Obvious belief elicitation

Elias Tsakas

Obvious belief elicitation

RM/19/001

GSBE

Obvious belief elicitation[*]

Elias Tsakas[†]

Maastricht University

January 2019

Abstract

An investigator is interested in arbitrarily approximating a subject's latent beliefs in obviously dominant strategies (Li, 2017). We prove that Karni's ascending mechanism (Karni, 2009) does not have an obviously dominant strategy. Thus, we introduce the novel descending Karni mechanism which always has obviously dominant strategies. Furthermore, under the assumption that the subject chooses an obviously dominant strategy, the true beliefs can be approximated with arbitrary precision with our mechanism. All our results hold for a very broad class of likelihood relations, going well beyond those that are represented by probabilistic beliefs.

Keywords: Belief elicitation, Karni mechanism, obvious dominance.
JEL codes: C90, C91, D81, D82, D83.

1. Introduction

Eliciting subjective beliefs is a widely-studied problem both in economics as well as in many other disciplines (Manski, 2004). Traditional examples of mechanisms that are commonly used for this purpose include the proper scoring rule method (Brier, 1950; Savage, 1971) and the lotteries method (Kadane and Winkler, 1988). More recently the family of Karni mechanisms has attracted attention in the literature (Karni, 2009). These mechanisms resemble the structure of well-known auctions.

A minimal property that we would typically like our mechanism to satisfy is that it is (weakly) dominant for experimental subjects to report truthfully. Unfortunately, despite the theoretical appeal, dominant strategies are sometimes difficult to identify in practice (especially if the mechanism is presented in its normal form), and a fortiori they are not always chosen by experimental subjects. For instance, the dominant strategy in second price auctions is selected less often than in strategically-equivalent ascending clock auctions (Kagel et al., 1987). In an attempt to explain this behavioral irregularity, Li (2017) proposed a refinement of the dominance criterion. In particular, a dominant strategy is said to be *obviously dominant* if at every information set that it reaches, the best outcome following a deviation at this information set is not superior to the worst outcome that can be achieved by sticking to said strategy from this information set onwards. He then provided experimental evidence showing that obviously dominant strategies are chosen more frequently than merely dominant ones (in mechanisms with the same normal form). Thus, our general research question in this note is *whether we can construct mechanisms that truthfully elicit subjective beliefs in obviously dominant strategies.*

[†]Department of Economics (AE1), Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Homepage: www.elias-tsakas.com; E-mail: e.tsakas@maastrichtuniversity.nl

We first observe that the basic static Karni mechanism does not have obviously dominant strategies. In fact, static mechanisms rarely do (Section 4.3), suggesting that we should probably focus on dynamic ones. To the best of our knowledge, the only existing candidate is Karni's ascending mechanism, which resembles an ascending clock auction and is strategically equivalent to his basic static mechanism (Karni, 2009).[1] However, as we prove, the ascending Karni mechanism does not have an obviously dominant strategy (Theorem 1). Thus, we introduce a variant, which we call the descending Karni mechanism, and we prove that an obviously dominant strategy always exists under our novel descending mechanism. Moreover, with our mechanism, under the assumption that the subject chooses an obviously dominant strategy, the investigator can elicit the subject's true beliefs with arbitrary precision (Theorem 2).

Notably, our results hold for a very broad set of subject's preferences, going well beyond expected utility theory (see Section 4.1 for a discussion). Indeed, we model subjective uncertainty with a likelihood relation which is not necessarily represented by a probability measure (Koopman, 1940; de Finetti, 1949). Within such a model, given that probabilities are not well defined, we need to generalize the notion of "approximating the subject's beliefs". Indeed, we first fix a sequence of events which the subject deems increasingly more likely, e.g., a sequence of nested events. Then, our mechanism identifies two consecutive events in this sequence such that, the subject deems the likelihood of our main event of interest to be between the respective likelihoods of these two events. In the special case of probabilistic beliefs, said sequence of events could comprise of objective lotteries with increasing number of winning tickets, thus eliciting an interval of probabilities within which the subjective probability of our event of interest will belong.

The paper is structured as follows: In Section 2 we introduce the model and the two mechanisms. In Section 3 we present the solution concepts and our two results. Section 4 contains a discussion. All proofs are relegated to the Appendix.

2. Model and mechanisms

There is an algebra $\mathcal{A}$ of events. A male subject's (latent) likelihood relation $\succeq$ is modelled by a weak order over $\mathcal{A}$. Each event $A \in \mathcal{A}$ is associated with the act that pays a fixed prize (desired by the subject) when A occurs and pays nothing otherwise. In this sense, whenever the subject receives the corresponding act, we simply say that he receives A. Hence, it is the case that A is deemed (weakly) more likely than B if and only if the act that is associated with A is (weakly) preferred to the act associated with B.

