Aggregating Inconclusive Evidence*

Gabrielle Gayer
Bar Ilan University

Ehud Lehrer
Tel Aviv University

Dotan Persitz
Tel Aviv University

October 18, 2021

Abstract

A decision maker wishes to determine whether her prior probability is consistent with the beliefs of several advisors, that are given in the form of capacities. We provide a necessary and sufficient condition for the compatibility of such a prior probability with those capacities. The condition states that the expected value of any (positive) random variable according to the prior probability is greater or equal to the weighted average of the concave integral of that random variable with respect to the different capacities. We demonstrate the usefulness of this result in a setting where an administrator is required to form a frequency distribution based on several data sets that may include inconclusive observations.

Keywords: Cooperative Games, Aggregating Data Sets, Concave Integrals.
JEL Classification: C71, D81.

1 Motivation

Consider the problem of a health authority that must make a recommendation on the composition of viruses in the influenza vaccine. The recommendation is based on the health authority's forecast regarding the viruses that are most likely to spread in the upcoming season. There are several health centers in different regions that collect data on patients in an attempt to diagnose their viruses. Vaccines are known to vary in their effectiveness across seasons.[1] Naturally, the health authority is able to justify its recommendation if it is supported by the data collected by the health centers.

More generally, managers, both in the civil and in the private sectors, must often operate under uncertainty, bearing in mind that they may be held accountable for their decisions. Therefore, it is essential for them to be able to prove that the probability that underlined their decisions was based on all available information of all possible sources.

To study this problem, we consider an administrator who forms a probability over the possible states of nature. In addition, there is a group of practitioners who collect relevant information on the matter under consideration. The practitioners transfer their raw data to the administrator. The administrator will be able to establish that her probability is well-founded if it is supported by the information available to each of the practitioners.

*A previous version was titled "Aggregating Non-Additive Beliefs".

[1]The seasonal influenza vaccine is designed to protect against the three or four influenza viruses that are most likely to spread and cause illness during the upcoming flu season. Twice a year, the World Health Organization provides recommendations on the composition of the influenza vaccine (in February for the Northern Hemisphere's vaccine and in September for the Southern Hemisphere's vaccine). More than 100 national influenza centers in over 100 countries conduct year-round surveillance for influenza that involves receiving and testing thousands of influenza virus samples from patients and report their results to the World Health Organization. See, for example, Osterholm et al. (2012) for an account of the effectiveness of these vaccines.

Suppose that the practitioners' information is given in the form of data sets containing evidence about the states of nature that possibly occurred in each observation. In some observations it may be known which state of nature was realized, yet in others the outcome could be ambiguous. For example, a case wherein the physician can perfectly diagnose the patient's condition corresponds to a single virus, whereas the outcome of cases that are only partially diagnosed contains several possible viruses.

An observation is inconclusive if the practitioner cannot attribute it to a single state of nature, but only to a subset of states of nature, namely events. A data set with inconclusive observations induces a characteristic function that assigns each event with the number of times it was known to have occurred in the data set. That is, the event $\{\omega_1, \omega_2\}$ is assigned with the number of patients for whom the practitioner's diagnosis is inconclusive - she is able to narrow the set of possible viruses down to ω_1 and ω_2, but can not determine which one is the correct diagnosis.

A frequency in the core of the resulting cooperative game is a possible realization of the outcomes' distribution that is consistent with these data. If the data set includes no inconclusive evidence there will be a single frequency in the core. However, if the data set contains some inconclusive evidence then there will be several frequencies in the core, each resolving the ambiguity differently.

The administrator could argue that her probability is well-founded if the associated frequency can be decomposed into frequencies in the cores of the corresponding data set based cooperative games. The result stated in Proposition 2 provides a sufficient and necessary condition for the existence of such a justification for the administrator's probability. The condition can be interpreted as testing the consistency of the administrator's probability against the raw data in every weighted combination of events.

Jaffray (1991), Gonzales and Jaffray (1998) and Arad and Gayer (2012), had already pointed out that imprecise statistical data generate ambiguity that is incorporated into beliefs. In our setting, a practitioner's characteristic function, after applying the appropriate transformation, becomes a special case of non-additive probabilities, also known as capacities. Capacities allow individuals to express their perceived ambiguity in the problem at hand.[2] Lehrer (2009) introduced the concept of concave integral as a method of ambiguity adverse individuals to evaluate alternatives under uncertainty based on hard facts only (the events that are known to have occurred). Here, in order to justify her probability, the administrator needs to establish that her probability is supported by the non-additive probabilities of the practitioners insofar as it is a weighted average of probabilities in the cores of their capacities.

Our general result (Proposition 1) states that given a set of capacities on the same set of states of nature (each with a non-empty core), a prior probability can be represented as a weighted average of probabilities in the cores of the these capacities if and only if for any positive random variable Y, the average expected value of Y according to the concave integral across capacities is bounded from above by the expected value of Y with respect to the prior probability. The weights, which are fixed, can represent the experience, the quality, the political power, or the influence of the practitioners.[3] Proposition 2 is a special case of this result where the information of the practitioners is given in the form of data sets and the weights are set to be proportional to the quality of each practitioner's data set.

In Section 2 we introduce the necessary and sufficient condition for the aggregation of

[2]See Schmeidler (1989), for a characterization of a decision maker with non-additive beliefs.

[3]There is a large body of literature that studies the opposite problem of how to aggregate several prior probabilities into a single probability. Genest and Zidek (1986), Cooke et al. (1991) and Clemen and Winkler (1999), among others, review different aggregation methods.

capacities evaluated according to the concave integral. In Section 3 we present the data sets' setting and provide the result on the aggregation of data sets. In Section 4 we conclude. All proofs are relegated to the appendix.

2 Aggregating Concave Integrals

Preliminaries

Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature ($n \geqslant 2$) and let Σ be an algebra of subsets of Ω called events which is given by the power set 2^{Ω}. Capacities are functions $v : \Omega \to \mathbb{R}_+$ that satisfy (i) no empty events ($v(\emptyset) = 0$) (ii) finiteness ($v(\Omega)$ is finite) and (iii) monotonicity ($S \subseteq T \Rightarrow v(S) \leq v(T)$).[4]

Concave integrals are integrals over capacities used to evaluate acts in a setting with non-additive beliefs. The Concave integral was introduced by Lehrer (2009) (it was behaviorally characterized in Lehrer and Teper (2015) and later generalized in Even and Lehrer (2014)) to allow for the expression of ambiguity aversion even when capacities are not convex.[5]

The concave integral of a finite non-negative random variable Y over the capacity v is given by $\int^{cav} Y dv = \min_{f \in F_v}\{f(Y)\}$ where F_v is the set of all concave and homogeneous of degree one functions $f : \mathbb{R}^n_+ \to \mathbb{R}$ such that $\forall B \in \Sigma : f(\chi^B) \geq v(B)$ where $\chi^B \in \{0,1\}^n$ denotes the indicator vector of B ($\chi^B_i = 1$ if $i \in B$ and $\chi^B_i = 0$ otherwise).[6,7]

A decomposition of vector Y is $\alpha_Y : \Sigma \to \mathbb{R}_+$ such that $\sum_{B \in \Sigma} \alpha_Y(B)\chi^B = Y$. Denote the set of all decompositions of Y by $D(Y)$ and the optimal decomposition of Y relative to capacity v by $\alpha^\star_Y = \arg\max_{\alpha_Y \in D(Y)} \left\{ \sum_{B \in \Sigma} \alpha_Y(B)v(B) \right\}$.[8] Lemma 1(i) in Lehrer (2009) states that $\int^{cav} Y dv = \sum_{B \in \Sigma} \alpha^\star_Y(B)v(B)$, namely, that the concave integral can be expressed as a linear combination of the capacities where the weights are the corresponding optimal decomposition elements.

