What money can't buy: allocations with priority lists, lotteries and queues

Daniele Condorelli

d-condorelli@northwestern.edu

Abstract

I study the welfare optimal allocation of a number of identical and indivisible objects to a set of heterogeneous risk-neutral agents under the hypothesis that money is not available. Agents have independent private values, which represent the maximum time that they are willing to wait in line to obtain a good. A priority list, which ranks agents according to their expected values, is optimal when hazard rates of the distributions of values are increasing. Queues, which allocates the object to those who wait in line the longest, are optimal in a symmetric setting with decreasing hazard rates.

JEL: D45, D82, H42. Keywords: rationing; queues; priority lists; lotteries.

Northwestern University and *University College London*. This version: May 2009; First Version: November 2005. I thank Philippe Jehiel for advice throughout the project. I also thank Mark Armstrong, Gary Becker, Martin Cripps, Paul Milgrom, Andras Niedermayer, Rakesh Vohra and Jidong Zhou for useful discussions. Write to: Daniele Condorelli, Kellogg School of Management, Evanston, 60208, IL, USA; Tel: +1-773-886-0209; Fax: +1-847-491-2530

1 Introduction

There are things that money cannot buy. Scarce medical resources are assigned through priority lists almost everywhere around the world. A fraction of green cards in the US is allocated by lottery. When demand exceeds supply, a large number of goods are rationed using queues, rather than by rising their prices to clear the market. More generally, queues, lotteries and priority lists are widely adopted mechanisms, as opposed to markets, for the allocation of public resources (e.g. goods, subsidies or services), in both developed and developing countries.[1]

For example, in February 2009 about 1,200 EU farming grants, worth up to 5,000 pounds each, were allocated on a first-come first-served basis to those who turned up in person at government buildings in Norther Ireland. Some farmers queued since Sunday ahead of Tuesday's morning opening. Another example is provided by city housing programs in New York, intended to provide affordable homes to middle and low-income New Yorkers. The city housing plan calls for 165,000 new housing units between 2003 and 2013. Lotteries are used to allocate the units that become available (for sale and for rent) at below market prices. Income requirements are set for participation in the lotteries.

The two examples highlight the trade-off that arise when the allocation is non monetary. Waiting in line generates a deadweight loss. However, queues may serve as screening devices, if the resources at stake are likely to generate higher value in the hands of those who are willing to wait in line the most for them. Instead, priority lists and lotteries cause no loss of time, but do not perform a fine screening. Therefore, a question arises, about how to design an optimal mechanism in this environment.

To address this issue, I investigate the design of mechanisms for the allocation of a number of identical indivisible objects to a set of heterogenous risk-neutral agents. The agents have unitary demand and independent private values. The novelty of my approach is that values represent the max-

[1] There is an extensive literature surveying and discussing institutional details of non-market allocations. For example, see Calabresi and Bobbit (1978), Elster (1992) and (1989), Okun (1975) and Walzer (1983).

imum times that agents are willing to wait to obtain a good. In contrast to the standard transferable utility case the designer faces a trade-off between allocative efficiency and cost minimization.[2] Increasing the efficiency of the allocation requires that agents are screened according to their private valuations. However, eliciting private information is costly. My main contribution is to characterize the set of *ex-ante Pareto optimal direct allocation mechanisms* (Proposition 1), and show how optimal mechanisms can be *implemented through queuing games, priority lists and lotteries* (Proposition 2).

It turns out that if all the hazard rates of the prior distributions of values are monotonically increasing (e.g. values comes from a normal, uniform, etc.), then the optimal mechanism does not exploit any private information and takes the form of a priority list (or a lottery, if agents are ex-ante identical). That is, goods are allocated to the agents which have, ex-ante, the highest expected values. The use of lotteries and priority lists (often in the form of point systems) is widespread. Conventional wisdom attributes their success to their fairness properties. A different rationale is provided here in terms of efficiency. These mechanisms prevent agents from engaging in wasteful rent seeking activities. On the contrary, full screening of private information is optimal, if, and only if, all hazard rates are monotonically decreasing. In the symmetric case, a standard queue approximately implements the optimal mechanism. In general, when hazard rates are not monotonic, the optimal mechanism may require both screening and pooling of values. In this case, implementation in the symmetric case is obtained by using a queue where the set of possible arrival time is restricted in order to induce pooling of values in equilibrium. It is remarkable that, because when the support of valuations is bounded the hazard rate must be increasing in a neighborhood of its upper bound, it is generically optimal to put a cap on queues. That is preventing people from joining the queue too early by assigning the same priority to all those who arrive earlier than the appropriately defined time threshold.

The feature that the designer is not able to condition the allocation of the goods on the willingness to pay of the agents is not explicitly modeled and will be taken as exogenous. However, there are several reasons why it

[2] In a setting with monetary values and no budget constraints it is possible to achieve the first best outcome by using a Vickrey mechanism and redistributing expected payments.

may not be optimal to allocate scarce resources to those that are able to pay the most for them, even when the designer only cares for the welfare of the agents involved.[3] First, as in the two examples above, agents may have a common monetary value for the goods, but still have heterogenous values in terms, for example, of the time that they are willing to wait in line. Second, one can argue that markets are not equitable. In fact, agents may be severely budget constrained relative to the expected price of the good. Therefore, their willingness to pay may be different from their ability to pay.[4] Third, because allocating the goods on the basis of the ability to pay may contrast our sense of justice, it can create externalities which are difficult to internalize.[5]

A brief review of the existing literature concludes this section. Holt and Sherman (1982) provided the first game theoretic analysis of queues. They study equilibria of specific queuing games with incomplete information. The works of Taylor, Tsui and Zhu (2003) and Koh, Yang and Zhu (2006) compare, computationally, the relative performance of queues and lotteries under specific value distributions. In contrast to the papers above, I provide, within the same environment, a closed form solution to the problem of designing an optimal mechanism. Closely related environments, where agent compete for a set of homogenous goods by engaging in costly effort or money burning, are studied in a number of papers in the context of different applications: McAfee and McMillan (1992), Chakravarty and Kaplan (2006), Yoon (2009), and Hartline and Roughgarden (2008). My results are significantly more gen-

[3]If an institution has a direct interest in screening individuals on the basis of specific characteristics a distribution based on the willingness to pay may not be optimal (e.g. academic prizes are not awarded to those that promise to pay the most for them). For more on this see Condorelli (2008).

