American Economic Review 2020, 110(2): 596–627
https://doi.org/10.1257/aer.20171553

Strategyproof Choice of Social Acts†

By Eric Bahel and Yves Sprumont*

We model uncertain social prospects as acts mapping states of nature to (social) outcomes. A social choice function (or SCF) assigns an act to each profile of subjective expected utility preferences over acts. An SCF is strategyproof if no agent ever has an incentive to misrepresent her beliefs about the states of nature or her valuation of the outcomes. It is unanimous if it picks the feasible act that all agents find best whenever such an act exists. We offer a characterization of the class of strategyproof and unanimous SCFs in two settings. In the setting where all acts are feasible, the chosen act must yield the favorite outcome of some (possibly different) agent in every state of nature. The set of states in which an agent's favorite outcome is selected may vary with the reported belief profile; it is the union of all states assigned to her by a collection of constant, bilaterally dictatorial, or bilaterally consensual assignment rules. In a setting where each state of nature defines a possibly different subset of available outcomes, bilaterally dictatorial or consensual rules can only be used to assign control rights over states characterized by identical sets of available outcomes. (*JEL* D71, D81, R53)

A. *The Problem*

Group decisions are often made under conditions of uncertainty.[1] Nations choose domestic and foreign policies, firms make investment choices, academic departments face recruitment decisions: in all of these examples, a given choice may yield different outcomes whose relative likelihood cannot be assessed objectively. Following Savage (1954), we model such uncertain prospects as *acts*, namely, mappings from a set of relevant states of nature into a set of conceivable outcomes. Because these outcomes matter to all members of the group, we speak of *social* acts.

* Bahel: Department of Economics, Virginia Polytechnic Institute and State University (email: erbahel@vt.edu); Sprumont: Département de Sciences Économiques and CIREQ, Université de Montréal (email: yves.sprumont@umontreal.ca). Roland Bénabou was the coeditor for this article. This paper supersedes an earlier discussion paper by Sprumont entitled "Strategy-Proof Choice of Acts: A Preliminary Study." It evolved from numerous discussions with D. Mishra and A. Sen, and, later, with S. Horan. We are immensely grateful to them. We thank four referees for very useful comments and suggestions. We also benefited from remarks by J. Alcalde-Unzu, M. Amarante, S. Barberà, J. C. Bobda, A. Bogomolnaia, C. Chambers, B. Feunou, H. Haller, M. Le Breton, J. Leroux, J. Massó, H. Moulin, M. Pycia, and N. Tideman. The research reported in this paper was supported by a FRQSC grant.
† Go to https://doi.org/10.1257/aer.20171553 to visit the article page for additional materials and author disclosure statements.

[1] The term *uncertainty* refers to contexts where there exists no objective probability measure giving the likelihood of the random events, as in the case of a horse race. Our work does not examine decision making under risk, where an objective probability measure gives the likelihood of every random event, as in the case of a coin flip.

The choice of a social act ought to be based on the preferences of the members of the group: the government of a democratic country should serve the interests of its citizens, the manager of a firm is appointed to make decisions on behalf of its shareholders, and a department head should take her colleagues' opinions into account. The resulting collective decision mechanism may therefore be modeled as a *social choice function* (or SCF) that asks the group members (henceforth called the agents) to report their preferences over acts, and recommends an act for every conceivable preference profile.[2]

Under the assumptions of Savage's theory, an agent's preferences over acts are summarized by the state-independent valuation she attaches to each conceivable outcome and her subjective beliefs about the likelihood of the various states of nature: she compares acts according to the *subjective expected utility* (SEU) they yield to her. Since preferences are typically private information, it is important that the SCF be incentive-compatible: all agents should always find it best to report their preferences truthfully. The purpose of this paper is to understand and describe such incentive-compatible SCFs. We focus on the specific property of *strategyproofness*, which requires that reporting one's true preferences be a dominant strategy. In the current context, this means that an agent should never benefit from misrepresenting the utility she attaches to the outcomes or her beliefs about the states of nature.

Incentive compatibility is a fundamental concern in social choice theory and SEU preferences are the cornerstone of individual decision theory. Yet, somewhat surprisingly, the problem of describing the class of strategyproof SCFs when agents have SEU preferences has been overlooked by the literature. The Gibbard-Satterthwaite theorem (Gibbard 1973, Satterthwaite 1975) does not apply because it requires the full domain of preferences over acts, whereas SEU preferences constitute only a small subset of this full domain. As we shall see, under uncertainty and SEU preferences, non-dictatorial strategyproof SCFs do indeed exist. Some of these SCFs use the differences in subjective beliefs to choose the social act. This is impossible in the case of pure risk where all subjective beliefs must coincide with the same objective probability distribution over states of nature. Under risk, strategyproofness (along with the mild condition of unanimity) entails that SCFs must be convex combinations of dictatorial SCFs, as shown by Hylland (1980): see part E of the introduction for a detailed description.

B. *Our Contribution*

We conduct our analysis of strategyproof SCFs in two settings. Our baseline model assumes no constraint on the set of acts that the social planner may choose from. Given a set of states of nature Ω and a set of outcomes X, all mappings from Ω to X are deemed feasible: there is uncertainty, but it does not affect the range of available opportunities. This model is analyzed in Sections I to IV.

[2] Our framework, like other models of voting, rules out any reallocation of private resources. There is no private wealth, and an SCF must choose social acts whose outcomes are purely public alternatives. This restriction is in the tradition of social choice theory. It is motivated by the observation that modern societies do not mix money with voting and, ultimately, by the desire to give all citizens, regardless of their wealth, an equal weight in the public decision process.

Most real-life collective decision problems under uncertainty do involve some sort of feasibility constraint. In our second model, which is studied in Section V, there is a possibly different subset of available outcomes in every state of nature, and a feasible act must yield in every state of nature an outcome that is available in this state.

Although we are primarily interested in strategyproofness, we impose, for tractability reasons, the auxiliary requirement of *unanimity*: if a feasible act is unanimously preferred to all other feasible acts, it should be chosen. This is a very weak and natural condition; it would indeed be hard to envision a society rejecting a course of action that every single constituent deems optimal. Unanimity is much weaker than *ex post efficiency* (asking that the chosen act should yield a Pareto-optimal outcome in every state of nature), which in turn is weaker than *ex ante efficiency* (asking that the social act itself should be Pareto-optimal).

In the baseline model, the class of strategyproof and unanimous SCFs may be described as follows. Each agent is assigned an event (i.e., a subset of states of nature) in which the outcome of the social act will be her favorite outcome. The assignment of control rights over states must be independent of the reported valuations but need not be constant: it may vary with the reported beliefs. We call every such SCF a *top selection*. Three basic types of such "assignment rules" are allowed under our axioms.

Under a *constant* rule, control rights over states are assigned to agents in a way that is fixed and independent of the beliefs. Under a *bilaterally dictatorial* rule, one agent chooses from an exogenous menu of events the one she considers most likely; and the complementary event is assigned to some other prespecified agent. Under a *bilaterally consensual* assignment rule, the state space is exogenously partitioned into two events that are tentatively assigned to two prespecified agents who may however agree to swap these events. Theorem 1 asserts that every strategyproof and unanimous SCF is a *locally bilateral top selection*, that is, a top selection generated by an assignment rule which is the union of constant, bilaterally dictatorial, or bilaterally consensual "sub-rules" defined over the cells of an arbitrary and exogenous partition of the state space. The name "locally bilateral top selection" emphasizes a surprising property common to all strategyproof and unanimous SCFs: in any given state of nature, the outcome of the social act depends on the preferences of at most two agents.

Note that *both* components of the preferences of *all* agents, tastes and beliefs, may be used to choose the social act. These components play *separate* roles: tastes determine which social outcomes may obtain and beliefs are used to determine in which states of nature these outcomes are picked. In addition, it comes from Theorem 1 that every strategyproof and ex ante efficient SCF is dictatorial: there is an agent whose favorite act is chosen at every preference profile. In the constrained model (discussed in Section V), the class of strategyproof and unanimous SCFs is narrower. As in the baseline model, each agent is assigned control over an event. In each state in that event, the chosen act yields that agent's favorite outcome *within the subset available in that state*. Again, the assignment of control rights over states can only vary with the beliefs. But it must now obey restrictions that are more stringent than in the baseline model. Theorem 2 asserts that bilaterally dictatorial or consensual rules can only be used to assign control rights over states characterized by *identical* sets of available outcomes. Strategyproof and unanimous SCFs may still be described as "locally bilateral constrained-top selections," but the only part of the beliefs that can be exploited is the one pertaining to outcome-irrelevant uncertainty. In particular, if all states generate distinct sets of available outcomes, then their assignment must be completely independent of the agents' beliefs.

For expositional convenience, we assume that an agent never views distinct events as equally likely. This is not very restrictive: as explained in Section VI, indifference (between events or outcomes) can be dealt with by combining our locally bilateral top selections with some tie-breaking rules (see footnote 15).

C. *An Example: Location of a Public Facility*

Consider a city council that must choose where to build a public facility such as an airport. The two models studied in this paper correspond to the following two variants of the problem faced by the council.

Extrinsic Uncertainty and the Baseline Model.—The first variant, which corresponds to our baseline model, involves no *intrinsic* uncertainty. The airport can be built at any of the four locations a, b, c, d. Everything relevant is known about the candidate sites. The set of possible outcomes is $X = \{a, b, c, d\}$ and a choice must be made from that set. Different city councillors have different preferences over a, b, c, d, and the mayor needs to elicit these preferences before making a decision.

Aware of the impossibility to construct a satisfactory revelation mechanism (the Gibbard-Satterthwaite theorem), the mayor suggests to make the decision contingent upon the resolution of some extrinsic uncertainty, say, a horse race: the chosen location will depend on which horse wins the race. This effectively turns the objects of the collective choice into uncertain prospects. The council's task is no longer to choose a site but to decide how the selected site will depend on the winner of the race. The set of relevant states of nature Ω consists of the possible winners of the horse race, and the council chooses an act, i.e., a mapping from Ω to $X = \{a, b, c, d\}$. All acts are feasible because the result of the horse race does not affect the viability of the various sites.

Introducing extrinsic uncertainty into the collective choice process allows the mayor to escape the Gibbard-Satterthwaite impossibility and split the decision power among the city councillors. Theorem 1 tells us exactly to what extent he may do so. To illustrate some of the strategyproof SCFs identified in Theorem 1, let us further assume that there are just two council members, Ann and Bob, and that the extrinsic uncertainty is described by a two-horse race: horse 1 wins in state ω_1, and horse 2 wins in state ω_2. There are thus 16 possible acts; a typical example is (a, b), under which the airport is located at a if horse 1 wins and at b if horse 2 does.

An example of a *constant* assignment rule would be to assign state ω_1 to Ann and state ω_2 to Bob. The council's choice function based on this assignment rule only depends on the agents' preferences regarding the sites a, b, c, d. Ann gets to pick the airport location if horse 1 wins the race, and it is Bob's choice to make if horse 2 wins. Thus, if Ann's favorite site is a and Bob's is b, the council chooses the social act (a, b): and so on for the remaining preference profiles.

An example of a *bilaterally dictatorial* rule would be to assign to Ann whichever state she finds more likely, and the other state to Bob. The social act chosen by the SCF based on this assignment rule depends on the valuations attached by Ann and

Bob to the sites a, b, c, d, as well as on Ann's belief about the result of the horse race. If, for instance, Ann's favorite site is a and Bob's is b, then the social act (a, b) will be selected if Ann believes that horse 1 will win the race, but (b, a) will be selected if she believes that horse 2 will win.

As a last example, consider the *bilaterally consensual* assignment rule where the state of nature ω_1 is assigned by default to Ann, and ω_2 to Bob. The council's choice function will then depend upon both agents' beliefs and favorite sites. For instance, if Ann's favorite site is a and Bob's is b then the council will select the social act (b, a) if Ann believes that horse 2 will win and Bob believes that horse 1 will win; and (a, b) at all other belief profiles.

More flexible choice functions are possible when there are more states of nature. Theorem 1 asserts that every strategyproof and unanimous SCF must be a locally bilateral top selection, which obtains by patching together procedures that are of one of the three varieties just described.

