Aggregating Inconclusive Data Sets*

Gabrielle Gayer
Bar Ilan University

Ehud Lehrer
Durham University

Dotan Persitz
Tel Aviv University

May 2, 2024

Abstract

An administrator is provided with data collected by several practitioners. These data may include inconclusive observations. The administrator is required to form a frequency distribution on the states of nature that would be approved by external auditors as long as it is compatible with the available information. We state a novel result on the compatibility of a probability with a finite set of capacities. We use this result to provide necessary and sufficient conditions for the compatibility of the administrator's frequency distribution with the data collected by the practitioners, according to two auditing criteria.

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

1 Introduction

Consider the problem confronting 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 those viruses that are most likely to spread in the upcoming season. In an attempt to diagnose these viruses, health centers in different regions collect data on patients. Since vaccines are known to vary in their effectiveness across seasons[1] the health authority wishes to be able to justify its recommendation. 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, often must operate under conditions of uncertainty. Therefore, it is essential that they be able to prove that the probability underlying their decisions is based on all available information from all possible sources.

To study this problem, we consider an administrator who forms a probability over the possible states of nature. In addition, a group of practitioners collect relevant information on the matter under consideration and transfer their raw data to the administrator. If the probability formed by the administrator is supported by the information obtained from each of the practitioners, she should be able to establish that it is well-founded.

*A previous version was titled "Aggregating Non-Additive Beliefs". We thank two anonymous referees and an anonymous Advisory Editor for comments that improved this paper considerably.

[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, in February for the northern hemisphere vaccine and in September for the southern hemisphere vaccine, the World Health Organization provides recommendations on the composition of the influenza vaccine. More than 100 national influenza centers in over 100 countries conduct year-round surveillance of influenza that involves receiving and testing thousands of influenza virus samples from patients and reporting 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 raw data sets containing evidence about the states of nature that may have occurred. In some observations the state of nature that was realized may be known, while in others the outcome may be ambiguous. Cases in which a physician can perfectly diagnose a patient's condition correspond to evidence of a single virus, whereas partially diagnosed cases indicate that an unknown virus, one out of several possibilities, is the cause. An observation is considered 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 *raw data set* induces a characteristic function that assigns to each event the frequency of observations for which this exact event is known to have occurred. For example, event $\{\omega_1, \omega_2\}$ is assigned a number corresponding to the patients with an inconclusive diagnosis – the practitioner is only able to narrow the set of possible viruses down to ω_1 and ω_2, but cannot determine which was actually present.

A *processed data set* induces a different characteristic function that assigns to each event the total number of observations for which a subset of states included in this event is known to have occurred. That is, event $\{\omega_1, \omega_2\}$ is assigned a number corresponding to the number of patients with either a conclusive diagnosis $\{\omega_1\}$ or a conclusive diagnosis $\{\omega_2\}$ or an inconclusive diagnosis $\{\omega_1, \omega_2\}$.

A justification of the administrator for her chosen probability distribution will depend on the type of auditor she needs to convince. The first type of auditor bases its approval on the raw data set while the second type bases its approval on the processed data set. Any frequency that does not belong to the core of the corresponding (raw or processed data sets) cooperative game, must assign to some collection of states a frequency which is too low given the information included in the data set. That is, this frequency is not a possible realization in view of the available data. However, a frequency in the core is a possible realization of the distribution of outcomes that is consistent with these data. If the data set (raw or processed) includes no inconclusive evidence, there will be a single frequency in the core. If the data set (raw or processed) contains inconclusive evidence, then, there will be several frequencies in the core, each of which resolves the ambiguity differently.

When there are multiple practitioners, the administrator is able to 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. Verifying whether the administrator's probability meets this requirement is relatively simple when the induced cooperative games are all convex as in the case of the processed data set auditors (see Lemma 1). Proposition 2 provides a necessary and sufficient condition for the existence of such a decomposition of the administrator's probability for the raw data set auditors. This condition can be interpreted as testing the consistency of the administrator's frequency against every weighted combination of events of the raw data.

The processed data set auditors can be viewed as more sophisticated than the raw data set auditors, approving only those frequency distributions that are consistent with every aspect of the available data. A comparison between the characterization of probability distributions compatible with processed and raw data sets reveals a fundamental distinction in the functions assigned to their corresponding auditors. Processed data set auditors bear the setup cost of processing the data and subsequently can determine whether the administrator's frequency distribution should be accepted by verifying a finite set of conditions. In contrast, raw data set auditors do not incur any setup costs, but their scope is limited to rejecting a proposed frequency distribution. This limitation arises from the fact that verifying the compatibility of the probability distribution with the raw data set requires an infinite number of examinations. Since raw data set auditors can only perform a finite number of checks they are restricted to rejecting proposed frequency distributions, much like statistical tests that can only refute the null hypothesis, but cannot accept it. In Appendix B we show that when a property termed Uniform Optimal Decomposition across Practitioners is fulfilled, the number of required examinations becomes finite and, therefore, raw data set auditors can accept probability distributions.

Jaffray (1991), Gonzales and Jaffray (1998) and Arad and Gayer (2012) have already pointed out that imprecise statistical data generate ambiguity that is incorporated in 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 the perceived ambiguity in the problem at hand.[2] Lehrer (2009) introduced the concept of the concave integral as a method in which ambiguity-averse individuals can evaluate alternatives in conditions of 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 it 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 these capacities if and only if for any positive random variable Y, the weighted average of the 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, quality, political power, or influence incorporated into the capacities. Proposition 2, which has already been mentioned above, is a special case wherein practitioners' information is given in the form of data sets and the weights are proportional to the number of observations in the practitioners' data sets.

