase.org/abstract?id=oai:arXiv.org:0804.2097`, 2008.

Hillman, A.L. and Swan, P.L., "Participation Rules for Pareto-Optimal Clubs," *Journal of Public Economics*, 1983, 20 (1), 55 – 76.

Holt, C.A. and Sherman, R. "Waiting-line auctions," *Journal of Political Economy*, 1982, 90 (2), 280–294.

Hoppe, H., Moldovanu, B. and Sela, A., "The Theory of Assortative Matching Based on Costly Signals," *Review of Economics Studies*, January 2009, 76 (1), 253 – 281.

Koh, W.T.H., Yang, Z. and Zhu, L.,"Lottery Rather than Waiting-line Auction," *Social Choice and Welfare*, 2006, 27 (2), 289 – 310.

Ledyard, J. and Palfrey, T., "A General Characterization of Interim Efficient Mechanisms for Independent Linear Environments,"*Journal of Economic Theory,* March 2007, 133(1), 441 – 466.

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

McAfee, R.P. and Miller A.D., "Unscheduled Appointments," *Working paper:* `http://www.mcafee.cc`, March 2009.

Milgrom, P. "Putting Auction Theory to Work," *Cambridge University Press*, 2004.

Moldovanu, B. and Sela, A., "The Optimal Allocation of Prizes in Contests," *American Economic Review*, June 2001, 91 (3), 542– 558.

Mucha, P., "Lottery underway for Phillies play-off tix," Philadelphia Inquirer, September 8, 2009, `http://www.philly.com/philly/news/breaking/20090908_Lottery_underway_for_Phillies_`

Myerson, R.B., "Optimal Auction Design," *Mathematics of Operations Research*, June 1981, 6 (1), 58 – 73.

Roth, A., Sönmez, T. and Ünver, M. U., "Kidney Exchange," *Quarterly Journal of Economics,* May 2004, 119 (2), 457 – 488.

Roth, A., "Repugnance as a Constraint on Markets," *The Journal of Economic Perspectives,* Summer 2007, 21 (3), 37 – 58.

Sönmez, T. and Ünver, M. U., "Course Bidding at Business Schools," Boston College Working Papers in Economics 618, 2005, *International Economic Review*, forthcoming.

Suen, W., "Rationing and rent dissipation in the presence of heterogeneous individuals," *Journal of Political Economy*, 1989, 97 (6), 1384–1394.

Taylor, G.A., Tsui, K.K.K. and Zhu, L., "Lottery or waiting-line auction?," *Journal of Public Economics*, 2003, 87 (5), 1313 – 1334.

Yoon, K., "Mechanism Design with Expenditure Consideration," Institute of Economic Research- Korean University Working Paper, No. 0903., 2009.

6   Appendix

6.1   Proof of Lemma 1

First, satisfying the first-order condition of the agents problem $\pi_{\widetilde{\theta}}(\theta, \theta) = 0$, having $\pi(0, 0) \geq 0$ and limiting $W(\widetilde{\theta})$ and $e(\widetilde{\theta})$ to be increasing is sufficient to satisfy incentive compatibility and individual rationality, since the single-crossing property is satisfied by our assumption of $v'(\theta)/v(\theta) > g'(\theta)/g(\theta)$. Second, we can also take advantage of the first-order condition and use the envelope theorem to find the agents' surpluses. Let $\pi(\theta) \stackrel{def}{=} \max_{\widetilde{\theta}} \pi(\theta, \widetilde{\theta})$. Now by the envelope theorem, we have

$$\pi'(\theta) = \pi_\theta(\theta, \theta) + \pi_{\widetilde{\theta}}(\theta, \theta) = \pi_\theta(\theta, \theta) = W(\theta)v'(\theta) - e(\theta) \cdot g'(\theta).$$

Therefore,

$$\pi(\theta) = \int_0^\theta \pi'(\hat{\theta}) d\hat{\theta} + \pi(0) = \int_0^\theta \left[W(\hat{\theta})v'(\hat{\theta}) - e(\hat{\theta})g'(\hat{\theta}) \right] d\hat{\theta} + \pi(0).$$

As mentioned before, the designer cares only about the total expected utility of the agents subject to feasibility (all collected $e(\theta)$ are wasted) and has payoff:

$$N \cdot E[\pi(\theta, \theta)] \;=\; N \int_0^1 \pi(\theta) d\theta = N \int_0^1 \int_0^\theta \left[W(\hat\theta) v'(\hat\theta) - e(\hat\theta) g'(\hat\theta) \right] d\hat\theta d\theta + N\pi(0) = (3)$$
$$N \int_0^1 \left[(1 - \theta) \left(W(\theta) v'(\theta) - e(\theta) g'(\theta) \right) \right] d\theta + N\pi(0).$$

(The last step is using integration by parts.)

