Direct Preference Elicitation Over Public Outcomes

Greg Leo*

Abstract

I apply implementation theory to study the design of experiments that elicit subject preferences over public outcomes. Any experiment that can infer preferences based on dominant strategies is outcome equivalent to a direct preference elicitation experiment where subjects directly state their preferences, and the outcome is chosen according to a strategy-proof social choice function. While all strategy-proof social choice functions make truthful reporting a dominant strategy, the strength of incentives differ. I demonstrate that a design called Random Dictator over Pairs [RDP] provides the strongest incentives to truthfully report an entire rank-ordering for any set of subject beliefs. Furthermore, it is the only design that is also robust to intransitive preferences.

I Introduction

Lab experiments are often used by economists to measure subject preferences over various kinds of outcomes. These outcomes may be purely private as in the study of risk preferences or time preferences[1]. The focus of this paper is eliciting preferences when the outcomes are public. In addition to the extensive literature on social preference and distributional preference[2], examples include preferences over who contributes to a public good (Bergstrom et al., 2019), or which set of rules to use in playing a game (Dal Bó et al., 2015).

The goal of the experimenter is to truthfully elicit unknown preferences. Unknown preferences also play a central role in mechanism design. A mechanism designer's goal is to implement a particular set of contingent outcomes (also known as a *social choice function*[3]) as the equilibrium of a game. However, when a social choice function can be implemented as the equilibrium of *any* game, it can be implemented as a truthful equilibrium of a game that *directly elicits preferences* and chooses the desired contingent outcome. This is the celebrated *revelation principle* (Gibbard, 1973).

This fact is relevant to the experiment designer as well. Any implementable social choice function can be used to truthfully and *directly* elicit unknown preferences. When used in the context of experiment design, I refer to implementable[4] social choice functions as *direct elicitation methods.* However, while any direct elicitation method makes truth-telling

*Leo: Vanderbilt University Department of Economics. VU Station B #351819 2301 Vanderbilt Place Nashville, TN 37235-1819. g.leo@vanderbilt.edu

[1]See for instance Kagel et al. (1995); Holt et al. (2002); Andersen et al. (2008); Andreoni and Sprenger (2012a,b).

[2]See, for instance, Andreoni and Miller (2002); Engelmann and Strobel (2004); Fisman et al. (2007); Chen and Li (2009).

[3]A mapping from preference profiles into outcomes.

[4]Throughout the paper, I focus on strategy-proof implementation.

a dominant strategy, the strength of incentives may differ. For instance, while randomly choosing an outcome is a strategy-proof SCF, the incentives are weak. Subjects are indifferent between submitting any preference ranking. *Random Dictator* (choosing a subject at random and implementing their favorite outcome) provides strong incentives for each player to correctly rank their favorite outcome but are otherwise weak.

The experiment designer's problem in this setting is to choose the particular direct elicitation method that provides the most desirable set of incentives. I suggest a solution to this problem when the experimenter has no strong prior about subject preferences and beliefs. The result is a generalized experiment design for eliciting ordinal preferences over public outcomes.

The proposed design is *Random Dictator over Pairs [RDP]*. Subjects submit rank-ordered lists over public outcomes. One subject and pair of outcomes are chosen randomly (with uniform probability). The outcome implemented is the one that the chosen subject ranks highest from the chosen pair. *RDP* has the property that it elicits truthful preferences regardless of subjects' beliefs (it is strategy-proof). Among all direct elicitation methods, it provides the strongest incentives against non-truthfully swapping the even most weakly incentivized pair of outcomes over all players and possible preference profiles.

I also discuss *RDP* modifications to deal with instances where the experimenter is interested in providing stronger incentives for subjects to reveal preferences near the top of their rank-order list. This analysis demonstrates a significant non-linear trade-off between incentives' strength and how those incentives are spread over the rank-order list. *RDP* settles this in one extreme, spreading the incentives as evenly as possible to collect information about the entire rank order. A classic alternative, **Random Dictator [*RD*]**, settles this at the other extreme, concentrating incentives at the top of the rank-order list but providing much stronger incentives there. I propose a generalized design that nests both *RDP and RD.* I also show that *RDP* is robust to preference intransitivities.