Intrinsic Uncertainty and the Constrained Model.—The second variant of the airport location problem corresponds to our constrained model. This variant involves genuine intrinsic uncertainty. Factors outside of the council's control determine which candidate sites are actually viable. A team of engineers must conduct geotechnical feasibility studies at the respective sites; legal aspects and indemnization questions associated with each site must be assessed by the council's lawyers; environmental impacts have to be estimated for some of the sites and public consultations may have to be held. Depending on the results of all these preliminary feasibility studies, a particular state of nature ω realizes and defines which subset of sites $X_\omega \subseteq \{a, b, c, d\}$ are viable.

The council must make a choice before the uncertainty is resolved. To secure approval and funding from the state authorities, it must write a report proposing a contingent plan that specifies which site will be chosen depending on the results of the feasibility studies. The council's contingent plan is again an act. In this variant, however, an act is feasible if and only if it specifies in every state of nature a site that is viable in this state. Thus, the set of feasible acts is $\prod_{\omega \in \Omega} X_\omega$ and the council must make a choice from that set.

In this variant of the problem, strategyproofness and unanimity again force us to split the state space into an event that falls under Ann's control, and its complement that falls under Bob's control. In each state ω assigned to Ann (Bob), the social act now picks Ann's (Bob's) favorite outcome *in the set* X_ω. To see why the assignment of states to Ann and Bob must obey more stringent restrictions than in the baseline model, suppose that the viable sites form either the set $X_{\omega_1} = \{a, b, c\}$ or the set $X_{\omega_2} = \{a, b, d\}$. Consider a bilaterally dictatorial rule under which Ann is assigned whichever state she finds more likely and Bob gets the other state. At a profile where Ann finds ω_1 more likely than ω_2 and ranks the outcomes in the order d, b, a, c, whereas Bob's favorite outcome is b, the chosen act is (b, b) because b is Ann's favorite in X_{ω_1} and Bob's favorite in X_{ω_2}. But if Ann reports instead that she finds ω_1 *less* likely than ω_2 (and keeps reporting the ranking d, b, a, c), she induces the choice of the act (b, d), which is better for her. The resulting SCF is not strategyproof.

In this example, the two states of nature define different subsets of sites $(X_{\omega_1} \neq X_{\omega_2})$. According to Theorem 2, the assignment of control rights to

Ann and Bob must therefore be exogenous. We discuss in Section VB an example of constrained problem where the assignment may depend upon the belief profile.

D. *Discussion of the Assumptions and Reinterpretation*

The Baseline Model.—As suggested by the illustration in the previous subsection, our two models have different interpretations: they describe collective decision problems under extrinsic versus intrinsic uncertainty. However, it is important to note that they are formally nested: the baseline model is just the particular case of the constrained model where $X_\omega = X$ in every state of nature ω. This is what makes the baseline model useful from a methodological viewpoint. Indeed, proving the results for the constrained model requires only minor adjustments to the proofs for the baseline model.

Even though the assumption that all acts are feasible is unrealistic in most applications involving intrinsic uncertainty, we provide here a reinterpretation of our baseline model under which that assumption *is* reasonable. Suppose that the agents must choose a *sequence* of sure outcomes rather than an uncertain prospect. Reinterpret Ω as a set of *time periods* rather than states of nature. In such a context, it is often natural to assume that the same outcomes are available in all periods. As an illustration, consider the collective choice problem of scheduling a series of meetings. A group of colleagues have to meet once every week during the next six weeks: for each of these weeks they must pick the day on which they will meet. Thus, the set of periods is $\Omega = \{1,2,3,4,5,6\}$, the set of available outcomes is $X = \{Mo, Tu, We, Th, Fr\}$, and the group must choose a sequence of meeting days $f : \Omega \to X$. Constant sequences such as $f = (Mo, Mo, Mo, Mo, Mo, Mo)$ (under which all meetings are held on Monday) are possible, but there is no reason a priori to rule out any particular sequence: the set of feasible sequences is unconstrained. Furthermore, it is natural to assume that preferences over such sequences have the SEU structure postulated in our model: each agent attaches a value to each day of the week,[3] as well as a nonnegative weight to each of the six coming weeks;[4] and she ranks sequences according to the weighted sum of values they yield.

The Constrained Model.—Our constrained model assumes that the set of feasible acts is a Cartesian product. This assumption is natural in a number of contexts. For instance, the standard formulation of general equilibrium theory under uncertainty (Debreu 1959, ch. 7) supposes that individual endowments are state-dependent and the set of social outcomes available in each state consists of all allocations of the collective endowment in that state.

Planning under stochastic budget is another class of problems to which our constrained model naturally applies. Imagine for example a cooperative of farmers who would like to acquire one among a few new types of farming machines for the next production cycles. Each farmer has her own preferences over these types of

[3] Equivalently, each agent attaches an opportunity cost (of attending the meeting) to every day of the week and seeks to minimize a weighted sum of the opportunity costs generated by the sequence of meetings.

[4] These weights are no longer interpreted as probabilities but rather express the relative importance of the respective meetings (or time periods) in the agent's mind. For example, an agent who believes that the first meeting is unimportant will attach a low weight to the first week.

machines, which use different technologies and come at different costs. The farmers decide to invest a given fraction of their profits into the acquisition of this new machine, but of course they do not know in advance which level of profit will be achieved. Their budget set (i.e., the types of machine that they can afford given the realized profit) is thus uncertain; and each farmer has to form her own belief about what types will be affordable. The cooperative hence needs to adopt a contingent plan, which in the language of our model is an act specifying which type of machine the group will purchase for every possible level of profit.

The Cartesian product assumption is essential to our analysis. Of course, many problems involve constraints that do *not* define a product set of feasible acts. Insurance problems are the quintessential example. Consider a couple owning a house of value v. In state of nature ω_1, no fire occurs during the coming year; in state ω_2, a fire occurs and destroys the house. The couple may purchase home insurance for the year at a price of t, which amounts to choosing the act $(v - t, v - t)$. On the other hand, not buying the insurance is equivalent to choosing the act $(v, 0)$. The set of outcomes that may arise in state ω_1 is $\{v - t, v\}$; and the set of outcomes that may arise in state ω_2 is $\{v - t, 0\}$. But the set of feasible acts is *not* the Cartesian product $\{v - t, v\} \times \{v - t, 0\}$. Indeed, note that the act $(v, v - t)$ is not feasible, although it belongs to this product set. Such problems with a non-Cartesian constraint fall outside the scope of our analysis.

E. *Related Work*

The literature related to the current paper may be divided into five strands. The first strand lies at the intersection of statistics and experimental psychology. It is concerned with the problem of eliciting an agent's assessment of the likelihood of uncertain events. The best known incentive-compatible elicitation procedures are the proper scoring rules of McCarthy (1956) and Savage (1971): see Gneiting and Raftery (2007) for a survey of the large literature on the topic. Other procedures include de Finetti's (1974) promissory notes method and Ducharme and Donnell's (1973) reservation probability mechanism. Variants of the latter mechanism appear in Grether (1981), Allen (1987), Holt (2006), and Karni (2009).

The second and most closely related strand studies strategyproofness in the context of risk, that is, when society chooses (social) lotteries rather than acts. The seminal contribution is due to Gibbard (1977), which analyzes mechanisms asking agents to report their preferences over sure outcomes only. Hylland (1980); Dutta, Peters, and Sen (2007, 2008); and Nandeibam (2013) allow agents to report full-fledged von Neumann-Morgenstern preferences over lotteries. A central finding in this literature is that every strategyproof and unanimous SCF is a random dictatorship.[5] Unanimity requires that the chosen lottery attach probability one to an outcome that is everyone's favorite. A random dictatorship selects each agent's most preferred outcome with a probability that does not depend on the reported preference profile.

[5] This result relies on the assumption that all orderings of the sure outcomes are allowed. In settings where further restrictions are imposed on preferences (such as private good allocation problems), the class of strategyproof and unanimous SCFs is typically richer.

The third strand of related work discusses social choice from a Cartesian product set. When individual preferences over such a set are separable *and* form a suitably rich domain, strategyproof and unanimous SCFs are products of strategyproof "sub-rules" defined on the marginal profiles of preferences over the components of the social alternatives: see Le Breton and Sen (1999); Border and Jordan (1983); Barberà, Sonnenschein, and Zhou (1991); and Barberà, Gul, and Stacchetti (1993). This decomposition property does not hold in our setting. The reason is that, although the set of all acts is a Cartesian product and subjective expected utility preferences are separable, they do not form a rich domain.[6] This lack of richness makes it possible to define strategyproof SCFs under which beliefs affect the states where an agent's top outcome is selected.

Fourth, there is a literature on collective decision problems that are repeated over time, such as the problem of repeatedly choosing a meeting date. It is assumed in this literature that preferences over sequences of decisions are represented by *unweighted* sums of utilities; and the goal is then to find *Bayesian* incentive-compatible mechanisms. A number of authors have proposed mechanisms that link collective decisions over time. This allows the planner to exploit the cardinal information contained in the per-period utilities and improve upon the outcome arising from repeated application of a one-period mechanism: see, for instance, Casella (2005), Jackson and Sonnenschein (2007), Casella and Gelman (2008), and Hortala-Vallve (2012). Of course, such Bayesian incentive-compatible mechanisms can only be constructed if the agents and the planner have statistical information about the distribution of preferences: an assumption that our approach dispenses with.

Finally, let us mention that the issue of preference aggregation under uncertainty has received a good deal of attention: see Hylland and Zeckhauser (1979); Mongin (1995); Gilboa, Samet, and Schmeidler (2004); Chambers and Hayashi (2006); and Gilboa, Samuelson, and Schmeidler (2014), among others. This literature, which is normative in nature, is not concerned with the incentive-compatibility issue and is therefore only tangentially related to our work. It shows that aggregation of preferences under uncertainty is problematic; it also questions the desirability of ex ante efficiency (and proposes weakened versions) when individual beliefs differ.

I. The Baseline Model

There is a group $N = \{1, \ldots, n\}$ of at least two agents. Uncertainty is modeled by a finite set of possible states of nature Ω of cardinality at least two. Each state of nature $\omega \in \Omega$ is interpreted as a full resolution of the relevant uncertainty; it is "a description of the world so complete that, if true and known, the consequences of every action would be known" (Arrow 1971, p. 45). Sets of states of nature are called events.

The model is completed by specifying a finite set X of possible outcomes. These outcomes describe what may happen to the agents. To emphasize that they are of

[6] The richness condition requires that, for any collection of admissible preferences over the *components* of the social alternatives, there exists a preference over the social alternatives themselves which induces marginal preferences over components coinciding with the ones in that collection. Since in our setting all state-contingent preferences over outcomes induced by a subjective expected utility preference over acts are identical, Le Breton and Sen's condition is violated.

interest to all members of the group, we often speak of *social* outcomes. The cardinality of X is at least three.

The collective problem is that of choosing a (*social*) *act*. An act is a function f from Ω to X determining an outcome in each possible state of nature. It should be understood as a "social course of action": once the course of action f has been decided by the group, nature selects a state ω, and this yields a unique realized social outcome $f(\omega)$. It may be helpful to think of an act as a collection of outcomes indexed by states of nature. For instance, if there are two states ω, ω' in Ω and three possible outcomes a, b, c in X, the nine possible acts are $(a,a), (a,b), (a,c)$, $(b,a), \ldots, (c,c)$. The act (a,b) describes a course of action under which outcome a obtains if the true state is ω and b obtains if the true state is ω'.

Agents have preferences over acts. The fundamental assumption maintained throughout the paper is that every agent i's preference ordering $\succcurlyeq_i$ is of the *subjective expected utility* (SEU) type. This means that there exist a (state-independent) valuation function v_i over the set of outcomes and a subjective probability measure p_i over the set of events such that

$$f \succcurlyeq_i g \Leftrightarrow E_{v_i}^{p_i}(f) \geq E_{v_i}^{p_i}(g)$$

for any two acts f, g, where $E_{v_i}^{p_i}(h) := \sum_{\omega \in \Omega} p_i(\omega)\, v_i(h(\omega))$ for any act h. In words: an agent compares social acts according to the subjective expected utility they yield to her. We emphasize that the valuation of an outcome is assumed to be independent of the realized state of nature.