Finally, we can eliminate $e(\theta)$ from (3) since $e(\theta)$ is dictated in the first-order condition $\pi_{\tilde\theta}(\theta, \theta) = 0$ by $W(\theta)$:

$$e'(\theta) = W'(\theta) \frac{v(\theta)}{g(\theta)}.$$

Hence,

$$e(\theta) = \int_0^\theta W'(\hat\theta) \frac{v(\hat\theta)}{g(\hat\theta)} d\hat\theta + e(0). \tag{4}$$

(Note the designer would always want to set $e(\underline\theta) = 0$.)[37] Now by eliminating $e(\theta)$, the designer's payoff now becomes

$$N \int_0^1 \left[(1 - \theta) \left(W(\theta) v'(\theta) - g'(\theta) \int_0^\theta W'(\hat\theta) \frac{v(\hat\theta)}{g(\hat\theta)} d\hat\theta \right) \right] d\theta + N\pi(0).$$

Integrating by parts of the second expression yields:

$$N \int_0^1 (1 - \theta) W(\theta) v'(\theta) d\theta - N \int_0^1 (1 - \theta) g'(\theta) d\theta \int_0^1 W'(\hat\theta) \frac{v(\hat\theta)}{g(\hat\theta)} d\hat\theta + \tag{5}$$
$$N \int_0^1 \left[W'(\theta) \frac{v(\theta)}{g(\theta)} \int_0^\theta \left(1 - \hat\theta \right) g'(\hat\theta) d\hat\theta \right] d\theta + N\pi(0).$$

We can rewrite this as

$$N \int_0^1 W(\theta)(1 - \theta) v'(\theta) d\theta - N \int_0^1 \left[W'(\theta) \frac{v(\theta)}{g(\theta)} \left(\int_\theta^1 (1 - \hat\theta) g'(\hat\theta) d\hat\theta \right) \right] d\theta + N\pi(0). \tag{6}$$

---

[37] While both $W(\theta)$ and $e(\theta)$ may have step increases, this can be technically solved by using Dirac delltas and our results will hold. See how this technique is applied to example 7 in equation 2 .

Again doing integration by parts yields

$$
N \int_0^1 W(\theta) \left[(1-\theta)\, v'(\theta) + d\left\{ \frac{v(\theta)}{g(\theta)} \left(\int_\theta^1 \left(1-\hat\theta\right) g'(\hat\theta) d\hat\theta \right) \right\} \right] d\theta +
$$
$$
N\, W(0) \left[v(0) + \frac{v(0)}{g(0)} \left(\int_0^1 \left(1-\hat\theta\right) g'(\hat\theta) d\hat\theta \right) \right] - N\, W(1)\ell_1
$$

where $\ell_1 = \lim_{\theta\to 1} \left\{ \frac{v(\theta)}{g(\theta)} \left(\int_\theta^1 (1-\hat\theta) g'(\hat\theta) d\hat\theta \right) \right\}$.

Substituting $\left(\frac{v(\theta)}{g(\theta)} \right)' \left(\int_\theta^1 \left(1-\hat\theta\right) g'(\hat\theta) d\hat\theta \right) - \left(\frac{v(\theta)}{g(\theta)} \right) (1-\theta)\, g'(\theta)$ for $d\left\{ \frac{v(\theta)}{g(\theta)} \left(\int_\theta^1 \left(1-\hat\theta\right) g'(\hat\theta) d\hat\theta \right) \right\}$ and then replacing $\int_\theta^1 \left(1-\hat\theta\right) g'(\hat\theta) d\hat\theta$ and $\int_0^1 \left(1-\hat\theta\right) g'(\hat\theta) d\hat\theta$ by $\theta g(\theta) - g(\theta) + \int_\theta^1 g(\hat\theta) d\hat\theta$ and $-g(0) + \int_0^1 g(\hat\theta) d\hat\theta$, respectively, yields the lemma.