[4]A related argument applies when the allocation is made on the basis of the willingness to wait in line (e.g. old people may find it more difficult to wait in line). However, men are endowed in principle with the same amount of time and, moreover, there is substantial evidence that queues are preferred to markets on grounds of fairness (see Kahneman, Knetsch and Thaler (1986)).

[5]For example, during the US Civil War one could avoid serving in the Union army by paying a certain amount of money. This appears to have been the cause for a number of riots. See Calabresi and Bobbit (1978) for an elaboration of this argument. Note that if externalities are present, the designer also needs to ban resale.

eral than those of the first three papers, as I consider an asymmetric setting and adopt much weaker restrictions on the distributions of values (i.e. non monotone hazard rates). Hartline and Roughgarden (2008) contains a result which was obtained independently and is very similar to my Proposition 1. In contrast to my analysis, the authors discuss applications to computer science and do not deal with the practical implementation of the optimal mechanism.

The rest of the paper is organized as follows; section 2 presents the model; section 3 describes the optimal direct mechanism; section 4 describes the practical implementation of the optimal direct mechanism; section 5 concludes the paper.

2 The Model

Let $N = \{1, \ldots, n\}$ represent the set of agents. There are $m < n$ identical goods to allocate. Agents are risk neutral and demands one good only. Each agent has *private valuation* for the good, $v_i \in V_i \equiv [\underline{v_i}, \overline{v_i})$ with $\underline{v} < \overline{v} \leq \infty$. Let $\boldsymbol{V} = V_1 \times \cdots \times V_n$, $\boldsymbol{v} = (v_1, \ldots, v_n)$ and $\boldsymbol{v_{-i}} = \{v_j : j \in N \setminus i\}$. The private valuation of someone represents the maximum time (measured in hours, minutes, etc.) that he is willing to wait in line to obtain the good. Preferences are quasilinear with respect to time. Furthermore, I assume that all agents suffer a disutility, normalized to zero, if they do not obtain the good. Therefore, an agent that obtains a good with probability $0 \leq p_i \leq 1$ and sustain a time-cost (i.e. waits in line an amount of time) equal to $c_i \geq 0$ has utility $p_i v_i - c_i$.[6] Unitary demand is modeled by setting a maximum utility of $v_i - c_i$ even if the agents receives more than one object. Each agent possesses a set of *observable characteristics*, which are common knowledge. The agent i's observable characteristics determine the beliefs that other agents hold about his private value. Observable characteristics are summarized by a continuous cumulative distribution function F_i, with support in V_i, and density f_i. Individual values are assumed to be stochastically independent.

[6] Assuming that $p_i v_i - f_i(c_i)$ and f_i is increasing and commonly known (e.g. when f_i represents a wage schedule) would not provide more generality to the model.

The task of the designer is to define an allocation mechanism whose equilibrium outcome maximizes the weighted sum of the agents' ex-ante expected utilities. An *allocation mechanism* can be any game, in which agents play under incomplete information about their opponents' valuations, whose outcome (for each profile of actions) consist of (i) a probability distribution for the m goods over the n agents and (ii) a vector of non-negative costs (i.e. the amounts of time that agents wait in line). Participation in the mechanism must be voluntary. Therefore any feasible allocation mechanism must allow agents to opt out, obtaining a payoff equal to zero. I ask that the designer maximizes the weighted sum of the agents' ex-ante expected utilities because, as illustrated in Holmstrom and Myerson's (1983), a mechanism whose equilibrium maximizes a weighted sum of agents' ex-ante expected utilities is *ex-ante incentive efficient.* That is, it is Pareto efficient within the set of incentive compatible mechanisms. Therefore, no other incentive compatible mechanism can be found that makes everyone better off prior to the realization of private values.[7]

3 Optimal Direct Mechanism Design

In this section I appeal to the revelation principle and, without loss of generality, search for an optimal mechanism within the class of equivalent incentive compatible direct allocation mechanisms. A *direct allocation mechanisms*, $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$, is a mapping providing an outcome (a distribution of goods and costs) for each profile of reports from the agents. Therefore, it is a set of functions $\left\{p_i : \boldsymbol{V} \to [0,1]\ ;\ c_i : \boldsymbol{V} \to \mathrm{I\!R}^+\right\}_{i=1}^n$ such that for all $\boldsymbol{v} \in \boldsymbol{V}$ the condition $\sum_{i=1}^n p_i(\boldsymbol{v}) \leq m$ holds. In playing mechanism $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ the ex-post utility to player i from announcing s_i when its true value is v_i, while all other players announce $\boldsymbol{v_{-i}}$ is $v_i p_i(s_i, \boldsymbol{v_{-i}}) - c_i(s_i, \boldsymbol{v_{-i}})$. Assuming that opponents are truthful, the expected utility at the interim stage is: $U_i(v_i, s_i) = v_i E_{\boldsymbol{v_{-i}}}\left[p_i(s_i, \boldsymbol{v_{-i}})\right] - E_{\boldsymbol{v_{-i}}}\left[c_i(s_i, \boldsymbol{v_{-i}})\right]$. A direct allocation mechanism $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ is *incentive compatible* if, and only if, for all i and v_i,

[7] If a mechanism is ex-ante incentive efficient, there cannot be any other mechanism that would surely be better for all, even after they receive their private information. This fact reflects decreasing insurance opportunities for the agents as more information is released.