This method of aggregating probabilities using a weighted average is known in statistics as the linear opinion pool (Stone (1961)). The literature on decision-making under conditions of uncertainty provides an axiomatic foundation for the linear opinion pool. Gajdos and Vergnaud (2013), Crès et al. (2011) and Basili and Chateauneuf (2020) characterize the preference of a decision-maker who consults with experts to form a belief over an uncertain environment. The experts' beliefs are assumed to be non-additive (represented by capacities or sets of probabilities), reflecting their perception of ambiguity in the problem under consideration. The proposed decision rules give rise to a belief that is a weighted average of experts' beliefs (which may contain either multiple probabilities as in Gajdos and Vergnaud (2013) and Crès et al. (2011) or a single probability as in Basili and Chateauneuf (2020)). While this literature aims at convincing decision makers to adopt beliefs that are weighted average of the practitioners' beliefs, we are interested in decision makers that formed their beliefs independently and provide a test that examines whether their beliefs satisfy this property.

Dempster and Shafer's theory of evidence (Dempster (1967, 1968), Shafer (1976)) is an alternative approach to the linear opinion rule for merging beliefs derived from different sources of evidence, known as Dempster's rule of combination. However, for the rule to yield meaningful results, it is essential that the evidence from separate sources does not exhibit significant contradictions.[3] Consequently, Dempster's rule appears to be more suitable for situations where evidence is collected for specific problems that naturally do not involve conflicting in-

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

[3] In fact, in the extreme case of complete conflict, Dempster's rule is not applicable at all.

formation,[4] in contrast to statistical data. This distinction sets it apart from our framework, which deals with data sets comprising distinct records that have the potential to contradict one another.

In Section 2 we present the setting for the data sets and the auditors and prepare the ground for our main result. In Section 3 we introduce the necessary and sufficient condition for the aggregation of capacities evaluated according to the concave integral (Proposition 1). In Section 4 we use Proposition 1 to provide necessary and sufficient conditions for the compatibility of the administrator's frequency distribution with the data collected by the practitioners, according to two auditing criteria. We discuss the characteristics of each approach and suggest a compromise. Section 5 concludes. All proofs are relegated to the appendix.

2 Aggregating Data Sets

A Single Data Set

Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature ($n \geqslant 2$). A data set is a sequence of T observations, indexed by $i \in \{1, ..., T\}$, denoted by $D = (B_1, ..., B_T)$ where $B_i \in 2^{\Omega} \setminus \{\varnothing, \Omega\}$.[5] The event B_i represents the set of all states that may have occurred in observation i. Event B_i is a singleton when it is clear which state of nature occurred in observation i. However, observations are assigned to non-singleton events when it is not clear which specific state of nature within that event has 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.

A characteristic function is $v : 2^{\Omega} \to \mathbb{R}$ such that $v(\emptyset) = 0$.[6] We denote the cooperative game induced by the characteristic function v by $G = (\Omega, v)$.

We assume that data sets are cross-sectional (i.e. the order of observations does not affect the inference) and therefore we can describe the raw data set with T observations in a characteristic function form. Let $V : 2^{\Omega} \to \mathbb{N}$ be a function such that (i) For every event $B \subset \Omega$, $V(B)$ is the number of occurrences of B in raw data set D and (ii) $V(\Omega) = T$. A *raw data set in a characteristic function form* is the cooperative game $G^V = (\Omega; V)$.

Processing the raw data yields a function $U : 2^{\Omega} \to \mathbb{N}$ such that (i) For every event $B \subset \Omega$, $U(B) = \sum_{b \subseteq B} V(b)$ is the number of occurrences of B and its subsets ($b \subseteq B$) in raw data set D and (ii) $U(\Omega) = T$. A *processed data set in a characteristic function form* is the cooperative game $G^U = (\Omega; U)$.

The mass function in Dempster and Shafer's theory of evidence (Dempster (1967, 1968), Shafer (1976)), also known as a basic probability assignment, is closely related to our raw data set. It represents the degree of belief or support for each hypothesis based on the observed evidence. Our processed data set which is derived from processing the raw data set, is associated with Dempster and Shafer's belief function.[7]

[4] To illustrate, Shafer's leading example throughout his book (Shafer, 1976) revolves around Sherlock Holmes gathering various pieces of evidence related to a single crime.

[5] 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.

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

[7] Technically, the raw data values need to be normalized such that their sum equals 1 to qualify as a valid mass function. Furthermore, for the processed data to be regarded as a belief function, the mass assigned to the entire state space in the raw data set must be zero.

The Administrator

Allowing for inconclusive evidence introduces ambiguity into the decision problem faced by the administrator. Choosing the composition of viruses in the influenza vaccine requires an exact assessment of the frequencies of the various types of viruses in the population. The administrator resolves the ambiguity generated by the inconclusive evidence by forming a frequency distribution of the T observations, that is, a vector $X \in \mathbb{R}^n_+$ where X_i is the frequency of ω_i such that $\sum_{i=1}^n X_i = T$.[8] The goal of the administrator is to get the approval of the auditors to the use of this vector of frequencies.

The Auditors

We distinguish between two types of auditors. The first type of auditors verifies that, for every event B, the sum of frequencies assigned by the administrator to the states included in B is greater than or equal to the number of occurrences of the exact event B, namely, that for every non empty $B \subset \Omega$, $\sum_{i \in B} X_i \geq V(B)$. That is, these auditors approve X if and only if $X \in C(G^V)$ where $C(G)$ denotes the core of the cooperative game $G = (\Omega; V)$. We refer to these auditors as *raw data set auditors*.

The second type of auditors verifies that, for every event B, the sum of frequencies assigned by the administrator to the states included in B is greater than or equal to the total number of occurrences of events that are subsets of B. These auditors, which we refer to as *processed data set auditors*, check that for every non empty $B \subset \Omega$, $\sum_{i \in B} X_i \geq U(B)$. That is, these auditors approve X if and only if $X \in C(G^U)$.[9]

It is straightforward to show that every frequency distribution that satisfies the processed data set auditors' conditions, also satisfies the raw data set auditors' conditions, but not vice versa. Consequently, the raw data set auditors can be viewed as less sophisticated than the processed data set auditors.