We fix an arbitrary event $E \in \mathcal{A}$ and consider a sequence of events $A_1, \dots, A_T \in \mathcal{A}$ such that:

(L_1) $A_T \succ \dots \succ A_{t+1} \succ A_t \succ \dots \succ A_1$,

(L_2) $A_T \succeq E \succeq A_1$.

We denote the class of likelihood relations that satisfy $(L_1) - (L_2)$ by $\mathfrak{L}$. Although the subject's likelihood relation is private information (i.e., $\succeq$ is latent), the investigator is assumed to know that it satisfies the previous two conditions (i.e., $\succeq \in \mathfrak{L}$).[2] Hence for every likelihood relation $\succeq \in \mathfrak{L}$ there exists some $t \in \{1, \dots, T-1\}$ such that

$$A_{t+1} \succeq E \succeq A_t. \tag{1}$$

---

[1]In the original paper, the mechanism is defined in continuous time. Since we are mostly interested in actual implementation, we focus on the discrete version of the mechanism (see Section 4.2 for a discussion).

[2]This assumption is easily justified when $\mathcal{A}$ is an algebra over the state space Ω such that $A_1 = \emptyset$ and $A_T = \Omega$, with $(A_t)_{t=1}^{T}$ being a sequence of nested events and the subject's preferences satisfying monotonicity (see Section 4.1).

A female investigator is then interested in eliciting such $t \in \{1, \dots, T-1\}$. In other words, she is interested in learning how likely E is, in relation to the events $A_1, \dots, A_T$.

We consider two mechanisms, the *ascending Karni mechanism* (Karni, 2009) and the novel *descending Karni mechanism*, both of which bear similarity to clock auctions. Assume that the set

$$\mathcal{T} := \{1, \dots, T\}$$

is used as a time index. At the initial history the investigator (playing as nature) draws some $s \in \mathcal{T}$, henceforth called the *stopping time*, which is unknown to the subject. It is known to the subject that all stopping times in $\mathcal{T}$ occur with positive probability (without being necessarily equally likely though).[3] The subject is then sequentially presented with *exit opportunities*, until right before the stopping time is reached. In the ascending mechanism, exit opportunities start at $t = 1$ and count upwards till $t = s-1$, whereas in the descending mechanism they start at $t = T$ and count downwards till $t = s+1$. Thus, non-terminal histories are identified by pairs of the form (t, s), with t being the current *exit opportunity* (which is always observed by the subject) and s being the *stopping time* (which, as we have already mentioned, is not observed before having been reached). There are two available actions at each non-terminal history in each mechanism, viz., stay (s) and exit (e). Let us describe the two mechanisms in detail.

Definition 1. (ASCENDING KARNI MECHANISM). The first exit opportunity is $t = 1$. The set of non-terminal histories is $\mathcal{H}_a := \{(t, s) \in \mathcal{T}^2 : t < s\}$, with each information set being identified by the respective exit opportunity in $\mathcal{T}_a := \{1, \dots, T-1\}$, i.e., the information set $\{(t, t+1), \dots, (t, T)\}$ is identified by $t \in \mathcal{T}_a$. The histories are ordered as follows:

- THE SUBJECT EXITS: if the subject exits at a non-terminal history (t, s), the game ends and he receives A_s.

- THE SUBJECT STAYS AND A STOPPING TIME IS NOT IMMEDIATELY REACHED: if the subject stays at (t, s) with $t < s - 1$, then the non-terminal history $(t + 1, s)$ is reached.

- THE SUBJECT STAYS AND A STOPPING TIME IS IMMEDIATELY REACHED: if the subject stays at $(s - 1, s)$, then the game ends and he receives E.

A *pure strategy* (for the subject) in the ascending mechanism is a mapping $\sigma : \mathcal{T}_a \to \{s, e\}$, with $\mathbb{S}_a$ denoting the set of all pure strategies. For a strategy $\sigma \in \mathbb{S}_a$, we define the corresponding *exit time*,

$$t_a(\sigma) := \min \left\{ \{t \in \mathcal{T}_a : \sigma(t) = e\} \cup \{T\} \right\}.$$

The condition $t_a(\sigma) = T$ implies that the subject does not take any exit opportunity, thus choosing to stay until a stopping time is reached. Then, we define the *reachable exit opportunities* by

$$\mathcal{T}_a(\sigma) := \{t \in \mathcal{T}_a : t \le t_a(\sigma)\}$$

for a strategy $s \in \mathbb{S}_a$.

Definition 2. (DESCENDING KARNI MECHANISM). The first exit opportunity is $t = T$. The set of non-terminal histories is $\mathcal{H}_d := \{(t, s) \in \mathcal{T}^2 : t > s\}$, with each information set being identified by the respective exit opportunity in $\mathcal{T}_d := \{2, \dots, T\}$, i.e., the information set $\{(t, 1), \dots, (t, t-1)\}$ is identified by $t \in \mathcal{T}_d$. The histories are ordered as follows:

---

[3] For our analysis the actual probability of each stopping time being drawn is irrelevant as long as they all occur with positive probability.