A capacity v induces the cooperative game $G = (\Omega; v)$. We denote the core of the cooperative game $G = (\Omega; v)$ by $C(v)$ and the vector of length n that all its elements are ones (zeros) by 1_n (0_n). By the definition of concave integrals $v(\Omega) \leq \int^{cav} 1_n dv$. Hence, since $D(1_n)$ is the set of all balancing weights, by the Shapley-Bondareva Theorem (Bondareva (1963) and Shapley (1967)),

Remark 1. *Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature and let $\Sigma = 2^{\Omega}$. Let v be a capacity on Ω.*

1. *$C(v)$ is non-empty if and only if $v(\Omega) = \int^{cav} 1_n dv$.*
2. *$C(v)$ is empty if and only if $v(\Omega) < \int^{cav} 1_n dv$.*

[4] When more convenient, we slightly abuse notation by treating v as a length 2^n vector.

[5] A capacity v is convex if $v(S) + v(T) \leq v(S \cup T) + v(S \cap T)$ for every two events S and T. In the Choquet expected utility model (Schmeidler (1989)) ambiguity aversion corresponds to convex capacities, for which the concave integral and the Choquet integral imply the same preferences over random variables (Lehrer (2009)).

[6] Note that this definition does not require v to be monotonic. Since the data sets in our framework may induce non-monotonic characteristic functions we prove that the concave integral can also operate on non-monotonic capacities.

[7] We use χ to denote the indicator matrix where the columns are the 2^n indicator vectors. Also, all vectors are defined to be column vectors. Row vectors are denoted by the superscript 'T'.

[8] For the non-emptyness of $D(Y)$ consider α such that $\forall i \in \{1, \ldots, n\} : \alpha(\{\omega_i\}) = Y_i$ and for every non-singleton $B \in \Sigma$, $\alpha(B) = 0$. $\alpha \in D(Y)$ for every Y.

Aggregation of Concave Integrals

We now extend the framework to allow for several capacities. Let, $\mathscr{V} = \{v_1, \ldots, v_m\}$, be a set of m capacities on Ω and denote $V(\Omega) = \sum_{j=1}^{m} v_j(\Omega)$. These capacities represent the (non-normalized) beliefs of m advisors who evaluate the expectation of random variables according to the concave integral. An m-Multi-Game $\bar{G}$ is the pair $\bar{G} = (\Omega; \mathscr{V})$. We denote the single cooperative game that is defined by the j^{th} characteristic function of the Multi-Game $\bar{G}$ by $\bar{G}_j = (\Omega; v_j)$.

A decision maker wishes to establish that the probability that underlies her decisions is admissible. Let $X \in \mathbb{R}^n_+$, such that $\sum_{i=1}^{n} X_i = \sum_{j=1}^{m} v_j(\Omega)$, be the decision maker's non-normalized probability also known as a charge. We say that the decision maker's probability belongs to the core of a multi-game induced by $\mathscr{V}$ ($X \in C(\bar{G})$) if there are m finite non-negative vectors $X^1, ..., X^m$ such that $\forall j : X^j \in C(\bar{G}_j)$ and $\sum_{j=1}^{m} X^j = X$ (see Section 4 in Gayer and Persitz (2016) for a discussion on this solution concept). In this case the decision maker could claim that her probability is well-founded as it is supported by the beliefs of the advisors. Proposition 1 is a novel result on the aggregation of concave integrals that presents a condition that can help the decision maker to corroborate her claim.

Proposition 1. *Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature and let $\Sigma = 2^{\Omega}$. Let $\mathscr{V} = \{v_1, \ldots, v_m\}$ be a set of m capacities on Ω. Let $X \in \mathbb{R}^n_+$ be a non-normalized probability on Ω for the set $\mathscr{V}$. $X \in C(\bar{G})$ if and only if every random variable $Y \in R^n_+ : \sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$*

To sketch the proof, if X is in the core of the multi-game induced by $\mathscr{V}$ it is a sum of members of the cores of each game in $\mathscr{V}$ ($X = \sum_{j=1}^{m} X^j$ where $\forall j : X^j \in C(\bar{G}_j)$). We first show for each capacity $v_j \in \mathscr{V}$ that the expectation of any $Y \in R^n_+$ according to a vector in the core of that game is larger or equal to its expectation according to the respective concave integral. Then summing over all capacities entails that the expectation of Y according to X must be larger or equal to the sum of the concave integrals of Y over all $v_j \in \mathscr{V}$.

If X in not in the core of the multi-game induced by $\mathscr{V}$ we construct Ys that do not satisfy $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$. In case the core of the multi-game induced by $\mathscr{V}$ is non-empty, a hyperplane separation theorem guarantees the existence of a vector Z that separates X and $C(\bar{G})$. We use this vector to construct a sequence of vectors, $Z^c = \frac{1}{c} \times [Z + c]$, that goes to 1_n when c goes to infinity. Then we show that $\sum_{j=1}^{m} \int^{cav} Z^c dv_j = \sum_{j=1}^{m} \min_{w^j \in C(\bar{G}_j)} \{ w^{j^T} \cdot Z^c \} > X^T \cdot Z^c$. In case the core of the multi-game induced by $\mathscr{V}$ is empty, Remark 1 implies that $\sum_{j=1}^{m} \int^{cav} 1_n dv_j > X^T \cdot 1_n$.

To better understand Proposition 1, note that the decision maker's non-normalized probability X can be normalized by $V(\Omega)$ to become a probability, $X^\star = \frac{1}{V(\Omega)} X$, and the advisors' capacities can be normalized by $v_j(\Omega)$ to become their non-additive beliefs, $v^\star_j = \frac{v_j}{v_j(\Omega)}$. In this case Proposition 1 can be restated using the above terminology: There exist m vectors $X^{\star j} \in C(\bar{G}^\star_j)$ such that $X^\star = \alpha_j X^{\star j}$ where $\alpha_j = \frac{v_j(\Omega)}{V(\Omega)}$ if and only if for every random variable $Y \in R^n_+ : \sum_{j=1}^{m} \alpha_j \int^{cav} Y dv^*_j \leq X^{\star T} \cdot Y$. Thus, the decision maker can prove that her probability is a weighted average of probabilities in the respective cores of the advisors' non-additive beliefs if and only if her evaluation of any random variable Y ($X^{\star T} \cdot Y$) is higher or equal to the weighted average of those of the advisors ($\int^{cav} Y dv^*_j$).

Even and Lehrer (2014) showed that the expectation of Y according to the concave integral is (weakly) higher than that according to the Choquet integral. Hence, if the advisors were to use the Choquet integral to evaluate random variables instead of the concave integral, showing that the decision maker's evaluation of Y is higher than the weighted average of the advisors' evaluations would be insufficient to prove that the decision maker's probability is supported by the non-additive beliefs of the advisors.