Lemma 1. *Let V be a raw data set in a characteristic function form and let U be the corresponding processed data set in a characteristic function form.*[10]

1. *G^U is a convex cooperative game.*[11]
2. *$C(G^U)$ is non-empty.*
3. *$C(G^U) \subseteq C(G^V)$.*

[8] We assume that the administrator has a single additive probability distribution. However, we could easily extend the model so that her belief is represented by a capacity with a non-empty core. In this case it seems natural to apply the consistency condition imposed by the auditors to each probability in the core of the administrator's capacity. Yet, we would have had to determine whether or not this capacity is justifiable if only a subset of the probabilities in the core of the administrator's belief satisfy this condition. We prefer to avoid this issue by restricting the administrator's beliefs to be additive.

[9] For simplicity, we abstract from the fact that the realized outcomes must be natural numbers.

[10] The lemma's proof of claims 1 and 2 hinges on the definitions of the raw and processed data sets. An alternative approach is to show that U is a belief function, which is known to induce a convex game whose core is non-empty.

[11] A cooperative game $G = (\Omega, v)$ is convex if $v(S) + v(T) \leq v(S \cup T) + v(S \cap T)$ for every two events S and T.

Multiple Data Sets

Notation

Let $\mathscr{V} = \{V_1, V_2, \ldots, V_m\}$ be a collection of characteristic functions on Ω. 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 multi-Game $\bar{G}$ by $\bar{G}_j = (\Omega; V_j)$.

Let $X \in \mathbb{R}^n_+$, such that $\sum_{i=1}^n X_i = \sum_{j=1}^m V_j(\Omega)$. X belongs to the core of multi-Game $\bar{G}$ ($X \in C(\bar{G})$) if there are m finite non-negative vectors $X^1, ..., X^m$ such that $\forall j : X^j \in C(\bar{G}_j)$ and $\sum_{j=1}^m X^j = X$.

Raw 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. Denote by $\mathscr{V} = \{V_1, V_2, \ldots, V_m\}$ the collection of raw data sets in a characteristic function form where V_i corresponds to data set D_i. Let $\bar{G}^{\mathscr{V}} = (\Omega; \mathscr{V})$ be the raw data multi-game.

The administrator wishes to establish that the frequency underlying her decisions is credible. Let $X \in \mathbb{R}^n_+$, such that $\sum_{i=1}^n X_i = \sum_{j=1}^m V_j(\Omega)$, be the administrator's aggregated frequency distribution (i.e. non-normalized probability, a charge).

We refer to $C(\bar{G}^{\mathscr{V}})$ as the *raw data core* of $\mathscr{D}$. Every element $X \in C(\bar{G}^{\mathscr{V}})$, allows the administrator to prove to the raw data set auditors that her probability, the relative frequencies of X, is well-founded as it is supported by the information of the practitioners.

We define the processed data multi-game $\bar{G}^{\mathscr{U}} = (\Omega; \mathscr{U})$ where $\mathscr{U} = \{U_1, U_2, \ldots, U_m\}$ and U_i is the processed data set in a characteristic function form that corresponds to data set D_i. We refer to $C(\bar{G}^{\mathscr{U}})$ as the *processed data core* of $\mathscr{D}$. Every element $X \in C(\bar{G}^{\mathscr{U}})$, allows the administrator to prove to the processed data set auditors that her probability is well-founded.

Main Result

Our main result, formally stated in Section 4, provides a necessary and sufficient condition for the decomposition of an aggregated frequency distribution on $T = \sum_{i=1}^m T_i$ observations into m frequency distributions such that the first is compatible with the requirements of a raw data set auditor regarding the first data set, the second is compatible with his requirements regarding the second data set and so on. In order to state the result we first provide a general result on information aggregation under ambiguity and then we apply it to our setting.

3 Aggregating Concave Integrals

Concave Integrals

Definitions

Capacities are functions $v : 2^\Omega \to \mathbb{R}_+$ that satisfy (i) no empty events ($v(\emptyset) = 0$) (ii) finiteness ($v(\Omega)$ is finite) and (iii) monotonicity ($S \subseteq T \Rightarrow v(S) \leq v(T)$). Concave integrals are integrals over capacities that are used to evaluate acts in settings with non-additive beliefs. Concave integrals were introduced by Lehrer (2009) (and later generalized in Even and Lehrer (2014)) to allow for aversion to ambiguity even when capacities are not convex.[12]

[12] In the Choquet expected utility model (Schmeidler (1989)) aversion to ambiguity corresponds to convex capacities for which the concave integral and the Choquet integral imply the same preferences over random variables (Lehrer (2009)).

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 functions of degree one $f : \mathbb{R}^n_+ \to \mathbb{R}$ such that $\forall B \in 2^\Omega : f(\chi^B) \geq v(B)$ where $\chi^B \in \{0,1\}^n$ denotes the indicator vector of B ($\chi^B_i = 1$ if $\omega_i \in B$ and $\chi^B_i = 0$ otherwise).[13,14]

A decomposition of vector Y is $\alpha_Y : 2^\Omega \to \mathbb{R}_+$ such that $\sum_{B \in 2^\Omega} \alpha_Y(B)\chi^B = Y$. Denote the set of all decompositions of Y by $D(Y)$ and the optimal decomposition of Y relative to capacity v by $\alpha^\star_{Y,v} = \arg\max_{\alpha_Y \in D(Y)} \left\{ \sum_{B \in 2^\Omega} \alpha_Y(B) v(B) \right\}$.[15] Lemma 1(i) in Lehrer (2009) states that $\int^{cav} Y dv = \sum_{B \in 2^\Omega} \alpha^\star_{Y,v}(B) v(B)$, namely, the concave integral can be expressed as a linear combination of the capacities of events where the weights are the corresponding optimal decomposition elements. The following remark on the concave integrals of indicator vectors will become useful.