$U_i(v_i, v_i) = \max_{s_i \in V} U_i(v_i, s_i) \geq 0$. That is, a mechanism is incentive compatible if truthful reporting is an equilibrium and everyone obtains a payoff higher than zero. The next lemma offers a tractable characterization of incentive compatible direct mechanisms in terms of $\boldsymbol{p}$ only. To simplify the notation write: $P_i(v_i) = E_{\boldsymbol{v_{-i}}}[p_i(v_i, \boldsymbol{v_{-i}})]$ and $C_i(v_i) = E_{\boldsymbol{v_{-i}}}[c_i(v_i, \boldsymbol{v_{-i}})]$.

Lemma 1. *A direct allocation mechanism $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ is incentive compatible if, and only if, for all i and $v_i \in V_i$:*

$$\forall v^* \in V_i : v_i \geq v^* \quad P_i(v_i) \geq P_i(v^*) \tag{1a}$$

$$C_i(v_i) = v_i P_i(v_i) - \int_{\underline{v}_i}^{v_i} P_i(x)dx \tag{1b}$$

The proof is well known and it is omitted (see e.g. Myerson (1981)). Note that $C_i(\underline{v}_i) \leq 0$ is necessary for incentive compatibility. Furthermore, because agents cannot receive positive transfers we must have that $C_i(\underline{v}_i) \geq 0$. This implies that $C_i(\underline{v}_i) = 0$ for all i. One special cost rule that satisfies (1b) for any $\boldsymbol{p}$ is the *canonical cost rule*:

$$c_i(\boldsymbol{v}) = p_i(\boldsymbol{v})v_i - \int_{\underline{v_i}}^{v_i} p_i(x_i, v_{-i})dx_i \tag{2}$$

According to this cost rule, only the winners, or those who participate in a lottery, sustain a cost, equal to the expected minimum value they could have, and still obtain the good under the allocation rule.[8] Observe that with the canonical cost rule agents have a dominant strategy to report their values truthfully (see Myerson (1981)).

There is a one to one mapping between the outcome of an incentive compatible direct mechanism and the mechanism itself. Therefore, a direct

[8] To understand the meaning of *expected minimum value*, suppose that there is only one good available and agent i has the highest value v_i, greater than some other value v''. Assume that $\boldsymbol{p}$ assigns the goods to agents with the highest values but values in $[v', v'')$ are pooled. Furthermore, assume the remaining $n-1$ agents all have values in that region. The payment from i will be equal to $\frac{1}{n}v' + \frac{n-1}{n}v''$. In fact, the minimum value that i *could* have, and still obtain a good is v' with probability $1/n$ (the probability that he would win the lottery against the other players if he played v') and it is v'' with the probability $(n-1)/n$ (which is when he would lose the lottery).

allocation mechanism $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ is ex-ante incentive efficient (i.e *optimal*) if it is incentive compatible and, for some set of non-negative Pareto weights $w_1, \ldots, w_n$, maximizes $\mathrm{E}_{\boldsymbol{v}} \{\sum_{i=1}^{n} w_i[v_i p_i(\boldsymbol{v}) - c_i(\boldsymbol{v})]\}$. It is easy to see that, if incentive constraints were not an issue, a *first best* allocation would assign the goods to the agents with the highest weighted values at no cost. It is a consequence of Lemma 1 that any first best allocation (i.e. unconstrained efficient) is not implementable, unless the allocation is dictatorial. That is, the weights of precisely $n - m$ agents are set equal to zero.

The building blocks to construct an optimal direct mechanism under incomplete information are the *priority functions*. These assign to each agent a unique *priority level* for each reported value. The construction of the priority functions follows the *ironing technique*, as developed in Myerson (1981). Define, for all $x \in [0, 1]$:

$$H_i(x) = \int_0^x \frac{1-z}{f(F^{-1}(z))} dz \quad G_i(x) = \operatorname{conv} \langle H_i(x) \rangle \quad g_i(x) = G_i'(x)$$

Here, $\operatorname{conv} \langle \cdot \rangle$ stands for the convex hull of the function.[9] Where the derivative of $G_i(v)$ is not defined, we extend it using the right or left derivative. The *priority function* λ_i for agent i is:

$$\lambda_i(v_i) = w_i g_i(F_i(v_i)) \tag{3}$$

The next proposition characterizes the optimal direct mechanisms. The statement is straightforward, but formalizing it requires a lot of notation. Therefore I relegate in the appendix the formal statement and the proof.

Proposition 1. *Fix $F_1, \ldots, F_n$ and $w_1, \ldots, w_n$. In the optimal direct mechanism agents report their values and the designer implements the following outcome. The allocation rule* **p** *is: the m agents with values that achieve the highest priority levels, as defined in (3), obtain the goods and ties in priority are broken by an equal chance lottery. The cost rule* **c** *is defined in (2).*

Suppose that *all hazard rates of the distributions of values are monotonically non-decreasing*. Then, $\lambda_i(v_i) = w_i \mathrm{E}[v_i]$ for all $v_i \in V_i$. In fact, H_i is concave and therefore G_i is a straight line going from $H_i(0) = 0$ to

[9] $G_i(x)$ is the highest convex function such that $G_i(x) \leq H_i(x)\ \forall x$.

$H_i(1) = \mathrm{E}[v_i]$. The optimal mechanism is a *priority list*. All agents are ranked according to their weighted expected value. Goods are assigned to the agents with the highest rankings, until all goods are allocated. Lotteries resolve ties. Agents do not need to sustain any cost because the allocation is built only on the basis of the observable characteristics (i.e. $F_1, \ldots, F_n$). The expected total value achieved by the *priority list* is equal to the sum of the m highest expected values, within the n agents. Clearly, if all agents are *ex-ante symmetric* and have equal weightings, then the optimal mechanism is an *equal chance lottery.*