Finally, a technically useful implication of Proposition 1 is that it provides an upper bound on the sum of concave integrals in case the core of the multi-game is non-empty. That is, for every random variable $Y \in R^n_+ : \sum_{j=1}^{m} \int^{cav} Y dv_j \leq \min_{X \in C(\bar{G})} X^T \cdot Y$.

3 Aggregating Datasets

A Single Data Set

A data set is a sequence of T observations, indexed by $i \in \{1,...,T\}$, denoted by $D = (B_1,...,B_T)$ where $B_i \in \Sigma \backslash \{\varnothing, \Omega\}$.[9] The event B_i represents the set of all states that may have occurred in observation i. When the state of nature that occurred in observation i is clear, B_i is a singleton. However, some observations may be assigned to non-singleton events when it is not clear which specific state of nature within that event had actually occurred. Following the example of health centers, in certain cases it may be known that a patient was infected with a type C virus (ruling out other types), but it is unknown which sub type infected the patient.

We assume that the data set is cross-sectional, meaning that the order of observations does not affect the inference. Therefore, we can also describe the data set in a characteristic function form. A *data set in a disaggregated characteristic function form* is the cooperative game $G^V = (\Omega; V)$ where $V : \Sigma \to \mathbb{N}$ be a function such that (i) For every event $B \subset \Omega$, $V(B)$ is the number of occurrences of B in data set D and (ii) $V(\Omega) = \arg\min\{[0,T] | C(G^V) \neq \emptyset\}$, so that $V(\Omega)$ is the lowest value that ensues a non-empty core.[10] Note that by Remark 1.1 and (ii): $V(\Omega) = \int^{cav} 1_n dV \leq T$.

Compatible Frequency Distributions

We say that $\bar{X}$ is compatible with the data set in a disaggregated characteristic function form G^V if $\bar{X} \in C(G^V)$. A frequency distribution of T observations that is compatible with data set D is a vector $X \in \mathbb{R}^n_+$ such that $\sum_{i=1}^n X_i = T$ and $\forall i \in N : x_i \geq \bar{x}_i$.

That is, X does not understate the number of observations assigned to event B according to V ($\forall B \subseteq \Omega : \sum_{\omega_i \in B} X_i \geq V(B)$). Note, however, that weak compatibility does not require the frequency distribution to resolve inconclusive observations "consistently". Consider a data set of 3 observations on 3 states of nature where $V(\{1\}) = 1, V(\{2\}) = 1, V(\{3\}) = 0, V(\{1,2\}) = 1$, $V(\{1,3\}) = 0$ and $V(\{2,3\}) = 0$. While the frequency distribution $X = (1,1,1)$ is weakly compatible with the data set, $V(\{1\}) + V(\{2\}) + V(\{1,2\}) > X_1 + X_2$. The total frequency attributed to states 1 and 2 understates the total number of observations assigned to the relevant events - $\{1\}$, $\{2\}$ and $\{1,2\}$. The frequency attributed to state 3 is erroneously overstated since no observation is assigned to an event that includes state 3.

A frequency distribution X is strongly compatible with data set D if $X \in C(G^U)$. That is,

[9] Excluding $B = \varnothing$ implies that an event that is known to have occurred cannot be empty. Excluding $B = \Omega$ implies that we ignore cases that add no information.

[10] Since $T = \sum_{S \subset N} V(S)$ the set $\{[0,T] | C(G^V) \neq \emptyset\}$ is non empty.

X does not understate the number of observations assigned to all events that are subsets of B. In fact, inconclusive observations are resolved "consistently" by assigning each observation to one of the states in the corresponding event (or to a mix of states in this event). Continuing the example above, the corresponding aggregated characteristic function is $U(\{1\}) = 1$, $U(\{2\}) = 1$, $U(\{3\}) = 0$, $U(\{1,2\}) = 3$, $U(\{1,3\}) = 1$ and $U(\{2,3\}) = 1$. The frequency distribution $X = (1,1,1)$ is not strongly compatible with the data set, yet, $Y = (1.5,1.5,0)$ is strongly compatible with it since the inconclusive observation of the event $\{1,2\}$ is attributed to both state $\{1\}$ and state $\{2\}$ equally.

Multiple Data Sets

Let $\mathscr{D} = \{D_1, D_2, \ldots, D_m\}$ be a collection of m data sets with $T = \sum_{i=1}^{m} T_i$ observations. Let $\bar{G}^{\mathscr{V}} = (\Omega; \mathscr{V})$ be the disaggregated multi game where $\mathscr{V} = \{V_1, V_2, \ldots, V_m\}$ and V_i is a disaggregated characteristic function that corresponds to data set D_i. We refer to $C(\bar{G}^{\mathscr{V}})$ as the weakly compatible core of $\mathscr{D}$. Similarly, we define the aggregated multi game $\bar{G}^{\mathscr{U}} = (\Omega; \mathscr{U})$ where $\mathscr{U} = \{U_1, U_2, \ldots, U_m\}$ and U_i is an aggregated characteristic function that corresponds to data set D_i. We refer to $C(\bar{G}^{\mathscr{U}})$ as the strongly compatible core of $\mathscr{D}$.

Systems of Weights

We generalize the balancing weights of Bondareva (1963) and Shapley (1967) to account for systems of weights whose total weights may differ across states of nature. Let $F : \Sigma \to \mathbb{R}_+$ be a system of weights. The vector of weights induced by F is denoted by $W^F = \sum_{B \in \Sigma} F(B) \chi^B$. We say that F_1 and F_2 are W-equivalent if $W^{F_1} = W^{F_2}$. The W-equivalence relation induces a partition on the set of all systems of weights.[11] Let us denote the set of all W-equivalence classes by Γ. For every class $\gamma \in \Gamma$, the maximal γ-weighted sum of a data set in a disaggregated characteristic function form V is $T_V^{W^\gamma} \equiv \max\limits_{F \in \gamma} \sum\limits_{B \in \Sigma} F(B)V(B)$ and the maximal γ-weighted sum of a data set in an aggregated characteristic function form U is $T_U^{W^\gamma} \equiv \max\limits_{F \in \gamma} \sum\limits_{B \in \Sigma} F(B)U(B)$.

Compatibility Conditions

We provide a necessary and sufficient condition for the decomposition of an aggregate frequency distribution into m frequency distributions such that the first is weakly compatible with the first data set, the second is weakly compatible with the second data set and so on.[12]

Proposition 2. *Let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be a finite set of states of nature and let $\Sigma = 2^\Omega$. Let $X \in \mathbb{R}^n_+$ be an aggregate frequency distribution vector for the set of m data sets $\mathscr{D}$. $X \in C(\bar{G}^{\mathscr{V}})$ if and only if every class $\gamma \in \Gamma$ satisfies $\sum\limits_{V_j \in \mathscr{V}} T_{V_j}^{W^\gamma} \leq \sum\limits_{i \in \{1,\ldots,n\}} W_i^\gamma X_i$.*

Proposition 2 is almost a direct application of Proposition 1. The only difference is that Proposition 1 requires $\mathscr{V}$ to be a set of capacities while in Proposition 2 $\mathscr{V}$ is a set of data sets which may be non-monotonic. In the proof we show that replacing the data sets by their monotonic covers, generates capacities with the same cores for which the concave integrals are well defined. In addition, since strong compatibility implies weak compatibility (by Lemma 1),

[11] The set of all the balancing weights of Bondareva (1963) and Shapley (1967) is identical to the class of functions F such that W^F is the vector of ones.