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

1. *Let $Y \in R^n_+$. If $\alpha^\star_{Y,v}(B) > 0$ then $v(B) = \int^{cav} \chi^B dv$.*
2. *Let $G = (\Omega; v)$ be the cooperative game induced by capacity v.*
   - *(a) $v(\Omega) \leq \int^{cav} \chi^\Omega dv$.*
   - *(b) $C(G)$ is non-empty if and only if $v(\Omega) = \int^{cav} \chi^\Omega dv$.*
   - *(c) $C(G)$ is empty if and only if $v(\Omega) < \int^{cav} \chi^\Omega dv$.*

Multiple Concave Integrals

Novel Result

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 $V(\Omega) = \sum_{j=1}^m v_j(\Omega)$. Let $X \in \mathbb{R}^n_+$, such that $\sum_{i=1}^n X_i = \sum_{j=1}^m v_j(\Omega)$, be a non-normalized probability for the set $\mathscr{V}$. A non additive probability (or normalized capacity) is a monotonic capacity for which the value of Ω equals 1. To convert capacities into non additive probabilities, let the practitioners' normalized capacities (their non-additive beliefs) be $\hat{v}_j = \frac{v_j}{v_j(\Omega)}$ and $\hat{\mathscr{V}} = \{\hat{v}_1, \ldots, \hat{v}_m\}$. In addition, denote $\hat{X} = \frac{1}{V(\Omega)} X$ and $\hat{X}^j = \frac{1}{v_j(\Omega)} X^j$, so that $\hat{X}$ becomes a probability on Ω. Finally, denote $\beta_j = \frac{v_j(\Omega)}{V(\Omega)}$ so that $\hat{X} = \sum_{j=1}^m \beta_j \hat{X}^j$.

Proposition 1 is a novel result that characterizes the core of the multi-game induced by $\mathscr{V}$. It shows that the probability of the decision maker is a β-weighted average of probabilities in the respective cores of the individuals' non-additive beliefs if and only if the evaluation of any positive random variable Yaccording to the probability ($\hat{X}^T \cdot Y$) is higher than or equal to the β-weighted average of the evaluations based on the individual capacities ($\int^{cav} Y d\hat{v}_j$).

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

[14] We abuse notation by using χ to denote both the set of indicator vectors and 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'.

[15]To show the non-emptiness of $D(Y)$ consider α such that $\forall i \in \{1, \ldots, n\} : \alpha(\{\omega_i\}) = Y_i$ and for every non-singleton $B \in 2^\Omega$, $\alpha(B) = 0$. $\alpha \in D(Y)$ for every Y.

Proposition 1. *Let $\hat{X} \in \mathbb{R}^n_+$ be a probability on Ω. There exist m vectors $\hat{X}^j \in C(\hat{\bar{G}}_j)$ such that $\hat{X} = \sum_{j=1}^m \beta_j \hat{X}^j$ if and only if for every random variable $Y \in R^n_+$: $\sum_{j=1}^{m} \beta_j \int^{cav} Y d\hat{v}_j \leq \hat{X}^T \cdot Y$.*

To sketch the proof, recall that if X is in the core of the multi-game induced by $\mathscr{V}$, it is a sum of members in the cores of each game in $\mathscr{V}$. 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 greater than or equal to its expectation according to the respective concave integral. Then summing over all capacities implies that the expectation of Y according to X must be greater than or equal to the sum of the concave integrals of Y over all $v_j \in \mathscr{V}$.

If X is not in the core of the multi-game induced by $\mathscr{V}$ we can construct a violation to the condition on the sum of concave integrals. Either the core of the multi-game induced by $\mathscr{V}$ is non-empty, in which case, we use a hyperplane separation theorem to construct a violating example or the core of the multi-game induced by $\mathscr{V}$ is empty and then Remark 1.2c implies that χ^Ω violates the condition.

The final step of the proof is to convert the condition into terms of probabilities and non-additive capacities.

Remarks

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

In addition, note that it is possible to extend the decision maker's probability to contain multiple priors by applying the condition in Proposition 1 to each of them separately (see the discussion in Footnote 8).

Finally, a technically useful implication of the proof 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 $Y \in R^n_+$: $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq \min_{X \in C(\bar{G})} X^T \cdot Y$.

4 Back to Aggregating Data Sets

Main Result: Raw Data Compatibility

Now we can present our main result on the aggregation of data sets.

Proposition 2. *Let $X \in \mathbb{R}^n_+$ be an aggregated frequency distribution on $T = \sum_{i=1}^m T_i$ observations for the set of m data sets $\mathscr{D}$. $X \in C(\bar{G}^{\mathscr{V}})$ if and only if every random variable $Y \in R^n_+$ satisfies $\sum_{V_j \in \mathscr{V}} \int^{cav} Y dV_j \leq Y \cdot X$.*

Proposition 2 is almost a direct application of the novel Proposition 1.[16] The main difference is that Proposition 1 requires $\mathscr{V}$ to be a set of capacities while in Proposition 2 $\mathscr{V}$ is a set of raw data sets which may not be monotonic. In the proof we show that replacing the

[16] A slight modification of this result was used to prove Proposition 3 in Gayer and Persitz (2016) (see p. 948).

raw data sets by their monotonic covers generates capacities with the same cores but whose concave integrals are well-defined.