6.2 Proof of Proposition 1

The proof follows the method of Myerson (1981, section 6) with the necessary extension to deal with $z(\theta)$ and $\mathcal{C}(\theta)$ at the endpoints of 0 and 1. For notational simplicity, denote $z'(0)$ as $\lim_{\theta\searrow 0} z'(\theta)$, $z'(1)$ as $\lim_{\theta\nearrow 1} z'(\theta)$ and $\mathcal{C}'(1)$ as $\lim_{\theta\nearrow 1} \mathcal{C}'(\theta)$. We can rewrite the surplus (divide by N) from Lemma 1 using $z(\theta)$ and $\mathcal{C}(\theta)$ as follows.

$$
\int_0^1 W(\theta) \left(\frac{v(\theta)}{g(\theta)} \right)' \left(\int_\theta^1 g(\hat\theta) d\hat\theta \right) d\theta + W(0) \frac{v(0)}{g(0)} \int_0^1 g(\theta) d\theta - W(1)\ell_1 =
$$
$$
\int_0^1 W(\theta) z'(\theta) d\theta + W(0) \frac{v(0)}{g(0)} \int_0^1 g(\theta) d\theta - W(1)\max\{\ell_1, 0\} - W(1)\min\{\ell_1, 0\} =
$$

$$
\int_0^1 W(\theta)\mathcal{C}'(\theta) d\theta + \int_0^1 W(\theta)\left(z'(\theta) - \mathcal{C}'(\theta) \right) d\theta +
$$
$$
W(0) \frac{v(0)}{g(0)} \int_0^1 g(\theta) d\theta - W(1)\max\{\ell_1, 0\} - W(1)\min\{\ell_1, 0\} =
$$

$$
\int_0^1 W(\theta)\mathcal{C}'(\theta) d\theta - W(1)\min\{\ell_1, 0\} - \int_0^1 W'(\theta)(z(\theta) - \mathcal{C}(\theta)) d\theta. \tag{7}
$$

33

Let us call the $\widehat{W}(\theta)$ resulting probability of receiving the object from the mechanism described in the Proposition. We will first show that the last expression of (7) is maximized by $\widehat{W}(\theta)$. Notice that since $z(\theta) \geq \mathcal{C}(\theta)$, for any weakly increasing W, we have

$$\int_0^1 W'(\theta)(z(\theta) - \mathcal{C}(\theta))d\theta \geq 0.$$

By integration by parts, this then implies

$$\int_0^1 W(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta + W(0)\frac{v(0)}{g(0)}\int_0^1 g(\theta)d\theta - W(1)\max\{\ell_1, 0\} \leq 0.$$

Therefore, if we show that

$$\int_0^1 \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta + \widehat{W}(0)\frac{v(0)}{g(0)}\int_0^1 g(\theta)d\theta - \widehat{W}(1)\max\{\ell_1, 0\} = 0, \qquad (8)$$

then we prove that the last expression of (7) is maximized by $\widehat{W}(\theta)$. Take θ_1 and θ_2 where $0 \leq \theta_1 \leq \theta_2 \leq 1$, $z(\theta_1) = \mathcal{C}(\theta_1)$, and $z(\theta_2) = \mathcal{C}(\theta_2)$. If such a θ_1 and θ_2 exist, then we will show that $\int_{\theta_1}^{\theta_2} \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta = 0$. Take any $[\theta_1', \theta_2'] \subset [\theta_1, \theta_2]$, where $z(\theta) = \mathcal{C}(\theta)$ for all $\theta \in [\theta_1', \theta_2']$. Then, for all $\theta \in (\theta_1', \theta_2')$, we have $z'(\theta) = \mathcal{C}'(\theta)$. Hence, $\int_{\theta_1'}^{\theta_2'} \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta = 0$. Now the remaining regions in $[\theta_1', \theta_2'] \subset [\theta_1, \theta_2]$ have $z(\theta) > \mathcal{C}(\theta)$ for all $\theta \in (\theta_1', \theta_2')$. This implies $\widehat{W}'(\theta) = 0$ within them since if $z(\theta) > \mathcal{C}(\theta)$, we have $\mathcal{C}''(\theta) = 0$ (implying a lottery within that region). However, across these region $\int_{\theta_1'}^{\theta_2'} (z'(\theta) - \mathcal{C}'(\theta))\, d\theta = 0$, since z and $\mathcal{C}$ are equal at the endpoints. This implies $\int_{\theta_1'}^{\theta_2'} \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta = 0$ and that overall $\int_{\theta_1}^{\theta_2} \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta = 0$.