[12] This result was used to prove Proposition 3 in Gayer and Persitz (2016) (see p. 948).

the condition stated in Proposition 2 is a necessary (though insufficient) condition for strong compatibility.

To demonstrate how Proposition 2 can be utilized to determine whether a frequency distribution is compatible with the available data sets, consider the following example with two practitioners and three states of nature. Practitioner 1's data set contains 3 observations that are all inconclusive, each containing a pair of states - the first observation includes states 1 and 2, the second observation includes states 1 and 3 and the third includes states 2 and 3. Practitioner 2's data set contains two observations, where the first observation is conclusive, containing only state 1, while the second observation is inconclusive including states 2 and 3.[13] An administrator's frequency distribution that assigns one observation to state 1 and four observations to state 3 ($X = (1,0,4)$) can be falsified with the help of Proposition 2. To see this take the class γ such that $w^\gamma = (1,1,0)$. First, $T_{v_1}^{w^\gamma} \geq 1$ on account of the system of weights F such that $F(\{1,2\}) = 1$ and $F(B) = 0$ otherwise, $\sum_{B\in\Sigma} F(B)v_1(B) = 1$. Moreover, $T_{v_2}^{w^\gamma} \geq 1$ since for a system of weights F such that $F(\{1\}) = F(\{2\}) = 1$ and $F(B) = 0$ otherwise, $\sum_{B\in\Sigma} F(B)v_2(B) = 1$. However, $w^\gamma \cdot (1,0,4) = 1 < T_{v_1}^{w^\gamma} + T_{v_2}^{w^\gamma}$, thus, by Proposition 2, the administrator's frequency distribution if found to be incompatible with the data sets collected by the practitioners.

The data sets that were considered in Proposition 2 were given in a disaggregated characteristic function form. Proposition 2 presents a condition that if satisfied provides support to the claim that a frequency distribution is weakly compatible with the available data sets. One can apply an adequate version of Proposition 2 to $\mathscr{U}$ to understand if a frequency distribution is strongly compatible with the available data sets. Alternatively, recall that Lemma 1 states that cooperative games induced by data sets in an aggregated characteristic function form are convex. Therefore, by Dragan et al. (1989) (see also Footnote 18 in Gayer and Persitz (2016)), $C(\bar{G}^{\mathscr{U}}) = C(\sum_{U\in\mathscr{U}} G^U)$. Therefore, strong compatibility can be established by verifying that a frequency distribution is in the core of $C(\sum_{U\in\mathscr{U}} G^U)$.

Aggregated Characteristic Function Form

It will also be useful to describe the data set in a corresponding aggregated form. Let $U : \Sigma \to \mathbb{N}$ be a function such that $\forall B \subseteq \Omega : U(B) = \sum_{b\subseteq B} V(b)$. That is, $U(B)$ is the total number of occurrences of B and its subsets in data set D. A *data set in an aggregated characteristic function form* is the cooperative game $G^U = (\Omega; U)$.[14]

Lemma 1. *Let V be a data set in a disaggregated characteristic function form and let U be the corresponding data set in an aggregated characteristic function form.*

1. *G^U is a convex cooperative game.*

2. *$C(G^U)$ is non empty.*

3. *$C(G^U) \subseteq C(G^V)$.*

4 Concluding Remarks

We provide a necessary and sufficient condition for the compatibility of a non-normalized probability distribution with a set of given capacities. We demonstrate the usefulness of this

[13]Formally, $v_1(\{1\}) = v_1(\{2\}) = v_1(\{3\}) = 0$, $v_1(\{1,2\}) = v_1(\{1,3\}) = v_1(\{2,3\}) = 1$, $v_1\{1,2,3\} = 3$, $v_2(\{1\}) = 1$, $v_2(\{2\}) = v_2(\{3\}) = 0$, $v_2(\{1,2\}) = v_2(\{1,3\}) = 0$, $v_2(\{2,3\}) = 1$ and $v_2\{1,2,3\} = 2$.

[14]Note that formally V is the Möbius transform of U.

result, by applying it to a setting where an administrator is required to form a frequency distribution based on several data sets that may include inconclusive observations.

References

Arad, Ayala and Gabrielle Gayer (2012) "Imprecise data sets as a source of ambiguity: A model and experimental evidence," *Management Science*, Vol. 58, No. 1, pp. 188–202.

Bondareva, Olga N. (1963) "Some applications of linear programming methods to the theory of cooperative games," *Problemy Kybernetiki*, Vol. 10, pp. 119–139. (in Russian).

Clemen, Robert T and Robert L Winkler (1999) "Combining probability distributions from experts in risk analysis," *Risk analysis*, Vol. 19, No. 2, pp. 187–203.

Cooke, Roger et al. (1991) *Experts in uncertainty: opinion and subjective probability in science*: Oxford University Press on Demand.

Dragan, Irinel, Jos Potters, and Stef Tijs (1989) "Superadditivity for Solutions of Coalitional Games," *Libertas Mathematica*, Vol. 9, pp. 101–110.

Even, Yaarit and Ehud Lehrer (2014) "Decomposition-integral: unifying Choquet and the concave integrals," *Economic Theory*, Vol. 56, No. 1, pp. 33–58.

Gayer, Gabrielle and Dotan Persitz (2016) "Negotiation across multiple issues," *Theoretical Economics*, Vol. 11, No. 3, pp. 937–969.

Genest, Christian and James V Zidek (1986) "Combining probability distributions: A critique and an annotated bibliography," *Statistical Science*, Vol. 1, No. 1, pp. 114–135.

Gonzales, Christophe and Jean-Yves Jaffray (1998) "Imprecise sampling and direct decision making," *Annals of Operations Research*, Vol. 80, pp. 207–235.

Jaffray, Jean-Yves (1991) "Belief functions, convex capacities, and decision making," in *Mathematical Psychology*: Springer, pp. 127–134.

Lehrer, Ehud (2009) "A new integral for capacities," *Economic Theory*, Vol. 39, No. 1, pp. 157–176.

Lehrer, Ehud and Roee Teper (2015) "Subjective independence and concave expected utility," *Journal of Economic Theory*, Vol. 158, pp. 33–53.

Osterholm, Michael T, Nicholas S Kelley, Alfred Sommer, and Edward A Belongia (2012) "Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis," *The Lancet infectious diseases*, Vol. 12, No. 1, pp. 36–44.

Schmeidler, David (1989) "Subjective probability and expected utility without additivity," *Econometrica: Journal of the Econometric Society*, pp. 571–587.

Shapley, Lloyd S. (1967) "On balanced sets and cores," *Naval Research Logistics Quarterly*, Vol. 14, pp. 453–460.

——— (1971/72) "Cores of convex games," *International Journal of Game Theory*, Vol. 1, pp. 11–26.

Appendix

The following results are used in the proof of Proposition 1.