Next, consider the opposite case, where *hazard rates are all monotonically decreasing.* Then $\lambda(v_i) = w_i \frac{1-F_i(v_i)}{f_i(v_i)}$. In fact, G_i is convex and so $G_i(x) = H_i(x)$ for all $x \in [0, 1]$. If agents are *ex-ante identical* and weighted equally, the optimal mechanism is a *full screening mechanism.* That is, goods are allocated to the agents with the highest realized values. In fact, $\lambda_i(v)$ is a strictly increasing function and for all i and j and v, $\lambda_i(v) = \lambda_j(v)$. Here, screening takes place up to the point where the identities of the agents that have the highest m values are known for sure. Therefore, the expected cost of screening is equal to $mE[v_{(n-m)}]$ (where $v_{(n-m)}$ indicates the $n-m$ highest value out of a sample of n independent extractions from v). The total value is given by $\sum_{z=1}^{m} E[v_{(z)}] - mE[v_{(n-m)}]$. If agents are *ex-ante asymmetric* (and the planner adopts equal weights), then the optimal mechanism will tend to be biased in favor of the agents that appear to have the strongest claims.[10]

Finally, suppose that *agents are symmetric but hazard rates are not monotonic.* In this case there may be intervals in the type space where the priority functions are constant (i.e. agents with different values receive the same priority) and other areas where these are increasing (i.e. agents with different values receive different priority). This is illustrated by the following example, which concludes the section.

[10]More precisely the mechanism will favor those agents whose distribution hazard-rate dominates those of the other agents. In fact, agent i hazard-rate dominates that of agent j if his hazard rate is always lower than that of j. Because under a monotonically decreasing hazard rate we have that $\lambda_i(v) = \frac{1-F_i(v)}{f_i(v)}$ for all v it follows that, fixing v, $\lambda_j(v) < \lambda_i(v)$. Therefore, i will get the good, even if, ex post, both i and j have the same value for it. Recall that in the optimal auction problem the designer discriminates in favor of the weakest bidder in order to extract higher payments from the strongest one.

Example: Consider the problem of distributing m tests for a rare but dangerous disease to a population of $n > m$ ex-ante identical agents, whose surplus is equally weighted by the designer. The disease can be successfully treated once discovered, but is otherwise fatal. The occurrence of the disease is highly correlated with lifestyle (e.g. sexual behavior, alcohol consumption). Therefore, potential individual benefits from taking the test, measured as the likelihood of having contracted the disease, depend on private information.[11] Everyone believes that individuals' values, measured as the willingness to spend time in line, have been independently drawn from a *piecewise uniform bimodal* distribution:

$$f(v) = \begin{cases} \frac{7}{10} & \text{if } v \in [0,1] \\ \frac{1}{10} & \text{if } v \in (1,2] \\ \frac{2}{10} & \text{if } v \in (2,3] \end{cases}$$

The optimal mechanism computed according to Proposition 1, assigns priority as follows:

$$\lambda_i(v) = \begin{cases} 0.65 & \text{if } v \in [0,1) \\ 1.16 & \text{if } v \in (1,3] \end{cases}$$

In other words, it pools agents with values in $[0,1)$ and agents with values in $[1,3]$, but screens between the two intervals. Therefore, screening is limited to discovering whether an agent belongs to the first or the second interval.

Agents that declare a lower value obtain lower priority but are not required to sustain a cost, even if they obtain a good. Agents that declare a value above 1 get priority in the allocation, but they are required to sustain a positive cost if they obtain a good. This cost must be such that an agent with value 1 is indifferent about declaring a value of 0 or a value of 1.[12] ★

[11] In a more general heterogeneous agents formulation of this example the expected benefits may depend on a combination of private and public information.

[12] If, for example , $n = 2$ and $m = 1$, the expected cost sustained by a player with a value above 1 must be set by the designer equal to: $\frac{2+\Pr\{v<1\}}{2+2\Pr\{v<1\}} = \frac{27}{34}$.

4 Practical Implementation

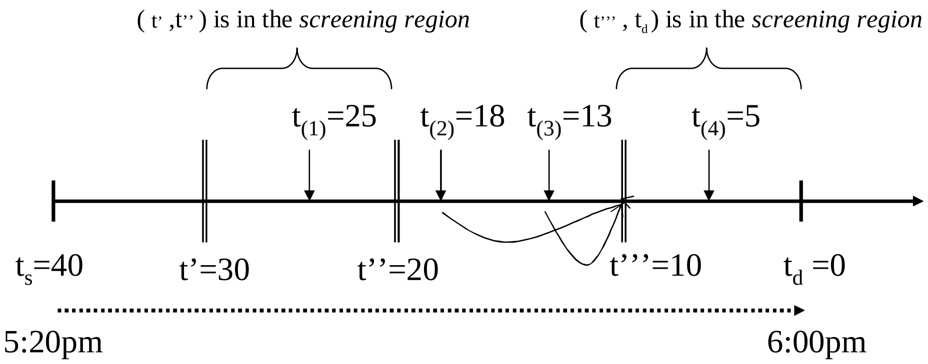
The aim of this section is to design simple mechanisms suitable for practical applications that will implement optimal outcomes when the distribution of values is known. Proposition 1 shows that if all hazard rates are monotonically non decreasing, then the optimal mechanism is a priority list, based only on observable characteristics, or an equal chance lottery (under full symmetry). Practical implementation in this case is straightforward.

When the condition above is not met it seems natural to study implementation via *queuing games*. In a queuing game the designer announces at some time t_s that the m goods will be distributed at some future date, t_d. At t_s, the designer also announces the *separating region*, T (i.e. a closed subset of the time interval $[t_s, t_d]$, which contains t_d). Thereafter, agents independently (i.e. without being able to monitor each other) decide if and when to join the queue. Agents can join the queue only once but can leave at no cost at any point in time. Upon joining the queue agents can see who is already in line. In general, the agent that arrives earlier gets priority, and ties are broken instantaneously via equal chance lotteries. However, anyone arriving outside the separating region T is counted as having joined the queue at a later time, equal to the closest subsequent point in T. [13] As agents are indifferent about when they have to wait in line, I normalize $t_d = 0$ and assume that the designer chooses t_s instead. Moreover, I can re-label the possible arrival times in terms of the *unit of times* that an agent should wait if he had to stand in line from that arrival time until t_d.[14] Figure 1 depicts, with an example, the timing and rules of a queuing game. Observe that depicted arrival times are not the equilibrium ones.

