INCENTIVES IN EXPERIMENTS WITH OBJECTIVE LOTTERIES[†]

YARON AZRIELI, CHRISTOPHER P. CHAMBERS, AND PAUL J. HEALY[*]

ABSTRACT. When subjects in an experiment are given multiple decisions, their choices in one decision may be distorted by the choices made in others. An experiment's payment mechanism is incentive compatible if no such distortions occur. Azrieli et al. (2014) provide two characterizations of incentive compatible mechanisms in a general decision-theoretic framework when subjects view random payments as Savage-style acts. Here, we consider the case where subjects view gambles as objective lotteries. Assuming that subjects' preferences over lotteries respect first order stochastic dominance, and that all such preferences are admissible, we completely characterize the set of incentive compatible payment mechanisms. In particular, paying for one randomly-chosen problem—the *Random Problem Selection* (RPS) mechanism—is incentive compatible. We show that the set of incentive compatible mechanisms is strictly larger than in the acts case, and we discuss these new incentive compatible mechanisms in detail.

Keywords: Experimental design; decision theory, mechanism design.

JEL Classification: D81; D84.

May 23, 2014

I Introduction

The power of the experimental method lies in the researcher's ability to control the environment. In economics, this control enables us to observe choice in isolation, abstracting away the complexities of the field. When subjects are paid for multiple decisions in a single experiment, however, control may be lost because the payment from one decision may impact subjects' choices in another. In other words, there may be complementarities that distort incentives. We say that such an experiment is not Incentive Compatible (IC). Without incentive compatibility, researchers may be led to incorrect conclusions about subjects' preferences. The design of the payment mechanism – how subjects' choices in the experiment are mapped into the rewards they obtain – is therefore a crucial part of the experimental design.

The incentives problem that may arise by paying for multiple decisions was famously illustrated by Savage (1954) in his "hot man" example. We assume that a hot man has well-defined rational preferences over the objects {swim, shower, beer}. In eliciting these preferences, one must offer multiple choices. For example, to understand the ranking between swim and shower, one must offer a choice from the set of bundles {swim, shower}. Now, consider what happens when trying to also understand the preference between shower and beer. One must offer a choice from the set {shower, beer}. Overall, the two choice problems generate an actual choice from the set of bundles {(swim, shower), (swim, beer), (shower, shower), (shower, beer)}. If the "true" preference over single items is swim > beer > shower, and if a swim together with a shower is preferred to a shower together with a beer, then the individual will choose swim from the first set and shower from the second, so that it appears *as if* the ranking were swim > shower > beer.

A proposed solution to this problem, due to Allais (1953), is to pay for one randomly-selected decision—what we call the Random Problem Selection (RPS) mechanism.[1] Although this apparently solves the complementarities problem, there are examples of preferences for which the RPS mechanism is not IC (Holt, 1986, e.g.). Thus, exact conditions under which this mechanism is IC are not well understood. Neither is it known whether other mechanisms can be used to guarantee truthful revelation of choices in experiments with multiple decisions.

---

[1]This is often called the Random Lottery Incentive Mechanism (RLIM). We choose RPS to stay consistent with our terminology in the companion paper Azrieli et al. (2014), where the RLIM terminology is not appropriate.

In a companion paper (Azrieli et al., 2014), we filled this gap by studying experiment incentives in a general decision-theoretic framework in which gambles are modeled as Savage-style acts. Assuming event-wise monotonicity (dominated gambles are never chosen) and nothing else, we show that the RPS mechanism is, in practice, the unique incentive compatible mechanism. There can be certain, contrived settings in which other non-RPS mechanisms can be incentive compatible, but these settings are rarely observed in practice.

In this paper we extend the (Azrieli et al., 2014) result to the case of objective lotteries. Specifically, we develop a choice-based theoretical framework that allows us to carefully study incentive compatibility in the classical environment of objective lotteries. We model an experiment as a list of decision problems, i.e., a list of sets of choice objects from which a subject should choose (as in the above "hot man" example). The researcher may use an objective randomization device to determine the payoff that the subject obtains as a function of his choices.[2] Such a payment mechanism is IC if the lottery obtained by truthfully choosing the favorite item at each decision problem is preferred by the subject to any other lottery he can get by misrepresenting his preferences.

Implicit in the above definition of incentive compatibility is that the subject enters the lab with a well-defined preference relation over the choice objects in the experiment, as well as with an extension of this preference to a preference relation over lotteries over these objects. The question whether a particular RPS mechanism (or any other mechanism) is IC or not is therefore a question about the relation between the underlying preference over choice objects and the extensions of this preference to lotteries which the experimenter deems admissible.

It is quite easy to see that if every extension to lotteries is admissible, then there is no way to incentivize multiple choices.[3] A natural restriction on extensions to lotteries is that they satisfy a monotonicity property with respect to First Order Stochastic Dominance (FOSD), relative to the underlying preference. Formally, an extension is monotonic if lottery f is preferred to lottery g whenever f dominates g in the sense of FOSD. We show in Theorem 1 that, as long as admissible extensions are monotonic, any RPS mechanism is IC. In other words, if a subject's preferences are such that he never prefers a dominated gamble, then any RPS mechanism provides him the right incentives to truthfully reveal his favorite element in each decision problem. It is also clear,

---

[2]Note that an RPS mechanism describes a particular way in which choices are mapped into lotteries over choice objects.

[3]See Azrieli et al. (2014) for the details.

though we do not formalize it, that monotonicity on a restricted domain is a necessary condition for incentive compatibility of an RPS mechanism.

Having established the importance of monotonicity, and given the normative appeal of this property, we then characterize the class of all IC mechanisms in a given experiment assuming that all monotonic extensions are admissible. That is, we take the view point of an experimenter who believes that his subjects' preferences respect FOSD, but is not willing to make any additional assumption on their preferences. The main result of this paper, Theorem 2, shows that, in a certain sense, any IC mechanism resembles an RPS mechanism, but that the class of IC mechanisms may extend beyond the class of RPS mechanisms in several ways.

To understand how incentive compatibility can extend beyond the RPS mechanism, consider the following example. Let $D_1 = \{x, y\}$, $D_2 = \{y, z\}$ and $D_3 = \{x, z\}$ be the three decision problems in some experiment. Now, for every (strict) preference over $\{x, y, z\}$, if the subject truthfully announces his choices, then his favorite alternative from the set $E = \{x, y, z\}$ will also be revealed. Below we will call sets with this property *surely identified* sets. We can include such sets in the support of the distribution that describes the mechanism, treating them as additional (hypothetical) decision problems. For instance, in the above example consider the distribution λ over subsets of $\{x, y, z\}$ given by $\lambda(D_1) = \lambda(D_2) = \lambda(D_3) = 0.3$ and $\lambda(E) = 0.1$. Then the induced mechanism pays with probability 0.7 the revealed most preferred alternative and with probability 0.3 the revealed second most preferred alternative.[4] Thus, the lottery induced by truth-telling strictly dominates any lottery obtained by reporting other messages, so the mechanism is incentive compatible.[5]

Still, we can generalize even further by allowing λ to put negative weight on some of the sets. For instance, set $\lambda(D_1) = \lambda(D_2) = \lambda(D_3) = 0.4$ and $\lambda(\{x, y, z\}) = -0.2$. This induces the mechanism that pays with probability 0.6 the revealed most preferred alternative and with probability 0.4 the revealed second most preferred alternative, so incentive compatibility is maintained. However, if we choose the weights to be $\lambda(D_1) = \lambda(D_2) = \lambda(D_3) = 0.6$ and $\lambda(\{x, y, z\}) = -0.8$, then the resulting mechanism will not be incentive compatible, since the revealed second-best alternative is chosen with higher

---

[4]Indeed, the most preferred element will be chosen at two of the actual decision problems, and the experimenter can then infer that it is optimal in E, so the total weight assigned to this element is 0.7. The second best will be the choice at one decision problem and be assigned a weight of 0.3. The least favorite gets a weight of 0.