This paper is organized as follows, section II presents the general theoretical environment, and examples of direct elicitation methods. Section III proves the optimality of *RDP* and discusses extensions. Section IV extends the analysis to intransitive preferences. Section V provides applications of the results in this paper and concludes.

II Environment

A Environment and Definitions

V is a finite set of public alternatives. n players each have preferences over V- a total-ordering represented by the ranking P. $\mathbf{P}$ is ranking-profile, a vector of n rankings $\mathbf{P} = (P_1, P_2, ..., P_n)$. $\left(\tilde{P}, \mathbf{P}_{-i}\right)$ is a ranking-profile in which all but the ith element are identical to that in $\mathbf{P}$ and the ith element is $\tilde{P}$. $\mathbf{P}^{ky}$ is a profile generated from $\mathbf{P}$ by raising the outcome y up by one position in k's ranking.

Each player assigns a cardinal utility to every member of V represented by the Bernoulli utility function $U(x)$ where $x \in V$. This is extended to a VNM utility function by assigning for any lottery ρ over V, $U(\rho) = \sum_{x \in V} U(x)\rho(x)$.

A **social-choice function** d is a assigns each $\mathbf{P}$ a lottery over V denoted $d(P)$. A strategy-proof social-choice function is one for which there is no U $\mathbf{P}$,$\tilde{P}$ such that: $U\left(d\left(\tilde{P}, \mathbf{P}_{-i}\right)\right) > U(d(P, \mathbf{P}_{-i}))$.

If the experiment designer is interested in truthfully and robustly (regarding the player's beliefs) eliciting preferences, then any strategy-proof social choice function is sufficient. Designing an experiment in this environment amounts to choosing a strategy-proof social choice function. Because of this equivalence, I refer to strategy-proof social choice functions as **Direct Elicitation Methods.** Below are several examples of direct elicitation methods.

Random Choice *[RC]: An outcome is chosen randomly.*

Random Dictator *[RD]: A single player i is chosen randomly. The implemented outcome is the one i ranks first.*

Majority over Random Pairs *[MRP]: A pair of outcomes are chosen at random. The implemented outcome is the one that a majority rank higher. (With ties broken randomly.)*

Random Dictator over Pairs *[RDP]: A pair of outcomes and a single player i are chosen randomly. The implemented outcome is the one that i ranks higher from the pair.*

These examples demonstrate the potential incentive shortcomings of alternative elicitation methods. *Random dictator* provides strong incentives for each player to list their top outcome truthfully. However, the incentives for listing outcomes below the top are weak. Players are indifferent between submitting any ranking which correctly positions their top choice. In *Majority over Random Pairs,* reporting truthfully is also dominant. However, the strength of incentives depends on beliefs. For instance, if a player thinks everyone else will submit the same ranking, they are indifferent between submitting all rankings.

RDP appears to overcome both of these shortcomings. As long as players are not indifferent between outcomes, they are not indifferent between submitting any rankings, regardless of their beliefs. Informally, *RDP* appears to be a desirable solution when the goal is to elicit the full ranking robustly. To formalize this, I start by providing Gibbard's characterization of strategy-proof social choice functions *(Gibbard, 1977)*.

B Characterizing Direct Elicitation Methods

The following two requirements are necessary and sufficient for a social choice function to be strategy-proof *(Gibbard, 1977)*:

Pairwise-Responsive: *social choice functions for which one player swapping two adjacent outcomes affects only the probabilities of those outcomes.*

Non-Perverse**:** *social choice functions for which swapping a single outcome upwards in a player's preferences cannot decrease the probability of that outcome being chosen.*

Theorem 1. *Gibbard (1977) Characterization*: a social choice function is strategy-proof if and only if it is pairwise responsive and non-perverse.

The next result provides the ingredients for building strategy-proof social choice functions. They are all mixtures[5] of two simple types of rules:

[5]Formally d is a mixture of k social choice functions $\{d_1, d_2, ...d_k\}$ if there are positive weights $(w_1, w_2, ...w_k)$ summing to 1 such that $d = w_1 d_1 + w_2 d_2 + ...w_k d_k$

Unilateral rules: *social choice functions that depend only on a single player's preferences.*

Duple rules: *social choice functions that choose between a pair of outcomes.*

Theorem 2. *Gibbard (1977) Mixture*: a social choice function is strategy-proof if and only if it is a mixture of pairwise responsive and non-perverse social choice functions that are all either unilateral or duple rules.