[13] For example (see Figure 1), if the designer announces at 5:20pm that she will distribute goods at 6pm and that the screening regions comprises the interval between 5:30pm and 5:40pm and the interval between 5:50pm and 6:00pm, then anyone arriving between 5:40pm and 5:50pm will be counted as if he joined exactly at 5:50pm.

[14] For example (see Figure 1), by saying that $t_s = 40$, that $T = [0, 10] \cup [20, 30]$ and that agent $i = 4$ arrives at $t_4 = 5$, I mean that the distribution time has been set at 40 units of time (i.e minutes in this example) after t_s, that the screening region comprises arrival times that are at most 10 units of time earlier than t_d or between 20 and 30 units of time from t_d, and that agent 4 decides to join the queue 5 units of time earlier than t_d.

Example: 2 goods allocated to 4 agents. *Distribution time* at 6:00pm and *screening region* is [5:30 , 5:40] and [5:50 , 6].



Outcome: one good is allocated to (1); the remaining good is allocated to (2) or (3) via an equal chance lottery.

Hereinafter, I will restrict a ttention to the case where agents are *ex-ante symmetric*, the support of possible valuations is *bounded*, and the planner *treats the agents equally* (i.e. for all i, $w_i = 1/n$ and $F_i = F$ for some F with support in $V = [\underline{v}, \overline{v}]$, where $\underline{v} < \overline{v} < \infty$).[15] Under the stated hypotheses, without loss of generality, it is possible to eliminate one variable from the implementation problem by setting $t_s = \overline{v}$.

Therefore, for each initial distribution of values, implementation requires identifying a queuing game, defined only by T, with an equilibrium such that every agent obtains the same interim payoff that he would obtain in the

[15] If the space of possible valuations is unbounded and $\lambda(x)$ is strictly increasing in some interval $[x, \infty)$, then implementation in the queueing game above can only be approximate. In fact, the designer would need to set an infinitely far off distribution time. Implementation of the asymmetric case appears difficult in this symmetric queuing game where those who arrive first must get priority and agents that do not get a good do not stay in line.

optimal direct mechanism. The following statements, which restrict agents' equilibrium behavior in the queueing game, are easily verifiable:

1. everyone joins the queue at some point (but possibly at t_d);

2. no one joins the queue outside the separating region T;

3. every agent that, upon joining the queue, observes that he has no chance of obtaining a good, immediately leaves the queue;

4. each agent joins the queue at a time such that he would be willing to wait until t_d in order to obtain a good for sure;

5. every agent that, upon joining the queue, expects to obtain a good, remains in line until t_d.

Hence, a Bayesian equilibrium in pure strategies in the queuing game defined by T can be fully characterized by a set of functions: $t_i : V_i \to T$ with $i = 1, \ldots, n$. A strategy t_i for agent i maps his possible values into arrival times within the separating region.

To achieve practical implementation, I start from the optimal direct mechanism $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ and find a permissible cost rule $\hat{\boldsymbol{c}}$ that satisfies (1b) and where only the winners pay a cost, independent of the realized values of the opponents. To do this I first need to resolve the uncertainty relative to lotteries that may arise in pooling regions. At this purpose, I associate to $\boldsymbol{p}$ the random vector $\boldsymbol{\ell}^{\boldsymbol{p}}$. Finally, call $t(v_i)$ the cost in terms of waiting time that would be assigned according to the new cost rule $\hat{\boldsymbol{c}}$ by the optimal direct mechanism to an agent with a given value if he gets a good. Implementation works by including in the separating region T all those arrival times that imply an amount of waiting time, such that, in the optimal direct mechanism with the new cost rule, there is some agent with some value that would be required to wait for that amount of time in the case that he gets a good. I have the following proposition. The proof is in the appendix.

Proposition 2. *Fix $F(\cdot), m, n$ and let $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ be an optimal mechanism according to Proposition 1. Define the random vector $(\ell_1^{\boldsymbol{p}}(\boldsymbol{v}), \ldots, \ell_n^{\boldsymbol{p}}(\boldsymbol{v}))$ as a vector of 1 and 0, where* 1 *at position i indicates the award of an object to* *agent i while 0 indicates that i will not obtain a good. Construct its probability function by defining the marginal distributions as $Prob\{\ell_i^{\mathbf{p}}(\boldsymbol{v}) = 1\} = p_i(\boldsymbol{v})$ and fixing, for each $\boldsymbol{v}$ and i, $\sum_{i=1}^{n} \ell_i^{\mathbf{p}}(\boldsymbol{v}) = m$. Define, for all $v_i \in [\underline{v}, \overline{v}]$,*

$$t(v_i) = E_{\boldsymbol{v}_{-i}}[v_i - \frac{\int_{\underline{v}}^{v_i} p_i(x_i, v_{-i})dx_i}{p_i(\boldsymbol{v})} \mid \ell_i^{\mathbf{p}}(\boldsymbol{v}) = 1]$$

Finally, define the separating region T as the image of $t(\cdot)$ over $[\underline{v}, \overline{v}]$,

$$T \equiv \{x \in [\underline{v}, \overline{v}] : t(v) = x \quad \textit{for some} \quad v \in [\underline{v}, \overline{v}]\}$$

For any $\boldsymbol{v}$, everyone with value v_i arriving at $t(v_i)$ in the queuing game defined by the designer, combined together with statements 1-5 above, is a Bayesian equilibrium that implements the optimal direct mechanism.