[5]In the discussion here we assume that the subject choices are consistent with some strict ordering of the elements. In our formal treatment we also deal with the issue of 'non-rationalizable messages', i.e. choices that cannot be rationalized by any strict ordering.

probability than the revealed first-best. Thus, some restrictions must be placed on λ in order for incentive compatibility to hold in the resulting mechanism. Theorem 2 shows that, in any experiment, any IC mechanism can be represented by a particular λ as above, and precisely describes the restrictions on λ that guarantee incentive compatibility. We note that in most experiments there are not many overlaps between the various decision problems; whenever this is the case (for example when the decision problems are pairwise disjoint) it follows from our characterization that the only IC mechanisms are RPS.

It is illuminating to compare this characterization to the one obtained in the companion paper Azrieli et al. (2014). In that work we characterize incentive compatibility of experiments under monotonicity, but when mechanisms map choices to *acts* instead of objective lotteries. Monotonicity in that framework means that if the subject is made better off *conditional* on any given state occurring, then he is made better off overall. The acts framework allows for more general extensions of preferences: Subjects may have their own subjective beliefs about the likelihood of different outcomes of the randomization device, or they may even have preferences which are not probabilistically sophisticated (Machina and Schmeidler, 1992) (e.g., they may be uncertainty averse). The assumption of the current paper that subjects view payments as lotteries can be thought of as an additional restriction on the set of admissible extensions. Since the experimenter can use this additional knowledge about extensions to construct IC mechanisms, one would expect that the class of IC mechanisms will be larger in the case of lotteries. In Section V we show that this is indeed the case: If a mechanism is IC in the acts framework, and one puts some (full-support) distribution over the state space of the randomization device, then the resulting lottery mechanism is IC. However, there are IC mechanisms in the lotteries environment that cannot be generated by any IC acts mechanism; in fact, these are exactly the mechanisms that have negative weights on some sets in their representation.

Finally, in Section VI we consider the particular case of experiments in which the choice objects are themselves lotteries over money. In this set-up an RPS mechanism generates a compound lottery, where in the first 'upper' stage a decision problem is randomly chosen for payment, and in the second 'lower' stage a dollar amount is randomly chosen according to the lottery that the subject chose in the realized decision problem of the first stage. Examples in the literature (e.g., Holt (1986)) show that if the subject reduces compound lotteries according to the laws of probability and has Rank-Dependent Utility (RDU) preferences over lotteries over money, then the RPS may not be IC. Our

framework and results make it easy to see the source of the failure: Reduction of compound lotteries together with monotonicity imply the independence axiom. Since RDU preferences typically violate independence, if one assumes reduction then it must be the case that monotonicity does not hold. Our Theorem 1 cannot be applied then, and the RPS may not be IC. In fact, we show also that in examples of the above type if subjects reduce compound lotteries and if all RDU preferences are admissible then no IC mechanism exists.

II THE FRAMEWORK

There is a finite set X of *choice objects*. The decision maker (also called the *subject*) has a *strict* preference relation $\succeq$ over X, i.e., $\succeq$ is a complete, transitive and antisymmetric binary relation. For any $x \in X$, let $L(x,\succeq) = \{y \in X \: : \: x \succeq y\}$ and $U(x,\succeq) = \{y \in X \: : \: y \succeq x\}$ be the lower- and upper-contour sets of x according to $\succeq$, respectively. The $\succeq$-*dominant* element of any set $E \subseteq X$ is denoted by $\text{dom}_\succeq(E)$. That is, $\text{dom}_\succeq(E)$ is the (unique) element of X satisfying $\text{dom}_\succeq(E) \in E$ and $\text{dom}_\succeq(E) \succeq y$ for every $y \in E$.

The researcher has an exogenously-given list of k decision problems, denoted $D = (D_1,\ldots,D_k)$, where $D_i \subseteq X$ for each $i \in \{1,\ldots,k\}$. Let $\mathscr{D} = \{D_1,\ldots,D_k\}$ represent the set of decision problems. We assume throughout that each $D_i \in \mathscr{D}$ is non-trivial, meaning $|D_i| > 1$, and that the same decision problem does not appear more than once, meaning $D_i \neq D_j$ whenever $i \neq j$. These assumptions are made only to simplify notation and can easily be relaxed.

The subject is asked to choose an element from each D_i. The announced choice vector (or, the subject's message) is denoted by $m = (m_1,\ldots,m_k)$. The space of all possible messages is $M = \times_i D_i$. For each $i \in \{1,\ldots,k\}$, let $\mu_i(\succeq) = \text{dom}_\succeq(D_i)$ be the $\succeq$-dominant element of D_i, and denote $\mu(\succeq) = (\mu_1(\succeq),\ldots,\mu_k(\succeq))$. We refer to $\mu(\succeq)$ as the *truthful* message for $\succeq$.

We assume that an objective randomization device can be used to determine payoffs, so that payments are given by lotteries. Denote by $\Delta(X)$ the set of all probability distributions over X. If $f \in \Delta(X)$ then $f(x)$ is the probability with which $x \in X$ is selected according to f. A *(payment) mechanism* $\varphi : M \to \Delta(X)$ takes the announced choice $m \in M$ and awards the subject with the lottery $\varphi(m) \in \Delta(X)$. Thus, $\varphi(m)(x)$ denotes the probability with which x is awarded when the decision maker announces m.

We refer to the pair (D,φ) as an *experiment*; D completely specifies the choices the subject must face, and φ describes how they are paid for those choices. Since D determines the domain of a mechanism, there is little distinguishing an experiment (D,φ)

from its associated mechanism φ; when it causes no confusion, we refer to experiments and mechanisms interchangeably.

We assume that the subject's preferences $\succeq$ extend to the space of lotteries $\Delta(X)$. An extension of $\succeq$ to $\Delta(X)$ is denoted by $\succeq^*$, and we assume that any admissible extension is complete and transitive. The asymmetric part of $\succeq^*$ is denoted by $\succ^*$. An extension $\succeq^*$ is assumed to agree with $\succeq$ on the space of degenerate lotteries. We let $\mathscr{E}(\succeq)$ denote the set of admissible extensions of $\succeq$.

Definition 1 (Incentive Compatibility). A mechanism φ is incentive compatible with respect to $\mathscr{E}$ if, for every preference $\succeq$, every extension $\succeq^* \in \mathscr{E}(\succeq)$, and every $m \neq \mu(\succeq)$, we have that $\varphi(\mu(\succeq)) \succ^* \varphi(m)$.

In other words, incentive compatible experiments induce the subject to announce truthfully, treating each decision problem as though it were in isolation. Note that whether or not a mechanism (or experiment) is incentive compatible depends crucially on $\mathscr{E}$. When there is no confusion, we drop the reference to $\mathscr{E}$ and simply refer to φ as incentive compatible.

Without making further assumptions on the correspondence $\mathscr{E}$, there do not exist incentive compatible mechanisms when the number of decision problems is $k \geq 2$ (see the companion paper Azrieli et al. (2014)). For the most part of this paper we assume that extensions $\succeq^*$ respect first-order stochastic dominance with respect to the underlying preference $\succeq$.

Definition 2 (First-Order Stochastic Dominance). Fix $\succeq$. The lottery f *dominates* the lottery g with respect to $\succeq$ (denoted $f \sqsupseteq g$) if, for every $x \in X$,

$$\sum_{\{x' \in X : x' \succeq x\}} f(x') \geq \sum_{\{x' \in X : x' \succeq x\}} g(x').$$

If there is strict inequality for at least one x then we say f *strictly dominates* g with respect to $\succeq$ ($f \sqsupset g$).[6]

Definition 3 (Monotonic Extension). An extension $\succeq^*$ of $\succeq$ is monotonic if $f \sqsupseteq g$ implies $f \succeq^* g$ and $f \sqsupset g$ implies $f \succ^* g$. The collection of all monotonic extensions of $\succeq$ is denoted by $\mathscr{E}^{\mathrm{mon}}(\succeq)$.

The following simple lemmas will be useful for some of the following results. The proofs are omitted.