Referring back to the examples, *Random Choice* is mixture over unilateral rules where each outcome set is a singleton. *Random Dictator* is a mixture of unilateral rules where each outcome set is the entire set V. *Majority over Random Pairs* is a mixture of the duple rules which choose between each possible pair of outcomes using the majority rule (a pairwise responsive and non-perverse rule). *Random Dictator over Pairs* is a mixture of unilateral rules where each outcome set is a pair.

Other, more exotic rules exist in the set as well. For instance: randomly choose a subset of A. If the subset is a pair, choose by majority rule, otherwise choose by Random Dictator. Any mixtures of the rules mentioned above are also a direct elicitation methods.

III Choosing an Elicitation Method

A Optimal Robust Elicitation

Since the designer is neutral about players' cardinal utilities, their objective function is ordinal. Let $d_y(P)$ denote the probability assigned by $d(P)$ to outcome $y \in V$ and let $\epsilon_k^y(d, \mathbf{P}) = d_y\left(\mathbf{P}^{ky}\right) - d_y(\mathbf{P})$ for any y that is not ranked at the top of k's list. This is the probability increase for outcome y if player k moves y up by one position in their ranking. Another way to think about ϵ is the probability that a player will receive their less desired outcome from a pair if they accidentally swaps that pair in their reported order. Higher values of ϵ represent stronger incentives. For instance, if d is *random choice* $\epsilon_k^y(d, \mathbf{P}) = 0$ everywhere, since preference reports are ignored. In a scenario with only one player and two outcomes where d chooses the player's favorite reported outcome $\epsilon_k^y(d, \mathbf{P}) = 1$. Falsely swapping preferences will always result in an undesirable outcome. With n players, where d is *Random Dictator*, $\epsilon_k^y(d_l, \mathbf{P}) = \frac{1}{n}$ for y that is ranked second in a player's list, but 0 elsewhere.

When the designer wishes to elicit a full ranking and has no strong prior about the distribution of $\mathbf{P}$ or player's beliefs about $\mathbf{P}$, it is reasonable to focus on the worst-case scenario incentives. In this case, the designer's objective is to maximize the minimum value of ϵ_k^y over all players, outcomes, and preference profiles. That is, the designer's goal is to find a rule d^* which solves: $Max_{d\in\mathcal{D}}\left[Min_{k,y,\mathbf{P}}\epsilon_k^y(d, \mathbf{P})\right]$. I refer to any rule which solves this problem *as* ***Robustly Optimal***. Below, I show that *RDP* is a robustly optimal rule. The proof is rather straight-forward. The following lemma establishes an upper bound for $Min_{k,y,\mathbf{P}}\epsilon_k^y(d, \mathbf{P})$, which I then show to be met by *RDP*.

Lemma. $Min_{k,y,\mathbf{P}}\epsilon_k^y(d, \mathbf{P}) \leq \frac{2}{n(m-1)m}$.

Proof. Denote $Min_{k,y,\mathbf{P}}\epsilon_k^y(d,\mathbf{P})=\tilde{\epsilon}$ and suppose $\tilde{\epsilon}>\frac{2}{n(m-1)m}$ and so any $\epsilon_k^y(d,\mathbf{P})>\frac{2}{n(m-1)m}$. Start from a profile $\mathbf{P_0}$ where player's rank the outcomes identically such that the ranking of any player is $(y_1,y_2,y_3,...,y_m)$. Now raise y_m to the top of each player's ranking to get $\mathbf{P}_1$. In making this change, y_m has been raised by $m-1$ positions in each of n player's rankings, increasing the probability of the outcome y_m by at least $(m-1)n$ times $\tilde{\epsilon}$. Thus, $d_{y_m}(\mathbf{P}_1)>\frac{2}{n(m-1)m}(m-1)n=\frac{2}{(m-1)m}(m-1)$. Now starting from $\mathbf{P}_1$, raise y_{m-1} up in every player's ranking until it is below y_m. Let $\mathbf{P}_2$ denote the new profile of rankings. By pairwise-responsiveness, $d_{y_m}(\mathbf{P}_2)=d_{y_m}(\mathbf{P}_1)$. By an analogous argument to that above, $d_{y_{m-1}}(,\mathbf{P}_2)>\frac{2}{n(m-1)m}(m-2)n=\frac{2}{(m-1)m}(m-2)$. Continuing in this way for each successive $y_{m-2},y_{m-3},...,y_1$. The final profile $\mathbf{P}_m$ has the outcomes reversed with respect to the original $\mathbf{P_0}$.