A few remarks are due. First, note that the formulation is general, as a lottery can be implemented by including only the distribution time in the screening region. Second, while the optimal direct mechanism is ex-post implementable, the queuing game does not admit a dominant strategy for the agents. Third, because the support of possible values is bounded, the hazard rate of the distribution of values will be increasing in a neighborhood of $\overline{v_i}$, implying that the separating region will always exclude arrival times close to $\overline{v_i}$. This means that enforcing a cap on the queue (i.e. a limit on the advance with which people can join the line) is always beneficial for welfare. Finally, it is interesting that this implementation method can be used more generally in mechanism design. In fact, Proposition 2 allows first-price implementation (i.e. only winners pay their bids) of Myerson's optimal auction when agents are symmetric but the regularity condition on the distributions of values is not met.

To conclude this section, I provide an example that clarifies the construction of the optimal, direct and indirect, mechanism in a symmetric environment where hazard rates are not monotonic.

Example: A government must distribute m food stamps to n citizens of a given town, which are treated as ex-ante identical. Individual values, i.e. the willingness to spend time in line to obtain the subsidy, are distributed according to $F_k(v) = v^k$ with $v \in [0, 1]$ and known $k \in (0, 1]$. Distributions in this family are everywhere decreasing (density is L-shaped), and distributions with a higher k first order stochastically dominate distributions with lower k (when $k = 1$ this is a uniform distribution). The inverse hazard rate is increasing in $(0,(1-k)^{\frac{1}{k}})$ and decreasing in $((1-k)^{\frac{1}{k}},1)$. Therefore, as k increases, the interval where the inverse hazard rate is decreasing gets larger (it always contains the upper bound of the support). Let $p(k) \in [0,1]$ be the unique p that solve $\frac{(1-p)p^{\frac{1-k}{k}}}{k} = \frac{k}{(1+k)(1-p)} - \frac{(1+k-p)p^{1/k}}{(1+k)(1-p)}$ for given k. Then, the optimal mechanism, assigns priorities as follows:

$$\lambda_i(v) = \begin{cases} v^{1-k}(1-v^k)k^{-1} & \text{if } \; v \in [0, p(k)^{\frac{1}{k}}) \\ p(k)^{1-k}[1-p(k)^k]k^{-1} & \text{if } \; v \in [p(k)^{\frac{1}{k}}, 1] \end{cases}$$

It can be shown that both $p(k)$ and $F_k^{-1}(p(k)) = p(k)^{1/k}$ are decreasing in k. Therefore, as expected, as k increases and the density puts more weight on high types, the optimal mechanism increases the share of pooling in the space of values. A profile of costs can be constructed that implements the allocation rule.

Turning to practical implementation, assume for simplicity that $n = 2$ and $m = 1$. The candidate symmetric equilibrium strategy is:[16]

$$t(v_i) = \begin{cases} E[v_j \mid v_j < v_i] & \text{if } \; v_i \leq p(k)^{1/k} \\ E[v_j \mid v_j < p(k)^{1/k}]\frac{2p(k)}{1+p(k)} + \frac{(1-p(k))p(k)^{1/k}}{1+p(k)} & \text{if } \; v_i > p(k)^{1/k} \end{cases}$$

The queuing game is as follows: set $t_s = 1$ and set the screening region as the image of the function t, i.e $T \equiv [0, t(x)]$. Note that when $k = 1$ this becomes a lottery, as the only arrival time in T is the distribution time.★

[16] We have that $\bar{c}_i(v_i, v_j) = 0$ if $\ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 0$ and $\bar{c}_i(v_i, v_j) = \min\{v_j, x\}$ if $\ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1$. Note that $\ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 0$ if $v_i < v_j$ and $v_i < p(k)^{1/k}$ or if $v_i > p(k)^{1/k}$ and $v_j > p(k)^{1/k}$ but the equal chance lottery favors agent j, while it is equal to 1 otherwise. Taking account of expectations produces the following ($\hat{c}_i(v_i, v_j) = 0$ if $\ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 0$):

$$\hat{c}_i(v_i, v_j) = \begin{cases} E[v_j \mid v_j < v_i] & \text{if } \; \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1,\, v_i < p(k)^{1/k} \\ \frac{E[v_j \mid v_j < p(k)^{1/k}]p(k)}{\frac{1}{2}(1-p(k))+p(k)} + \frac{\frac{1}{2}(1-p(k))}{\frac{1}{2}(1-p(k))+p(k)}p(k)^{1/k} & \text{if } \; \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1,\, v_i \geq p(k)^{1/k} \end{cases}$$

5 Conclusion

In this paper I have shown how to construct a mechanism for the efficient allocation of a set of scarce resources in an environment where money cannot be used to transfer utility but agents can still signal their value for the good at stake by engaging in wasteful activities, like waiting in line. In this context, selecting an allocation mechanism involves a trade-off between allocative efficiency and cost minimization that is not present when information about individual values can be obtained without waste of resources. The mechanisms I obtain are practically implementable, in the sense that they work essentially as priority lists, lotteries or queues, where the set of possible arrival times is appropriately restricted.

From a positive perspective, this paper suggests that one reason behind the success of different mechanisms in different environments may be their ability to balance the need to achieve an efficient allocation and the cost of wasteful rent-seeking activities that agents perform in order to secure an award. One normative implication, that is robust to different specifications of the distribution of beliefs, is that standard queues, often observed in practice, are rarely optimal. It always pays to put a cap on the queue in such a way to prevent agents to join the queue too early.

References

Calabresi, Guido and Philip Bobbit, *Tragic Choices*, New York: W. W. Norton & Company, 1978.

Chakravarty, Surajeet and Todd Kaplan, "Manna from Heaven or Forty Years in the Desert: Optimal Allocation without Transfer Payments," *mimeo, University of Exeter*, 2006.

Condorelli, Daniele, "Value, Willingness to Pay and the Allocation of Scarce Resources," *mimeo, Northwestern University*, 2008.

Elster, Jon, *Solomonic Judgements*, Cambridge, UK: Cambridge University Press, 1989.

——, *Local Justice*, New York: Russel Sage Foundation, 1992.

Hartline, Jason and Tim Roughgarden, "Optimal Mechanism Design and Money Burning," *STOC 2008 Conference proceedings*, 2008.

Holmstrom, Bengt and Roger Myerson, "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 1983, *51*, 1799–1819.