- THE SUBJECT EXITS: if the subject exits at a non-terminal history (t, s), the game ends and he receives E.

- THE SUBJECT STAYS AND A STOPPING TIME IS NOT IMMEDIATELY REACHED: if the subject stays at (t, s) with $t > s + 1$, then the non-terminal history $(t - 1, s)$ is reached.

- THE SUBJECT STAYS AND A STOPPING TIME IS IMMEDIATELY REACHED: if the subject stays at $(s - 1, s)$, then the game ends and he receives A_s.

Similarly to the previous case, a *pure strategy* (for the subject) in the descending mechanism is a mapping $\sigma : \mathcal{T}_d \to \{\mathtt{s}, \mathtt{e}\}$, with $\mathbb{S}_d$ denoting the set of all pure strategies. For a strategy $\sigma \in \mathbb{S}_d$, we define the corresponding *exit time*,

$$t_d(\sigma) := \max \left\{ \, \{ t \in \mathcal{T}_d : \sigma(t) = \mathtt{e} \} \cup \{ 1 \} \, \right\} .$$

The condition $t_a(\sigma) = 1$ says that the subject does not take any exit opportunity, thus choosing to stay until nature stops the clock. Then, we define the *reachable exit opportunities* by

$$\mathcal{T}_d(\sigma) := \{ t \in \mathcal{T}_d : t \geq t_d(\sigma) \}$$

for a strategy $s \in \mathbb{S}_d$.

(a) The ascending mechanism. (b) The descending mechanism.
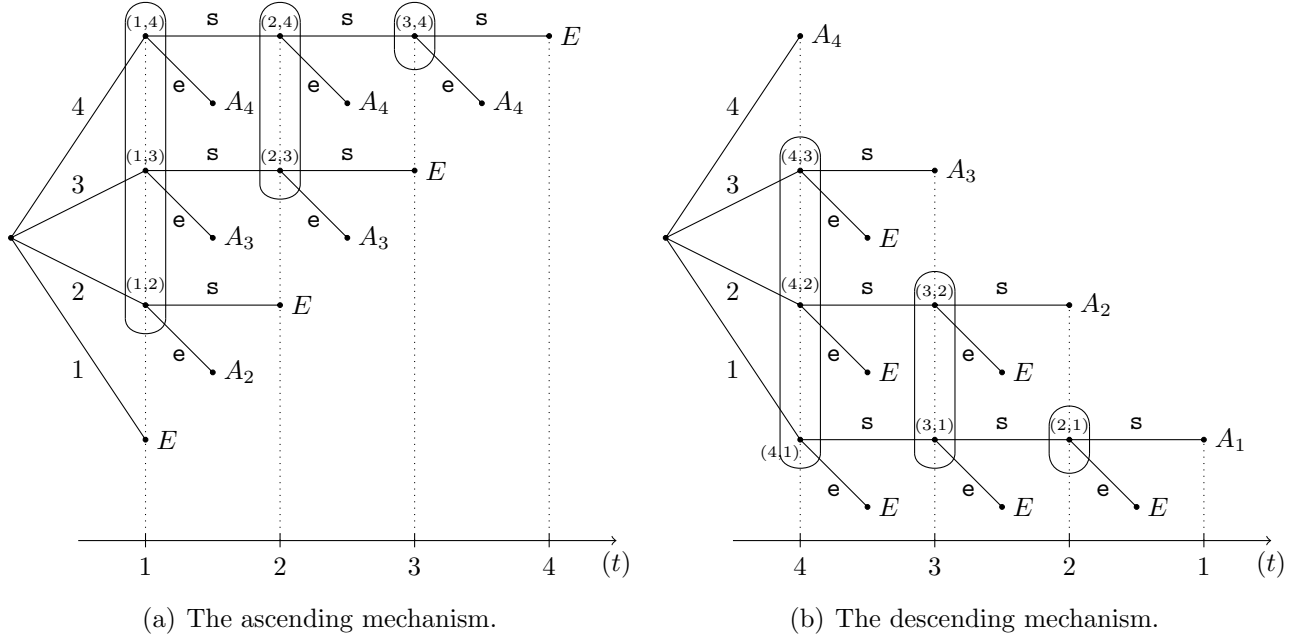
Figure 1: The two Karni mechanisms with $\mathcal{T} = \{1, 2, 3, 4\}$.