Usefulness of Proposition 2

To demonstrate how Proposition 2 can be utilized to determine whether a frequency distribution is compatible with the available raw 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 one containing a pair of states - the first observation includes States 1 and 2, the second observation includes States 1 and 3, and the third observation includes States 2 and 3. Practitioner 2's data set contains two observations: the first observation is conclusive, containing only State 1, while the second observation is inconclusive including States 2 and 3. 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 $v_2(\{1,2,3\}) = 2$. 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, consider the random variable $Y = (1,1,0)$, then $\int^{cav} Y dv_1 = 1 * v_1(1,2) = 1$, and $\int^{cav} Y dv_2 = 1 * v_2(\{1\}) + 1 * v_2(\{2\}) = 1$, however $Y * X = 1$. Thus, by Proposition 2, the administrator's frequency distribution is found to be incompatible with the data sets according to the raw data set auditors (and therefore also according to the processed data auditors).

Proposition 2 necessitates conducting an infinite number of examinations (one for each random variable Y) to ascertain the compatibility of the probability distribution with the raw data set. Consequently, raw data set auditors who are constrained to executing only a finite number of checks are limited to rejecting proposed frequency distributions, as demonstrated in the example above. In Appendix B we study a special class of $\mathscr{V}$ that satisfies a property termed Uniform Optimal Decomposition across Practitioners, which requires only a finite number of examinations to determine the compatibility of the probability distribution with the raw data set. In this case raw data set auditors can accept a proposed probability distribution.

Processed Data Compatibility

Proposition 2 presents a condition that if satisfied, a raw data set auditor would accept the suggested frequency distribution as compatible with the available raw data sets. One can apply an adequate version of Proposition 2 to $\mathscr{U}$ to understand whether a frequency distribution would be found acceptable by the processed data set auditors as well. In fact, since inclusion in the processed data core implies inclusion in the raw data core (by Lemma 1), the condition stated in Proposition 2 is a necessary (though insufficient) condition for an aggregate frequency to belong to $C(\bar{G}^{\mathscr{U}})$. Alternatively, recall that Lemma 1 states that the cooperative game induced by a processed data set is convex. By Dragan et al. (1989) (see also Footnote 18 in Gayer and Persitz (2016)), $C(\bar{G}^{\mathscr{U}}) = C(\sum_{U_i \in \mathscr{U}} G^{U_i})$, and therefore, compatibility with processed data set auditors can be established simply by verifying that the frequency distribution is in the core of the summation game, $\sum_{U_i \in \mathscr{U}} G^{U_i}$.

Raw Data Set Auditors vs. Processed Data Set Auditors

If there are no inconclusive observations the auditing criteria of the raw and the processed data set auditors coincide. Therefore, the difference between these two types lies within the permissible resolution of inconclusive evidence. To demonstrate, consider a raw 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 $V(\{2,3\}) = 0$. Raw data set auditors would approve the frequency distribution $X = (1,1,1)$ since $V(\{1,2\}) < X_1 + X_2$. However, the total frequency attributed to States 1 and 2 understates the total number of observations assigned to the relevant events - $\{1\}$, $\{2\}$ and $\{1,2\}$ - since $V(\{1\}) + V(\{2\}) + V(\{1,2\}) > X_1 + X_2$. Therefore, $X = (1,1,1)$ would be disapproved by the processed data set auditors. For the same reason, the processed data set auditors would claim that the frequency attributed to State 3 is erroneously overstated since no observation is assigned to an event that includes State 3. On the other hand, the frequency distribution $Y = (1.5, 1.5, 0)$ would be approved by the processed data set auditors (and therefore also by the raw data set auditors). Here the inconclusive observation of the event $\{1,2\}$ is equally attributed to States 1 and 2, but not to State 3.

Advanced Raw Data Set Auditors

One partial, yet relatively simple, remedy to the overly simplified nature of the raw data set auditors is to add to the condition in Proposition 2 the requirement that no state of nature is assigned a number larger than the total number of observations that were attributed to events that contain it, namely, $\forall \omega_i \in \Omega : X_i \leq \sum_{S \subset \Omega \setminus \{\omega_i\}} V(S \cup \{\omega_i\})$. It is easy to see that this additional requirement rules out $X = (1,1,1)$ as a compatible frequency distribution in the above-mentioned example.

Clearly, every frequency distribution approved by these *advanced raw data set auditors* is approved also by the standard raw data set auditors. In addition, every frequency distribution approved by the processed data set auditors is approved by these advanced raw data set auditors.[17] Therefore, the set of frequency distributions that the advanced raw data set auditors find compatible falls between that of the standard raw data set auditors and the processed data set auditors.[18]

5 Concluding Remarks

We provide a necessary and sufficient condition for the compatibility of a probability distribution with a set of given non-additive beliefs. The condition is that for any positive random variable Y, a given weighted average of the expected value of Y according to the concave integrals across capacities is bounded from above by the expected value of Y with respect to the prior probability.

This result is applied to a setting wherein an administrator is provided with data collected by several practitioners. These data may include inconclusive observations and therefore give rise to many possible frequency distributions. The administrator is required to form a frequency

[17]To see this, recall that every frequency distribution approved by a processed data set auditor is also approved by a standard raw data set auditor (Lemma 1). In addition, note that $X \in C(G^U)$ implies that for every state ω_i: $\sum_{\omega_j \in \Omega \setminus \{\omega_i\}} X_j \geq U(\Omega \setminus \{\omega_i\})$. Therefore, for every state ω_i: $\sum_{\omega_j \in \Omega \setminus \{\omega_i\}} X_j \geq \sum_{S \subseteq \Omega \setminus \{\omega_i\}} V(S)$. This implies that for every state ω_i: $\sum_{\omega_j \in \Omega} X_j - \sum_{\omega_j \in \Omega \setminus \{\omega_i\}} X_j \leq V(\Omega) - \sum_{S \subseteq \Omega \setminus \{\omega_i\}} V(S)$. That is, for every state ω_i: $X_i \leq \sum_{S \subset \Omega \setminus \{\omega_i\}} V(S \cup \{\omega_i\})$ so the additional requirement is satisfied. This means that every frequency distribution approved by a processed data set auditor is also approved by an advanced raw data set auditor.