Holt, Charles and Roger Sherman, "Waiting-Line Auctions," *Journal of Political Economy*, April 1982, *90* (2), 280–294.

Kahneman, Daniel, Jack Knetsch, and Richard Thaler, "Fairness as a Constraint on Profit-seeking: Entitlements in the Market," *American Economic Review*, September 1986, *76* (4), 728–741.

Koh, Winston, Zhenlin Yang, and Lijing Zhu, "Lottery rather than waiting-line auction," *Social Choice and Welfare*, 2006, *27*, 289–310.

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

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

Okun, Arthur M., *Equality and Efficiency: The Big Tradeoff*, Washington: Brookings Institution Press, 1975.

Taylor, Grant, Kevin Tsui, and Lijing Zhu, "Lottery or waiting-line auction?," *Journal of Public Economics*, 2003, *87*, 1313–1334.

Walzer, Michael, *Spheres of Justice*, New York: Basic Books, 1983.

Yoon, Kiho, "Mechanism Design with Expenditure Consideration," *mimeo, Korea University*, 2009.

Appendix

Proposition 1 For each $\boldsymbol{v}$, the set of agents is partitioned in a chain of ordered sets, $M_1(\boldsymbol{v}), M_2(\boldsymbol{v}), \ldots$ according to their priority levels. Formally, set $M_0(\mathbf{v}) \equiv \emptyset$ and define $M_x(\mathbf{v})$ recursively as follows:

$$M_{x+1}(\boldsymbol{v}) \equiv \left\{ i \in N \setminus \bigcup_{z \leq x} M_z \mid \lambda_i(v_i) = \max_{j \in \{N \setminus \cup_{z \leq x} M_z\}} \lambda_j(v_j) \right\}$$

Define $I_j(\boldsymbol{v})$ as the set of agents with the highest priority levels, up to those included in $M_j(\boldsymbol{v})$: $I_j(\boldsymbol{v}) = \{i \in \bigcup_{z \leq j} M_z(\boldsymbol{v})\}$.

Let $|X|$ denote the cardinality of an arbitrary set X. Pick the highest natural number s such that $|I_s(\boldsymbol{v})| \leq m$. Call $k = m - |I_s(\boldsymbol{v})|$, and $r = |M_{s+1}(\boldsymbol{v})|$. An incentive compatible symmetric direct allocation mechanism $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ maximizes (7) if, and only if, $\forall i \in N$, $\forall \mathbf{v} \in V^n$: [17]

$$p_i(\boldsymbol{v}) = \begin{cases} 1 & \text{if } \ i \in I_s(\boldsymbol{v}) \\ k/r & \text{if } \ i \in M_{s+1}(\boldsymbol{v}) \\ 0 & \text{otherwise} \end{cases}$$

The cost rule can be any set of functions $\boldsymbol{c}$ such that for all i and $v \in V_i$:

$$C_i(v_i) = v_i P_i(v_i) - \int_{\underline{v_i}}^{v_i} P_i(x) dx$$

Proof: As a first step towards the solution to the problem, let us rewrite the objective function using Lemma 1 to substitute for the cost functions:

$$\mathrm{E}_{\boldsymbol{v}} \left\{ \sum_{i=1}^{n} w_i [v_i p_i(\boldsymbol{v}) - c_i(\boldsymbol{v})] \right\} = \sum_{i=1}^{n} w_i \{ \int_{\underline{v_i}}^{\overline{v_i}} (\int_{\underline{v}}^{v_i} P_i(x) dx) dF_i(v_i) \}$$

By changing the order of integration, and integrating by parts:

$$\mathrm{E}_{\boldsymbol{v}} \left[\sum_{i=1}^{n} w_i p_i(\boldsymbol{v}) \frac{1 - F_i(v_i)}{f_i(t_i)} \right]$$

[17] I restrict attention to symmetric mechanisms. This is without loss of generality here because considering asymmetric mechanisms will not improve on the symmetric solution obtained.

The designer's problem is now the following:

$$\max_{p_i:\boldsymbol{V}\to[0,1]\ i=1,\dots,n} \mathrm{E}_{\boldsymbol{v}}\left[\sum_{i=1}^{n} w_i p_i(\boldsymbol{v})\frac{1-F_i(v_i)}{f_i(v_i)}\right]$$

subject to:

$$\sum_{i=1}^{n} p_i(\boldsymbol{v}) \leq m \ \forall \boldsymbol{v} \in \boldsymbol{V}$$

$$P_i(v) \geq P_i(v^*) \ \forall i \in N, \quad \forall v, v^* \in V_i : v \geq v^*$$

It can readily be seen that the candidate solution satisfies the first constraint above, and that $\boldsymbol{c}$ satisfies (1b). To prove that (1a) is also satisfied, note that $\lambda_i(v)$ is the derivative of a convex function and therefore it is monotonically increasing. Then, $\forall \boldsymbol{v_{-i}}\ p(\boldsymbol{v})$ is increasing in v_i, which implies that $P_i(\cdot)$ is also increasing. Now, let us sum and subtract $\lambda_i(v_i)/w_i$ inside the objective function and rewrite it to obtain:

$$\sum_{i=1}^{n} \mathrm{E}_{v_i}\left\{P_i(v_i)\lambda_i(v_i) + P_i(v_i)\left[w_i\frac{1-F_i(v_i)}{f_i(v_i)} - \lambda_i(v_i)\right]\right\}$$

Consider the second term of this expression for every i:

$$w_i \int_{\underline{v_i}}^{\overline{v_i}} P_i(v_i)\left[\frac{1-F_i(v_i)}{f_i(v_i)} - g_i(F_i(v_i))\right] f(v_i) dv_i$$

Integrating by parts:

$$w_i P_i(v_i)\left[H_i(F_i(v_i)) - G_i(F_i(v_i))\right] \Big|_{\underline{v_i}}^{\overline{v_i}} - \\ w_i \int_{\underline{v_i}}^{\overline{v_i}} \left[H_i(F_i(v_i)) - G_i(F_i(v_i))\right] dP_i(v_i))$$