Under the final profile of rankings, $d_{y_m}(\mathbf{P}_m)>\frac{2}{(m-1)m}(m-1)$, $d_{y_{m-1}}(\mathbf{P}_m)>\frac{2}{(m-1)m}(m-2),d_{y_{m-2}}(P_m)>\frac{2}{(m-1)m}(m-3),...,d_{y_1}(\mathbf{P}_m)>0$. Summing over these inequalities yields $\sum_{i\in V}d_i(\mathbf{P}_m)>\frac{2}{(m-1)m}\sum_{i=1}^m(m-i)=\frac{2}{(m-1)m}\frac{1}{2}(m-1)m=1$ implying the probability that some outcome in V is chosen is strictly greater than 1, a contradiction. □

Proposition 3. *(RDP is Robustly Optimal.)*

Proof. In RDP, raising an outcome y up one position only changes the outcome if the player is randomly chosen as the dictator, which happens with probability $\frac{1}{n}$. Conditional on being chosen, raising y up one position only changes the outcome when the pair of chosen outcomes is y and the outcome immediately below y (previously above y) since this is the only outcome for which y's relative position has changed. This probability these two outcomes are chosen is $\frac{2}{m(m-1)}$. The change in probability for increasing y by one position is $\epsilon_k^y(d,\mathbf{P})=\frac{2}{nm(m-1)}$ for any y,$\mathbf{P}$, k. Thus, the $Min_{k,y,\mathbf{P}}\epsilon_k^y(d,\mathbf{P})=\frac{2}{n(m-1)m}$. Since RDP attains the upper bound established above, it is robustly optimal. □

B Comparison to Voting Under Uniform Preferences

Under particular preference distributions, alternative direct elicitation procedures may offer better incentives. For instance, under majority rule, every player who believes they will be pivotal with a high probability faces strong incentives to correctly rank the two relevant outcomes. When n is odd, a player who always believes they will be pivotal under random majority rule over random pairs has $\epsilon_k^y(d,\mathbf{P})=\frac{2}{(m-1)m}$ (n time larger than under *RDP*). This requires very particular beliefs, and for any player who is not pivotal in a vote, $\epsilon_k^y(d,\mathbf{P})=0$.

Under uniform preferences, the probability a player is pivotal for a particular pair of outcomes (when n is odd) is $\binom{n-1}{\frac{n-1}{2}}\left(\frac{1}{2}\right)^{n-1}$. This gives $\epsilon_k^y(d,\mathbf{P})=\frac{2}{(m-1)m}\binom{n-1}{\frac{n-1}{2}}\left(\frac{1}{2}\right)^{n-1}$. Thus under uniform preferences, majority rule provides better incentives than RDP as long as $\binom{n-1}{\frac{n-1}{2}}\left(\frac{1}{2}\right)^{n-1}>\frac{1}{n}$ which is true when $n>1$.

$$\binom{n-1}{\frac{n-1}{2}}(p)^{\frac{n-1}{2}}(1-p)^{\frac{n-1}{2}}=\frac{1}{n}$$

IV Intransitive Preferences

The analysis above assumes a cardinal utility function. However, if preferences over outcomes are not transitive, then this assumption is invalid. One popular means for testing the assumption of transitivity is through violations of the Generalized Axiom of Revealed Preference. In this section, I extend the analysis of procedures to elicit potentially intransitive preference relations.