Henceforth, we denote by $\mathbb{S}$ the set of strategies when we wish to speak of both mechanisms simultaneously, with the understanding that each mechanism should be considered with its corresponding definitions. Likewise, $t(\sigma)$ and $\mathcal{T}(\sigma)$ are the general definitions of the exit time and the reachable exit opportunities for a strategy $\sigma \in \mathbb{S}$. Moreover, for each $t \in \mathcal{T}$, we denote by $\sigma_t \in \mathbb{S}$ a typical strategy such that $t(\sigma_t) = t$.

Let $O(\sigma | t, s)$ denote the event that the subject receives by playing σ from (t, s) onwards conditional on (t, s) having been reached, with $O(\sigma | s)$ being the event that he receives by σ at the first information set, i.e., in the ascending mechanism $O(\sigma | s) := O(\sigma | 1, s)$ and in the descending mechanism $O(\sigma | s) := O(\sigma | T, s)$.

3. Solution concepts and results

We consider the following solution concepts:

Definition 3. (DOMINANCE). A pure strategy $\sigma \in \mathbb{S}$ is $\succeq$-*dominant* if for all $s \in \mathcal{T}$ and all $\sigma' \in \mathbb{S}$,

$$O(\sigma|s) \succeq O(\sigma'|s). \tag{2}$$

The set of all $\succeq$-dominant strategies is denoted by $\mathbb{D}_\succeq \subseteq \mathbb{S}$.

Definition 4. (OBVIOUS DOMINANCE). A pure strategy $\sigma \in \mathbb{S}$ is $\succeq$-*obviously dominant* if for each $t \in \mathcal{T}(\sigma)$, for each pair of non-terminal histories (t, s) and (t, s'), and each $\sigma' \in \mathbb{S}$ such that $t \in \mathcal{T}(\sigma')$ and $\sigma'(t) \neq \sigma(t)$

$$O(\sigma|t, s) \succeq O(\sigma'|t, s'). \tag{3}$$

The set of all $\succeq$-obviously dominant strategies is denoted by $\mathbb{OD}_\succeq \subseteq \mathbb{S}$.

Dominance is the usual requirement that σ performs at least as well as any alternative strategy no matter which final offer nature has chosen. Note that we are not comparing payoffs given each state of nature, but rather given each stopping time. Obvious dominance is stronger, requiring that at each information set of non-terminal histories that can be reached under σ, the worst case from proceeding with σ is at least as desirable as the best case that might result from deviating to another strategy σ' that also reaches this information set and prescribes a different action at this same information set.

Define the set of exit times that can be observed under the assumptions that the subject chooses dominant and obviously dominant strategies respectively:

$$\begin{aligned}
\mathcal{T}_\mathbb{D}^\succeq &:= \{t(\sigma) \mid \sigma \in \mathbb{D}_\succeq\}, \\
\mathcal{T}_\mathbb{OD}^\succeq &:= \{t(\sigma) \mid \sigma \in \mathbb{OD}_\succeq\}.
\end{aligned}$$

Definition 5. (DOMINANCE ELICITING MECHANISM). A mechanism is *dominance eliciting* if exactly one of the following two conditions holds:

(i) $\emptyset \neq \mathcal{T}_\mathbb{D}^\succeq \subseteq \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}$ for every $\succeq\ \in \mathfrak{L}$.

(ii) $\emptyset \neq \mathcal{T}_\mathbb{D}^\succeq \subseteq \{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\}$ for every $\succeq\ \in \mathfrak{L}$.

A mechanism is *fully dominance eliciting* whenever the respective weak inclusion holds with equality.

Definition 6. (OBVIOUS DOMINANCE ELICITING MECHANISM). A mechanism is *obvious dominance eliciting* if exactly one of the following two conditions holds:

(i) $\emptyset \neq \mathcal{T}_\mathbb{OD}^\succeq \subseteq \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}\}$ for every $\succeq\ \in \mathfrak{L}$.

(ii) $\emptyset \neq \mathcal{T}_\mathbb{OD}^\succeq \subseteq \{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\}$ for every $\succeq\ \in \mathfrak{L}$.

A mechanism is *fully obvious dominance eliciting* whenever the respective weak inclusion holds with equality.

Thus, a mechanism is dominance eliciting, if a dominant strategy exists, and moreover – assuming that the subject surely plays a dominant strategy – if the investigator observes an exit time t then she learns that either $A_{t+1} \succeq E \succeq A_t$ or $A_t \succeq E \succeq A_{t-1}$. Notice the order of the quantifiers, viz., either the first condition holds for all likelihood relations or the second one does for all likelihood relations, rather than for all likelihood relations one of the two conditions holding. The mechanism

is fully dominance eliciting if the converse holds too, i.e., the only strategies that are dominant are those satisfying the respective condition. For an obvious dominance eliciting mechanism, the previous conditions are true under the considerably stronger assumption that a subject is sure to play an obviously dominant strategy. Therefore, an obvious dominance eliciting (resp., fully obvious dominance eliciting) mechanism is dominance eliciting (resp., fully dominance eliciting), but not necessarily vice versa. Then, we are ready to present our two results:

Theorem 1. *The ascending Karni mechanism is fully dominance eliciting, but not obvious dominance eliciting.*

SKETCH OF THE PROOF (FOR $T = 4$). We consider the mechanism depicted in Figure 1(a), while the full proof for arbitrary $T > 0$ is relegated to the Appendix.