---

[6]Because it will always be obvious, we use a notation which suppresses the dependence of $\sqsupseteq$ and $\sqsupset$ on $\succeq$.

Lemma 1. Assume $\mathscr{E}(\succeq) \subseteq \mathscr{E}^{\mathrm{mon}}(\succeq)$ for every $\succeq$. If for every $\succeq$ and every $m \neq \mu(\succeq)$ we have that $\varphi(\mu(\succeq)) \sqsupset \varphi(m)$, then φ is incentive compatible with respect to $\mathscr{E}$.

Lemma 2. A mechanism φ is incentive compatible with respect to $\mathscr{E}^{\mathrm{mon}}$ if and only if, for every preference $\succeq$ and every $m \neq \mu(\succeq)$, $\varphi(\mu(\succeq)) \sqsupset \varphi(m)$.

Remark. The conclusion of Lemma 2 continues to hold even if the set of admissible extensions $\mathscr{E}(\succeq)$ does not contain all monotonic extensions, as long as this set is "sufficiently rich". For example, if any expected utility extension is admissible then the lemma is still true.

III THE RANDOM PROBLEM SELECTION MECHANISM

A common payment mechanism is one in which a single decision problem is randomly selected, and the subject is paid with his choice at that problem. We call such a mechanism a *Random Problem-Selection* (RPS) Mechanism. Formally,

Definition 4 (Random Problem-Selection Mechanism). A mechanism φ is a *random problem-selection mechanism* (RPS) if there exists a full-support probability distribution λ over $\mathscr{D}$ such that for every alternative $x \in X$,

$$(1) \qquad \varphi(m)(x) = \sum_{\{1 \leq i \leq k \ : \ m_i = \{x\}\}} \lambda(D_i).$$

Theorem 1. If $\mathscr{E}(\succeq) \subseteq \mathscr{E}^{\mathrm{mon}}(\succeq)$ for every $\succeq$ and φ is an RPS mechanism, then φ is incentive compatible with respect to $\mathscr{E}$.

Proof. Follows immediately from Lemma 1, and from the obvious fact that lying in an RPS mechanism induces a lottery that is strictly dominated by the lottery obtained by truth-telling. $\square$

IV CHARACTERIZATION OF INCENTIVE COMPATIBLE MECHANISMS

In this section we provide a complete characterization of incentive compatible mechanisms when all monotonic extensions are admissible (or when the set of admissible extensions is sufficiently rich, see the Remark after Lemma 2). Recall that, by Lemma 2, incentive compatibility in this set-up is equivalent to the property that the lottery obtained by truth-telling strictly dominates any lottery that can be obtained by lying.

The example in the introduction illustrates how incentive compatibility can extend beyond the RPS mechanism. We now introduce notation and definitions required to formally state and prove the characterization result.

Surely Identified Sets

Let $M_R = \{m \in M : (\exists \succeq)\ m = \mu(\succeq)\}$ be the set of *rationalizable* messages. $M_{NR} = M \setminus M_R$ is then defined as the set of non-rationalizable messages.

Definition 5. Fix any rationalizable message $m = (m_1, \ldots, m_k) \in M_R$. For every $x, y \in X$, say that x is *directly revealed preferred* to y under m if there is $1 \le i \le k$ such that $m_i = x$ and $y \in D_i$, or if $x = y$. Denote the transitive closure of this relation by $R(m)$, and say that x is *revealed preferred* to y under choices m if $xR(m)y$.

The relation $R(m)$ is reflexive, transitive and antisymmetric since $m \in M_R$, but it need not be complete. Denote by $L(x, m) = \{y \in X : xR(m)y\}$ and $U(x, m) = \{y \in X : yR(m)x\}$ the sets of elements that are revealed to be worse than x and better than x under choices m, respectively. Clearly, $L(x, m) \subseteq L(x, \succeq)$ and $U(x, m) \subseteq U(x, \succeq)$ when $m = \mu(\succeq)$, with strict inclusions for some x when $R(m)$ is not a complete relation.

Let $\mathrm{dom}_m(E)$ be the $R(m)$-*dominant* element of E, if one exists. Notice that if $m = \mu(\succeq)$, then either $\mathrm{dom}_m(E)$ does not exist or else $\mathrm{dom}_m(E) = \mathrm{dom}_{\succeq}(E)$.

Definition 6 (Surely Identified Sets). A non-empty set $E \subseteq X$ is *surely identified* *(SI)* if, for every $m \in M_R$, $\mathrm{dom}_m(E)$ exists. In other words, E is SI if, for any order $\succeq$, the message $m = \mu(\succeq)$ identifies the most-preferred element of E, so that $\mathrm{dom}_m(E) = \mathrm{dom}_{\succeq}(E)$.

Let $SI(\mathscr{D})$ be the collection of sets surely identified from the given set of decision problems $\mathscr{D}$.[7] Obviously, any D_i is in $SI(\mathscr{D})$, but there may be other sets in $SI(\mathscr{D})$ (as in the example in the introduction). A characterization of surely identified is given by the following lemma, whose proof can be found in Azrieli et al. (2014).

Lemma 3. $E \in SI(\mathscr{D})$ if and only if E is either a singleton, or for every pair $\{x, y\} \subseteq E$, there exists $D \in \mathscr{D}$ for which $\{x, y\} \subseteq D \subseteq E$.

Weighted Set-Selection Mechanisms

We can now define a generalization of RPS mechanisms called weighted set-selection (WSS) mechanisms.

Definition 7 (Weighted Set-Selection Mechanisms). A mechanism $\varphi : M \to \Delta(X)$ is a *weighted set-selection mechanism* (WSS) if there exists some $\lambda : SI(\mathscr{D}) \to \mathbb{R}$ such that

---

[7]Recall that $\mathscr{D} = \{D_1, \ldots, D_k\}$ is the collection of decision problems, while $D = (D_1, \ldots, D_k)$ is the ordered list of decision problems.

for every rationalizable $m \in M_R$ and every alternative $x \in X$,

$$\varphi(m)(x) = \sum_{\{E \,:\, \text{dom}_m(E)=\{x\}\}} \lambda(E).$$

The requirement that $\varphi(m)$ be a well-defined lottery places some restrictions on the weighting function λ. For example, it cannot put negative weight on any singleton set: If $\lambda(\{x\}) < 0$ and there is no other $E \in SI(\mathscr{D})$ for which $\text{dom}_m(E) = \{x\}$, then $\varphi(m)(x) = \lambda(\{x\}) < 0$, which is forbidden. Furthermore, we have that

$$\sum_x \varphi(m)(x) = \sum_x \left[\sum_{\{E \,:\, \text{dom}_m(E)=\{x\} \}} \lambda(E) \right]$$
$$= \sum_E \lambda(E),$$

so it must be that $\sum_E \lambda(E) = 1$. These observations prove the following lemma.

Lemma 4. A weighted set-selection mechanism must be associated with a weighting function λ that satisfies

(1) $\sum_E \lambda(E) = 1$, and
(2) $\lambda(\{x\}) \geq 0$ for every $x \in X$.

Remark. A weighted set-selection mechanism uniquely determines the vector λ that represents it. That is, if λ and λ' are two different weighting vectors then the corresponding mechanisms φ and φ' differ on M_R. This can be seen by considering a minimal (with respect to inclusion) SI set E for which $\lambda(E) \neq \lambda(E')$, and an order $\succeq$ which ranks all elements of E below every other element of X. The top element of E according to $\succeq$ is chosen with different probabilities under φ and φ' when the choices are $\mu(\succeq)$.

Positivity on Switches

Recall the example in the introduction which shows that not any weighting function λ induces an incentive compatible mechanism. We now formalize a condition on λ which is precisely what's needed to guarantee incentive compatibility of the associated mechanism. We start with the following definition.

Definition 8 (Switch Test Set). Let $x, y \in X$ and $A \subseteq X \setminus \{x, y\}$. A surely identified set $E \in SI(\mathscr{D})$ is a *switch test set* for x and y against A if $\{x, y\} \subseteq E \subseteq A \cup \{x, y\}$. Let $T(x, y, A)$ denote the collection of switch test sets for x and y against A.