If there is no point $\theta \in (0, 1)$ where $z(\theta) = \mathcal{C}(\theta)$, then $\widehat{W}$ is a lottery and constant. Equation (8) follows. We now know that equation (8) holds if there exists a particular θ_1 and θ_2 that satisfies our original conditions $0 \leq \theta_1 \leq \theta_2 \leq 1$, $z(\theta_1) = \mathcal{C}(\theta_1)$, and $z(\theta_2) = \mathcal{C}(\theta_2)$, yet also satisfy $\int_0^{\theta_1} \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta + \widehat{W}(0)\frac{v(0)}{g(0)}\int_0^1 g(\theta)d\theta = 0$ and $\int_{\theta_2}^1 \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta - \widehat{W}(1)\max\{\ell_1, 0\} = 0$.

Now if $\frac{v(0)}{g(0)}\int_0^1 g(\theta)d\theta = 0$, then the first new condition is satisfied since we can set $\theta_1 = 0$.

34

If $\frac{v(0)}{g(0)} \int_0^1 g(\theta)d\theta > 0$, then set θ_1 as $\min\{\theta > 0 | z(\theta) = \mathcal{C}(\theta)\}$. Note that $\theta_1 > 0$. For all $\theta \in (0, \theta_1)$, we have $\widehat{W}'(\theta) = 0$ since $z(\theta) \neq \mathcal{C}(\theta)$. Then, the first condition $\int_0^{\theta_1} \widehat{W}(\theta)(z'(\theta) - \mathcal{C}'(\theta))d\theta + \widehat{W}(0)\frac{v(0)}{g(0)} \int_0^1 g(\theta)d\theta = \int_0^{\theta_1} W'(\theta)(z(\theta) - \mathcal{C}(\theta))d\theta = 0$ is satisfied. If $\ell_1 \leq 0$, then we can set $\theta_2 = 1$, and the second condition is satisfied. If $\ell_1 > 0$, then we can define θ_2 as $\max\{\theta > 0 | z(\theta) = \mathcal{C}(\theta)\}$ and proceed in a similar manner.

Now we are left to show that the first two expressions of equation (8), $\int_0^1 W(\theta)\mathcal{C}'(\theta)d\theta - W(1)\min\{\ell_1, 0\}$, are also maximized by $\widehat{W}(\theta)$. Whenever $\mathcal{C}'(\theta_2) > \mathcal{C}'(\theta_1)$, which can only happen if $\theta_2 > \theta_1$, then at some point in between, $\mathcal{C}'' > 0$. This then implies that $\widehat{W}(\theta_2)$ is maximally higher than $\widehat{W}(\theta_1)$. Finally, $W(1)\min\{\ell_1, 0\}$ is weakly positive and maximized by highest possible $W(1)$ if strictly positive, which happens with $\widehat{W}(1)$.

6.3 Proof of Proposition 3 (ii)

We start by showing that if $v(1)$ is bounded, then $\ell_1 \geq 0$. Since $v(1)$ is finite, $\lim_{\theta \to 1}(\theta - 1)v(\theta) = 0$. Moreover, since $\frac{v(\theta)}{g(\theta)} \int_\theta^1 g(\hat{\theta})d\hat{\theta} \geq 0$, we have $\ell_1 \geq 0$. This shows that the third part of the expression in Lemma 1 is maximized with a lottery.