First, assume that E *is ranked strictly between two of our events*, e.g., let $A_3 \succ E \succ A_2$. In this case, the only dominant strategy is to exit at $t = 2$, i.e., $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{2\} = \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}$. Second, assume that E *is equally likely as one of our events*, e.g., let $E \sim A_2$. In this case, the two dominant strategies are to exit at $t = 1$ or to exit at $t = 2$, i.e., $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{1, 2\} = \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}$. Hence, Condition (i) in Definition 5 holds with equality for every $\succeq \in \mathfrak{L}$, and therefore the ascending Karni mechanism is fully dominance eliciting.

Now, going back to the case where $A_3 \succ E \succ A_2$, observe that the unique dominant strategy is not obviously dominant. Indeed, conditional on the current information set ($t = 2$), the best case scenario if the subject exits at $t = 3$ is A_4 which is superior to the worst case scenario if he exits at $t = 2$ which is A_3. Therefore, the set of obviously dominant strategies is empty for some $\succeq \in \mathfrak{L}$, implying that the mechanism does not satisfy either of the two conditions of Definition 6, and therefore it is not obvious dominance eliciting. □

Theorem 2. *The descending Karni mechanism is fully obvious dominance eliciting.*

SKETCH OF THE PROOF (FOR $T = 4$). We now switch attention to the mechanism depicted in Figure 1(b), while the full proof for arbitrary $T > 0$ is relegated to the Appendix.

First, assume that E *is ranked strictly between two of our events*, e.g., let $A_3 \succ E \succ A_2$. In this case the only obviously dominant strategy is to exit at $t = 3$, i.e., $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{3\} = \{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\}$. Second, assume that E *is equally likely as one of our events*, e.g., let $E \sim A_3$. In this case, the two obviously dominant strategies are to exit at $t = 3$ or to exit at $t = 4$, i.e., $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{3, 4\} = \{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\}$. Hence, Condition (ii) in Definition 5 holds with equality for every $\succeq \in \mathfrak{L}$, and therefore the ascending Karni mechanism is fully obvious dominance eliciting. □

4. Discussion

4.1. Preference specifications

Throughout the paper we have considered a very broad family of preferences. Formally, there exists a set of outcomes X and a measurable state space $(\Omega, \mathcal{A})$. The subject has complete and transitive preferences $\trianglerighteq$ over the set of $\mathcal{A}$-measurable acts in X^{Ω}. As usual, with slight abuse of notation, we write $x \trianglerighteq y$ whenever the constant act that pays x at every state is weakly preferred to the constant act that pays y at every state.

For two outcomes $x, y \in X$ and an event $A \in \mathcal{A}$, we denote by $x_A y$ the act that pays x if A occurs and pays y otherwise. Then, for $x \trianglerighteq y$ and $x' \trianglerighteq y'$, we further postulate $x_A y \trianglerighteq x_B y$ if and only if $x'_A y' \trianglerighteq x'_B y'$, viz., we postulate Savage's (1954) weak comparative probability. Then, the likelihood relation $\succeq$ is obtained as follows: for $x \trianglerighteq y$ and any two $A, B \in \mathcal{A}$, it is the case that $A \succeq B$ if

and only if $x_A y \trianglerighteq x_B y$. While this axiom is not in principle needed for our results, we impose it so that the likelihood relation does not depend on the prize, which is quite natural to assume. In other words, the likelihood relation is the primitive that we would like to elicit, and the choice of the prize is inconsequential for the elicited qualitative beliefs.

Moreover, although not formally needed for our results, we can naturally further assume monotonicity with respect to set inclusion, i.e., if $A \supseteq B$ then $A \succeq B$. This is a weakening of Anscombe and Aumann's (1963) monotonicity. In this case, we can select any nested sequence of events,

$$\Omega =: A_T \supset \cdots \supset A_t \supset \cdots \supset A_1 := \emptyset,$$

and the two conditions $(L_1) - (L_2)$ hold trivially for every $E \in \mathcal{A}$.

The most common special case of our preferences is a preference relation represented by an expected utility function, in which case $\succeq$ is represented by a (unique probability) measure $\mu \in \Delta(\Omega, \mathcal{A})$, such that $A \succeq B$ if and only if $\mu(A) \geq \mu(B)$.