To see why switch test sets are important, as well as the reason for the name, consider an order $\succeq$ with $L(y, \succeq) = A \cup \{y\}$ and $L(x, \succeq) = A \cup \{x, y\}$ (i.e., y is ranked immediately

above A and x immediately above y). Let $\succeq^{xy}$ be the order obtained from $\succeq$ by switching the ranking of x and y while keeping all other elements in place. Then E is a switch test set for x and y against A if and only if the revealed most preferred element of E under $\succeq$ is x and under $\succeq^{xy}$ is y. Thus, it is exactly on switch test sets where the switch in the order of x and y will be revealed. For incentive compatibility to hold, the total weight assigned by λ to the collection of switch test sets $T(x, y, A)$ should be strictly positive. This would guarantee that, when comparing the lotteries $\mu(\succeq)$ and $\mu(\succeq^{xy})$, the probability of x is higher in the former while that of y is higher in the latter.

Definition 9 (**Switch Positivity**)**.** A weighted set-selection mechanism φ (with associated weighting vector λ) satisfies *switch positivity* if, for every $x, y \in X$ and $A \subseteq X \setminus \{x, y\}$ such that $T(x, y, A) \neq \emptyset$, it holds that

$$\sum_{\{E \in T(x, y, A)\}} \lambda(E) > 0.$$

Remark. If the collection $T(x, y, A)$ is not empty, then it contains at least one of the decision problems in $\mathscr{D}$. Indeed, $E \in T(x, y, A)$ means that $\{x, y\} \subseteq E \subseteq A \cup \{x, y\}$. Since E is surely identified, Lemma 3 implies that there is $D \in \mathscr{D}$ such that $\{x, y\} \subseteq D \subseteq E$. It follows that $D \in T(x, y, A)$ as well.

Dealing With Non-Rationalizable Messages

To achieve incentive compatibility with respect to $\hat{\mathscr{E}}^{\mathrm{mon}}$ of a WSS mechanism, we must ensure that non-rationalizable messages are always dominated by the truthful message. Let us go back to the example of $X = \{x, y, z\}$ and $D = (\{x, y\}, \{y, z\}, \{z, x\})$. Figure I illustrates the restrictions on lotteries obtained by non-rationalizable messages in this example. Consider the preference $\succeq$ where $x \succeq y \succeq z$ and the point $\varphi(\mu(\succeq))$, denoted simply as xyz in the figure. If a lottery f is to be dominated by $\varphi(\mu(\succeq))$ then it must be that $\varphi(\mu(\succeq))(x) \geq f(x)$ and $\varphi(\mu(\succeq))(z) \leq f(z)$. These inequalities are represented by the two dashed lines in the figure, with the light gray area showing all lotteries dominated by $\varphi(\mu(\succeq))$ for preferences $\succeq$. Non-rationalizable messages m should map to a lottery $\varphi(m)$ that is dominated by $\varphi(\mu(\succeq))$ for *every* preference $\succeq$. In the figure, this is the dark gray region labeled Φ_{NR}. The six vertices must be excluded from Φ_{NR} because incentive compatibility with respect to $\hat{\mathscr{E}}^{\mathrm{mon}}$ requires that all non-rationalizable messages be strictly dominated; if they map to the same lottery as the truthful message, then only weak dominance is achieved.
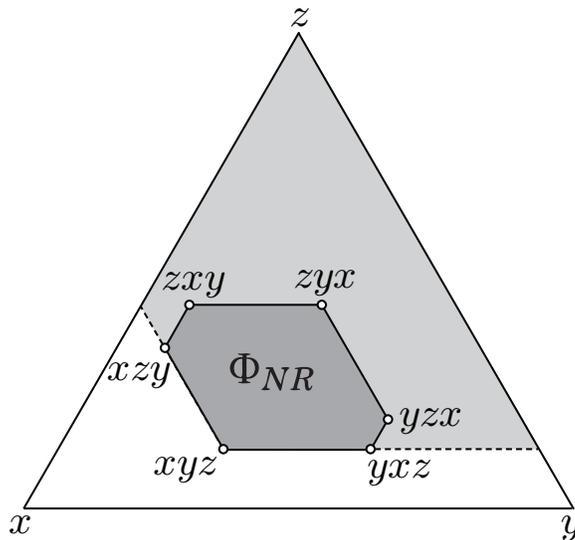
FIGURE I. In $\Delta(X)$, the lotteries stochastically dominated by the point $\varphi(\mu(\succeq))$ when $x > y > z$ (labeled xyz) are shown in light gray. The lotteries that are dominated by the truthful announcement for every $\succeq$ are shown in dark gray (Φ_{NR}).

Formally, let $\varphi(M_R)$ be the set of lotteries that can be obtained by announcing any rationalizable message and $co(\varphi(M_R))$ be the convex hull of that set. We denote $\Phi_{NR} = co(\varphi(M_R)) \setminus \varphi(M_R)$.

A cursory interpretation of Figure I suggests (incorrectly) that incentive compatibility with respect to $\hat{\mathscr{E}}^{mon}$ is characterized by Φ_{NR} having a 'hexagonal' shape, with sides parallel to the sides of $\Delta(X)$ (generalized appropriately for higher dimensions). Unfortunately, this is only true when rationalizable messages always fully reveal the preference relation (any pair of elements is a decision problem). If there are some preferences $\succeq$ and $\succeq'$ where $\mu(\succeq) = \mu(\succeq')$, then the hexagonal requirement no longer applies since the two preferences must map to the same lottery. The resulting shape changes.[8] This complication makes a purely geometric interpretation of incentive compatibility much less tractable; instead we present our characterization in terms of the weighting vector λ.

---

[8] For example, in Figure I, if $D = (\{x,y\}, \{y,z\})$ then $\mu(yxz) = \mu(yzx)$ and $\mu(xzy) = \mu(zxy)$, so $\phi(M_R)$ has only four distinct points. In that case, Φ_{NR} is rectangular. If $D = (\{x,y\})$ then Φ_{NR} is a horizontal line segment. If $D = (\{x,y,z\})$ then Φ_{NR} is a triangle.

The Characterization Theorem

Theorem 2. A mechanism $\varphi : M \to \Delta(X)$ is incentive compatible with respect to $\hat{\mathscr{E}}^{\mathrm{mon}}$ if and only if it is a weighted set-selection mechanism such that

(1) φ satisfies switch positivity; and

(2) if $m \in M_{NR}$ then $\varphi(m) \in \Phi_{NR}$.

The proof of this theorem—provided in the appendix—proceeds in three steps: First, we characterize a set of restrictions on the lotteries $\varphi(m)$ that are equivalent to incentive compatibility. Second, we show how an incentive compatible φ can be represented via a supermodular capacity, and how the restrictions on lotteries imposed by incentive compatibility translate into certain restrictions on that capacity. Third, we show that the capacity can be translated into a weighting vector λ—so that φ is in fact a weighted set-selection mechanism—and how the restrictions on the capacity imply that φ must satisfy switch positivity. Finally, we 'close the loop' by proving that any weighted set-selection mechanism satisfying these two conditions is in fact incentive compatible.

V Lotteries Versus Acts: A Comparison of Characterizations

In this section we compare IC mechanisms in the objective lotteries framework of the current paper to IC mechanisms in the more general acts set-up of the companion paper Azrieli et al. (2014). As explained in the introduction, we would like to show that in a certain sense there are more IC mechanisms when subjects view the randomization device as generating objective lotteries.

To formalize this idea, we must be able to compare directly a mechanism ϕ in the acts framework to a mechanism φ in the lotteries framework.[9] We say that (Ω, μ, ϕ) *generates* φ if, for each $m \in M$ and $x \in X$,

$$\varphi(m)(x) = \mu\left(\{\omega \in \Omega : \phi(m)(\omega) = x\}\right).$$

If the above equality holds for every rationalizable message $m \in M_R$ then we say that (Ω, μ, ϕ) generates φ on rationalizable messages.