We now show that if $v(1)$ is bounded and $g(1) > 0$, the $\lim_{\theta \to 1} \left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right) = 0$. By the mean value theorem, $\lim_{\theta \to 1} \left(\frac{v(\theta)}{g(\theta)}\right)' \left(\int_\theta^1 g(\hat{\theta})d\hat{\theta}\right) = \lim_{\theta \to 1} \left(\frac{v(\theta)}{g(\theta)}\right)' (1 - \theta)g(\theta)$. Since our assumptions imply that $\left(\frac{v(\theta)}{g(\theta)}\right)' > 0$, a bounded $v(1)$ implies that g is bounded from above. Hence, we only need to show that $\lim_{\theta \to 1} \left(\frac{v(\theta)}{g(\theta)}\right)' (1 - \theta) = 0$. Denote $z(v/g) = \theta^{-1}(v/g)$ and $x = v/g$. Since v and g are analytic functions, $z(x)$ is also an analytic function. We also have $z(x)$ strictly increasing on $(\underline{x}, \overline{x})$ with $z(\overline{x}) = 1$. Since $v(1)$ is bounded, $g(1) > 0$, it must be that $\overline{x}$ is finite. We will now show that

$$\lim_{x \to \overline{x}} \frac{1 - z(x)}{z'(x)} = 0.$$

The proof is by contradiction. Assume not. If so, for this limit to be strictly positive we

35

must have $\lim_{x \to \overline{x}} z'(x) = 0$. By L'Hopital's rule, we have

$$\lim_{x \to \overline{x}} \frac{1 - z(x)}{z'(x)} = \lim_{x \to \overline{x}} \frac{-z'(x)}{z''(x)} > 0.$$

Hence, for the same reasons, we must have $\lim_{x \to \overline{x}} z''(x) = 0$. We can repeat this indefinitely, we must have all derivatives equal 0 at $\overline{x}$. Therefore, the Taylor series expansion $z(x)$ at $\overline{x}$ cannot be equal to the function in the neighborhood of $\overline{x}$. This provides a contradiction that $z(x)$ is an analytic function. Notice that $\lim_{\theta \to 1} (1 - \theta) \left(\frac{v(\theta)}{g(\theta)} \right)' = \lim_{x \to \overline{x}} \frac{1 - z(x)}{z'(x)} = 0$.

Since $\lim_{\theta \to 1} \left(\frac{v(\theta)}{g(\theta)} \right)' \left(\int_\theta^1 g(\hat{\theta}) d\hat{\theta} \right) = 0$, we must have a $\theta' < 1$ such that $\left(\frac{v(\theta)}{g(\theta)} \right)' \left(\int_\theta^1 g(\hat{\theta}) d\hat{\theta} \right)$ is decreasing for all $\theta > \theta'$. Hence, it would create the highest surplus if there was a lottery starting weakly below θ' and including 1, that is, goods will be first given randomly to those with types within this region. Since $g(\theta)$ is bounded, $\int_\theta^1 g(\hat{\theta}) d\hat{\theta}$ is bounded. Since $v(1)$ is bounded, $g(1) > 0$, and $\left(\frac{v(\theta)}{g(\theta)} \right)' > 0$, we have $\frac{v(\theta)}{g(\theta)}$ bounded. Together these imply that $\int_{\theta_1}^{\theta_2} \left(\frac{v(\theta)}{g(\theta)} \right)' \left(\int_\theta^1 g(\hat{\theta}) d\hat{\theta} \right) d\theta$ is bounded. This implies that the average virtual surplus in the higher end lottery is bounded. Denote B as this bound. For any $\theta'' < \theta'$ that is not part of this higher-end lottery. The average surplus of a region including θ'' and going to the bottom of the higher-end lottery is lower than the average surplus in the lottery. However, if we set $\alpha > B / \int_0^1 g(\hat{\theta}) d\hat{\theta}$, then for $\widehat{v}(\theta) = v(\theta) + \alpha g(\theta)$ the average surplus including 0 and going to the bottom end of the higher-end lottery must be higher than B. Hence, a lottery overall is optimal.

Geometrically, the slope of $z(\theta)$ is not only bounded, but flat as it nears 1. Increasing $v(\theta)$ to $\widehat{v}(\theta)$ simply shifts $z(\theta)$ up by a constant. Since it is weakly increasing for all $\theta < 1$, for a large enough shift we have $\mathcal{C}(\theta)$ simply be a line from $z(0)$ to $z(1)$ and hence a lottery is an optimal mechanism.