A Ordinal Environment

The objects of interest in this section are binary preference relations. A profile of preference relations for all players is denoted $\succ$ and $\succ_{-i}$ is given the usual meaning. Let $\mathcal{S}_i$ be the set of all asymmetric relations on V (the set of outcomes) and $\mathcal{S} = \times_{i-1}^{n} \mathcal{S}_i$. An **ordinal social-choice function** d assigns every $\succ \in \mathcal{S}$ a lottery over V denoted $d(\succ)$.

Familiar social-choice functions in this ordinal environment must be suitably redefined. For instance, in the cardinal environment, *Random Dictator* chooses the maximal element from a player's preference ordering. However, a maximal element might not exist with intransitive preferences, such as in example 5 below. One way to generalize Random Dictator to the ordinal environment is to choose a dictator and an element from the top-cycle of that dictator's preferences. A top-cycle is a set of outcomes $T \subseteq V$ such that for all $x \in T$ and all $y \in V/T$, $x \succ_i y$. If $\succ_i$ is transitive, then the set T is a singleton, the maximal element. In example 5 below, I refer to this social choice function as *ordinal random-dictator.*

There is a further complication in extending preferences about V to preferences about lotteries over V and defining a suitable notion of strategy-proofness. Here, I focus on generalizing the conditions which are necessary and sufficient for strategy-proofness in the cardinal setting.

Ordinal Pairwise-Responsive: *social choice functions for which one player swapping a single pair $x \succ_i y$ to $y \succ_i x$ affects only the probabilities of x and y.*

Ordinal Non-Perverse**:** *social choice functions for which one player swapping a single pair $x \succ_i y$ to $y \succ_i x$ does not strictly decrease the probability of y.*

It is noted that when restricted to transitive preference orderings, these requirements are identical to the necessary and sufficient conditions for strategy-proofness from Gibbard (1977). Instead of defining and characterizing an ordinal notion of strategy-proofness, I define ordinal strategy-proof using these two generalized properties.

Definition 4. An ordinal social choice function d is **Ordinal Strategy-Proof** if and only if it is *ordinal pairwise-responsive* and *ordinal non-perverse.*

Example 5. (*Ordinal Random Dictator* is not Ordinal Strategy-Proof).

Suppose $n = 2$ and $V = \{x, y, z\}$. Player one has cyclic preferences $x \succ_1 y$, $y \succ_1 z$, $z \succ_1 x$ and player two has transitive preferences $x \succ_2 y$, $y \succ_2 z$, $x \succ_2 z$. Under truthful reporting, x, y, z are all in the top-cycle of $1's$ preferences. Thus, if 1 is chosen as the dictator, one of these three will be chosen at random. On the other hand, x is the only element of 2's top-cycle. Thus, under truthful reporting, the probability that x is chosen is $\frac{2}{3}$, the probability y is chosen is $\frac{1}{6}$, and the probability z is chosen is $\frac{1}{6}$.

If player 1 instead reports the profile $\tilde{\succ}_1$: $x\tilde{\succ}_1 y$, $y\tilde{\succ}_1 z$, $x\tilde{\succ}_1 z$ (by swapping only the relation on the pair x, z) then the probability x is chosen is 1. Since the probability that y is chosen also changes, this demonstrates that the rule is not ordinal pariwise-responsive.

Definition 6. Ordinal Random Dictator over Pairs [*ORDP*].

Ordinal Random Dictator over Pairs randomly selects an $i \in \{1, ..., n\}$ and pair of (non-identical) outcomes x, y. It implements the outcome (either x or y) that is preferred according to reported preferences $\succ_i$.

Proposition 7. *ORDP is Ordinal Strategy-Proof.*

Proof. Let $p_{i,\{x,y\}}$ *be the probability that* i *is chosen as dictator and* $\{x, y\}$ *as the random pair. For a particular* i *and* $\succ$ *assume without loss of generality that* $x \succ_i y$. *If the relation on* x *and* y *is swapped by player* i *then the probability that* y *is chosen increases by* $p_{i,\{x,y\}}$. *The probability* x *is chosen decreases by the same amount. Thus, the rule is ordinal non-perverse. Since no other outcomes are affected, the rule is ordinal pairwise-responsive.* □

Thus, in addition to being robust in cardinal environments, *RDP* provides robust incentives against falsely changing elements of the binary relation $\succ_i$ even when that relation is intransitive. However, to learn about potentially intransitive preferences, the binary relation $\succ_i$ must be elicited, rather than the ranking P_i.