[18]This set, however, need not coincide with the standard raw data set auditors' or with the processed data set auditors' set of compatible frequency distributions. For example, suppose that there are four states of nature and four observations that correspond to $v(\{1\}) = 1$, $v(\{4\}) = 1$, $v(\{1,2\}) = 1$ and $v(\{3,4\}) = 1$ while all other events were not observed. The frequency distribution $(3,0,0,1)$ is only compatible according to the standard raw data set auditors, whereas the frequency distribution $(2,1,0,1)$ is compatible also according to the more advanced raw data set auditors, but not by the processed data set auditors. Finally, the frequency distribution $(1,1,1,1)$ is compatible according to all three types of auditors.

distribution on the states of nature that would be approved by external auditors. We consider two types of auditors, the first type relies on raw data in its inspection, whereas the second type relies on processed data. We provide necessary and sufficient conditions for the compatibility of the administrator frequency distribution with the data collected by the practitioners, according to both auditing criteria.

A comparison between the characterization of probability distributions compatible with processed and raw data sets reveals a fundamental distinction in the functions assigned to their corresponding auditors. Processed data set auditors bear the setup cost of processing the data and subsequently can determine whether the administrator's frequency distribution should be accepted by verifying a finite set of conditions. In contrast, raw data set auditors do not incur any setup costs, but their scope is limited to rejecting a proposed frequency distribution. This limitation arises from the fact that verifying the compatibility of the probability distribution with the raw data set requires an infinite number of examinations (one for each random variable Y). As a result, raw data set auditors can only perform a finite number of checks and are therefore restricted to rejecting proposed frequency distributions.

We conclude by proposing a third approach to audit the administrator's frequency distribution. The approach taken by the advanced raw data set auditors can be viewed as a reasonable compromise between the costly approach of the processed data set auditors and the over permissiveness of the raw data set auditors.

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.

Basili, Marcello and Alain Chateauneuf (2020) "Aggregation of experts' opinions and conditional consensus opinion by the Steiner point," *International Journal of Approximate Reasoning*, Vol. 123, pp. 17–25.

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).

Crès, Hervé, Itzhak Gilboa, and Nicolas Vieille (2011) "Aggregation of multiple prior opinions," *Journal of Economic Theory*, Vol. 146, No. 6, pp. 2563–2582.

Dempster, Arthur P (1967) "Upper and lower probabilities induced by a multivalued mapping," *The Annals of Mathematical Statistics*, Vol. 38, No. 2, pp. 325–339.

——— (1968) "A generalization of Bayesian inference," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 30, No. 2, pp. 205–232.

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.

Gajdos, Thibault and Jean-Christophe Vergnaud (2013) "Decisions with conflicting and imprecise information," *Social Choice and Welfare*, Vol. 41, No. 2, pp. 427–452.

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

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.

Lovász, László (1983) "Submodular functions and convexity," *Mathematical Programming The State of the Art: Bonn 1982*, pp. 235–257.

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.

Shafer, Glenn (1976) *A mathematical theory of evidence*, Vol. 42: Princeton university press.

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.

Stone, Mervyn (1961) "The linear opinion pool," *Ann. Math. Statist*, Vol. 32, pp. 1339–1342.

Appendix

A Proofs

A.1 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 $U(S\cap T)$ is the number of observations in D that are assigned to events that are subsets of both S and T. $U(S\cup T)$ 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 $S\cap T$. Hence, $U(S\cup T)\geq U(S)+U(T)-U(S\cap T)$ 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. □

A.2 Proof of Remark 1

Proof. First, note that $v(B) \leq \int^{cav} \chi^B dv$ since the decomposition where $\alpha_{\chi^B,v}(b) = 0$ for every $b \subset B$ and $\alpha_{\chi^B,v}(B) = 1$ generates a value of $v(B)$ and there might be decompositions of χ^B that generate higher values.

Next, suppose $v(B) < \int^{cav} \chi^B dv$. Then, (i) $\alpha^{\star}_{\chi^B,v}(B) < 1$ and (ii)

$$v(B) < \sum_{b \subset B} \alpha^{\star}_{\chi^B,v}(b)v(b) + \alpha^{\star}_{\chi^B,v}(B)v(B)$$

Hence,

$$v(B) < \frac{1}{1-\alpha^{\star}_{\chi^B,v}(B)} \sum_{b \subset B} \alpha^{\star}_{\chi^B,v}(b)v(b)$$

If $\alpha^{\star}_{\chi^B,v}(B) > 0$ then the decomposition where $\alpha^{\star\star}_{\chi^B,v}(b) = \frac{\alpha^{\star}_{\chi^B,v}(b)}{1-\alpha^{\star}_{\chi^B,v}(B)}$ for all $b \subset B$ and $\alpha^{\star\star}_{\chi^B,v}(B) = 0$ achieves a strictly higher value than $\alpha^{\star}_{\chi^B,v}$, in contradiction to its optimality.
Hence, $v(B) < \int^{cav} \chi^B dv$ implies $\alpha^{\star}_{\chi^B,v}(B) = 0$.