Lemma 2

Lemma 2. *Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature and let $\Sigma = 2^{\Omega}$. Let v be a capacity on Ω and let Y be a finite non-negative random variable on Ω. Denote $\hat{H} = \{h \in \mathbb{R}^n_+ | \forall B \in \Sigma : \sum_{\omega_i \in B} h_i \geq v(B)\}$ and the set of its extreme points by H.*[15] *Then,*

$$\int^{cav} Y dv = \min_{h \in H} h^T \cdot Y.$$

Proof. By the definitions of concave integral and optimal decomposition,

$$\int^{cav} Y dv = \max_{\alpha : \Sigma \to \mathbb{R}_+} \left\{ \sum_{B \in \Sigma} \alpha(B) v(B) \middle| \sum_{B \in \Sigma} \alpha(B) \chi^B = Y, \forall B \in \Sigma : \alpha(B) \geq 0 \right\}$$

Since v and Y are finite and since $D(Y)$ is non-empty, there is a solution to the maximization problem. Therefore, by the general strong duality theorem, the dual has the same solution. Hence,

$$\int^{cav} Y dv = \min_{h \in \mathbb{R}^n_+} \left\{ h^T \cdot X | \forall B \in \Sigma : \sum_{\omega_i \in B} h_i \geq v(B) \right\} = \min_{h \in \hat{H}} h^T \cdot Y$$

$\hat{H}$ is non empty[16] and convex.[17] Since $h^T \cdot Y$ is a linear function of h and since $\hat{H}$ is convex, the minimum of $h^T \cdot Y$ is achieved on the extreme points of $\hat{H}$. Thus, $\int^{cav} Y dv = \min_{h \in H} h^T \cdot Y$. □

Lemma 3

Lemma 3. *Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature and let $\Sigma = 2^{\Omega}$. Let v be a capacity on Ω. If $C(v)$ is non empty there is a neighborhood U of 1_n such that every non-negative random variable $Y \in U$ on Ω satisfies*

$$\int^{cav} Y dv = \min_{c \in C(v)} c^T \cdot Y.$$

Proof. First note that $C(v) = \{h \in \hat{H} | \sum_{\omega_i \in \Omega} h_i = v(\Omega)\} \neq \emptyset$. Therefore, $\min_{c \in C(v)} c^T \cdot Y \geq \min_{h \in \hat{H}} h^T \cdot Y$. Moreover, by the proof of Lemma 2, $\min_{c \in C(v)} c^T \cdot X \geq \min_{h \in H} h^T \cdot Y = \int^{cav} Y dv$.

Suppose, to the contrary, that there is a sequence Y_t that converges to 1_n and each element satisfies $\min_{c \in C(v)} c^T \cdot Y_t > \int^{cav} Y_t dv$.

[15] $h \in \hat{H}$ is an extreme point of $\hat{H}$ if there are no $\tilde{h}, \tilde{\tilde{h}} \in \hat{H}$ and $\lambda \in (0,1)$ such that $h = \lambda \tilde{h} + (1-\lambda)\tilde{\tilde{h}}$.

[16] For example, by the monotonicity of capacities, $(v(\Omega), \ldots, v(\Omega))' \in \hat{H}$.

[17] $h, \bar{h} \in \hat{H}$ implies that $\chi' h \geq v$ and $\chi' \bar{h} \geq v$ and therefore for every $\lambda \in [0,1]$:

$$\chi'(\lambda h + (1-\lambda)\bar{h}) = \lambda \chi' h + (1-\lambda)\chi' \bar{h} \geq \lambda v + (1-\lambda) v = v$$

Since by Lemma 2, for every t, $\min_{h \in H} h^T \cdot Y_t = \int^{cav} Y_t dv$ it must be that for every t, $h^t \in H \backslash C(v)$ where $h^t = \arg\min_{h \in H} h^T \cdot Y_t$. In particular, since $h^t \in H \backslash C(v)$ then $h^{t^T} \cdot 1_n > v(\Omega)$.

Let us consider the sequence $\int^{cav} Y_t dv$. Since (i) H is finite (ii) The elements of H are finite (iii) Y_t is a sequence of finite elements and (iv) $\int^{cav} Y_t dv = \min_{h \in H} h^T \cdot Y_t$, the sequence $\int^{cav} Y_t dv$ is bounded.

Thus, $\int^{cav} Y_t dv$ has a convergent subsequence $\int^{cav} Y_s dv$. The limit of this subsequence is $\lim_{s \to \infty} \int^{cav} Y_s dv = \int^{cav} \lim_{s \to \infty} Y_s dv = \int^{cav} 1_n dv = v(\Omega)$, the last equality is due to $C(v)$ being non empty and Remark 1.1.

Since $\int^{cav} Y_s dv$ converges, every of its subsequences is also convergent, and to the same limit. Since H is finite, at least one such subsequence is $\int^{cav} Y_r dv$ such that Y_r converges to 1_n and all its elements correspond to the same h^r. For this subsequence $\lim_{r \to \infty} \int^{cav} Y_r dv = \lim_{r \to \infty} h^{r^T} \cdot Y_r = h^{r^T} \cdot \{\lim_{r \to \infty} Y_r\} = h^{r^T} \cdot 1_n$. Hence, $h^{r^T} \cdot 1_n = v(\Omega)$. Contradiction.

Hence, there is a neighborhood U of 1_n such that every non-negative random variable $Y \in U$ satisfies $\int^{cav} X dv = \min_{c \in C(v)} c^T \cdot Y$. □

Lemma 4

Lemma 4. *Let $\bar{G}$ be an m-Multi-Game, $\bar{G} = (\Omega; \mathscr{V})$. Then, $C(\bar{G})$ is a closed and convex set.*

Proof. For every $v_j \in \mathscr{V}$, $C(v_j)$ is compact since (i) A set of vectors that satisfies a set of weak linear inequalities is closed (recall that the empty set is closed) and (ii) A set of non-negative vectors that satisfy efficiency is bounded (recall that the capacities are non-negative). Since $C(\bar{G})$ is the sum of compact individual cores, it is also compact. Thus, $C(\bar{G})$ is a closed set.

To show that $C(\bar{G})$ is a convex set, let $Z, \hat{Z} \in C(\bar{G})$. First, for every $\lambda \in [0,1]$ we get that $\lambda Z + (1-\lambda)\hat{Z}$ is a non normalized probability vector for the set $\mathscr{V}$ since

$$\sum_{i=1}^{n} (\lambda Z + (1-\lambda)\hat{Z})_i = \sum_{i=1}^{n} \lambda Z_i + (1-\lambda)\hat{Z}_i = \lambda \sum_{i=1}^{n} Z_i + (1-\lambda) \sum_{i=1}^{n} \hat{Z}_i$$

$$= \lambda \sum_{j=1}^{m} v_j(\Omega) + (1-\lambda) \sum_{j=1}^{m} v_j(\Omega) = \sum_{j=1}^{m} v_j(\Omega)$$