The possibility of intransitive preferences further restricts the set of strategyproof mechanisms. When $\succ_i$ is restricted to be a well-ordering for each player, strategy-proof rules are mixtures of unilateral and duple rules. However, when$\succ_i$ may be intransitive, the set of strategy-proof mechanisms are equivalent to mixtures of *only* duple rules.

Proposition 8. *A Mechanism is Ordinal Strategy-Proof if and only if it is a mixture over Duple Rules.*

The proof that a duple rule is strategy-proof is a straight-forward extension of Proposition 7. Thus, this proof focuses on proving that Ordinal Strategy-Proof implies a mixture over duple rules.

Let $\tilde{\succ}_i^S$ be the restriction of i's preferences to the set $S \subseteq V$. The proof below focuses only on what happens when a player changes reported preferences among three potential outcomes. Let $a, b, c \in V$. For convenience, I represent a the preference restriction $\succ_i^{\{a,b,c\}}$using binary strings as follows. The first element of the string represents the direction of i's preference between a and b. The value of the first element is 1 if $a \succ_i b$ and 0 if $a \prec_i b$. The second value is 1 if $b \succ_i c$ and 0 if $b \prec_i c$. The third value is 1 if $a \succ_i c$ and 0 if $a \prec_i c$. Using this representation, there are eight possible preference orderings player i can have over the three outcomes.

Consider the probability that the mechanism assigns to outcomes a, b and c for a fixed $\succ_i^{V/\{a,b,c\}}$ and $\succ_{-i}$. Since d must be ordinal pairwise-responsive, changing only $\succ_i^{\{a,b,c\}}$leaves the probabilities of outcomes other than a, b, c fixed. The following table provides, for each possible $\succ_i^{\{a,b,c\}}$a probability that each outcome a, b, c is chosen.

$\succ_i^{\{a,b,c\}}$	$Pr\ a$	b	c
$0,0,0$	a_1	b_1	c_1
$0,0,1$	a_2	b_2	c_2
$0,1,0$	a_3	b_3	c_3
$0,1,1$	a_4	b_4	c_4
$1,0,0$	a_5	b_5	c_5
$1,0,1$	a_6	b_6	c_6
$1,1,0$	a_7	b_7	c_7
$1,1,1$	a_8	b_8	c_8

Table 1: Probabilities of each outcome conditional on reported preference relation.

Pairwise-responsiveness also places other restrictions on the mechanism. For instance, $a_1 = a_3$ since $0,0,0$ and $0,1,0$ differ only by the preference ordering between outcomes b and c. Applying this logic to the grid provides the following:

$\succ_i^{\{a,b,c\}}$	$Pr\ a$	b	c
$0,0,0$	a_1	b_1	c_1
$0,0,1$	a_2	b_1	c_2
$0,1,0$	a_1	b_2	c_3
$0,1,1$	a_2	b_2	c_4
$1,0,0$	a_3	b_3	c_1
$1,0,1$	a_4	b_3	c_2
$1,1,0$	a_3	b_4	c_3
$1,1,1$	a_4	b_4	c_4

Table 2: Probabilities of each outcome conditional on reported preference relation after applying pairwise-responsiveness assumption.

By pairwise-responsiveness, the sum of each row must be the same. Thus, $a_1 + b_1 + c_1 = a_4 + b_3 + c_2$. Similarly, by pairwise responsiveness, $b_1 = b_3 + (a_3 - a_1)$. From these two equalities,

$a_4 = a_1 + (a_3 - a_1) + (a_2 - a_1)$. Letting $\alpha = a_2 - a_1$ and $\beta = a_3 - a_1$. The grid can be re-written as follows:

$\succ_i^{\{a,b,c\}}$	$Pr\ a$	b	c
$0,0,0$	a_1	$b_3 + \alpha$	$c_4 + \beta + \gamma$
$0,0,1$	$a_1 + \beta$	$b_3 + \alpha$	$c_4 + \gamma$
$0,1,0$	a_1	$b_3 + \gamma$	$c_4 + \beta$
$0,1,1$	$a_1 + \beta$	$b_3 + \gamma$	c_4
$1,0,0$	$a_1 + \alpha$	b_3	$c_4 + \beta + \gamma$
$1,0,1$	$a_1 + \alpha + \beta$	b_3	$c_4 + \gamma$
$1,1,0$	$a_1 + \alpha$	$b_3 + \alpha + \gamma$	$c_4 + \beta$
$1,1,1$	$a_1 + \alpha + \beta$	$b_3 + \alpha + \gamma$	c_4

Table 3: Probabilities of each outcome are consistent with a mixture of duple rules.