4.2. Discrete versus continuous mechanisms

Contrary to original continuous (descending) mechanism (Karni, 2009), here we consider the discrete counterpart. The reason for this slight departure is twofold. First, theoretically there are strategies that it is not clear how they should be implemented, e.g., consider the continuous ascending mechanism and suppose that σ is such that the subject stays at every $t \in [0, \tau]$ and exits at every $t \in (\tau, T]$. Then, the subject's exit time is not well-defined, as $\{t \in [0, T] : \sigma(t) = \mathbf{e}\}$ does not achieve a minimum. The second reason is more of practical nature. It is very hard to imagine how one would implement continuous strategies in the lab. In either case our theoretical results suggest that – even with discrete mechanisms – we can approximate the subject's beliefs with arbitrary precision. Indeed, we can arbitrarily choose a "discrete ruler" to measure the beliefs, viz., the sequence of events $A_1, \ldots, A_T$. Notably the ruler can consist of events that are assigned an objective probability, as in the example in the introduction. In this case, we can put E in a probability scale with arbitrary precision.

4.3. Static Karni mechanism

As we have already mentioned in the introduction, static mechanisms rarely have an obviously dominant strategy. Let us illustrate that this is the case for the static counterpart of each of the two dynamic mechanisms that we consider in this paper.

Starting with the ascending mechanism, the corresponding static mechanism would be defined as follows: the investigator randomly draws some some $s \in \mathcal{T}$, while the subject chooses some $t \in \mathcal{T}_a$. The payoff are defined analogously to the dynamic mechanism, i.e., the subject receives A_s is $t < s$, and he receives E if $t \geq s$. On the other hand, the corresponding static mechanism to our descending Karni mechanism is defined as follows: the investigator again randomly draws some some $s \in \mathcal{T}$, while the subject chooses some $t \in \mathcal{T}_d$. The subject receives E if $t > s$, and he receives A_s if $t \leq s$. A strategy t is obviously dominant if the worst case scenario from choosing t is at least as good as the best case scenario from choosing any other $\tau \neq t$.

It is not difficult to verify that there is no such strategy for any of the two mechanisms. To see this observe that a strategy is obviously dominant in the corresponding static mechanism if and only if the aforementioned condition holds at the first information set of the corresponding dynamic mechanism.

4.4. Multiple events

In his original paper, Karni (2009) mentions that his (continuous static) mechanism can be also used to elicit a multinomial distribution (i.e., subjective beliefs about multiple events) in dominant strategies, by repeatedly applying it for each event separately. This approach does not directly extend to elicitation in obviously dominant strategies, even if we use our novel descending mechanism.

In order to do so, we would need to carefully select the stakes that we pay for each of the events whose subjective probability we want to elicit. Take for instance a partition $\mathcal{E} = \{E_1, E_2, E_3\}$ of events, and assume that the subject has probabilistic beliefs about the events in $\mathcal{E}$. The investigator is interested in approximating these beliefs with error no larger than $1/T$. To this end, she first runs the descending mechanism to elicit the probability that the subject attaches to E_1, using a sequence of events $A_1^1, \ldots, A_T^1$ such that A_t^1 is a lottery that yields the prize with probability t/T. Subsequently, she runs again the descending mechanism to elicit the probability that the subject attaches to E_2, using a sequence of events $A_1^2, \ldots, A_T^2$ such that A_t^2 is a lottery that yields the prize with probability t/T. That way, the obviously dominant strategy in the first mechanism is not affected by the existence of the second mechanism, i.e., the two mechanisms are effectively independent.

A. Proofs

A.1. Proof of Theorem 1

Lemma A1. *If $A_{t+1} \succ E \succ A_t$ for some $t \in \{1, \ldots, T-1\}$, then $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t\}$.*

PROOF. *Downward deviations:* Take some $\tau < t$ and consider the corresponding strategy σ_τ. Take a stopping time $s \in \mathcal{T}$, and consider the following three cases:

(i) $s \leq \tau < t$: Then, we obtain $O(\sigma_t|s) = O(\sigma_\tau|s) = E$.

(ii) $\tau < s \leq t$: Then, we obtain $O(\sigma_t|s) = E \succ A_t \succeq A_s = O(\sigma_\tau|s)$.

(ii) $\tau < t < s$: Then, we obtain $O(\sigma_t|s) = O(\sigma_\tau|s) = A_s$.

Hence, $O(\sigma_t|s) \succeq O(\sigma_\tau|s)$ for all $s \in \mathcal{T}$, with the preference relation being strict for some $s \in \mathcal{T}$.

Upward deviations: Take some $\tau < t$ and consider the corresponding strategy σ_τ. Once again, take a stopping time $s \in \mathcal{T}$, and consider the following three cases:

(i) $s \leq t < \tau$: Then, we obtain $O(\sigma_t|s) = O(\sigma_\tau|s) = E$.