In addition, since $Z, \hat{Z} \in C(\bar{G})$ there exist $2m$ vectors $Z^1, \ldots, Z^m$ and $\hat{Z}^1, \ldots, \hat{Z}^m$ such that $\forall j$: $Z^j \in C(v_j)$, $\hat{Z}^j \in C(v_j)$ and $\sum_{j=1}^{m} Z^j = Z$ and $\sum_{j=1}^{m} \hat{Z}^j = \hat{Z}$. By the convexity of the core of a single game $\forall \lambda \in [0,1]$, $\forall j$: $\lambda Z^j + (1-\lambda)\hat{Z}^j \in C(v_j)$. These vectors sum to $\lambda Z + (1-\lambda)\hat{Z}$ since,

$$\sum_{j=1}^{m} \lambda Z^j + (1-\lambda)\hat{Z}^j = \lambda \sum_{j=1}^{m} Z^j + (1-\lambda) \sum_{j=1}^{m} \hat{Z}^j = \lambda Z + (1-\lambda)\hat{Z}$$

Hence, $\lambda Z + (1-\lambda)\hat{Z} \in C(\bar{G})$. Thus, $C(\bar{G})$ is a convex set. □

Proof of Proposition 1

Proof. First suppose that $X \in C(\bar{G})$. Then, X is a non normalized probability vector on Ω for the set $\mathscr{V}$ and there are m finite non-negative random variables $X^1, \ldots, X^m$ on Ω such that $\forall v_j \in \mathscr{V} : X^j \in C(v_j)$ and $\sum_{j=1}^{m} X^j = X$.

Recall that for every $Y \in R_+^n$, $D(Y)$ denotes the non-empty set of all decompositions. For every random variable $Y \in R^n_+$, for every decomposition $\alpha_Y \in D(Y)$ and for every capacity $v_j \in \mathscr{V}$ we get

$$X^{j^T} \cdot Y = X^{j^T} \cdot \left[\sum_{B \in \Sigma} \left[\alpha_Y(B) \times \chi^B \right] \right] = \sum_{B \in \Sigma} \left[X^{j^T} \cdot \left[\alpha_Y(B) \times \chi^B \right] \right] =$$

$$\sum_{B \in \Sigma} \left[\alpha_Y(B) \times \left[X^{j^T} \cdot \chi^B \right] \right] = \sum_{B \in \Sigma} \left[\alpha_Y(B) \times \sum_{\omega_i \in B} X_i^j \right] \geq \sum_{B \in \Sigma} \left[\alpha_Y(B) \times v_j(B) \right]$$

Where the first equality is by the definition of a decomposition and the final inequality is true since $\forall v_j \in \mathscr{V} : X^j \in C(v_j)$ implies that $\forall v_j \in \mathscr{V}, \forall B \in \Sigma : \sum_{\omega_i \in B} X_i^j \geq v_j(B)$.

In particular, for every random variable $Y \in R^n_+$ and for every capacity $v_j \in \mathscr{V}$, $X^{j^T} \cdot Y \geq \sum_{B \in \Sigma} \left[\alpha_Y^\star(B) \times v_j(B) \right]$. Hence, by Lemma 1(i) in Lehrer (2009), for every random variable $Y \in R^n_+$ and for every capacity $v_j \in \mathscr{V}$, $X^{j^T} \cdot Y \geq \int^{cav} Y dv_j$. Summing over all capacities, for every random variable $Y \in R^n_+$ we get $X^T \cdot Y = \sum\limits_{j=1}^m X^{j^T} \cdot Y = \sum\limits_{j=1}^m \left[X^{j^T} \cdot Y \right] \geq \sum\limits_{j=1}^m \int^{cav} Y dv_j$. Thus, $X \in C(\bar{G})$ implies that for every random variable $Y \in R^n_+$, $X^T \cdot Y \geq \sum_{j=1}^m \int^{cav} Y dv_j$.

Next suppose $X \notin C(\bar{G})$. Let us first attend to the case where $C(\bar{G})$ is non-empty.

Since $C(\bar{G})$ is closed and convex (by Lemma 4) and since a singleton is closed and convex, the separating hyperplane theorem guarantees that there is a vector $Z = (Z_1, ..., Z_n) \neq 0_n$ that separates X and $C(\bar{G})$. That is, there exists $Z \neq 0_n$ such that for every $w \in C(\bar{G})$, $X^T \cdot Z < w^T \cdot Z$. Thus, there exists $Z \neq 0_n$ such that $X^T \cdot Z < \min_{w \in C(\bar{G})} \{ w^T \cdot Z \}$.

For a positive constant c denote by Z^c the vector that has $Z_i^c = \frac{Z_i + c}{c}$ as a representative element. X and every member of $C(\bar{G})$ are non normalized probability vectors on Ω for the set $\mathscr{V}$. Therefore, for every $w \in C(\bar{G})$, $\sum_{i=1}^n X_i = \sum_{i=1}^n w_i = \sum_{j=1}^m v_j(\Omega)$. Hence,

$$X^T \cdot Z^c = \frac{1}{c} \times (X^T \cdot Z) + (X^T \cdot 1_n) = \frac{1}{c} \times (X^T \cdot Z) + (w^T \cdot 1_n) < \frac{1}{c} \times \min_{w \in C(\bar{G})} \{ w^T \cdot Z \} + (w^T \cdot 1_n) =$$

$$\min_{w \in C(\bar{G})} \{ \frac{1}{c} \times (w^T \cdot Z) \} + (w^T \cdot 1_n) = \min_{w \in C(\bar{G})} \{ \frac{1}{c} \times (w^T \cdot Z) + (w^T \cdot 1_n) \} = \min_{w \in C(\bar{G})} \{ w^T \cdot Z^c \}$$

Thus, for every positive constant c and for every $w \in C(\bar{G})$ we get $X^T \cdot Z^c < \min_{w \in C(\bar{G})} \{ w^T \cdot Z^c \}$.

Denote $w^\star = \arg\min_{w \in C(\bar{G})} \{ w^T \cdot Z^c \}$. Since $w^\star \in C(\bar{G})$ there exist $w^{1\star}, ..., w^{m\star}$ such that $\forall v_j \in \mathscr{V} : w^{j\star} \in C(v_j)$ and $\sum_{j=1}^m w^{j\star} = w^\star$. Moreover, $\forall v_j \in \mathscr{V} : w^{j\star} \in \arg\min_{w^j \in C(v_j)} \{ w^{j^T} \cdot Z^c \}$.[18] Therefore, for every positive constant c, $X^T \cdot Z^c < \sum_{j=1}^m \min_{w^j \in C(v_j)} \{ w^{j^T} \cdot Z^c \}$.