We further assume that the preference relation $\succcurlyeq_i$ is a *strict* ordering: agent i is never indifferent between two acts. This is a reasonable assumption given that the set of acts is finite. We make it for convenience; allowing for indifference would add considerable technical difficulties (which we explain in Section VI) but is unlikely to bring much insight. Without loss of generality, we normalize the valuation function v_i of each agent i so that the value of the worst outcome is 0 and the value of the best outcome is 1.

Imagine now a social planner in charge of choosing a social act on behalf of the members of the group. Because preferences are private information, the social planner needs to propose a mechanism by which she requests that information from each agent and specifies how the reported preferences will be used to choose the social act. Such a mechanism is called a social choice function. Formally, if $\mathcal{D}$ denotes the domain of strict SEU preferences described in the previous two paragraphs,[7] a *social choice function* (or SCF) is a function $\varphi : \mathcal{D}^N \to X^\Omega$ assigning an act to every profile of strict SEU preferences. It is convenient to identify an agent's preference with the valuation function and beliefs that generate it.[8] A profile of preferences in $\mathcal{D}^N$ may then be regarded as a list $(v,p) = ((v_1,p_1), \ldots, (v_n,p_n))$ of such valuation functions and beliefs, one for each agent, and $\varphi(v,p)$ is the social act chosen at the preference profile (v,p). We emphasize that this act is allowed to depend on the agents' beliefs as well as on the fine-grained "cardinal" information

[7] See the online Appendix for a formal definition of this domain.

[8] A probability measure p_i qualifies as a belief if and only if it is injective: agent i finds no two events equally likely. This requirement follows from our assumption that preferences over acts are strict.

contained in their valuation functions. In particular, it may change when an agent's valuation function is replaced with one that generates the same ranking of the outcomes but a different ordering of the acts. Thus, no information about individual preferences is a priori discarded.

Because preferences are private information, it is important that the SCF used by the social planner induces the agents to report these preferences truthfully. The specific notion of incentive-compatibility studied in this paper is strategyproofness. Write v_{-i} and p_{-i} for the valuation and belief sub-profiles obtained by deleting v_i from v and p_i from p, respectively. An SCF φ is *strategyproof* if, for all $i \in N$, all $(v,p) \in \mathcal{D}^N$, and all $(v_i',p_i') \in \mathcal{D}$,

$$E_{v_i}^{p_i}\big(\varphi(v,p)\big) \geq E_{v_i}^{p_i}\Big(\varphi\big((v_i',v_{-i}),(p_i',p_{-i})\big)\Big).$$

Strategyproofness means that distorting one's preference, by misrepresenting one's valuation function or one's beliefs, is never profitable.[9] At every profile (v,p) and for every agent i, any pair (w_i,q_i) representing the same ordering $\succcurlyeq_i$ as (v_i,p_i) is a dominant strategy in the preference revelation game generated by φ at (v,p).

Our ultimate objective is to understand the structure of strategyproof SCFs. To facilitate the analysis, however, we impose the very weak auxiliary requirement of unanimity. An SCF φ is *unanimous* if, for all $(v,p) \in \mathcal{D}^N$ and all $f \in X^\Omega$,

$$\left[E_{v_i}^{p_i}(f) \geq E_{v_i}^{p_i}(g) \text{ for all } i \in N \text{ and all } g \in X^\Omega\right] \Rightarrow \left[\varphi(v,p) = f\right].$$

Unanimity only says that if an act happens to be the favorite of all agents, the planner should select it.[10] This requirement imposes no restriction on the collective choice when the agents disagree about the best social act.

II. A Top Selection Lemma

As a preliminary step towards our characterization result, this section establishes a fundamental consequence of strategyproofness and unanimity. At every belief profile, each agent must be assigned a subset of states of nature in which the outcome of the social act will be her favorite outcome. This means that there is no room for compromising through acts which yield "middle-of-the road outcomes" that nobody likes or dislikes very much. At the same time, the result does leave open the possibility of avoiding dictatorship (by selecting the favorite outcomes of different agents in different states of nature) in the spirit of random dictatorships à la Hylland (1980).

[9] A strategyproof SCF may fail to give a *strict* incentive to report one's preference truthfully. Indeed, the direct revelation game generated by a strategyproof SCF φ at a given preference profile (v,p) may have other dominant-strategy equilibria than the true preference profile (v,p). However, *under our assumption of strict preferences*, the act $\varphi(v',p')$ selected by φ at *any* dominant-strategy equilibrium (v',p') of this direct revelation game is precisely the act $\varphi(v,p)$ selected at the true preference profile (v,p). Thus, if the planner cares about implementing a particular SCF (and is not interested per se in eliciting the agents' correct preferences), there is no loss of generality in focusing on strategyproofness rather than dominant-strategy implementation.

[10] In this baseline model, an agent's favorite act is necessarily constant: it yields the same outcome in every state. Unanimity therefore amounts to asking that if all agents agree on the best *outcome*, the planner should pick the constant act yielding that outcome in every state.

A precise formulation requires some terminology and notation. For every valuation function v_i, let $\tau(v_i)$ denote the unique *top* of v_i: this is the favorite outcome of agent i. Define an *assignment* (*of states to agents*) to be an n-component partition $\mathbf{A} = (A_1, \ldots, A_n)$ of the set of states Ω, and denote the set of all assignments by $\mathcal{S}$. Letting $\mathcal{P}$ denote the set of beliefs, define an *assignment rule* to be a function $s : \mathcal{P}^N \to \mathcal{S}$ associating with each belief profile p an assignment $s(p) = (s_1(p), \ldots, s_n(p))$. Here, $s_i(p)$ is the event assigned to agent i at p;[11] we call it i's *share*. This is the set of states of nature in which the outcome of the social act will be i's favorite outcome. Note that an agent's share may be empty. For every preference profile (v,p) and every state of nature ω, denote by $\varphi(v,p;\omega)$ the outcome that the act $\varphi(v,p)$ yields in state ω.

TOP SELECTION LEMMA: *If an SCF φ is strategyproof and unanimous*, *then there exists a unique assignment rule s such that*, *for all* $(v,p) \in \mathcal{D}^N$, $\omega \in \Omega$, *and* $i \in N$, *we have*

$$\omega \in s_i(p) \Rightarrow \varphi(v,p;\omega) = \tau(v_i). \tag{1}$$

We say that the assignment rule s is *associated* with (or *generates*) φ; and we call φ a *top selection*.

Note that this lemma really contains two statements. The first is that every strategyproof and unanimous SCF may only choose acts that yield in every state of nature some agent's top outcome. As a consequence, the top selection lemma forbids outcomes that are natural compromises. Recall the airport location problem described in part C of the introduction; and remember that an agent's top outcome yields a valuation of 1. Suppose that Ann and Bob agree that location b is very good but strongly disagree on the merits of a and c. Specifically, suppose that $v_{Ann}(a) = 1 > v_{Ann}(b) = 0.99 > v_{Ann}(c) = 0$ and $v_{Bob}(c) = 1 > v_{Bob}(b) = 0.99 > v_{Bob}(a) = 0$. Regardless of the beliefs, the social act cannot yield the natural compromise location b in any state of nature. One of the four acts $(a,a), (a,c), (c,a), (c,c)$ must be chosen by the council. Indeed, we show in the proof of the top selection lemma that choosing location b in any of the two states of nature would impose to do the same at valuation profiles where b is only marginally preferred to the respective worst locations of Ann and Bob, which then leads to a profitable manipulation by either Ann or Bob.

The second statement contained in the top selection lemma is that the set of states in which an agent's top outcome is selected may vary only with the profile of beliefs: the valuation profile v is not an argument of the function s. Therefore, the social act ultimately depends only upon the belief profile and the *tops* of the valuation functions.

Thus, the message of the lemma is one of simplicity and separability. A strategyproof and unanimous SCF must necessarily be *simple*. The ordering by the agents of the outcomes they do not rank at the top, hence, a fortiori, the cardinal valuations they attach to these outcomes, must be completely ignored. Simplicity is both

[11] It is more explicit (but also more cumbersome) to talk of assignment of *control rights over states* since agents do not "consume" states but are only granted the right to determine the social outcome in the states assigned to them. Both terminologies have merits, and we will use them interchangeably.

a weakness and a strength. Much relevant information must be discarded but the collective choice mechanism is easily applicable: there is no need for the agents to analyze their own preferences in any detail.

All strategyproof and unanimous SCFs must be *separable*: the agents' valuation functions determine *which outcomes* may obtain, and their beliefs determine *in which states* these outcomes do obtain. This structure stems from the separable nature of the SEU preferences themselves.

The proof of the top selection lemma is in online Appendix Section 2.A but it may be worth sketching the main lines of the argument here. We start by deriving the following three implications of strategyproofness (in Lemmas 1, 2, and 3):

(i) *Monotonicity*: if the chosen act changes when an agent's valuation of an outcome increases (all else equal), then her subjective probability that the social act picks that outcome also increases;

(ii) *Ordinality*: given a belief profile, the same social act must be chosen at all preference profiles generating the same profile of orderings over outcomes;

(iii) *Permutation Invariance*: permuting the utilities of two outcomes that are adjacent in the ranking generated by an agent's valuation function does not change the events where the remaining outcomes are selected.

The central part of the proof of the top selection lemma is Lemma 4. Using the three facts above, we establish the *Tops and Tops Only Property*: (i) in every state of nature, the social act chooses some agent's favorite outcome and (ii) this choice does not depend upon the profile of valuations of non-top outcomes. To prove Lemma 4, we first show that, at any given belief profile and for any profile of favorite outcomes, there is a valuation profile at which the chosen act yields some (possibly different) agent's top in every state of nature: this of course follows from unanimity when all agents have the same top, and we use an induction argument to show that it holds in general. We then show that the social act cannot change if the agents' valuations of non-top outcomes are modified.

Lemma 4 implies that, at each preference profile (v,p), a strategyproof and unanimous SCF assigns to each agent i a share $s_i(v,p)$ of the state space where her favorite outcome is selected. It also guarantees that this share $s_i(v,p)$ does not change with the profile of valuations of non-top outcomes, but still allows the possibility that $s_i(v,p)$ might differ from $s_i(v',p)$ if the profiles of *top* outcomes at v and v' differ. The last part of the proof of the top selection lemma rules out that possibility.

III. The Characterization Theorem

The top selection lemma is not a characterization result yet. While the SCF generated by an assignment rule is obviously unanimous, it need not be strategyproof. Our task is now to determine which assignment rules do indeed generate a strategyproof SCF.

It should be clear that if the agents' shares of the state space are fixed *independently* of the reported beliefs, then nobody has an incentive to manipulate the resulting SCF. Indeed, since beliefs are totally ignored by the social planner and only the tops of the valuation functions are used to choose the social act, the only potentially beneficial manipulations consist in misrepresenting one's favorite outcome. But an agent can only harm herself by doing so because the social outcome would be worse in those states that are assigned to her and would remain unchanged in the other states, thereby decreasing her overall subjective expected utility.

Top selections generated by constant assignment rules are akin to the random dictatorships identified by Hylland (1980) in the context of risky choices. The difference is that the dictator's identity is no longer decided by tossing a coin but is made conditional on the occurrence of an event which does not have an objective probability.

The outstanding question at this point is whether strategyproof and unanimous SCFs can be more sophisticated than these "random dictatorships." Is it possible to condition an agent's share of the state space on the reported belief profile without generating an incentive to manipulate the resulting SCF? Obviously, there must be restrictions on the assignment rule s. For instance, it is easy to see that an agent should never be able to obtain a (subjectively) more likely share of the state space by misreporting her beliefs. Otherwise, this agent would have an incentive to misreport her beliefs at any preference profile where an outcome she does not consider best is the favorite of all other agents. Is this necessary condition also sufficient to guarantee strategyproofness?[12] Can one describe explicitly the assignment rules that generate a strategyproof SCF? Our main theorem will answer these questions.

In a first stage, we begin by identifying three elementary types of assignment rules that do generate strategyproof SCFs. These SCFs should be regarded as "building blocks": in a second stage, we will explain how these building blocks can be combined to construct more complex strategyproof SCFs.