This table is a mixture of duple rules. For instance, with probability α, the mechanism chooses the pair a, b and assigns a if $a \succ b$ and b if $b \succ a$.

V Discussion

A *Applications*

Below is a brief selection of potential applications of the direct elicitation methods discussed here.

Preferences over Games

An application related to Dal Bó et al. (2015) regards eliciting preferences over which type of game players would like to engage. Dal Bó et al. (2015) have player's selecting between two games. They use both Random Dictator and majority rules. Since there are only two outcomes in their experiment, both are strategy-proof. RDP is a natural direct elicitation method to extend such an experiment to elicit player's full rank-order over three or more games.

On the other hand, *RDRS-l* would be a natural design if there are many possible games, and only a subset of the rank-ordering is of interest. For instance, if the experimenter is interested in learning about each player's top three games from a set of ten then *RDRS*-8 (choosing a random dictator's favorite outcome over a random set of 8 games) provides stronger incentives for correctly revealing the top three for each player while making them indifferent in ranking the bottom seven. As discussed above, since the ranking below the top three is irrelevant, it would be practical to ask players to rank only their top three.

Preferences Over Gender Composition of Tournament Teams

There is a growing literature about gender-differences in competitive environments (for a review, see Niederle and Vesterlund, 2011). Healy and Pate (2011) find that women prefer to compete in teams of two, while men prefer to compete as individuals in an adding task. Suppose an experimenter is interested in how men and women differ in their preferences over the gender compositions of teams they will compete in and compete against in tournaments involving teams of three. There are ten possible partitions of six players (three males and three females) into teams of three. The experimenter can use *RDP* to elicit each player's preferences over the ten possible match-ups and measure the differences in the types of teams men and women prefer to be on and the teams they prefer to compete against.

If it is of particular interest how the all-male against all-female team ranks, this can be included in all choice sets (as discussed in section B). To check for the demand effect caused by this asymmetry, it is advisable to check against a separate data-set which does not up-weight the same-gender matches.

B *Conclusion*

In this paper, I have provided several results useful to researchers designing experiments to elicit preferences over public outcomes. It is possible to directly elicit preferences using any strategy-proof social choice function (which I refer to as direct elicitation methods) as the outcome function. Under direct elicitation, the incentives for truthful reporting are transparent to subjects and easy to analyze for the designer. At times, it is desirable to elicit preferences within the context of a particular game. In those cases, the methods discussed here may not be appropriate.

VI Appendix

A Rank-Weighted Elicitation

RDP provides the strongest possible incentives across a player's full ranking of outcomes. In specific experimental contexts, the entire list may not be of interest. For instance, the experimenter may be most interested in a subject's favorite outcomes. As discussed above, *Random Dictator* provides strong incentives for players to truthfully report their favorite outcome but sacrifices incentives to truthfully report outcomes below the top. It is also possible to provide incentives that are "intermediate" to these two extremes by using choice sets that are larger than pairs (*RDP*) and smaller than the entire set V (*RD).*

Random Dictator over l-Restricted Sets *[RDRS-l]: A subset of V with cardinality l and a single player i are chosen randomly. The implemented outcome is the one that i ranks higher from the chosen set.*

This is a direct elicitation method since it is a mixture of pairwise-responsive and non-perverse unilateral rules. It nests both *RDP* and *RD.* Using a large l strengthens incentives at the top of player's rank-orders while using small l spreads the incentives. Where d_l is *RDRS-l,* and p is the location of y in P_k:

$$\epsilon_k^y(d_l, \mathbf{P}) = \frac{1}{n}\frac{(l)!\,(m-l)!}{m!}\binom{m-p}{l-2}$$