For every capacity $v_j \in \mathscr{V}$, let U_j be the neighborhood of 1_n that satisfies Lemma 3. That is, $\int^{cav} Y dv_j = \min_{c^j \in C(v_j)} \{ c^{j^T} \cdot Y \}$ for every non negative random variable $Y \in U_j$. Let $U = \cap_j U_j$. Therefore, $\int^{cav} Y dv_j = \min_{c^j \in C(v_j)} \{ c^{j^T} \cdot Y \}$ for every $v_j \in \mathscr{V}$ and every non negative

[18] To see that, suppose that $\exists v_j \in \mathscr{V}$ such that $w^{j\star} \in C(v_j)$ but $w^{j\star} \notin \arg\min_{w^j \in C(v_j)} \{ w^{j^T} \cdot Z^c \}$ while $w^{j\star\star} \in \arg\min_{w^j \in C(v_j)} \{ w^{j^T} \cdot Z^c \}$. Then,

$$w^{1\star^T} \cdot Z^c + \cdots + w^{j\star\star^T} \cdot Z^c + \cdots + w^{m\star^T} \cdot Z^c < w^{1\star^T} \cdot Z^c + \cdots + w^{j\star^T} \cdot Z^c + \cdots + w^{m\star^T} \cdot Z^c$$

Denote $\bar{w} = w^{1\star} + \cdots + w^{j\star\star} + \cdots + w^{m\star}$. Then $\bar{w} \in C(\bar{G})$ and $\bar{w}^T \cdot Z^c < w^{\star^T} \cdot Z^c$ in contradiction to $w^\star = \arg\min_{w \in C(\bar{G})} \{ w \cdot Z^c \}$. Hence, $\forall v_j \in \mathscr{V} : w^{j\star} \in \arg\min_{w^j \in C(v_j)} \{ w^j \cdot Z^c \}$.

random variable $Y \in U$. As a consequence, for every non negative random variable $Y \in U$, $\sum_{j=1}^{m} \int^{cav} Y dv_j = \sum_{j=1}^{m} \min_{c^j \in C(\bar{G}_j)} \{c^{j^T} \cdot Y\}$.

Note that (i) Z^c goes to 1_n when c goes to infinity; (ii) Z^c is non-negative for large enough c and (iii) the w^js are the minimizers of $\min_{c^j \in C(\bar{G}_j)} \{c^{j^T} \cdot Z^c\}$. Let c be large enough so that $Z^c \in U \cap \mathbb{R}_+^n$. Hence,

$$\sum_{j=1}^{m} \int^{cav} Z^c dv_j = \sum_{j=1}^{m} \min_{w^j \in C(v_j)} \{w^{j^T} \cdot Z^c\} > X^T \cdot Z^c$$

Thus, if $C(\bar{G})$ is non-empty, $X \notin C(\bar{G})$ implies that there exists $Y \in \mathbb{R}_+^n$ that does not satisfy $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$. That is, if $C(\bar{G})$ is non-empty and every $Y \in \mathbb{R}_+^n$ satisfies $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$ then $X \in C(\bar{G})$.

Finally, we attend to the case where $X \notin C(\bar{G})$ and $C(\bar{G})$ is empty. Consider $Y = 1_n$. Thus, $X^T \cdot Y = X^T \cdot 1_n = \sum_{i=1}^{n} X_i = \sum_{j=1}^{m} v_j(\Omega)$, where the final equality is true since X is non normalized probability vector on Ω for the set $\mathscr{V}$.

By definition, $C(\bar{G})$ is empty if and only if $\exists v_j \in \mathscr{V} : C(v_j) = \emptyset$. Then, by Remark 1.2, $v_j(\Omega) < \int^{cav} 1_n dv_j$. Moreover, by the same remark, $v_k(\Omega) \leq \int^{cav} 1_n dv_k$ for all $v_k \in \mathscr{V} \backslash \{v_j\}$. Therefore,

$$\sum_{j=1}^{m} \int^{cav} Y dv_j = \sum_{j=1}^{m} \int^{cav} 1_n dv_j > \sum_{j=1}^{m} v_j(\Omega) = X^T \cdot Y$$

Thus, if $C(\bar{G})$ is empty, for every non normalized probability vector on Ω for the set $\mathscr{V}$, $X \in \mathbb{R}_+^n$, there exists $Y \in \mathbb{R}_+^n$ such that $\sum_{j=1}^{m} \int^{cav} Y dv_j > X^T \cdot Y$. That is, if every $Y \in \mathbb{R}_+^n$ satisfies $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$ then $X \in C(\bar{G})$. That completes the proof. $\square$

Proof of Lemma 1

Proof. (i) $U(S)$ is the number of observations in D that are assigned to events that are subsets of S, $U(T)$ is the number of observations in D that are assigned to events that are subsets of T and $U(T \cap S)$ is the number of observations in D that are assigned to events that are subsets of both S and T. $U(T \cup S)$ is the number of observations in D that are assigned to events that are subsets of $S \cup T$, meaning it is at least the number of observations in D that are assigned to events that are either in S or in T excluding those in $T \cap S$. Hence, $U(S \cup T) \geq U(S) + U(T) - U(T \cap S)$ and therefore $U(S) + U(T) \leq U(S \cup T) + U(S \cap T)$. That is, G^U is a convex cooperative game.

(ii) $C(G^U)$ is non empty since the core of any convex cooperative game is non empty (see Shapley (1971/72)).

(iii) Recall that $\forall B \subset \Omega : U(B) = \sum_{b \subseteq B} V(b)$ and $U(\Omega) = V(\Omega)$. Let $X \in C(G^U)$ ($C(G^U)$ is non empty) then $\sum_{i=1}^{n} X_i = U(\Omega) = V(\Omega)$ and $\sum_{\omega_i \in B} X_i \geq U(B) = \sum_{b \subseteq B} V(b) \geq V(B)$. Thus, $X \in C(G^V)$. Hence, $C(G^U) \subseteq C(G^V)$ and $C(G^V)$ is non empty. $\square$

Lemma 5

Lemma 5. *Let V be a data set in a disaggregated characteristic function form and let G be the cooperative game induced by V. Let $\tilde{V}$ be the monotonic cover*[19] *of V and let $\tilde{G}$ be the cooperative game induced by $\tilde{V}$. Then, (i) $\tilde{V}$ is a capacity, (ii) $C(G) = C(\tilde{G})$, (iii) Let $Y \in \mathbb{R}^n_+$ be a finite non-negative random variable on Ω. Then, $\int^{cav} Y dV = \int^{cav} Y d\tilde{V}$.*

Proof. By the definition of V, we have $V(\emptyset) = 0$ and therefore also $\tilde{V}(\emptyset) = 0$. By the same definition we have (i) $V(\Omega)$ is the number of observations and (ii) $\tilde{V}(\Omega) = \max\{V(R)|R \subseteq \Omega\}$. Hence, $\tilde{V}(\Omega) = V(\Omega)$, that is $\tilde{V}(\Omega)$ is finite. Finally, by definition, a monotonic cover is monotonic. Thus, $\tilde{V}$ is a capacity.

Before we prove the second part, note that a data set in a disaggregated characteristic function form induces both a non-negative cooperative game and a non-negative monotonic cover. Therefore, the elements of $C(G)$ and $C(\tilde{G})$ must be non-negative.

First, since $\tilde{V}(\Omega) = V(\Omega)$, $\sum_{i\in\{1,\dots,n\}} x_i = V(\Omega)$ if and only if $\sum_{i\in\{1,\dots,n\}} x_i = \tilde{V}(\Omega)$.

Next, if $x \in C(\tilde{G})$ then $\forall B \subset \Omega : \sum_{\omega_i \in B} x_i \geq \tilde{V}(B)$. That is, if $x \in C(\tilde{G})$ then $\forall B \subset \Omega : \sum_{\omega_i\in B} x_i \geq \max\{V(R)|R \subseteq B\}$. In particular, if $x \in C(\tilde{G})$ then $\forall B \subset \Omega : \sum_{\omega_i\in B} x_i \geq V(B)$. Thus, if $x \in C(\tilde{G})$ then $x \in C(G)$.