Proposition 1. If ϕ is an incentive compatible act-mechanism (defined on some state space Ω), and μ is a full-support probability distribution over Ω, then the lotteries-mechanism φ generated by (Ω, μ, ϕ) is incentive compatible.

---

[9]In the acts framework (ignoring bundle payments; see Azrieli et al., 2014), a mechanism is a function $\phi : M \to X^{\Omega}$, where Ω is a finite state space and X^{Ω} is the set of all acts, i.e. mappings from Ω to X.

The proof of this proposition follows from the discussion in the introduction and is therefore omitted. The following example shows that there can be many act-mechanisms that generate a given lottery-mechanism, and that incentive compatibility is not necessarily preserved across frameworks.

Example 1. Let $X = \{x, y, z\}$. Suppose $k = 3$ with $D_1 = \{x, y\}$, $D_2 = \{y, z\}$, and $D_3 = \{z, x\}$.

In the lotteries framework, consider an RPS mechanism φ with $\lambda(D_i) = 1/3$ for each i. The subject receives their revealed-most-preferred element of X with probability 2/3 and their revealed-second-most-preferred element with probability 1/3. A non-rationalizable message results in the uniform lottery over X.

This mechanism can be generated by an RPS mechanism ϕ in the acts framework, where $\Omega = \{\omega_1, \omega_2, \omega_3\}$—each corresponding to a decision problem—and a distribution μ with $\mu(\omega_i) = 1/3$ for each i. Here, both mechanisms are incentive compatible in their respective frameworks.

But φ can also be generated by the following non-incentive-compatible act-mechanism ϕ and distribution μ: Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mu(\omega_i) = 1/3$ for each i. For rationalizable message m, set $\phi(m)(\omega_1) = \phi(m)(\omega_2) = \mathrm{dom}_m(X)$ and $\phi(m)(\omega_3)$ equal to the revealed-second-most-preferred element of X. For non-rationalizable message m set $\phi(m)(\omega_1) = x, \phi(m)(\omega_2) = y$, and $\phi(m)(\omega_3) = z$. This mechanism is not incentive compatible in the acts framework because beliefs are subjective: A subject who believes ω_3 will occur with high enough probability will prefer to reveal their true favorite element as if it were their second-most-preferred.

The following example demonstrates that there are incentive compatible lotteries-mechanisms that cannot be generated by incentive compatible acts-mechanisms (even when restricted to rationalizable messages).

Example 2. Let $X = \{x, y, z\}$. Suppose $k = 4$ with $D_1 = \{x, y\}$, $D_2 = \{y, z\}$, $D_3 = \{z, x\}$, and $D_4 = \{x, y, z\}$.

In the lotteries framework, consider a mechanism with $\lambda(D_i) = 0.4$ for each $i \in \{1, 2, 3\}$ and $\lambda(D_4) = -0.2$. This mechanism pays the revealed-most-preferred element of X with probability 0.6 and the revealed-second-most-preferred element with probability 0.4. Also, set $\varphi(m)$ to be the uniform distribution over X whenever m is non-rationalizable. By Theorem 2 this mechanism is incentive compatible.

However, this mechanism cannot be generated by any incentive compatible mechanism in the acts framework. To prove this, suppose that (Ω, μ, ϕ) generates φ, where ϕ incentive compatible. By Theorem 1 in Azrieli et al. (2014), each $\omega \in \Omega$ corresponds to

some decision problem (or to a singleton) and pays the selected item from that problem. Consider first $\succeq$ with $z \succ x \succ y$. Since $\varphi(\mu(\succeq))(x) = 0.4$, the set of ω's corresponding to D_1 or to the singleton $\{x\}$ must have μ-probability of 0.4. But, by a symmetric argument, the same is true for D_2 and $\{y\}$ and to D_3 and $\{z\}$. But then $\sum_\omega \mu(\omega) \geq 1.2$, a contradiction.

The difficulty in generating φ from an incentive compatible ϕ in Example 2 comes from λ assigning negative weights to certain SI sets. In fact, this exactly characterizes the cases where φ cannot be generated by an incentive compatible ϕ.

Proposition 2. Assume that φ is an incentive compatible lotteries-mechanism.

(1) If the associated weighting vector λ of φ is non-negative, then there exists an incentive compatible acts-mechanism ϕ (on some Ω) and a probability μ on Ω such that (Ω, μ, ϕ) generates φ on rationalizable messages.

(2) If the associated weighting vector λ of φ contains negative elements, then φ cannot be generated by any incentive compatible acts-mechanism ϕ (even when restricted to rationalizable messages).

For (1), the construction of the first mechanism in Example 1 can be generalized to any lotteries mechanism with non-negative λ to get a generating incentive compatible acts-mechanism. The proof of (2) is similar to the proof that φ cannot be generated by an incentive compatible acts-mechanism in Example 2. The details are omitted.

VI Choice from Lotteries, Reduction, and RDU Preferences

Many experimental tests of decision-theoretic models ask subjects to make choices from menus of lotteries whose outcomes are dollar payments. In this case, payments in the RPS mechanism represent two-stage lotteries, where the 'upper' stage refers to the random draw of a decision problem, and the 'lower' stage refers to the draw of a dollar amount according to the chosen lottery from that problem.

It is possible that the subject "reduces" compound lotteries into one-stage lotteries according to the laws of probability, and thus that his preferences over the space of single stage lotteries (over money) completely determine his preferences over compound lotteries. The following example—adapted from Holt (1986) and Cox et al. (2014)—shows that incentive compatibility of the RPS mechanism can fail when non-expected utility models (in the lower stage) are combined with the reduction of compound lotteries.

Definition 10 (Rank-Dependent Utility). A subject has rank-dependent utility (RDU) preferences if a simple lottery $f = (x_1, p_1; x_2, p_2; \ldots; x_n, p_n) \in \Delta(\mathbb{R})$ (with $x_1 < x_2 < \cdots < x_n$)

is evaluated according to the functional

$$U_q(f) = \sum_{s=1}^{n} u(x_s) \left[q(\sum_{r=1}^{s} p_r) - q(\sum_{r=1}^{s-1} p_r) \right],$$

where $u : \mathbb{R} \to \mathbb{R}$ is increasing, $q : [0,1] \to [0,1]$ is increasing, strictly concave over $(0, 0.5)$ and strictly convex over $(0.5, 1)$, $q(0) = 0$, and $q(1) = 1$.

Example 3. Let $l = (\$0, 1/2; \$3, 1/2)$ be an equiprobable lottery between \$0 and \$3, and consider $D_1 = \{l, \$1\}$ and $D_2 = \{l, \$2\}$. If a subject has rank-dependent utility with $u(x) = x^{3/4}$, $q(1/4) = 1/3$, $q(1/2) = 1/2$, and $q(3/4) = 2/3$, then $\$2 >^* l >^* \$1$. Thus, the truthful announcement is $m^* = (l, \$2)$. Now consider an RPS mechanism for lotteries that puts equal probability on each D_i being chosen. Announcing m^* gives the lottery $\varphi(m^*) = (\$0, 1/4; \$2, 1/2; \$3, 1/4)$, assuming compound lotteries are reduced to single-stage lotteries. Announcing $m' = (1, 2)$ gives the lottery $\varphi(m') = (\$1, 1/2; \$2, 1/2)$. Plugging in the values of u and q, we find that $U(\varphi(m')) > U(\varphi(m^*))$, so incentive compatibility is violated.

The reason the RPS mechanism fails in this example stems from the well-known fact that reduction of compound lotteries, when combined with our monotonicity assumption, implies the Von Neumann-Morgenstern independence axiom on the space of single stage lotteries.[10] Consequently, if a model of preferences violates independence (as the RDU model does), but the reduction of compound lotteries is assumed, then that model must violate monotonicity. Without monotonicity, Theorem 1 cannot be applied and there is no guarantee that the RPS mechanism is incentive compatible.