The three building blocks are:

(i) **SCFs Generated by Constant Assignment Rules:** The assignment of control rights over states is fixed exogenously; it does not vary with the reported beliefs. As argued at the start of Section III, this guarantees that the associated SCF, which is essentially a "random dictatorship," is strategyproof because misrepresenting one's favorite outcome is never profitable.

(ii) **SCFs Generated by Bilaterally Dictatorial Assignment Rules:** A bilaterally dictatorial assignment rule fixes *two* agents, say, 1 and 2, and assigns every state to one or the other of these two agents: this is the bilateral aspect of the rule. The identity of these agents is totally independent of the reported preference profile. The assignment of states is based on the beliefs of only *one* of the two agents: this is the dictatorial aspect of the rule. This agent, say, 1, is offered an exogenous menu of (possibly overlapping) events and is requested to identify

[12] When $n = 2$, it is easy to see that an SCF is strategyproof and unanimous if and only if it is generated by a strategyproof assignment rule. Since assignment rules are mathematically equivalent to mechanisms for allocating indivisible objects to agents with additively separable preferences over bundles of objects, Theorem 1 solves, as a by-product, the problem of characterizing all strategyproof such mechanisms: they are precisely the locally bilateral assignment rules defined below. This two-agent result was proved independently by Amanatidis et al. (2017), which however does not study at all the problem of choosing social outcomes under uncertainty, nor considers n-agent assignment rules.

the event she finds most likely in this menu. This event is assigned to her and the complementary event goes to agent 2. The corresponding SCF then picks 1's top outcome in the event she declared most likely, and 2's top outcome otherwise. In a sense, bilateral dictatorship is akin to the notion of delegation: given the exogenous menu of events, the dictator (agent 1) chooses the event she wants to have control rights over, thereby delegating to agent 2 the choice of the social outcome in the complementary event.

Agents 3 to n cannot manipulate such an SCF because it ignores their preferences altogether. Just as in a random dictatorship, agents 1 and 2 have no incentive to misrepresent their *valuation functions* because only their top is recorded and reporting a false top is never profitable under a top selection. Agent 2 cannot manipulate the SCF by misrepresenting her beliefs because they are simply ignored. Agent 1 has no incentive to do so either because the assignment of control rights over states is suitably responsive to the beliefs she reports. Choosing in the menu an event that does not have maximal likelihood for her would (i) decrease her chances to get her favorite outcome, and, because the rule is bilateral, (ii) increase the chances to get agent 2's favorite outcome: the overall subjective expected utility of agent 1 would decrease.

(iii) **SCFs Generated by Bilaterally Consensual Assignment Rules:** A bilaterally consensual assignment rule again fixes two agents, say, 1 and 2. The state space is also exogenously partitioned into two events. The first event is tentatively assigned to agent 1, and its complement is assigned to 2. However, if agent 1 reports that she finds the second event more likely than the first *and* agent 2 reports the opposite belief, they exchange events. This is the consensual aspect of the rule: in contrast to the bilaterally dictatorial rules, the beliefs of both agents 1 and 2 now affect the assignment. The associated SCF picks an agent's reported top outcome in every state that has been assigned to her. Remarkably, such SCFs Pareto-improve upon fixed random dictatorships à la Hylland by allowing the agents to determine the social outcome in those states that they subjectively find particularly likely. Disagreement in the players' preferences is essential to these Pareto improvements: welfare gains over random dictatorships occur only at those preference profiles where some agents (i) disagree on the best outcome and (ii) have opposite beliefs about the likelihood of their tentatively assigned events (see Section IVC for illustration).

For the same reason as before, agents 3 to n cannot manipulate such an SCF. Agents 1 and 2 have no manipulation opportunity either, because the SCF is a top selection and the assignment of control rights over states is responsive to their beliefs: given the belief reported by the other, an agent either (i) cannot affect the event she receives or (ii) is assigned a more likely event when she reports truthfully than when she does not.

Under the SCFs discussed so far, only two agents effectively take part in the collective choice process. We now explain how these elementary SCFs can be combined to allow *all* agents to affect the social act. The key to constructing such composite SCFs lies in the separable nature of SEU preferences. The idea is simply to "patch together" the outcomes delivered by different SCFs on disjoint events: this operation will preserve strategyproofness because SEU agents assess the value of such composite acts by summing the values of their components.

Specifically, let us fix an exogenous partition of the state space into a number of events. For each event belonging to that partition, let us also fix a (possibly different) pair of agents. On each event in the partition, we may now use an assignment "sub-rule" to assign the states belonging to that event *based on these agents' conditional beliefs over these states*. We then compute the overall event assigned to an agent by taking the union of the events assigned to her by all these assignment sub-rules. Finally, in every state of nature, the associated SCF chooses the favorite outcome of the agent to whom that state is assigned.

Here is an example of such a composite SCF. Expanding the airport location example sketched in the introduction, suppose there are three city councillors, Ann, Bob, and Chuck, and four possible winners of the horse race, $\omega_1, \omega_2, \omega_3$, and ω_4. The state space is partitioned into two components: $\{\omega_1, \omega_2\}$ and $\{\omega_3, \omega_4\}$. A bilaterally dictatorial sub-rule is used to assign the states in the first component: specifically, Ann is asked which of ω_1 or ω_2 she considers to be more likely; this state is assigned to her and the other state goes to Bob. A bilaterally consensual sub-rule is used to assign the states in the second component of the partition: by default, Bob is endowed with ω_4 and Chuck with ω_3; they are assigned these states unless both agree to swap them. The resulting aggregate assignment rule is

$$(2)\ s(p) = \begin{cases} (\omega_1; \omega_2\omega_3; \omega_4) & \text{if } p^1_{Ann} > p^2_{Ann} \text{ and } \left[p^3_{Bob} > p^4_{Bob} \text{ and } p^3_{Chuck} < p^4_{Chuck}\right] \\ (\omega_2; \omega_1\omega_3; \omega_4) & \text{if } p^1_{Ann} < p^2_{Ann} \text{ and } \left[p^3_{Bob} > p^4_{Bob} \text{ and } p^3_{Chuck} < p^4_{Chuck}\right] \\ (\omega_1; \omega_2\omega_4; \omega_3) & \text{if } p^1_{Ann} > p^2_{Ann} \text{ and } \left[p^3_{Bob} < p^4_{Bob} \text{ or } p^3_{Chuck} > p^4_{Chuck}\right] \\ (\omega_2; \omega_1\omega_4; \omega_3) & \text{if } p^1_{Ann} < p^2_{Ann} \text{ and } \left[p^3_{Bob} < p^4_{Bob} \text{ or } p^3_{Chuck} > p^4_{Chuck}\right], \end{cases}$$

where p^t_i is shorthand notation for $p_i(\omega_t)$ and curly brackets are omitted (for example, $\omega_2\omega_3$ stands for $\{\omega_2, \omega_3\}$). At every preference profile (v, p), the social act $\varphi(v, p)$ chosen by the associated SCF yields agent i's top $\tau(v_i)$ in all states in the event $s_i(p)$.

Whenever the assignment sub-rule used on a given event of the partition of the state space is constant, bilaterally dictatorial, or bilaterally consensual (as in the above example), it is easy to see that no agent can manipulate the resulting SCF by misrepresenting her beliefs about the states belonging to that event. This is because (i) this agent's expected utility *conditional on the considered event* cannot increase and (ii) her expected utility conditional on the complementary event is independent of the restriction of her beliefs to the considered event, implying that her overall expected utility cannot increase. It follows that a composite SCF based on a union of constant, bilaterally dictatorial, or bilaterally consensual assignment sub-rules is strategyproof.

Theorem 1 asserts that the converse statement is also true: every strategyproof and unanimous SCF is a top selection based on a union of constant, bilaterally dictatorial, or bilaterally consensual assignment sub-rules.

It is now time to state this result formally. Given a nonempty event Ω', define a belief on Ω' to be an injective probability measure (see footnote 8) p_i on the collection of subsets of Ω'; denote the set of all these beliefs by $\mathcal{P}(\Omega')$. An assignment of Ω' is an n-component partition of Ω', and the set of all such assignments is denoted by $\mathcal{S}(\Omega')$. An Ω'-assignment rule is a function $s : \mathcal{P}(\Omega')^N \rightarrow \mathcal{S}(\Omega')$. The building-block rules described earlier are formally defined as follows:[13]

(i) An Ω'-assignment rule s is *constant* if there exists an assignment $\mathbf{A}$ of Ω' such that $s(p) = \mathbf{A}$ for all $p \in \mathcal{P}(\Omega')^N$.

(ii) An Ω'-assignment rule s is (i,j)-*dictatorial* if there exists a proper covering $\mathcal{A}$ of Ω' such that[14]

$$\big(s_i(p), s_j(p)\big) = \Big(\underset{\mathcal{A}}{\operatorname{argmax}}\, p_i,\ \Omega' \backslash \underset{\mathcal{A}}{\operatorname{argmax}}\, p_i\Big)$$

for all $p \in \mathcal{P}(\Omega')^N$. A rule s is bilaterally dictatorial if it is (i,j)-dictatorial for some (unique) ordered pair of agents (i,j).

(iii) An Ω'-assignment rule s is (i,j)-*consensual* (with default $A \subset \Omega'$) if

$$\big(s_i(p), s_j(p)\big) = \begin{cases} (\Omega' \backslash A, A) & \text{if } p_i(\Omega' \backslash A) > p_i(A) \text{ and } p_j(A) > p_j(\Omega' \backslash A) \\ (A, \Omega' \backslash A) & \text{otherwise.} \end{cases}$$

A rule s is bilaterally consensual if it is (i,j)-consensual for some pair of agents (i,j).

Combining these building blocks, we say that an assignment rule $s : \mathcal{P}^N \rightarrow \mathcal{S}$ is *locally bilateral* if there is a partition $\{\Omega^t\}_{t=1}^T$ of Ω and, for each $t = 1, \ldots, T$, a constant, bilaterally dictatorial, or bilaterally consensual Ω^t-assignment rule s^t such that

$$s_i(p) = \bigcup_{t=1}^{T} s_i^t(p \mid \Omega^t) \tag{3}$$

for all $p \in \mathcal{P}^N$ and all $i \in N$, where $p \mid \Omega^t$ is the profile of conditional beliefs generated by p on Ω^t. This partition is called canonical if there is at most one t for which s^t is constant and, for each ordered pair of agents (i,j), at most one t for which s^t is (i,j)-dictatorial. Note that every locally bilateral assignment rule is associated with a unique canonical partition. If the assignment rule s associated with a strategyproof and unanimous SCF φ is locally bilateral, we call φ a *locally bilateral top selection*. The main result for the baseline model is stated below. An overview of the proof is available in Appendix 1 and the detailed proof is available in the online Appendix.

[13] We could equivalently redefine an Ω'-assignment rule as a function from $\mathcal{P}(\Omega)^N$ to $\mathcal{S}(\Omega')$ rather than a function from $\mathcal{P}(\Omega')^N$ to $\mathcal{S}(\Omega')$, and replace $\mathcal{P}(\Omega')^N$ with $\mathcal{P}(\Omega)^N$ in the three definitions below. The formulation we chose emphasizes the fact that an Ω'-assignment rule only uses information contained in the profile of conditional beliefs generated on Ω'. This formulation will prove more convenient in the proof of Theorem 1.

[14] By a proper covering of Ω' we mean a collection of non-nested (possibly overlapping) subsets of Ω' that cover Ω' and have an empty intersection. For instance, $\{\{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}\}$ is a proper covering of $\{\omega_1, \omega_2, \omega_3\}$.

Note that if s is an (i,j)-dictatorial rule, then $s_k(p) = \varnothing$ for all $k \neq i,j$ and all p. Moreover, because $\mathcal{A}$ is a proper covering of Ω', s is not constant and there is no ordered pair $(i',j') \neq (i,j)$ for which s is also (i',j')-dictatorial.