For the other direction, suppose $x \in C(G)$. Fix B and let $R \subset B$. Since $x \in C(G)$ then $\sum_{\omega_i\in R} x_i \geq V(R)$. Since x is non-negative, $\sum_{\omega_i\in B} x_i \geq \sum_{\omega_i\in R} x_i$. Therefore, $\sum_{\omega_i\in B} x_i \geq V(R)$. As a result, if $x \in C(G)$ then $\forall B \subset \Omega, \forall R \subseteq B : \sum_{\omega_i\in B} x_i \geq V(R)$. Thus, $\forall B \subset \Omega : \sum_{\omega_i\in B} x_i \geq \max\{V(R)|R \subseteq B\}$. Hence, $\forall B \subset \Omega : \sum_{\omega_i\in B} x_i \geq \tilde{V}(B)$. That is, if $x \in C(G)$ then $x \in C(\tilde{G})$.

It is left to be shown that for every finite non-negative random variable, $Y \in \mathbb{R}^n_+$, the concave integral is the same whether it is calculated directly over V or over its monotonic cover ($\int^{cav} Y dV = \int^{cav} Y d\tilde{V}$).

First, note that, by definition, for every $B \in \Sigma$ we have $\tilde{V}(B) \geq V(B)$.[20] Let $\alpha_Y \in D(Y)$. Then, $\sum_{B\in\Sigma} \alpha_Y(B)\tilde{V}(B) \geq \sum_{B\in\Sigma} \alpha_Y(B)V(B)$. Denote the optimal decomposition of Y relative to V by $\alpha^\star_Y$ and the optimal decomposition of Y relative to $\tilde{V}$ by $\tilde{\alpha}^\star_Y$. Thus,

$$\sum_{B\in\Sigma} \tilde{\alpha}^\star_Y(B)\tilde{V}(B) \geq \sum_{B\in\Sigma} \alpha^\star_Y(B)\tilde{V}(B) \geq \sum_{B\in\Sigma} \alpha^\star_Y(B)V(B)$$

Hence, $\int^{cav} X dV \leq \int^{cav} X d\tilde{V}$.

Finally, for every $B \subseteq \Omega$ denote by $S(B) = \arg\max_{R\subseteq B} V(R)$ the subset of B that determines $\tilde{V}(B)$.[21] Let $\beta : \Sigma \to \mathbb{R}_+$ be the following a system of weights,

$$\beta(R) = \sum_{\{B\in\Sigma|S(B)=R\}} \tilde{\alpha}^\star_Y(B) + \sum_{\{B\in\Sigma|S(B)=B\setminus R\}} \tilde{\alpha}^\star_Y(B)$$

The vector of weights induced by β is denoted by W^β.

$$W^\beta_i = \sum_{\{R\in\Sigma|\omega_i\in R\}} \beta(R) = \sum_{\{R\in\Sigma|\omega_i\in R\}} \sum_{\{B\in\Sigma|S(B)=R\}} \tilde{\alpha}^\star_Y(B) + \sum_{\{R\in\Sigma|\omega_i\in R\}} \sum_{\{B\in\Sigma|S(B)=B\setminus R\}} \tilde{\alpha}^\star_Y(B)$$

The first term on the right-hand-side is the sum of weights over all the events that include state

[19] The monotonic cover of $G = (\Omega, V)$ is $\tilde{G} = (\Omega, \tilde{V})$ such that $\forall B \subseteq \Omega : \tilde{V}(B) = \max_{R\subseteq B} V(R)$.

[20] This is close to the monotonicity with respect to capacities property stated in Section 11.1.2 of Lehrer (2009). It is not the same since V may be non-monotonic.

[21] In cases where there is more than one maximizer, we assume, with no loss of generality, that $S(B)$ is the first in some given list of subsets.

ω_i and were determined by an event that includes state ω_i. The second term on the right-hand-side is the sum of weights over all the events that include state ω_i and were determined by an event that does not include state ω_i. Hence, this can also be written as

$$W_i^\beta = \sum_{\{B\in\Sigma|\omega_i\in S(B)\}} \tilde{\alpha}_Y^\star(B) + \sum_{\{B\in\Sigma|\omega_i\in B\setminus S(B)\}} \tilde{\alpha}_Y^\star(B) = \sum_{\{B\in\Sigma|\omega_i\in B\}} \tilde{\alpha}_Y^\star(B) = Y_i$$

Thus, β is a decomposition of X.

Note that,

$$\int^{cav} Y d\tilde{V} = \sum_{B\in\Sigma} \tilde{\alpha}_Y^\star(B)\tilde{V}(B) = \sum_{R\in\Sigma} \sum_{\{B\in\Sigma|S(B)=R\}} \tilde{\alpha}_Y^\star(B)\tilde{V}(B) = \sum_{R\in\Sigma} \sum_{\{B\in\Sigma|S(B)=R\}} \tilde{\alpha}_Y^\star(B)V(R)$$

The second equality is true since every event B has a corresponding event $S(B)$ that determines it and the third is due to the definitions of monotonic cover and $S(B)$. Thus,

$$\int^{cav} Y dV \geq \sum_{R\in\Sigma} \beta(R)V(R) = \sum_{R\in\Sigma} \sum_{\{B\in\Sigma|S(B)=R\}} \tilde{\alpha}_Y^\star(B)V(R) + \sum_{R\in\Sigma} \sum_{\{B\in\Sigma|S(B)=B\setminus R\}} \tilde{\alpha}_Y^\star(B)V(R)$$

$$= \int^{cav} Y d\tilde{V} + \sum_{R\in\Sigma} \sum_{\{B\in\Sigma|S(B)=B\setminus R\}} \tilde{\alpha}_Y^\star(B)V(R) \geq \int^{cav} Y d\tilde{V}$$

The first inequality is due to β being a decomposition of Y (but not necessarily the optimal one). The next equality is by the definition of β while the following equality is due to the result above. The final inequality results from system of weights and data sets being non-negative. This completes the proof since $\int^{cav} Y d\tilde{V} = \int^{cav} Y dV$. □

Proof of Proposition 2[22]

Proof. Suppose $X \in \mathbb{R}_+^n$ is an aggregate frequency distribution vector for the set of m data sets $\mathscr{V}$. By Proposition 1, $X \in C(\bar{G})$ if and only if every random variable $Y \in R_+^n$ satisfies

$$\sum_{V_j\in\mathscr{V}} \int^{cav} Y dV_j \leq Y \cdot X$$

By the definition of T_V^γ, $\sum_{V_j\in\mathscr{V}} \int^{cav} Y dV_j = \sum_{V_j\in\mathscr{V}} T_V^Y$. Finally, by the definition of W-equivalence classes every random variable $Y \in R_+^n$ corresponds to a class $\gamma \in \Gamma$, therefore, the proof is complete. □

[22] A different proof of this result involves constructing a set of linear inequalities that characterize the set of all the decompositions of a given payoff vector such that all vectors belong to the cores of the respective individual games. Then, Farkas' Lemma (or equivalently, the hyperplane separation theorem) can be used to show that this set of inequalities has a solution if and only if the above condition is satisfied. We prefer the proof presented here due to the novel result on concave integrals.