In fact, we now show that for the decision problems in Example 3 there is *no* incentive compatible payment mechanism if any RDU preference is admissible. We need to slightly modify our framework in order to accommodate this example. Let $\Delta(\mathbb{R})$ be the set of all simple (finite support) lotteries on $\mathbb{R}$. A degenerate lottery that pays \$x with probability 1 is denoted by δ_x. The subject faces the two decision problems from Example 3, $D_1 = \{l, \delta_1\}$ and $D_2 = \{l, \delta_2\}$, where l is the lottery that pays 0 or 3 with equal probabilities.

We assume that the subject has RDU preferences over $\Delta(\mathbb{R})$, represented by a functional U_q (for some u and q) as in Definition 10. Notice that this allows the subject to be indifferent between some of the lotteries. We therefore need to modify the definition of incentive compatibility to allow for weak preferences. We do that by requiring that,

---

[10] For a detailed discussion of this result and other related issues see Segal (1990). Segal's axiom of "compound independence" is essentially the same as our monotonicity assumption.

whenever a message is truthful, the output of the mechanism is (weakly) preferred to any other possible outcome, with strict preference whenever the other message is not truthful.

Since we assume that the subject reduces compound lotteries, his preferences over the compound lotteries induced by the mechanism are already captured by his functional U_q. Thus, a mechanism can be described by a function $\varphi : M \to \Delta(\mathbb{R})$. Notice that we allow the mechanism to pay with arbitrary (simple) lotteries, not necessarily lotteries over $\{0, 1, 2, 3\}$.

Proposition 3. In the set-up described above, there exists no incentive compatible mechanism for the decision problems $D_1 = \{l, \delta_1\}$ and $D_2 = \{l, \delta_2\}$.

The proof of this proposition appears in the appendix. While the proposition is stated for a particular pair of decision problems, we believe that this impossibility result is typical, and that for most experiments there will be no incentive compatible mechanism. Thus, either reduction of compound lotteries or the domain of admissible preferences must be relaxed in order to get positive results.

VII Discussion

This paper focuses on the elicitation of multiple choices when an experimenter can randomize. From a theoretical perspective, our work is probably closest to the classic work of Gibbard (1977), who characterizes strategy-proof mechanisms in a probabilistic context when only ordinal preferences can be elicited. He characterizes these mechanisms as a kind of random-dictatorship, whereby a "dictator" is an agent that solely determines the outcome. Our paper is comparable to the special case in which there is only one agent present. Gibbard does not, however, uncover the special structure of these dictatorial mechanisms in the form that we uncover, presumably because his interest was in understanding the implications of strategy-proofness *across* agents. Another important difference between the papers is that in Gibbard's framework agents report their entire ranking over alternatives, while we consider the more general case in which the favorite alternatives in several subsets are reported.

APPENDIX A: PROOF OF THEOREM 2

A.1: Step 1: Restrictions on φ

Recall that $L(x, \succeq)$ and $U(x, \succeq)$ are the lower- and upper-contour sets of x according to $\succeq$, respectively. Let $r(x, \succeq) = |U(x, \succeq)|$ be the rank of x in $\succeq$. Two elements x, y are *adjacent* in $\succeq$ if $|r(x, \succeq) - r(y, \succeq)| = 1$. A *switch* of x, y in an order $\succeq$ is the replacement of the order of x, y, where x, y are adjacent in $\succeq$. Denote the obtained order by $\succeq^{xy}$.[11]

Lemma 5. φ is incentive compatible with respect to $\hat{\mathscr{E}}^{\mathrm{mon}}$ if and only if it has the following two properties:

(1) For every $\succeq$ and every x, y with $r(x, \succeq) = r(y, \succeq) - 1$,
   (a) $\varphi(\mu(\succeq))(z) = \varphi(\mu(\succeq^{xy}))(z)$ for every $z \neq x, y$.
   (b) $\varphi(\mu(\succeq))(x) > \varphi(\mu(\succeq^{xy}))(x)$ and $\varphi(\mu(\succeq))(y) < \varphi(\mu(\succeq^{xy}))(y)$ whenever
       $\mu(\succeq) \neq \mu(\succeq^{xy})$.
(2) $\varphi(m) \in \Phi_{NR}$ whenever $m \in M_{NR}$.

Proof. Assume φ is incentive compatible and fix some $\succeq$ and some x, y with $r(x, \succeq) = r(y, \succeq) - 1$. If $\mu(\succeq) = \mu(\succeq^{xy})$ then the conditions are trivially true. Now assume that they differ. Let $z \neq x, y$ be some other element of X. Assume first that $r(z, \succeq) < r(x, \succeq)$, so z is ranked above x (and y) according to $\succeq$. Incentive compatibility implies that $\varphi(\mu(\succeq))(U(z, \succeq)) \geq \varphi(\mu(\succeq^{xy}))(U(z, \succeq))$ and that $\varphi(\mu(\succeq^{xy}))(U(z, \succeq^{xy})) \geq \varphi(\mu(\succeq))(U(z, \succeq^{xy}))$. But since $U(z, \succeq) = U(z, \succeq^{xy})$ we get that they are equal, that is $\varphi(\mu(\succeq))(U(z, \succeq)) = \varphi(\mu(\succeq^{xy}))(U(z, \succeq))$. The same argument applies to any z ranked above x (according to $\succeq$), which proves that $\varphi(\mu(\succeq))(z) = \varphi(\mu(\succeq^{xy}))(z)$ for any such z. A similar argument proves the assertion for elements z ranked below y. It follows that we must have $\varphi(\mu(\succeq))(x) > \varphi(\mu(\succeq^{xy}))(x)$ (and therefore $\varphi(\mu(\succeq))(y) < \varphi(\mu(\succeq^{xy}))(y)$) in order for $\varphi(\mu(\succeq)) \sqsupset \varphi(\mu(\succeq^{xy}))$ to hold. This concludes the proof of property 1.

As for property 2, whenever $m \in M_{NR}$ incentive compatibility implies that $\varphi(\mu(\succeq)) \sqsupset \varphi(m)$ for every $\succeq$. First, this implies that $\varphi(m) \neq \varphi(\mu(\succeq))$ for every $\succeq$. Second, assume that $\varphi(m)$ is not in Φ_{NR}. Then by the separation theorem there is a vector $u \in \mathbb{R}^X$ such that $\sum_x u(x)\varphi(m)(x) > \sum_x u(x)\varphi(\mu(\succeq))(x)$ for every $\succeq$. By boundedness of the set Φ_{NR}, we can choose u such that $u(x) \neq u(y)$ whenever $x \neq y$. Let $\succeq_u$ be the order over X defined by u (formally, $u(x) > u(y)$ implies $x > y$). Then an expected utility maximizer with utilities $u(\cdot)$ prefers to report the non-rationalizable choices m over his true choices $\mu(\succeq_u)$. But

---

[11]Formally, $x \succeq y \iff y \succeq^{xy} x$ and, for all other $w, z \in X$, $w \succeq z \iff w \succeq^{xy} z$. Note that $\succeq^{xy}$ is only well-defined if x and y are adjacent in $\succeq$.

this means $\varphi(\succeq_u)$ does not first-order stochastically dominate $\varphi(m)$ according to $\succeq_u$, a contradiction.

Conversely, assume that properties 1 and 2 are satisfied. Fix some $\succeq$ and consider some rationalizable deviation $m \in M_R$, $m \neq \mu(\succeq)$. Let $\succeq' \in \mu^{-1}(m)$. Consider a minimal sequence of switches that starts at $\succeq$ and ends at $\succeq'$. This means that x and y are switched somewhere along the path if and only if $x \succ y$ but $y \succ' x$. Then property 1 implies that after any switch along the way we get a lottery that is dominated (relative to $\succeq$) by the previous one. This shows that $\varphi(\mu(\succeq))$ dominates $\varphi(\mu(\succeq')) = \varphi(m)$. Finally, if $m \in M_{NR}$ then by property 2 and the above argument $\varphi(m)$ is a convex combination of lotteries that are dominated (relative to $\succeq$) by $\varphi(\mu(\succeq))$, so it is dominated as well. This proves the lemma. $\qquad\square$

A.2: Step 2: Capacity Representation

A capacity is a set function $v : 2^X \to \mathbb{R}$ such that $v(\emptyset) = 0$. A capacity v is *normalized* if $v(X) = 1$ and *monotone* if $A \subseteq B$ implies $v(A) \leq v(B)$.