THEOREM 1 (Characterization): *An SCF* $\varphi : \mathcal{D}^N \to X^\Omega$ *is strategyproof and unanimous if and only if it is a locally bilateral top selection.*

The SCFs identified in Theorem 1 are potentially responsive to the preference information reported by *all* agents; in this respect they improve upon the dictatorial SCFs of the Gibbard-Satterthwaite theorem by sharing the decision power ex ante. Furthermore, some of these SCFs exploit both the information regarding the agents' valuations of the outcomes and the information regarding their beliefs: and the theorem identifies precisely how these beliefs may be used. Our framework allows exactly two canonical methods for eliciting beliefs in a way that produces a strategyproof SCF. Bilaterally dictatorial and bilaterally consensual assignment rules exploit beliefs in very different ways. The former rules allow the social planner to extract detailed information about the beliefs of a single agent; and their range (set of n-component partitions of the state space generated by the rule) may be large. The latter have only a binary range but allow the planner to exploit differences in beliefs between agents.

The SCFs in Theorem 1 are flexible enough to allow the choice of a potentially large array of acts. In this respect they improve upon the strategyproof procedure of majority voting between *two* given acts (which of course violates unanimity). For instance, when there are exactly as many agents as states, the SCF choosing the act which yields agent i's favorite outcome in state ω_i has *full range*: every act will be chosen at some preference profile.

Finally, as suggested by their name, all the SCFs in Theorem 1 are locally *bilateral*: for any given state of nature, the outcome of the social act only depends on the preferences of at most two agents. This remarkable consequence of strategyproofness and unanimity is specific to the problem we study; we are not aware of any similar result in the literature. The property is rather unexpected; it is an indirect consequence of the public nature of outcomes. To gain some intuition for the property, recall our earlier location example involving the three councillors Ann, Bob, and Chuck but suppose now that there are three states of nature. States are assigned to agents based on Ann's beliefs:

$$s(p_{Ann}, p_{Bob}, p_{Chuck}) = \begin{cases} (\omega_1, \omega_2, \omega_3) & \text{if } \operatorname{argmax}_\Omega p_{Ann} = \omega_1 \\ (\omega_2, \omega_3, \omega_1) & \text{if } \operatorname{argmax}_\Omega p_{Ann} = \omega_2. \\ (\omega_3, \omega_1, \omega_2) & \text{if } \operatorname{argmax}_\Omega p_{Ann} = \omega_3 \end{cases} \tag{4}$$

This is not a locally bilateral rule: for any given state and any of the *three* agents, there is a belief profile where that state is assigned to that agent. To see why the top selection SCF φ generated by s is not strategyproof, consider a preference profile (v, p) such that $p_{Ann}(\omega_1) = 0.52$, $p_{Ann}(\omega_2) = 0.12$, $p_{Ann}(\omega_3) = 0.36$, $v_{Ann}(\tau(v_{Bob})) = 1$, and $v_{Ann}(\tau(v_{Chuck})) = 0$. If all agents report their preferences truthfully, the selected act yields Ann's favorite outcome in state 1, Bob's in state 2, and Chuck's in state 3, leading to a subjective expected utility of 0.64 for Ann. If instead Ann reports the same valuation function but claims that she finds ω_3 to be the most probable state, the selected act yields Bob's favorite outcome in state 1, Chuck's in state 2, and Ann's in state 3, delivering to Ann a true subjective expected utility of 0.88, which is higher.

As a technical remark, note that our assumption $|X| \geq 3$ is needed for the result of Theorem 1. When there are only two outcomes and the number of agents is odd, majority voting between the two constant acts is a strategyproof and unanimous SCF, although it does not belong to the family described in Theorem 1.

IV. Further Implications of Theorem 1

This section provides additional implications of Theorem 1 and offers a procedure for testing whether a given rule is strategyproof and unanimous.

A. *Ex Ante Efficiency and Strategyproofness*

We start by emphasizing the following consequence of Theorem 1. As pointed out in the introduction, *every strategyproof and ex ante efficient SCF is dictatorial: there is an agent whose favorite act is chosen at every preference profile.* It is easy to check that every dictatorial SCF is ex ante efficient. To see why no other locally bilateral top selection satisfies this property, consider a profile (v,p) such that $\tau(v_i) \in \{a,c\}$ and $v_i(b) = 1 - \varepsilon$ for all $i \in N$. If φ is a non-dictatorial locally bilateral top selection then $\varphi(v,p)$ must be a binary act that yields a in some nonempty event and c in the complementary (and also nonempty) event. When ε is small enough, $\varphi(v,p)$ is Pareto-dominated by the constant act yielding the compromise outcome b in every state.

B. *Testing Whether an SCF Is a Locally Bilateral Top Selection*

Let us now turn to a practical question: when provided with a particular SCF, how does one check in a few steps whether it belongs to the family identified in Theorem 1? The *first* test consists in checking whether the considered SCF always selects top outcomes. An SCF fails this test as soon as there is a preference profile where the chosen act yields, in at least one state of nature, an outcome that is no agent's favorite. As explained in the previous paragraph, non-dictatorial ex ante efficient SCFs are of this type and, hence, they fail the first test.

For SCFs that pass this first hurdle, the *second* test requires that the set of states in which the top of a given agent is selected should not vary with the reported valuations. For instance, consider the SCF selecting agent 1's favorite act if a majority of agents value outcome a more than outcome b, and agent 2's favorite act otherwise. This SCF passes the first test but fails the second because the identity of the "dictator" (1 or 2) depends on the reported valuations. Any SCF passing these first two tests must be a top selection in the sense of (1); and one may then focus on its assignment rule for the remaining tests.

Given the SCF's assignment rule, the *third* test consists in checking that any given state is assigned to *at most two given agents* (independently of other agents' beliefs) as we vary the belief profile over the set of all profiles. For example, the assignment rule described in (4) fails this "bilaterality test" because each of the three agents receives ω_1 at some belief profile.

If an assignment rule passes this third hurdle, the recipient of any given state either (i) does not depend on the beliefs, or (ii) depends on the beliefs of one agent, or (iii) depends on the beliefs of exactly two agents. One can then partition the state space Ω and proceed with the final and *fourth* test as follows. The first cell of the partition gathers all states that are assigned *independently* of the reported beliefs. The second cell contains all states which are assigned to agent 1 or 2 depending *only on agent* 1*'s beliefs*: this is the cell associated with a sub-rule that should be $(1,2)$-dictatorial. The test to be performed here consists in checking that agent 1 indeed receives the event she finds most likely among all events she may receive in this cell. Likewise, for every ordered pair (i,j), all states that are assigned to i or j depending only on agent i's beliefs are gathered in a same cell, and a similar test is performed. Finally, all states that are assigned to i or j on the basis of the beliefs of *both i* and *j* are collected in a same cell and one must check that the assignment rule is indeed (i,j)-consensual on that cell.

As a simple example of assignment rule violating the fourth test, recall the version of the airport location problem where $N = \{Ann, Bob, Chuck\}$, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, and suppose that the assignment rule is as follows:

$$s(p_{Ann}, p_{Bob}, p_{Chuck}) = \begin{cases} (\omega_1, \omega_2, \omega_3) & \text{if } p^1_{Ann} < p^2_{Ann} \\ (\omega_2, \omega_1, \omega_3) & \text{if } p^2_{Ann} < p^1_{Ann}. \end{cases} \tag{5}$$

Observe that there are two cells in the resulting partition of Ω: the "constant cell" $\{\omega_3\}$ which is always assigned to Chuck; and the "(Ann, Bob)-dictatorial cell" $\{\omega_1, \omega_2\}$ where the dictator, Ann, *is never assigned the state she finds most likely*. Hence the SCF generated by this assignment rule is not strategyproof (although it passes the first three tests). Finally, note that the SCF generated by the assignment rule given in (2) passes all four tests and is therefore a locally bilateral top selection.

C. *Welfare Gains over Random Dictatorships*

As pointed out in Section III, SCFs based on consensual assignment rules Pareto-improve upon fixed random dictatorships à la Hylland. To illustrate the size of these gains and how they vary with the differences in the agents' beliefs, consider the following two-agent example.

Fix a random dictatorship φ^{rd} under which $A(nn)$ dictates the outcome in the exogenously fixed event E and $B(ob)$ dictates the outcome in the complementary event E^c, regardless of the beliefs. Let φ^{cons} be the consensual SCF which implements the above arrangement by default but allows the agents to consensually trade the events they were tentatively assigned.

Clearly, φ^{cons} may only improve upon φ^{rd} at those profiles where the consensual trade indeed occurs, that is, when $\pi_A := p_A(E) < 1/2 < p_B(E) =: \pi_B$. At every other profile, the SCFs φ^{rd} and φ^{cons} choose the same social act and hence generate the same welfare distribution. Write $v^B_A := v_A(\tau(v_B))$ and $v^A_B := v_B(\tau(v_A))$ to refer to the respective agents' valuations of the other's top outcome. At any profile $(v,\pi) = ((v_A, v_B),(\pi_A,\pi_B))$ such that $\pi_A < 1/2 < \pi_B$, the utility gains generated by switching from φ^{rd} to φ^{cons} are

$$\begin{aligned} U_A^{cons} - U_A^{rd} &= \left[(1-\pi_A)1 + \pi_A v_A^B\right] - \left[\pi_A 1 + (1-\pi_A) v_A^B\right] \\ &= (1-2\pi_A)(1-v_A^B) \end{aligned}$$

for Ann and

$$\begin{aligned} U_B^{cons} - U_B^{rd} &= \left[\pi_B 1 + (1-\pi_B) v_B^A\right] - \left[(1-\pi_B)1 + \pi_B v_B^A\right] \\ &= (2\pi_B - 1)(1-v_B^A) \end{aligned}$$

for Bob. The aggregate welfare gain is the sum of these utility gains,

$$G = (1-2\pi_A)(1-v_A^B) + (2\pi_B-1)(1-v_B^A).$$

Unless $v_A^B = v_B^A$, this gain need not be monotonically increasing in the belief difference $\pi_B - \pi_A$.

Keeping $v = (v_A, v_B)$ fixed, let us now assume that $\pi = (\pi_A, \pi_B)$ is a random variable. The aggregate welfare gain at v is the random variable

$$G = \begin{cases} (1-2\pi_A)(1-v_A^B) + (2\pi_B-1)(1-v_B^A) & \text{if } \pi_A < 1/2 < \pi_B \\ 0 & \text{otherwise.} \end{cases}$$

The expected value of this welfare gain is

$$\begin{aligned} E(G) = \Pr(\textit{trade})\Big[&(1-2E(\pi_A \,|\, \textit{trade}))(1-v_A^B) \\ &+ (2E(\pi_B \,|\, \textit{trade}) - 1)(1-v_B^A)\Big], \end{aligned} \tag{6}$$

where *trade* denotes the event $\pi_A < 1/2 < \pi_B$. Once again, it is not difficult to see from (6) that the expected welfare gain is generally not necessarily monotonic in the expected belief difference $E(\pi_B - \pi_A \,|\, \textit{trade})$. However, this monotonicity property is satisfied in the symmetric case where $v_A^B = v_B^A =: \alpha$. One can then see that (6) boils down to

$$E(G) = 2\Pr(\textit{trade})E(\pi_B - \pi_A \,|\, \textit{trade})(1-\alpha), \tag{7}$$

which is increasing in (and proportional to) the probability of trade $\Pr(\textit{trade})$ and the conditional expected belief difference $E(\pi_B - \pi_A \,|\, \textit{trade})$.

Moreover, the expected welfare gain is affected by the nature of the correlation between the beliefs. To illustrate this point, suppose that π_B is uniformly distributed on $[0,1]$ and

$$\pi_A = \lambda \pi_B + (1 - \lambda)(1 - \pi_B), \tag{8}$$

where $0 \leq \lambda \leq 1$. Observe that the beliefs are positively correlated if and only if $\lambda > 1/2$. Indeed, the covariance between π_A and π_B is $\mathrm{cov}(\pi_A, \pi_B) = E(\pi_A \pi_B) - E(\pi_A)E(\pi_B) = (2\lambda - 1)\mathrm{var}(\pi_B) = (1/12)(2\lambda - 1)$, which is positive if and only if $\lambda > 1/2$.

If $\mathrm{cov}(\pi_A, \pi_B) \geq 0$, then $\lambda \geq 1/2$ and we obtain from (8) that $\pi_A < 1/2 \Leftrightarrow \pi_B < 1/2$. It follows that $\Pr(\mathit{trade}) = 0$, hence also $E(G) = 0$: the expected welfare gain is nil when the beliefs are positively correlated or uncorrelated.