Consider the first term of the expression above. It is equal to zero: $H_i(\underline{v_i}) = G_i(\underline{v_i})$ and $H_i(\overline{v_i}) = G_i(\overline{v_i})$, because G_i is the convex hull of the continuous function H_i and thus they coincide at endpoints (the continuity of H_i follows from assuming an atomless f_i). The objective function becomes:

$$\sum_{i=1}^{n} \mathrm{E}_{\boldsymbol{v}}\left[p_i(\boldsymbol{v})\lambda_i(v_i)\right] - \sum_{i=1}^{n} w_i \int_{\underline{v_i}}^{\overline{v_i}} \left[H_i(F_i(v_i)) - G_i(F_i(v_i))\right] dP_i(v_i)$$

It is easy to see that the candidate solution $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ maximizes the first sum as it puts all the probability on the players for whom $\lambda_i(v_i)$ is maximal. To conclude the proof, I now show that the second term is equal to zero. It must always be non negative, as $\forall v \in [\underline{v_i}, \overline{v_i})$ $H_i \geq G_i$. That it is equal to zero, follows because G_i is the convex hull of H_i and so, whenever $H_i(F_i(v_i)) > G_i(F_i(v_i))$, then G_i must be linear. That is, if $G(x) < H(x)$, $G_i''(x) = g_i'(x) = 0$. Therefore, to conclude, $\lambda_i(v)$ will be a constant in a neighborhood of v_i, which implies that $P_i(v)$ will also be a constant. ▮

Proposition 2 Fix $F(\cdot), m, n$ and let $\langle \boldsymbol{p}, \boldsymbol{c} \rangle$ be an optimal mechanism according to Proposition 1. Define the random vector $(\ell_1^{\boldsymbol{p}}(\boldsymbol{v}), \ldots, \ell_n^{\boldsymbol{p}}(\boldsymbol{v}))$ as a vector of 1 and 0, where 1 at position i indicates the award of an object to agent i while 0 indicates that i will not obtain a good. Construct its probability function by defining the marginal distributions as $Prob\{\ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1\} = p_i(\boldsymbol{v})$ and fixing, for each $\boldsymbol{v}$ and i, $\sum_{i=1}^n \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = m$. Define, for all $v_i \in [\underline{v}, \overline{v}]$,

$$t(v_i) = E_{\boldsymbol{v}_{-i}}[v_i - \frac{\int_{\underline{v}}^{v_i} p_i(x_i, v_{-i}) dx_i}{p_i(\boldsymbol{v})} \mid \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1]$$

Finally, define the separating region T as the image of $t(\cdot)$ over $[\underline{v}, \overline{v}]$,

$$T \equiv \{x \in [\underline{v}, \overline{v}] : t(v) = x \quad \text{for some} \quad v \in [\underline{v}, \overline{v}]\}$$

For any $\boldsymbol{v}$, everyone with value v_i arriving at $t(v_i)$ in the queuing game defined by the designer, combined together with statements 1-5 above, is a Bayesian equilibrium that implements the optimal direct mechanism.

Proof: First, I show that if everyone follows the candidate equilibrium strategy in the queuing game, then the outcome of the optimal direct mechanism is implemented. Second, I show that, the candidate equilibrium is indeed an equilibrium. For any $\boldsymbol{v} \in V^n$, suppose that everyone arrives at $t(v_i)$. According to statements 3 and 5, an agent leaves immediately if he does not get a good, but otherwise remains in line until the distribution time. In the queuing game goods are allocated to the agents that join the queue earliest, within the set of arrival times T, and lotteries resolve ties immediately. Hence, if everyone uses arrival strategy $t(v_i)$ goods are allocated to those with the highest $t(v_i)$ and lotteries resolve ties. Those that do not get a good do not pay any cost. Consider now the optimal direct mechanism with the $\hat{\boldsymbol{c}}$ cost rule, under symmetry. As $Prob\{\ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1\} = p_i(\boldsymbol{v})$, it is easy to verify that it satisfies the requirement for incentive compatibility. Formally, this is defined as follows:

$$\bar{c}_i(v_i) = \begin{cases} 0 & \text{if } \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 0 \\ E_{\boldsymbol{v}_{-i}}[v_i - \frac{\int_{\underline{v_i}}^{v_i} p_i(x_i, v_{-i}) dx_i}{p_i(\boldsymbol{v})} \mid \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1] & \text{if } \ell_i^{\boldsymbol{p}}(\boldsymbol{v}) = 1 \end{cases}$$

Incentive compatibility (1a) requires that the priority $\lambda(v_i)$ in the assignment is non-decreasing in the value of each agent. Furthermore, incentive compatibility (1b) implies that agents obtaining the same priority bear the same interim cost (i.e. $\hat{C}(v_i)$). It follows that they also bear the same cost ex-post in the case of success (i.e. $\hat{c}(v_i)$ when $\ell^p(\boldsymbol{v}) = 1$), as the cost does not depend on the realized values of all other agents. As a consequence, $t(v_i)$ increases when $\lambda(v_i)$ increases and is constant otherwise. This proves, given the rules of the queuing game, that the strategy considered implements the same outcome as the optimal direct mechanism.

To conclude this proof, I need to show that this is indeed an equilibrium of the queuing game defined by the designer. Remember that i has no interest in arriving at any time outside the relevant set defined, which is the image of $t(\cdot)$. Then, if everyone plays according to $t(\cdot)$, an agent i with value v_i essentially faces the choice between arriving at $t(v_i)$, or mimicking what an agent with some other value would choose, according to the candidate equilibrium strategy. Therefore, the payoffs are the same as in the direct optimal mechanism. It follows that, if the agent chooses to use a strategy other than $t(v_i)$, then he will obtain the outcome assigned to an agent with a different value in the optimal direct mechanism. But this is not possible because the optimal direct mechanism is incentive compatible. ∎