Definition 11. A capacity v satisfies *switch positivity* if for every $x, y \in X$ and $A \subseteq X \setminus \{x, y\}$ the following holds: If $T(x, y, A) \neq \emptyset$ then $v(A \cup \{x, y\}) + v(A) > v(A \cup \{x\}) + v(A \cup \{y\})$; otherwise, $v(A \cup \{x, y\}) + v(A) = v(A \cup \{x\}) + v(A \cup \{y\})$.

If v satisfies switch positivity then it is *supermodular*, meaning $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$ for every $A, B \subseteq X$.

Lemma 6. If a mechanism φ is incentive compatible with respect to $\hat{\mathcal{E}}^{\mathrm{mon}}$ then there exists a normalized and monotone capacity v that satisfies switch positivity such that $\varphi(m)(x) = v(L(x, m)) - v(L(x, m) \setminus \{x\})$ for every $m \in M_R$ and every $x \in X$.

Proof. Given $A \subseteq X$, consider some order $\succeq$ which ranks A at the bottom. Define $v(A) = \varphi(\mu(\succeq))(A) := \sum_{x \in A} \varphi(\mu(\succeq))(x)$. Notice first that, under incentive compatibility, v is well-defined in the sense that it does not depend on the particular order $\succeq$ used. Indeed, this follows from property (1a) in Lemma 5. It is also clear that v is normalized and monotone.

We now claim that v satisfies switch positivity. To see this, take any x, y and $A \subseteq X \setminus \{x, y\}$. Consider some order $\succeq$ with $L(x, \succeq) = A \cup \{x\}$ and $L(y, \succeq) = A \cup \{x, y\}$. We have $v(A \cup \{x\}) = \varphi(\mu(\succeq))(A \cup \{x\})$ and $v(A \cup \{y\}) = \varphi(\mu(\succeq^{xy}))(A \cup \{y\})$, so

$$
\begin{aligned}
v(A \cup \{x\}) + v(A \cup \{y\}) &= \varphi(\mu(\succeq))(A \cup \{x\}) + \varphi(\mu(\succeq^{xy}))(A \cup \{y\}) \\
&= v(A) + \varphi(\mu(\succeq))(A) + \varphi(\mu(\succeq))(x) + \varphi(\mu(\succeq^{xy}))(y).
\end{aligned}
$$

Now, if $T(x, y, A) \neq \emptyset$ then $\mu(\succeq) \neq \mu(\succeq^{xy})$ (see Remark), so by property (1b) of Lemma 5 we have $\varphi(\mu(\succeq^{xy}))(y) < \varphi(\mu(\succeq))(y)$. Thus,

$$
\begin{aligned}
v(A \cup \{x\}) + v(A \cup \{y\}) &< v(A) + \varphi(\mu(\succeq))(A) + \varphi(\mu(\succeq))(x) + \varphi(\mu(\succeq))(y) \\
&= v(A) + v(A \cup \{x, y\}),
\end{aligned}
$$

as required. On the other hand, if $T(x, y, A) = \emptyset$ then $\mu(\succeq) = \mu(\succeq^{xy})$, so we get

$$
\begin{aligned}
v(A \cup \{x\}) + v(A \cup \{y\}) &= v(A) + \varphi(\mu(\succeq))(A) + \varphi(\mu(\succeq))(x) + \varphi(\mu(\succeq))(y) \\
&= v(A) + v(A \cup \{x, y\}).
\end{aligned}
$$

Finally, we need to show that $\varphi(m)(x) = v(L(x, m)) - v(L(x, m) \setminus \{x\})$ whenever $m \in M_R$. Fix $m \in M_R$ and some $x \in X$. We claim that there is $\succeq \in \mu^{-1}(m)$ such that $L(x, m) = L(x, \succeq)$. Since $L(x, m) \subseteq L(x, \succeq)$ for all $\succeq \in \mu^{-1}(m)$, it is sufficient to show that the reverse inclusion holds for some $\succeq \in \mu^{-1}(m)$. To see this, start with an arbitrary $\succeq \in \mu^{-1}(m)$ and consider the set $L(x, \succeq) \setminus L(x, m)$. If this set is empty we are done. Otherwise, take the highest ranked element (according to $\succeq$) in this set, say y. Then for any z ranked between y and x (including $z = x$) it cannot be that $zR(m)y$, so by a sequence of switches we can put y above x without changing the resulting choices. By repeating this procedure for every element in $L(x, \succeq) \setminus L(x, m)$ we get the desired order, say $\succeq'$. For this order we have

$$
\begin{aligned}
v(L(x, m)) - v(L(x, m) \setminus \{x\}) &= v(L(x, \succeq')) - v(L(x, \succeq') \setminus \{x\}) \\
&= \varphi(m)(L(x, \succeq')) - \varphi(m)(L(x, \succeq') \setminus \{x\}) \\
&= \varphi(m)(x),
\end{aligned}
$$

as needed. □

A.3: Step 3: Weighting Vector Representation

Lemma 7. Given φ, if there exists a normalized and monotone capacity v that satisfies switch positivity such that $\varphi(m)(x) = v(L(x, m)) - v(L(x, m) \setminus \{x\})$ for every $m \in M_R$ and every $x \in X$, then φ is a weighted set-selection mechanism that satisfies switch positivity.

Proof. Let v be a normalized and monotone capacity that satisfies switch positivity and represents φ as in the assertion of the lemma. As is well known (see Gilboa and Schmeidler, 1995, e.g.), any capacity can be uniquely represented as a linear combination of the 'unanimity capacities'. That is, there is a unique vector $\{\lambda(E)\}_{E \subseteq X, E \neq \emptyset}$ such that $v(A) = \sum_{E \subseteq A} \lambda(E)$ for every $A \subseteq X$.

We first show that if B is not SI then $\lambda(B) = 0$. If B is not SI then by Lemma 3 there are $x, y \in B$ such that for no $1 \le i \le k$ it holds that $\{x, y\} \subseteq D_i \subseteq B$. This in turn implies that $T(x, y, B \setminus \{x, y\})$ is empty (see Remark). Since v satisfies switch positivity we get that

$$\sum_{\{E \,:\, \{x,y\} \subseteq E \subseteq B\}} \lambda(E) = v(B) - v(B \setminus \{x\}) - v(B \setminus \{y\}) + v(B \setminus \{x, y\}) = 0.$$

But if $\{x, y\} \subseteq D_i \subseteq B$ for no i then for every set E in the sum on the left it is also true that $\{x, y\} \subseteq D_i \subseteq E$ for no i, which implies that every such E is not SI. By induction on the size of B we can therefore prove that $\lambda(B) = 0$.

We next check that the vector λ satisfies switch positivity. Take any $x, y \in X$ and $A \subseteq X \setminus \{x, y\}$. By the last paragraph,

$$\sum_{\{E \in T(x,y,A)\}} \lambda(E) = v(A \cup \{x, y\}) - v(A \cup \{x\}) - v(A \cup \{y\}) + v(A).$$

If $T(x, y, A) \neq \emptyset$ then since v satisfies switch positivity we have that $v(A \cup \{x, y\}) - v(A \cup \{x\}) - v(A \cup \{y\}) + v(A) > 0$. Thus, λ satisfies switch positivity.

The last thing to check is that λ in fact represents the weighted set-selection mechanism φ as in Definition 7. This follows from (for $m \in M_R$)

$$\varphi(m)(x) = v(L(x, m)) - v(L(x, m) \setminus \{x\}) = \sum_{\{E \,:\, x \in E \subseteq L(x,m)\}} \lambda(E) = \sum_{\{E \,:\, \text{dom}_m(E) = x\}} \lambda(E).$$

$\square$

A.4: Step 4: Weighted Set-Selection Mechanisms are Incentive Compatible

Lemma 8. If φ is a weighted set-selection mechanism that satisfies switch positivity and satisfies $\varphi(m) \in \Phi_{NR}$ whenever $m \in M_{NR}$ then φ is incentive compatible with respect to $\hat{\mathscr{E}}^{\text{mon}}$.