If $\mathrm{cov}(\pi_A, \pi_B) < 0$, then $\lambda < 1/2$ and (8) implies that $\pi_A < 1/2 \Leftrightarrow \pi_B > 1/2$, so that $\Pr(\mathit{trade}) = \Pr(\pi_A < 1/2 < \pi_B) = \Pr(\pi_B > 1/2) = 1/2$. Moreover, the conditional expected belief difference is $E(\pi_B - \pi_A | \mathit{trade}) = E(\pi_B - \pi_A | \pi_B > 1/2)$ $= E((1 - \lambda)(2\pi_B - 1) | \pi_B > 1/2) = (1/2)(1 - \lambda)$. From (7), the expected welfare gain is then

$$E(G) = \frac{1}{2}(1 - \lambda)(1 - \alpha) = \left(\frac{1}{4} - 3\mathrm{cov}(\pi_A, \pi_B)\right)(1 - \alpha).$$

Thus, when beliefs are negatively correlated, the expected welfare gain increases with the absolute value of the covariance between them.

V. Choice of Social Acts under Feasibility Constraints

This section describes how our analysis generalizes to contexts where not all acts are feasible. We focus on the specific model sketched in the introduction, where each state of nature defines a subset of available outcomes: an act is feasible if it yields in each state of nature an outcome which is available in that state.

A. *The Constrained Model*

As in our baseline model, there is a finite set of agents $N = \{1, \ldots, n\}$ with $n \geq 2$, a finite set of states of nature Ω with $|\Omega| \geq 2$, and a finite set of conceivable outcomes X. An act is a function $f : \Omega \to X$ and preferences over acts are strict orderings of the SEU type. The definition of the preference domain $\mathcal{D}$ is unchanged, and so is the definition of the set of beliefs $\mathcal{P}$.

We postulate that, in every state of nature $\omega \in \Omega$, a subset $X_\omega \subseteq X$ of outcomes are available, with

$$\text{(a)} \quad |X_\omega| \geq 3 \quad \text{for all } \omega \in \Omega \qquad \text{and} \qquad \text{(b)} \quad \bigcup_{\omega \in \Omega} X_\omega = X. \tag{9}$$

Equation 9(a) plays the same role as in Theorem 1. Dropping it would enlarge the set of strategyproof and unanimous SCFs by allowing, e.g., majority voting to decide the outcome in those states ω where $|X_\omega| = 2$. Equation 9(b) entails no loss of generality: if an outcome is available in no state of nature, we may as well delete it from the set of conceivable outcomes X.

We define an act to be *feasible* if and only if it belongs to $\prod_{\omega \in \Omega} X_\omega$, that is, a feasible act yields in any given state of nature an outcome that is available in this state. An SCF is now a mapping φ from $\mathcal{D}^N$ to the set of feasible acts $\prod_{\omega \in \Omega} X_\omega$. This constrained model reduces to our baseline model if $X_\omega = X$ for all $\omega \in \Omega$.

The definition of strategyproofness remains unchanged. The requirement of unanimity is adapted by replacing the set X^Ω by the set $\prod_{\omega \in \Omega} X_\omega$. Thus, an SCF $\varphi : \mathcal{D}^N \to \prod_{\omega \in \Omega} X_\omega$ is *unanimous* if, for all $(v,p) \in \mathcal{D}^N$ and all $f \in \prod_{\omega \in \Omega} X_\omega$,

$$\left[E_{v_i}^{p_i}(f) \geq E_{v_i}^{p_i}(g) \text{ for all } i \in N \text{ and all } g \in \prod_{\omega \in \Omega} X_\omega\right] \Rightarrow \left[\varphi(v,p) = f\right].$$

Just as in the baseline model, this axiom simply means that a feasible act should be selected whenever it is unanimously preferred to all other feasible acts. In contrast to the baseline model, this act need not be constant.

B. *Results*

The top selection lemma generalizes in the obvious way. In the constrained model, strategyproof and unanimous SCFs are now *constraine*d-top selections: in every state of nature, the selected act must yield an outcome that is some agent's favorite in the set of outcomes that are available in this state. Moreover, as in the baseline model, the assignment of states to agents may vary only with the reported beliefs. For any valuation function v_i and any state of nature ω, denote by $\tau_\omega(v_i)$ the unique maximizer (or *constrained top*) of v_i in X_ω. Recalling that $\mathcal{S}$ denotes the set of assignments of Ω and keeping the definition of an assignment rule unchanged, we have the following result.

CONSTRAINED-TOP SELECTION LEMMA: *If an SCF $\varphi : \mathcal{D}^N \to \prod_{\omega \in \Omega} X_\omega$ is strategyproof and unanimous, then there exists an assignment rule $s : \mathcal{P}^N \to \mathcal{S}$ such that, for all $(v,p) \in \mathcal{D}^N$, $\omega \in \Omega$, and $i \in N$, we have*

$$\omega \in s_i(p) \Rightarrow \varphi(v,p;\omega) = \tau_\omega(v_i). \tag{10}$$

We call φ a constrained-top selection.

Theorem 1 also generalizes. However, the presence of feasibility constraints on the admissible acts imposes additional restrictions on the assignment rules that generate strategyproof SCFs.

Recall that an assignment rule s is locally bilateral if there exists a (unique canonical) partition $\{\Omega^t\}_{t=1}^T$ of Ω and, for each $t = 1, \ldots, T$, a constant, bilaterally dictatorial, or bilaterally consensual Ω^t-assignment rule s^t, such that $s_i(p) = \bigcup_{t=1}^T s_i^t(p|\Omega^t)$ for all $p \in \mathcal{P}^N$ and all $i \in N$. We say that such a locally bilateral assignment rule s is *iso-constrained* if for all $t \in \{1, \ldots, T\}$ such that s^t is not a constant Ω^t-assignment rule,

$$\omega_1, \omega_2 \in \Omega^t \Rightarrow X_{\omega_1} = X_{\omega_2}.$$

This means that two states belonging to a cell on which s is not constant must generate the same set of available outcomes. We emphasize that this is not an additional assumption imposed on the constrained model; it is a property which, as we will prove, is satisfied by every assignment rule generating a strategyproof SCF.

To understand this property, recall the two-state example described in part B of the introduction, where $X_{\omega_1} = \{a,b,c\} \neq \{a,b,d\} = X_{\omega_2}$ and the assignment rule s is dictatorial: Ann is assigned control over whichever state she declares more likely and Bob is assigned control over the remaining state. This assignment rule is not iso-constrained because the assignment of ω_1, ω_2 varies with Ann's beliefs even though these states generate distinct sets of available outcomes. As we have seen, the resulting SCF is manipulable. If Ann's ranking of the outcomes is d,b,a,c and Bob's favorite is b, the social act is (b,b) when Ann finds ω_1 more likely than ω_2, and (b,d) otherwise. This gives an incentive to Ann to always report that ω_2 is more likely than ω_1, even when she actually finds ω_1 more likely than ω_2. This profitable manipulation is possible precisely because $X_{\omega_1} \neq X_{\omega_2}$.

Note that a similar manipulation would arise if s were consensual. Suppose, for instance, that ω_1 is assigned by default to Ann, ω_2 is assigned to Bob, and the agents are free to swap these states if they agree to do so. As before, suppose that Ann's ranking of the outcomes is d,b,a,c and Bob's favorite is b. If both agents find ω_1 more likely than ω_2, the chosen act is (b,b). By reporting that she finds ω_2 more likely than ω_1, Ann again induces the choice of (b,d), which she prefers.

As it turns out, a manipulation is possible if and only if s is not iso-constrained.

THEOREM 2 (Constrained Characterization): *An SCF $\varphi : \mathcal{D}^N \to \prod_{\omega\in\Omega} X_\omega$ is strategyproof and unanimous if and only if it is a constrained-top selection whose associated assignment rule s is locally bilateral and iso-constrained.*

The respective proofs of the Constrained-top selection lemma and Theorem 2 are given in online Appendix Section 2.E.

Theorem 2 reduces to Theorem 1 if $X_\omega = X$ for all $\omega \in \Omega$. In that case, every locally bilateral assignment rule is trivially iso-constrained and stating this compatibility requirement thus becomes superfluous.

An important corollary to Theorem 2 is that, *if $X_\omega \neq X_{\omega'}$ whenever $\omega \neq \omega'$, then every strategyproof and unanimous SCF is generated by a constant assignment rule*. The beliefs must be ignored if all states define different subsets of available outcomes.

Put differently, beliefs can only be exploited if they bear on outcome-irrelevant uncertainty. Such uncertainty need not be extrinsic to the problem, however. The following example of outcome-irrelevant *intrinsic* uncertainty illustrates this point. Consider again the airport location problem and suppose there is no uncertainty regarding locations a,b,c: these three sites are viable regardless of the state of nature. On the other hand suppose that the viability of site d is determined by two random and independent factors ϕ_1, ϕ_2, which take value 0 or 1, with the understanding that site d is available if and only if $\phi_1\phi_2 = 1$. For example, factor ϕ_1 may encode the result of the geotechnical studies on site d and ϕ_2 the denouement of the legal procedures aiming at expropriating private land owners occupying this site. Site d must receive a green light from both the engineers and the lawyers before it is deemed viable. Under this specification, the state space may be written as $\Omega = \{\omega_{00}, \omega_{01}, \omega_{10}, \omega_{11}\}$, where state $\omega_{kk'}$ $(k, k' = 0, 1)$ occurs if $\phi_1 = k$ and $\phi_2 = k'$. The sets of available sites for the respective states are given by $X_{\omega_{00}} = X_{\omega_{01}} = X_{\omega_{10}} = \{a, b, c\}$ and $X_{\omega_{11}} = \{a, b, c, d\}$. It follows from Theorem 2 that bilaterally dictatorial or consensual rules may be used to assign the control rights over the states $\omega_{00}, \omega_{01}, \omega_{10}$, essentially *allowing the councillors to bet on the factors that could make site d unavailable*; but state ω_{11} must be assigned independently of the beliefs.

The example above shows that intrinsic uncertainty may be used to assign states to agents. Of course, the identity of the agent whose favorite outcome is chosen can also be made conditional on the resolution of *extrinsic* uncertainty, provided that the latter induces a "weakly finer" state space than the "natural state space" associated with the collective choice problem. By natural state space, we mean the state space Ω^* in which every state of nature is precisely *characterized* by the set of outcomes that are available in it. In the example above, the natural state space is $\Omega^* = \{\omega_0, \omega_1\}$, with $X_{\omega_0} = \{a, b, c\}$ and $X_{\omega_1} = \{a, b, c, d\}$. Instead of making the identity of the agent choosing from X_{ω_0} conditional upon the results of the feasibility studies at site d, the planner could make it conditional on the result of, say, a three-horse race. Introducing this extrinsic uncertainty into the problem induces a new state space $\Omega = \{\omega_0^1, \omega_0^2, \omega_0^3, \omega_1\}$, where ω_0^h is the state of nature where site d is not viable *and* horse h wins the race. This state space refines Ω^* in the sense that each "natural state of nature" $\omega \in \Omega^*$ is associated with a unique nonempty event $E(\omega) \subseteq \Omega$ of the new state space in such a way that $X_{\omega'} = X_\omega$ for each $\omega' \in E(\omega)$ and $\Omega = \bigcup_{\omega \in \Omega^*} E(\omega)$. Bilaterally dictatorial or consensual rules may be used to assign the states $\omega_0^1, \omega_0^2, \omega_0^3$ because $X_{\omega_0^1} = X_{\omega_0^2} = X_{\omega_0^3} = \{a, b, c\}$.

Theorem 2 is also useful in the context of the meeting-scheduling application described in the introduction. Suppose that one particular day of a given week (say, Friday of the first week) is an official holiday. This may be modeled by assuming $X_1 = \{Mo, Tu, We, Th\}$, $X_2 = \cdots = X_6 = \{Mo, Tu, We, Th, Fr\}$. Theorem 2 implies that the control right over the first week must be assigned to an exogenously chosen agent, but the control rights over the remaining weeks may depend upon the subjective weights that the agents attach to these weeks.