(ii) $t < s \leq \tau$: Then, we obtain $O(\sigma_t|s) = A_s \succeq A_{t+1} \succ E = O(\sigma_\tau|s)$. The weak preference $A_s \succeq A_{t+1}$ follows from $s \geq t+1$, which is the case because $s \in \mathcal{T}$.

(ii) $t < \tau < s$: Then, we obtain $O(\sigma_t|s) = O(\sigma_\tau|s) = A_s$.

Once again, $O(\sigma_t|s) \succeq O(\sigma_\tau|s)$ for all $s \in \mathcal{T}$, with the preference relation being strict for some $s \in \mathcal{T}$.

Combine the previous two cases and the proof is completed. $\qquad \square$

Lemma A2. *If $A_1 \sim E$, then $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{1\}$. If $A_t \sim E$ for some $t \in \mathcal{T} \setminus \{1\}$, then $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t-1, t\}$.*

PROOF. *Downward deviations:* This part is relevant only for the case where $A_t \sim E$ for some $t \in \mathcal{T} \setminus \{1\}$. Take some $\tau < t$ and consider the corresponding strategy σ_τ. Take stopping time $s \in \mathcal{T}$, and consider the following three cases:

(i) $s \leq \tau < t$: Then, we obtain $O(\sigma_t|s) = O(\sigma_\tau|s) = E$.

(ii) $\tau < s \leq t$: Then, we obtain $O(\sigma_t|s) = E \sim A_t \succeq A_s = O(\sigma_\tau|s)$. If $\tau = t - 1$, then $t = s$, and therefore $A_t \sim A_s$. If on the other hand, $\tau < t - 1$, then for any $s \in \mathcal{T}$ with $\tau < s < t$ it is the case that $A_t \succ A_s$.

(ii) $\tau < t < s$: Then, we obtain $O(\sigma_t|s) = O(\sigma_\tau|s) = A_s$.

Hence, $O(\sigma_t|s) \succeq O(\sigma_\tau|s)$ for all $s \in \mathcal{T}$, with the preference relation being strict for some $s \in \mathcal{T}$ if and only if $\tau < t - 1$.

Upward deviations: This part is relevant for both cases, i.e., when $A_t \sim E$ for any $t \in \mathcal{T}$. Take some $\tau > t$ and consider the corresponding strategy σ_τ. Following identical steps to the respective part of Lemma A1, we show that $O(\sigma_t|s) \succeq O(\sigma_\tau|s)$ for all $s \in \mathcal{T}$, with the preference relation being strict for some $s \in \mathcal{T}$.

Combine the previous two cases to complete the proof. $\qquad\square$

DOMINANCE: If $\succeq\, \in \mathfrak{L}$ is such that $A_{t+1} \succ E \succ A_t$ then $\{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\} = \{t\}$, and thus by Lemma A1, we obtain $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}$. Second, suppose that $\succeq\, \in \mathfrak{L}$ is such that $A_1 \sim E$, in which case $\{t \in \mathcal{T} : A_2 \succeq E \succeq A_1\} = \{1\}$, and by Lemma A2, we obtain $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t \in \mathcal{T} : A_2 \succeq E \succeq A_1\}$. Finally, let $\succeq\, \in \mathfrak{L}$ be such that $A_t \sim E$ for some $t \in \mathcal{T} \setminus \{1\}$, in which case $\{t' \in \mathcal{T} : A_{t'+1} \succeq E \succeq A_{t'}\} = \{t - 1, t\}$. Hence, by Lemma A2, we obtain $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}$. Therefore, Condition (i) in Definition 5 is satisfied with equality, implying that the ascending Karni mechanism is fully dominance eliciting.

OBVIOUS DOMINANCE: Let $t < T - 1$ and consider some $\succeq\, \in \mathfrak{L}$ such that $A_{t+1} \succ E \succ A_t$, implying that $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t\}$ (by Lemma A1), and therefore $\mathcal{T}_{\mathbb{OD}}^{\succeq} \subseteq \{t\}$. Then, take $t < \tau \leq s < s'$, and observe that $O(\sigma_t|t, s) = A_s \prec A_{s'} = O(\sigma_\tau|t, s')$ for any σ_τ, implying that $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \emptyset$. Therefore, the ascending Karni mechanism is not obvious dominance eliciting.

A.2. Proof of Theorem 2

Lemma A3. *If $A_t \succ E \succ A_{t-1}$ for some $t \in \{2, \ldots, T\}$, then $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{t\}$.*

PROOF. *Upward deviations:* Take an arbitrary $\tau > t$, and consider a strategy $\sigma_\tau \in \mathbb{S}$ such that $t(\sigma_\tau) = \tau$. Take two non-terminal histories (τ, s) and (τ, s'). Then, consider the following two cases:

(i) $t \leq s$: Then, we obtain $O(\sigma_t|\tau, s) = A_s \succ E = O(\sigma_\tau|\tau, s')$.