Proof. To check that φ is incentive compatible we use Lemma 5. It follows immediately from the definition of a weighted set-selection mechanism that in a switch of adjacent two elements x, y the probability of any other element z being selected is not affected. Further, we have just showed above that the probability of x goes strictly up after a switch that increases the rank of x and changes the truthful message. This proves property 1 of Lemma 5. Property 2 of Lemma 5 is satisfied by assumption. $\square$

Appendix B: Proof of Proposition 3

It will be convenient to think of any lottery f as a function $f : \mathbb{R} \to \mathbb{R}$, with $f(x)$ being the probability assigned to x by f. For instance, the lottery l in the proposition is identified with the function $l(0) = l(3) = 1/2$ and $l(x) = 0$ otherwise. Addition of lotteries and multiplication of lotteries by scalars are performed pointwise (yielding functions which are not necessarily lotteries). The expected utility of a lottery f for a subject with utility function u (defined on dollar amounts) is denoted by $u \cdot f$.

The proof is broken into two claims:

Claim 1. If φ is incentive compatible then there is $a \in (0, 1/2]$ such that $\varphi(\delta_1, \delta_2) - \varphi(l, \delta_2) = -a\delta_0 + 2a\delta_1 - a\delta_3$.

Proof. Consider some strictly increasing utility function u such that $1/2u(0) + 1/2u(3) = u(1)$. Then an expected utility maximizer (a special case of RDU preferences) with utility function u is indifferent between l and δ_1, so both announcements (δ_1, δ_2) and (l, δ_2) are truthful. By incentive compatibility this decision maker must also be indifferent between $\varphi(\delta_1, \delta_2)$ and $\varphi(l, \delta_2)$, that is $u \cdot \varphi(\delta_1, \delta_2) = u \cdot \varphi(l, \delta_2)$.

First, it cannot be true that there is $x \in \mathbb{R} \setminus \{0, 1, 3\}$ such that $\varphi(\delta_1, \delta_2)(x) \neq \varphi(l, \delta_2)(x)$. Indeed, let u be as above, and change the values of u in a small neighborhood of x (small enough that it does not affect any other point in the support of the two lotteries) to get a new increasing function $\tilde{u}$ that still satisfies $1/2\tilde{u}(0) + 1/2\tilde{u}(3) = \tilde{u}(1)$. Then it cannot be true that both equalities $u \cdot \varphi(\delta_1, \delta_2) = u \cdot \varphi(l, \delta_2)$ and $\tilde{u} \cdot \varphi(\delta_1, \delta_2) = \tilde{u} \cdot \varphi(l, \delta_2)$ hold simultaneously, a contradiction.

Thus, there are $a, b, c \in \mathbb{R}$ such that $\varphi(\delta_1, \delta_2) - \varphi(l, \delta_2) = a\delta_0 + b\delta_1 + c\delta_3$. Consider a utility function u with $u(0) = 1, u(1) = 2$ and $u(3) = 3$. By the same argument as above we get that $u \cdot \varphi(\delta_1, \delta_2) = u \cdot \varphi(l, \delta_2)$, which implies $a + 2b + 3c = 0$. By considering another function u with $u(0) = 1, u(1) = 3$ and $u(3) = 5$ gives $a + 3b + 5c = 0$. Solving these two equations gives $b = -2a$ and $c = a$.

To conclude, we showed that $\varphi(\delta_1, \delta_2) - \varphi(l, \delta_2) = -a\delta_0 + 2a\delta_1 - a\delta_3$ for some $a \in \mathbb{R}$. The fact that $a > 0$ follows by looking at an expected utility maximizer with the utility function $u(x) = x$ (who strictly prefers the lottery l over δ_1). The fact that $a \leq 1/2$ is obvious. $\square$

The proof of the above claim only considered EU preferences. The following claim makes use of RDU preferences which do not satisfy independence.

Claim 2. If φ is incentive compatible then $\varphi(\delta_1, \delta_2) = \delta_1$ and $\varphi(l, \delta_2) = l$.

Proof. Let u be the function $u(x) = 0$ for $x < 1$, $u(1) = 1$, and $u(x) = 3$ for $x > 1$. Consider an RDU decision maker with utility function u and a probability weighting function q satisfying $q(0) = 0, q(1/2) = 1/2, q(1) = 1$. Any such decision maker prefers the lottery l over δ_1 and prefers δ_2 over l. Thus, for incentive compatibility, any such decision maker must prefer $\varphi(l, \delta_2)$ over $\varphi(\delta_1, \delta_2)$, that is $U_q(\varphi(l, \delta_2)) > U_q(\varphi(\delta_1, \delta_2))$.

By the definition of u we have that

$$U_q(h) = 1[q(H(1)) - q(H(1-))] + 3[1 - q(H(1))]$$

for every lottery h with cdf H. Denote by F the cdf of the lottery $\varphi(l, \delta_2)$ and by G the cdf of $\varphi(\delta_1, \delta_2)$. Then incentive compatibility requires that

$$(2) \qquad U_q(\varphi(l, \delta_2)) - U_q(\varphi(\delta_1, \delta_2)) = 2[q(G(1)) - q(F(1))] - [q(F(1-)) - q(G(1-))] > 0$$

for every q as above.

By the previous claim, there is $0 < a \le 1/2$ such that $G(1) - F(1) = F(1-) - G(1-) = a$. We claim now that incentive compatibility can only hold if $a = 1/2$., i.e. if $G(1) = 1, F(1) = F(1-) = 1/2$ and $G(1-) = 0$. Indeed, in any other case it is possible to choose q (strictly increasing, $q(0) = 0$, $q(1) = 1$, $q(1/2) = 1/2$) such that $q(G(1)) - q(F(1))$ is much smaller than $q(F(1-)) - q(G(1-))$, which violates (2). The precise construction of such q depends on which of the intervals $[0, 1/2]$ or $[1/2, 1]$ each one of these four numbers belongs to, but it is straightforward in all cases. For instance, assume that $G(1-) \in [0, 1/2]$ and $F(1-), F(1), G(1)$ are all in $(1/2, 1]$. Then we can find q such that $q(F(1-)), q(F(1))$ and $q(G(1))$ are all close to 1, while $q(G(1-))$ is at most $1/2$. Other cases are treated similarly.

$\square$

To conclude the proof, notice that the same arguments can be applied to the choices in the second decision problem. That is, incentive compatibility requires that $\varphi(l, l) = l$ and $\varphi(l, \delta_2) = \delta_2$. Since we cannot have $\varphi(l, \delta_2) = l$ and $\varphi(l, \delta_2) = \delta_2$ at the same time, we conclude that an incentive compatible mechanism does not exist.

REFERENCES

Allais, M., October 1953. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. Econometrica 21 (4), 503–546.

Azrieli, Y., Chambers, C. P., Healy, P. J., 2014. Incentives in experiments: A theoretical framework, Ohio State University Working Paper.

Cox, J., Sadiraj, V., Schmidt, U., 2014. Paradoxes and mechanisms for choice under risk, Georgia State University Experimental Economics Center Working Paper 2014-01.

Gibbard, A., 1977. Manipulation of schemes that mix voting with chance. Econometrica, 665–681.

Gilboa, I., Schmeidler, D., February 1995. Canonical representations of set functions 20 (1), 197–212.

Holt, C. A., 1986. Preference reversals and the independence axiom. American Economic Review 76, 508–515.

Machina, M. J., Schmeidler, D., 1992. A more robust definition of subjective probability. Econometrica, 745–780.

Savage, L. J., 1954. The Foundations of Statistics. John Wiley & Sons, New York, NY.

Segal, U., 1990. Two-stage lotteries without the reduction axiom. Econometrica, 349–377.