VI. Concluding Comments

We examined the problem of designing incentive-compatible collective choice mechanisms when agents have SEU preferences over uncertain social prospects modeled as Savage acts. We showed that strategyproof and unanimous SCFs are (possibly constrained) top selections generated by locally bilateral assignment rules. When all acts are feasible, the SCFs generated by all such assignment rules are strategyproof (Theorem 1). When the set of available outcomes varies with the state of nature, bilaterally dictatorial or consensual sub-rules can only be used to assign control rights over states characterized by identical sets of available outcomes (Theorem 2).

The assumption that preferences are strict orderings (which rules out indifference) could be dispensed with. Under the SCFs identified in our theorems, the act selected at a given preference profile yields some agent's top outcome in every state of nature. When this best outcome is not unique, one faces the problem of characterizing the tie-breaking rules (for choosing between multiple top outcomes) which guarantee that the resulting SCF is strategyproof.[15] Likewise, if an agent may assign the same subjective probability to two events, the assignment rules on which our SCFs are based are no longer well defined and one must characterize which refinements generate strategyproof SCFs. These are difficult but rather technical issues.

We conclude by mentioning some open problems.

(i) How should we choose between the social choice functions identified in our theorems? Assuming a given (say, uniform) distribution over the set of all preference profiles, one could search for SCFs that maximize some measure of expected welfare: such as the expected sum of normalized utilities. Alternatively, one could proceed axiomatically and impose properties that complement strategyproofness and unanimity. However, as a corollary to our theorems, it follows that anonymity is incompatible with the combination of strategyproofness and unanimity.[16]

(ii) Anonymous strategyproof SCFs deserve to be studied. If there is an odd number of agents, majority voting between two prespecified acts is clearly strategyproof and anonymous. But more general and flexible SCFs are possible in this class. Partition the state space into a collection of events. For each event specify two "sub-acts," that is, two mappings from that event into the set of outcomes, and apply *quota voting* to choose between these two sub-acts: the first sub-act is selected if it is preferred to the second by at least κ agents (with the quota κ given exogenously); otherwise the second sub-act is selected. Let then the chosen act be the concatenation of all the sub-acts selected (possibly with different quotas). The additive separability of SEU preferences guarantees that this SCF is strategyproof; it is also obviously anonymous.

(iii) It has long been recognized that the assumption of (state-independent) SEU preferences is unrealistic in many contexts: see Savage and Aumann (1987) for a discussion. Are our locally bilateral top selections well defined and strategyproof on some larger domains of preferences? It is straightforward to check that the class of *state-dependent* SEU preferences forms a rich domain of separable preferences in the sense of Le Breton and Sen (1999). Hence, their Theorem 4.1 implies that every strategyproof and unanimous SCF defined on that rich domain must be a top selection generated by some

[15] In the baseline model, a simple example of tie-breaking rule can be described as follows. Pick an exogenous ordering of the outcomes (say, $a, b, c, \ldots$) and an exogenous ordering of the states (say, $\omega_1, \omega_2, \omega_3, \ldots$). For every valuation v_i with multiple top outcomes, redefine $\tau(v_i)$ to be the unique top outcome of agent *i that comes first* in the ordering $a, b, c, \ldots$. To break ties between events E_1, E_2 such that $p_i(E_1) = p_i(E_2)$, proceed lexicographically: pick E_1 if either the first state in E_1 comes before the first state in E_2 (given the ordering $\omega_1, \omega_2, \omega_3, \ldots$), or the first states are the same in both events but the second state in E_1 comes before that in E_2, and so on. When indifferences are allowed, every SCF in the family described by Theorem 1 will remain strategyproof if it is coupled with a tie-breaking rule of the type just described. Determining which other tie-breaking rules generate strategyproof SCFs is an open problem.

[16] Anonymity means that permuting the preferences of any two agents should not affect the chosen act. This incompatibility contrasts with the results of the literature on social choice under risk where random dictatorship with equal probabilities satisfies all three axioms: see Hylland (1980).

constant assignment rule. The reason why top selections based on bilaterally dictatorial or bilaterally consensual assignment rules cannot be used is crystal clear. Such rules exploit the agents' beliefs to assign the control rights over states of nature. State-dependent SEU preferences, however, generally fail to induce well-defined beliefs: beliefs and state-dependent valuations can be identified jointly but not always separately.[17] Nevertheless, some *subclasses* of state-dependent SEU preferences do allow a separate identification of beliefs and state-dependent valuations. An example is the subclass axiomatized by Karni (1993). On any such subclass, the SCFs identified in our two theorems are well defined, strategyproof, and unanimous. An important open problem is the characterization of the largest such subclass.

(iv) In the constrained model of Section V, the set of feasible acts is a product set. As explained in part D of the introduction, some important real-life problems involve non-Cartesian feasibility constraints. In such problems, the class of strategyproof and unanimous SCFs will generally depend in a subtle way upon the particular nature of these feasibility constraints. These problems are beyond the scope of the current paper but our results should provide a good starting point for their study. A similar generalization to constrained sets of alternatives was successfully achieved in the literature on strategyproofness on rich domains of additively separable preferences originally defined over product sets: see, in particular, Barberà, Massó, and Neme (2005) and Reffgen and Svensson (2012).

(v) In many contexts, it will also be natural to impose restrictions on preferences over outcomes. This may enlarge the class of strategyproof and unanimous SCFs. An interesting case is that of shareholders of a firm choosing acts with monetary outcomes: the profits to be shared. Here all agents have the same monotonic preference ordering over outcomes but not necessarily the same valuation functions or the same beliefs. While the unconstrained problem is uninteresting: the constant act choosing the highest profit level in all states is dominant, the problem of choosing acts under non-Cartesian feasibility constraints is entirely nontrivial. Another interesting problem is the strategyproof allocation of an uncertain collective endowment: think of a cooperative of fishermen deciding on how to split the different types of fish they will catch. In this case the agents have to report (i) monotonic valuation functions over bundles of commodities and (ii) their beliefs about what endowment will be realized.

(vi) Our work has focused on the choice of acts whose outcomes are purely public alternatives. The question of finding strategyproof mechanisms allowing the reallocation of privately owned resources under SEU preferences remains unsolved. For example, the *competitive mechanism* (which assigns to each preference profile the competitive allocation of the corresponding exchange

[17] For a discussion on well-defined beliefs (under state-dependent SEU preferences) and the related literature, see Drèze and Rustichini (2004).

economy) is *not* strategyproof: this was shown by Hurwicz (1972) for classical preferences and the fact remains true for SEU preferences.

Appendix 1: Overview of the Proof of Theorem 1

The proof of Theorem 1 consists of a sequence of lemmas. We present the main ones here, explain their purpose, and provide a few hints regarding their proof. The explanations offered in this overview are not meant to be a formal argument; the detailed proofs are available in the online Appendix.

It is easy to check that every locally bilateral top selection is a strategyproof and unanimous SCF. In order to prove the converse statement, consider a strategyproof and unanimous SCF φ; and suppose that φ is generated by the assignment rule s. We must show that s is a the union of a collection of constant, bilaterally dictatorial, or bilaterally consensual assignment sub-rules.

I. *Super-Strategyproofness and Non-Bossiness*

The first step consists in establishing that s must be "super-strategyproof." The formal proof of this property is given in online Appendix Section 2.B. As argued at the beginning of Section III, an agent should not be able to manipulate her own share of the state space by misreporting her beliefs. But this is not enough. Agent 1 cares about her share $s_1(p)$ only to the extent that the social act will yield her top outcome $\tau(v_1)$ in the event $s_1(p)$. Remember that outcomes are public in nature. At a profile where agent 2's favorite social outcome coincides with 1's and every other agent's favorite outcome is agent 1's worst, agent 1 is therefore interested in maximizing the *joint* share of the state space assigned to 1 and 2, and the rule s must prevent 1 from manipulating this joint share. More generally, no agent should be able to increase her subjective assessment of the likelihood of the joint share of any *subset* of agents to which she belongs by misrepresenting her own beliefs.

Formally, given an assignment $\mathbf{A} = (A_1, \ldots, A_n) \in \mathcal{S}$ and a set $M \subseteq N$, define $A_M = \bigcup_{i \in M} A_i$. The assignment rule s is *super-strategyproof* if $p_i(s_M(p)) \geq p_i(s_M(p_i', p_{-i}))$ for all i, M such that $i \in M \subset N$, all $p \in \mathcal{P}^N$, and all $p_i' \in \mathcal{P}$. The formal result is the following.

SUPER-STRATEGYPROOFNESS LEMMA: *The assignment rule s is super-strategyproof.*

An immediate consequence is that s must be *non-bossy:* for all $i \in N, p \in \mathcal{P}^N$ and $p_i' \in \mathcal{P}$, $s_i(p) = s_i(p_i', p_{-i}) \Rightarrow s(p) = s(p_i', p_{-i})$. In words: no agent can affect another agent's share without affecting her own.

II. *Local Bilaterality*

Using this super-strategyproofness lemma (and the non-bossiness property it implies), the second step of the proof consists in showing that the *local* behavior of s is completely determined. The formal proof of this result is detailed in online Appendix Section 2.B.

The set of beliefs is endowed with a natural relation of adjacency: two beliefs are $\{A,B\}$-*adjacent* if the likelihood orderings they generate over events differ on a single unordered pair $\{A,B\}$ of disjoint nonempty events. For instance, if $\Omega = \{1,2,3\}$, any two beliefs p_i, q_i such that

$$p_i(\{1,2,3\}) > p_i(\{1,2\}) > p_i(\{1,3\}) > p_i(\{2,3\}) > p_i(\{1\}) > p_i(\{2\}) > p_i(\{3\}),$$

$$q_i(\{1,2,3\}) > q_i(\{1,3\}) > q_i(\{1,2\}) > q_i(\{2,3\}) > q_i(\{1\}) > q_i(\{3\}) > q_i(\{2\})$$

are $\{\{2\},\{3\}\}$-adjacent.

Fix an agent i and consider two belief profiles $p,(p_i',p_{-i})$ such that p_i' is $\{A,B\}$-adjacent to p_i. Suppose that $s_i(p) \neq s_i(p_i',p_{-i})$. Then, since p_i and p_i' agree on the ranking of all disjoint events other than A,B, the events $s_i(p)$ and $s_i(p_i',p_{-i})$ must obtain from one another precisely by exchanging A against B. Otherwise, one of the events $s_i(p)$, $s_i(p_i',p_{-i})$ will be considered more likely than the other under both p_i and p_i', leading to a profitable manipulation for agent i at either p or (p_i',p_{-i}).

Moreover, because s is super-strategyproof, one can show that when $s_i(p)$ and $s_i(p_i',p_{-i})$ differ, the shares of all but one of the agents other than i must remain unchanged. This leads to the following result.

LOCAL BILATERALITY LEMMA: *Let A,B be nonempty, disjoint events; and let $i \in N$, $p \in \mathcal{P}^N$, $p_i' \in \mathcal{P}$ be such that p_i,p_i' are $\{A,B\}$-adjacent and $p_i(A) > p_i(B)$. Then, either (i) $s(p) = s(p_i',p_{-i})$ or (ii) there exists $j \in N\backslash i$ such that*

$$s_i(p)\backslash s_i(p_i',p_{-i}) = A = s_j(p_i',p_{-i})\backslash s_j(p),$$

$$s_i(p_i',p_{-i})\backslash s_i(p) = B = s_j(p)\backslash s_j(p_i',p_{-i}),$$

$$s_k(p) = s_k(p_i',p_{-i}) \quad \textit{for all } k \in N\backslash\{i,j\}.$$

This lemma fully characterizes how the agents' shares of the state space are allowed to change as the reported beliefs move from one profile to an "adjacent profile." The rest of the proof consists in turning this local characterization into the global result stated in Theorem 1.

III. *A Preliminary Partition of the State Space*

At this point of the proof, the central difficulty consists in identifying the cells of the partition of the state space where s is constant, bilaterally dictatorial, or bilateral consensual. Moreover, on those cells where s is bilaterally dictatorial, we must construct the menu of events from which the local dictator is allowed to choose. Likewise, on those cells where s is bilaterally consensual, we must identify which pair of agents is decisive, which sub-assignment is chosen by default, and which one is chosen in case of agreement between the two agents involved.