(i) $s < t$: Then, we obtain $O(\sigma_t|\tau, s) = E = O(\sigma_\tau|\tau, s')$.

Downward deviations: Take an arbitrary $\tau < t$, and consider a strategy $\sigma_\tau \in \mathbb{S}$ such that $t(\sigma_\tau) = \tau$. Take two non-terminal histories (t, s) and (t, s'). Then, consider the following two cases:

(i) $\tau \leq s'$: Then, we obtain $O(\sigma_t|t, s) = E \succ A_{t-1} \succeq A_{s'} = O(\sigma_\tau|t, s')$.

(i) $s' < \tau$: Then, we obtain $O(\sigma_t|t, s) = E = O(\sigma_\tau|t, s')$.

Putting the previous two cases together implies that $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{t\}$. $\qquad\square$

Lemma A4. *If $A_T \sim E$, then $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{T\}$. If $A_t \sim E$ for some $t \in \mathcal{T} \setminus \{T\}$, then $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{t, t + 1\}$.*

PROOF. *Upward deviations:* This part is relevant only for the case where $A_t \sim E$ for some $t \in \mathcal{T}\setminus\{n\}$. Take some $\tau > t$ and consider the corresponding strategy σ_τ. Take two non-terminal histories (τ, s) and (τ, s'). Then, consider the following two cases:

9

(i) $t \leq s$: Then, we obtain $O(\sigma_t|\tau,s) = A_s \succeq A_t \sim E = O(\sigma_\tau|\tau,s')$, with the preference relation being strict for some non-terminal histories (τ,s) and (τ,s') if and only of $\tau > t + 1$.

(ii) $s < t$: Then, we obtain $O(\sigma_t|\tau,s) = E = O(\sigma_\tau|\tau,s')$.

Downward deviations: This part is relevant for both cases, i.e., when $A_t \sim E$ for any $t \in \mathcal{T}$. Take some $\tau < t$ and consider the corresponding strategy σ_τ. Following identical steps to the respective part of Lemma A3, we show that $O(\sigma_t|t,s) \succeq O(\sigma_\tau|t,s')$, with the inequality being strict for some non-terminal histories (t,s) and (t,s').

Put the previous two parts together and the result is proven. □

OBVIOUS DOMINANCE: By Lemmas A3 and A4, it is the case that $\mathcal{T}_{\mathbb{OD}}^{\succeq} \neq \emptyset$ for every $\succeq \, \in \mathfrak{L}$. Then, if $\succeq \, \in \mathfrak{L}$ is such that $A_{t+1} \succ E \succ A_t$ then $\{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\} = \{t\}$, and thus by Lemma A3, we obtain $\mathcal{T}_{\mathbb{OD}}^{\succeq} = \{t \in \mathcal{T} : A_{t+1} \succeq E \succeq A_t\}$. Second, suppose that $\succeq \, \in \mathfrak{L}$ is such that $A_T \sim E$, in which case $\{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\} = \{T\}$, and by Lemma A4, we obtain $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\}$. Finally, let $\succeq \, \in \mathfrak{L}$ be such that $A_t \sim E$ for some $t \in \mathcal{T} \setminus \{T\}$, in which case $\{t' \in \mathcal{T} : A_{t'+1} \succeq E \succeq A_{t'}\} = \{t, t+1\}$. Hence, by Lemma A4, we obtain $\mathcal{T}_{\mathbb{D}}^{\succeq} = \{t \in \mathcal{T} : A_t \succeq E \succeq A_{t-1}\}$. Therefore, Condition (ii) in Definition 6 is satisfied with equality, implying that the ascending Karni mechanism is fully obvious dominance eliciting.

References

ANSCOMBE, F.J. & AUMANN, R.J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics* 34, 199–205.

BRIER, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78, 1–3.

DE FINETTI, B. (1949). La "logica del plausible" secondo la concezione di Polya. *Atti della XLII Riunione, Societa Italiana per il Progresso delle Scienze*, 227–236.

KADANE, J. & WINKLER, R. (1988). Separating probability elicitation from utilities. *Journal of the American Statistical Association* 83, 357–363.

KAGEL, J., HARSTAD, R. & LEVIN, D. (1987). Information impact and allocation rules in auctions with affiliated private values: a laboratory study. *Econometrica* 55, 1275–1304.

KARNI, E. (2009). A mechanism for eliciting probabilities. *Econometrica* 77, 603–606.

KOOPMAN, B. (1940). The axioms and algebra of intuitive probability. *Annals of Mathematics* 41, 269–292.

LI, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review* 107, 3257–3287.

MANSKI, C.F. (2004). Measuring expectations. *Econometrica* 72, 1329–1376.

SAVAGE, L. (1954). *The foundations of statistics*. Wiley, NY: Dover Publications.

———— (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association* 66, 783–801.