An example of $m = 6$ and $n = 2$ is provided below. Notice that much stronger incentives can be provided when outcomes only near the top of each ranking are required. This is a result of the combinatorial nature of the problem. To get accurate information about an entire rank-order, all possible $\frac{m(m-1)}{2}$ pairs must be potentially sampled. However, to get accurate information only about the top of a player's list, only the entire set V needs to be sampled. Notice that in the table below with $m = 6$, $\frac{m(m-1)}{2} = 15$. The incentives provided at the top of the rank list for *RD* is exactly 15 times larger than the incentives at any position in *RDP.* This non-linear trade-off represents an important consideration in experimental design.

	$l = 2$	3	4	5	6
$p = 2$ (*RDP*)	0.022	0.022	0.022	0.022	0.022
$p = 3$	0.067	0.050	0.033	0.017	0
$p = 4$	0.133	0.067	0.022	0	0
$p = 5$	0.222	0.056	0	0	0
$p = 6$ (*RD*)	0.333	0	0	0	0

Table 4: ϵ_k^y for *RDRS-l* with $n = 3$, $m = 6$

A practical extension of this method, especially where m is large, is to ask players to rank only their top $(m-1)-l$ outcomes. Since in any random choice set, *RDRS-l* will always include at least one item in a player's top $(m-1)-l$, the way outcomes ranked below this are irrelevant.

B Outcome-Weighted Elicitation

In some contexts, the experimenter may be interested in where a particular outcome, or a subset of outcomes, falls in a player's ranking. Here, the experimenter would like to provide stronger incentives to correctly rank these outcomes. This is possible using *RD*P but weighting the random selection of choice-sets so that the outcomes of interest are included more often. For example, by always including particular outcome y in the choice-pair, $\epsilon_k^y(d, \mathbf{P}) = \frac{1}{n}\frac{1}{m-1}$. Thus, the incentives for correctly ranking y are $\frac{m}{2}$ times larger than *RDP*; however, as a trade-off, players are indifferent in how they position two adjacent outcomes that do not include y.

More complicated hybrids of outcome-weighted and rank-weighted methods are also possible. A practical concern for using outcome-weighted elicitation is that breaking the elicitation method's symmetry might create a demand effect, skewing how players rank up or down-weighted outcomes. However, t is possible to use standard *RDP* as a baseline to check for these effects.

References

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E.E., 2008. Eliciting risk and time preferences. Econometrica 76, 583–618.

Andreoni, J., Miller, J., 2002. Giving according to garp: An experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

Andreoni, J., Sprenger, C., 2012a. Estimating time preferences from convex budgets. The American Economic Review 102, 3333–3356.

Andreoni, J., Sprenger, C., 2012b. Risk preferences are not time preferences. The American Economic Review 102, 3357–3376.

Bergstrom, T., Garratt, R., Leo, G., 2019. Let me, or let george? motives of competing altruists. Games and Economic Behavior 118, 269–283.

Chen, Y., Li, S.X., 2009. Group identity and social preferences. The American Economic Review 99, 431–457.

Dal Bó, E., Dal Bó, P., Eyster, E., 2015. The Demand for Bad Policy When Voters Underappreciate Equilibrium Effects. Technical Report. Working Paper.

Engelmann, D., Strobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. The American Economic Review 94, 857–869.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. The American Economic Review 97, 1858–1876.

Gibbard, A., 1973. Manipulation of voting schemes: a general result. Econometrica: journal of the Econometric Society , 587–601.

Gibbard, A., 1977. Manipulation of schemes that mix voting with chance. Econometrica 45, 665–681. URL: http://www.jstor.org/stable/1911681.

Healy, A., Pate, J., 2011. Can teams help to close the gender competition gap? The Economic Journal 121, 1192–1204.

Holt, C.A., Laury, S.K., et al., 2002. Risk aversion and incentive effects. American economic review 92, 1644–1655.

Kagel, J.H., Roth, A.E., Hey, J.D., 1995. The handbook of experimental economics. Princeton university press Princeton.

Niederle, M., Vesterlund, L., 2011. Gender and competition. Annu. Rev. Econ. 3, 601–630.