With that goal in mind, online Appendix Section 2.C begins by defining a *preliminary* partition of the state space as follows. Let Ω_0 be the set of states whose assignment is constant, let Ω_1 be the set of states whose assignment varies with the belief of a single agent, and let Ω_2 be the set of states whose assignment varies with the beliefs of *at least* two agents. By definition, $\{\Omega_0, \Omega_1, \Omega_2\}$ is a partition of Ω. At this point, however, more than two agents may influence the assignment of any given state in Ω_2, and there is no restriction on the number of agents to whom a given state in $\Omega_1 \cup \Omega_2$ may be assigned.

IV. *Proving Bilateral Consensus on* Ω_2

The proof proceeds by considering the states in Ω_2 first. We show in online Appendix Section 2.C that each of these states may be assigned to exactly two distinct agents, and the assignment must be based on the beliefs of these two agents only. More specifically, states in Ω_2 must be assigned through bilateral consensus.

BILATERAL CONSENSUS LEMMA: *For every $\omega \in \Omega_2$ there exists a unique event $E^\omega \subseteq \Omega_2$ containing ω, and there exists a bilaterally consensual E^ω-assignment rule s^ω such that*

$$s_i(p) \cap E^\omega = s_i^\omega(p|E^\omega)$$

for all $p \in \mathcal{P}^N$ and $i \in N$.

This lemma fully determines the behavior of s on Ω_2. Indeed, for any two states $\omega, \omega' \in \Omega_2$, since there exist a bilaterally consensual E^ω-rule s^ω and a bilaterally consensual $E^{\omega'}$-rule $s^{\omega'}$ such that $s_i(p) \cap E^\omega = s_i^\omega(p|E^\omega)$ and $s_i(p) \cap E^{\omega'} = s_i^{\omega'}(p|E^{\omega'})$ for all $i \in N$, we must have either (i) $E^\omega = E^{\omega'}$ and $s^\omega = s^{\omega'}$, or (ii) $E^\omega \cap E^{\omega'} = \varnothing$. This means that there exists a partition $\{\Omega^t\}_{t=1}^{T_2}$ of Ω_2 and, for each $t = 1, \ldots, T_2$, a bilaterally consensual Ω^t-assignment rule s^t such that

$$s_i(p) \cap \Omega_2 = \bigcup_{t=1}^{T_2} s_i^t(p|\Omega^t)$$

for all $p \in \mathcal{P}^N$ and $i \in N$. This is precisely the structure of s on Ω_2 asserted in Theorem 1.

The proof of the bilateral consensus lemma works by "contagion" over the set of belief profiles.

Fix a state $\omega \in \Omega_2$. In online Appendix Section 2.C.1, we fix a profile π of beliefs on $\Omega \backslash \omega$, and we restrict our attention to the region $\mathcal{P}^N(\pi)$ of belief profiles on Ω which generate the same profile of likelihood orderings as π on the subsets of $\Omega \backslash \omega$. Using the local bilaterality lemma, we show that there exist two agents i,j and two disjoint events A, B, whose union contains ω, such that the restriction of s to $A \cup B$ coincides with the (i,j)-consensual $(A \cup B)$-assignment rule with default B *on the region* $\mathcal{P}^N(\pi)$. At that point, the agents i,j and the sets A, B may depend upon π.

In online Appendix Section 2.C.2, we consider belief profiles (π_k', π_{-k}) over $\Omega \backslash \omega$ where π_k' is adjacent to π_k for some agent $k \in N$. In a series of "contagion lemmas," we describe how the behavior of the restriction of s to $A \cup B$ on the region $\mathcal{P}^N(\pi_k', \pi_{-k})$ is linked to the behavior of its restriction to $A \cup B$ on $\mathcal{P}^N(\pi)$. Online Appendix 2.C.3

establishes more powerful contagion lemmas describing how the assignment of A, B on the region $\mathcal{P}^N(\pi)$ is linked with their assignment on non-adjacent regions.

Online Appendix Section 2.C.4 patches the pieces together. Using the contagion results of Sections 2.C.2 and 2.C.3 and the connectedness of the set of all beliefs on $\tilde{\Omega}$, we show that s is an (i,j)-consensual $(A \cup B)$-assignment rule with default B on the whole domain $\mathcal{P}^N$. The claim follows by setting $E^{\tilde{\omega}} = A \cup B$.

V. *Proving Bilateral Dictatorship on* Ω_1

Next, in online Appendix Section 2.D, we turn to the assignment of the states in Ω_1. Let Ω_{11} be the subset of those states in Ω_1 whose assignment varies with the beliefs of agent 1. We show that these states are assigned by bilateral dictatorship of agent 1.

BILATERAL DICTATORSHIP LEMMA: *There exist a set* $N_1 \subseteq N \backslash 1$, *a partition* $\{\Omega_{11}^j\}_{j \in N_1}$ *of* Ω_{11}, *and for each* $j \in N_1$ *a* $(1,j)$*-dictatorial* Ω_{11}^j*-assignment rule* s^j *such that*

$$s_i(p) \cap \Omega_{11} = \bigcup_{j \in N_1} s_i^j\left(p | \Omega_{11}^j\right) \tag{11}$$

for all $p \in \mathcal{P}^N$ *and* $i \in N$.

Let us outline the proof here. Consider the family of all subsets of Ω_{11} that are assigned to agent 1 at some belief profile. We begin by showing that $s_1(p) \cap \Omega_{11}$ maximizes p_1 over that family whenever p_1 is a so-called Ω_{11}-dominant belief: one in which only the probability differences between events in Ω_{11} are large. We then use the local bilaterality lemma to extend this observation to all belief profiles p. The next and crucial step consists in proving that every state in Ω_{11} can only be allocated to a single agent other than 1. The set Ω_{11} can therefore be partitioned into a collection of subsets $\{\Omega_{11}^j\}$ such that every state in Ω_{11}^j is allocated to either 1 or j, and super-strategyproofness can be used to show that $s_1(p) \cap \Omega_{11}^j$ maximizes p_1 over the family of all subsets of Ω_{11}^j that are assigned to agent 1 at some belief profile. The argument is completed by appealing to non-bossiness.

We have stated the bilateral dictatorship lemma for agent 1, but a corresponding lemma obviously holds for every agent. It now follows from these bilateral dictatorship lemmas, the bilateral consensus lemma, and the definition of Ω_0, that s is a union of constant, bilaterally dictatorial, or bilaterally consensual assignment sub-rules.

REFERENCES

Allen, Franklin. 1987. "Discovering Personal Probabilities When Utility Functions Are Unknown." *Management Science* **33 (4): 542–44.**

Amanatidis, Georgios, Georgios Birmpas, George Christodoulou, and Evangelos Markakis. 2017. "Truthful Allocation Mechanisms without Payments: Characterization and Implications on Fairness." *Proceedings of the 2017 ACM Conference on Economics and Computation*: 545–62.

Arrow, Kenneth J. 1971. *Essays in the Theory of Risk-Bearing*. Chicago: Markham Publishing Co.

Barberà, Salvador, Faruk Gul, and Ennio Stacchetti. 1993. "Generalized Median Voter Schemes and Committees." *Journal of Economic Theory* 61 (2): 262–89.
Barberà, Salvador, Jordi Masso, and Alejandro Neme. 2005. "Voting by Committees under Constraints." *Journal of Economic Theory* 122 (2): 185–205.
Barberà, Salvador, Hugo Sonnenschein, and Lin Zhou. 1991. "Voting by Committees." *Econometrica* 59 (3): 595–609.
Border, Kim C., and J. S. Jordan. 1983. "Straightforward Elections, Unanimity and Phantom Voters." *Review of Economic Studies* 50 (1): 153–70.
Casella, Alessandra. 2005. "Storable Votes." *Games and Economic Behavior* 51 (2): 391–419.
Casella, Alessandra, and Andrew Gelman. 2008. "A Simple Scheme to Improve the Efficiency of Referenda." *Journal of Public Economics* 92 (10–11): 2240–61.
Chambers, Christopher P., and Takashi Hayashi. 2006. "Preference Aggregation under Uncertainty: Savage vs. Pareto." *Games and Economic Behavior* 54 (2): 430–40.
de Finetti, Bruno. 1974. *Theory of Probability.* Vol. 1. New York: Wiley.
Debreu, Gerard. 1959. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium.* New Haven, CT: Yale University Press.
Drèze, Jacques H., and Aldo Rustichini. 2004. "State-Dependent Utility and Decision Theory." In *Handbook of Utility Theory,* Vol. 2, edited by Salvador Barberà, Peter J. Hammond, and Christian Seidl, 839–92. Dordrecht, Netherlands: Kluwer.
Ducharme, Wesley M., and Michael L. Donnell. 1973. "Intrasubject Comparison of Four Response Modes for 'Subjective Probability' Assessment." *Organizational Behavior and Human Performance* 10 (1): 108–17.
Dutta, Bhaskar, Hans Peters, and Arunava Sen. 2007. "Strategy-Proof Cardinal Decision Schemes." *Social Choice and Welfare* 28 (1): 163–79.
Dutta, Bhaskar, Hans Peters, and Arunava Sen. 2008. "Strategy-Proof Cardinal Decision Schemes: Erratum." *Social Choice and Welfare* 30 (4): 701–2.
Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41 (4): 587–601.
Gibbard, Allan. 1977. "Manipulation of Schemes That Mix Voting with Chance." *Econometrica* 45 (3): 665–81.
Gilboa, Itzhak, Dov Samet, and David Schmeidler. 2004. "Utilitarian Aggregation of Beliefs and Tastes." *Journal of Political Economy* 112 (4): 932–38.
Gilboa, Itzhak, Larry Samuelson, and David Schmeidler. 2014. "No-Betting-Pareto Dominance." *Econometrica* 82 (4): 1405–42.
Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102 (477): 359–78.
Grether, David M. 1981. "Financial Incentive Effects and Individual Decision-Making." California Institute of Technology Social Science Working Paper 401.
Holt, Charles A. 2006. *Markets, Games, and Strategic Behavior.* Boston: Pearson.
Hortala-Vallve, Rafael. 2012. "Qualitative Voting." *Journal of Theoretical Politics* 24 (4): 526–54.
Hurwicz, Leonid. 1972. "On Informally Decentralized Systems." In *Decision and Organization,* edited by C. B. McGuire and Roy Radner. Amsterdam: North-Holland.
Hylland, Aanund. 1980. "Strategy Proofness of Voting Procedures with Lotteries as Outcomes and Infinite Sets of Strategies." Unpublished.
Hylland, Aanund, and Richard J. Zeckhauser. 1979. "The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values." *Econometrica* 47 (6): 1321–36.
Jackson, Matthew O., and Hugo F. Sonnenschein. 2007. "Overcoming Incentive Constraints by Linking Decisions." *Econometrica* 75 (1): 241–57.
Karni, Edi. 1993. "Subjective Expected Utility Theory with State-Dependent Preferences." *Journal of Economic Theory* 60 (2): 428–38.
Karni, Edi. 2009. "A Mechanism for Eliciting Probabilities." *Econometrica* 77 (2): 603–6.
Le Breton, Michel, and Arunava Sen. 1999. "Separable Preferences, Strategyproofness, and Decomposability." *Econometrica* 67 (3): 605–28.
McCarthy, John. 1956. "Measures of the Value of Information." *Proceedings of the National Academy of Sciences* 42 (9): 654–55.
Mongin, Philippe. 1995. "Consistent Bayesian Aggregation." *Journal of Economic Theory* 66 (2): 313–51.
Nandeibam, Shasikanta. 2013. "The Structure of Decision Schemes with Cardinal Preferences." *Review of Economic Design* 17 (3): 205–38.
Reffgen, Alexander, and Lars-Gunnar Svensson. 2012. "Strategy-Proof Voting for Multiple Public Goods." *Theoretical Economics* 7 (3): 663–88.

Satterthwaite, Mark Allen. 1975. "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory* 10 (2): 187–217.

Savage, Leonard J. 1954. *The Foundations of Statistics.* New York: Wiley.

Savage, Leonard J. 1971. "Elicitation of Personal Probabilities and Expectations." *Journal of the American Statistical Association* 66 (336): 783–801.

Savage, Leonard J., and Robert Aumann. 1987. "Letters between Leonard Savage and Robert Aumann 1971." In *Essays on Economic Decisions under Uncertainty*, edited by Jacques Drèze, 76–81. Cambridge, UK: Cambridge University Press.