Now, suppose that $Y \in R^n_+$. By definition,

$$\int^{cav} Y dv = \sum_{b \not\subset B} \alpha^{\star}_{Y,v}(b)v(b) + \sum_{b \subset B} \alpha^{\star}_{Y,v}(b)v(b) + \alpha^{\star}_{Y,v}(B)v(B)$$

If $v(B) < \int^{cav} \chi^B dv$ then

$$\sum_{b \not\subset B} \alpha^{\star}_{Y,v}(b)v(b) + \sum_{b \subset B} \alpha^{\star}_{Y,v}(b)v(b) + \alpha^{\star}_{Y,v}(B)v(B) \leq \sum_{b \not\subset B} \alpha^{\star}_{Y,v}(b)v(b) + \sum_{b \subset B} (\alpha^{\star}_{Y,v}(b) + \alpha^{\star}_{Y,v}(B)\alpha^{\star\star}_{\chi^B,v}(b))v(b)$$

Since $\alpha^{\star}_{Y,v}$ is optimal, the two expressions must be equal. That is, $\alpha^{\star}_{Y,v}(B) = 0$. Hence, $v(B) < \int^{cav} \chi^B dv$ implies $\alpha^{\star}_{Y,v}(B) = 0$. Therefore $\alpha^{\star}_{Y,v}(B) > 0$ implies $v(B) = \int^{cav} \chi^B dv$.

The second part of the remark is almost trivial. The first assertion is a specific case of the first step of this proof. That is, since $v(B) \leq \int^{cav} \chi^B dv$ then, in particular, $v(\Omega) \leq \int^{cav} \chi^{\Omega} dv$. Next, since $D(\chi^{\Omega})$ is the set of all balancing weights, by the Shapley-Bondareva Theorem (Bondareva (1963) and Shapley (1967)) we get that $C(G)$ is non-empty if and only if $v(\Omega) = \int^{cav} \chi^{\Omega} dv$, while $C(G)$ is empty if and only if $v(\Omega) < \int^{cav} \chi^{\Omega} dv$. □

A.3 Lemma 2

Lemma 2. *Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature. Let v be a capacity on Ω and let Y be a finite non-negative random variable on Ω. Denote $\hat{H} = \{h \in \mathbb{R}^n_+ | \forall B \in 2^{\Omega} : \sum_{\omega_i \in B} h_i \geq v(B)\}$ and the set of its extreme points by H.[19] Then,*

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

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

$$\int^{cav} Y dv = \max_{\alpha: 2^n \to \mathbb{R}_+} \left\{ \sum_{B \in 2^n} \alpha(B)v(B) \middle| \sum_{B \in 2^{\Omega}} \alpha(B)\chi^B = Y, \forall B \in 2^{\Omega} : \alpha(B) \geq 0 \right\}$$

[19] $h \in \hat{H}$ 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 $h = \lambda \tilde{h} + (1-\lambda)\tilde{\tilde{h}}$.

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,

$$\int^{cav} Y dv = \min_{h \in \mathbb{R}^n_+} \left\{ h^T \cdot Y | \forall B \in 2^{\Omega} : \sum_{\omega_i \in B} h_i \geq v(B) \right\} = \min_{h \in \hat{H}} h^T \cdot Y$$

$\hat{H}$ is non empty[20] and convex.[21] 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, $\displaystyle\int^{cav} Y dv = \min_{h \in H} h^T \cdot Y$. □

A.4 Lemma 3

Lemma 3. *Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite set of states of nature. Let v be a capacity on Ω. If $C(v)$ is non empty there is a neighborhood U of 1_n (a length n vector of ones) 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, $\min_{c \in C(v)} c^T \cdot Y \geq \min_{h \in H} h^T \cdot Y = \int^{cav} Y dv$.

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

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 $h^{tT} \cdot 1_n > v(\Omega)$.

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 1.2b.

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 $\lim_{r \to \infty} \int^{cav} Y_r dv = \lim_{r \to \infty} h^{rT} \cdot Y_r = h^{rT} \cdot \{\lim_{r \to \infty} Y_r\} = h^{rT} \cdot 1_n$. Hence, $h^{rT} \cdot 1_n = v(\Omega)$. Contradiction.

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

A.5 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 $v_j \in \mathscr{V}$, $C(\bar{G}_j)$ 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

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

[21] $h, \bar{h} \in \hat{H}$ implies that $\forall B \in 2^{\Omega}$, $(\chi^B)^T h \geq v(B)$ and $(\chi^B)^T \bar{h} \geq v(B)$ and therefore for every $\lambda \in [0,1]$:

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

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 $\forall j$: $Z^j \in C(\bar{G}_j)$, $\hat{Z}^j \in C(\bar{G}_j)$ 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 $\forall \lambda \in [0,1]$, $\forall j$: $\lambda Z^j + (1-\lambda)\hat{Z}^j \in C(\bar{G}_j)$. 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. □

A.6 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 $\forall v_j \in \mathscr{V} : X^j \in C(\bar{G}_j)$ 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

$$X^{j^T} \cdot Y = X^{j^T} \cdot \left[\sum_{B \in 2^\Omega} \left[\alpha_Y(B) \times \chi^B\right]\right] = \sum_{B \in 2^\Omega}\left[X^{j^T} \cdot \left[\alpha_Y(B) \times \chi^B\right]\right] =$$

$$\sum_{B \in 2^\Omega}\left[\alpha_Y(B) \times \left[X^{j^T} \cdot \chi^B\right]\right] = \sum_{B \in 2^\Omega}\left[\alpha_Y(B) \times \sum_{\omega_i \in B} X^j_i\right] \geq \sum_{B \in 2^\Omega}\left[\alpha_Y(B) \times v_j(B)\right]$$

Where the first equality is by the definition of a decomposition and the final inequality is true since $\forall v_j \in \mathscr{V} : X^j \in C(\bar{G}_j)$ implies that $\forall v_j \in \mathscr{V}, \forall B \in 2^\Omega : \sum_{\omega_i \in B} X^j_i \geq v_j(B)$.

In particular, for every random variable $Y \in R^n_+$ and for every capacity $v_j \in \mathscr{V}$, $X^{j^T} \cdot Y \geq \sum_{B \in 2^\Omega}\left[\alpha^\star_{Y,v_j}(B) \times v_j(B)\right]$. 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 $X^T \cdot Y = \sum_{j=1}^{m} X^{j^T} \cdot Y = \sum_{j=1}^{m}\left[X^{j^T} \cdot Y\right] \geq \sum_{j=1}^{m}\int^{cav} Y dv_j$.

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 $Z = (Z_1, \ldots, Z_n) \neq 0_n$ 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 $w^{1\star}, \ldots, w^{m\star}$ such that $\forall v_j \in \mathscr{V} : w^{j\star} \in C(\bar{G}_j)$ and $\sum_{j=1}^m w^{j\star} = w^\star$. Moreover, $\forall v_j \in \mathscr{V} : w^{j\star} \in \arg\min_{w^j \in C(\bar{G}_j)} \{w^{j^T} \cdot Z^c\}$.[22] Therefore, for every positive constant c, $X^T \cdot Z^c < \sum_{j=1}^m \min_{w^j \in C(\bar{G}_j)} \{w^{j^T} \cdot Z^c\}$.

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

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

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 $\exists v_j \in \mathscr{V} : C(\bar{G}_j) = \emptyset$. Then, by Remark 1.2c, $v_j(\Omega) < \int^{cav} 1_n dv_j$. Moreover, by Remark 1.2a, $v_k(\Omega) \leq \int^{cav} 1_n dv_k$ for all $v_k \in \mathscr{V} \backslash \{v_j\}$. Therefore,

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

---

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

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 is, there exist m vectors $X^j \in C(\bar{G}_j)$ such that $X = \sum_{j=1}^{m} X^j$ if and only if every $Y \in \mathbb{R}^n_+$ satisfies $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$.

The final step of the proof is to normalize the capacities and X. First note that $\hat{X} = \frac{X}{V(\Omega)} = \frac{1}{V(\Omega)} \sum_{j=1}^{m} X^j = \sum_{j=1}^{m} \frac{1}{V(\Omega)} X^j = \sum_{j=1}^{m} \frac{v_j(\Omega)}{V(\Omega)} \frac{X^j}{v_j(\Omega)} = \sum_{j=1}^{m} \beta_j \hat{X}^j$. Thus, since $X^j \in C(\bar{G}_j)$ if and only if $\hat{X}^j \in C(\hat{\bar{G}}_j)$, we conclude that $X \in C(\bar{G})$ if and only if there exist m vectors $\hat{X}^j \in C(\hat{\bar{G}}_j)$ such that $\hat{X} = \sum_{j=1}^{m} \beta_j \hat{X}^j$.

Finally, $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$ if and only if $\frac{1}{V(\Omega)} \sum_{j=1}^{m} \int^{cav} Y dv_j \leq \frac{1}{V(\Omega)} X^T \cdot Y$. Since $\int^{cav} Y dv_j = \sum_{B \in 2^\Omega} \alpha^\star_{Y,v}(B) v(B)$ and since $\alpha^\star_{Y,v}(B) = \alpha^\star_{Y,\hat{v}}(B)$, we get that $\frac{1}{V(\Omega)} \sum_{j=1}^{m} \int^{cav} Y dv_j = \frac{1}{V(\Omega)} \sum_{j=1}^{m} v_j(\Omega) \int^{cav} Y d\hat{v}_j = \sum_{j=1}^{m} \beta_j \int^{cav} Y d\hat{v}_j$. In addition, $\frac{1}{V(\Omega)} X^T \cdot Y = \frac{1}{V(\Omega)} \sum_{j=1}^{m} (X^j)^T \cdot Y = \frac{1}{V(\Omega)} \sum_{j=1}^{m} v_j(\Omega) (\hat{X}^j)^T \cdot Y = \sum_{j=1}^{m} \beta_j (\hat{X}^j)^T \cdot Y = \hat{X}^T \cdot Y$ where the last equality uses $\hat{X} = \sum_{j=1}^{m} \beta_j \hat{X}^j$.

Therefore, we showed that there exist m vectors $\hat{X}^j \in C(\hat{\bar{G}}_j)$ such that $\hat{X} = \sum_{j=1}^{m} \beta_j \hat{X}^j$ if and only if for every random variable $Y \in R^n_+$: $\sum_{j=1}^{m} \beta_j \int^{cav} Y d\hat{v}_j \leq \hat{X}^T \cdot Y$. □

A.7 Lemma 5 [see also the discussion in Even and Lehrer (2014)]

Lemma 5. *Let V be a raw data set in a characteristic function form and let G be the cooperative game induced by V. Let $\tilde{V}$ be the monotonic cover*[23] *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 $\tilde{V}(\emptyset) = 0$. Since $\tilde{V}(\Omega) = \max\{V(R) | R \subseteq \Omega\}$ and $V(\Omega)$ is the number of observations we get that $\tilde{V}(\Omega) = V(\Omega)$, 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 raw data set in a 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, $\forall B \subset \Omega : \sum_{\omega_i \in B} x_i \geq$

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

$\max\{V(R)|R \subseteq B\}$. 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 $B \in 2^\Omega$ we have $\tilde{V}(B) \geq V(B)$.[24] Let $\alpha_Y \in D(Y)$. Then, $\sum_{B \in 2^\Omega} \alpha_Y(B)\tilde{V}(B) \geq \sum_{B \in 2^\Omega} \alpha_Y(B)V(B)$. 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,

$$\sum_{B \in 2^\Omega} \tilde{\alpha}^\star_Y(B)\tilde{V}(B) \geq \sum_{B \in 2^\Omega} \alpha^\star_Y(B)\tilde{V}(B) \geq \sum_{B \in 2^\Omega} \alpha^\star_Y(B)V(B)$$

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

Finally, for every $B \subseteq \Omega$ denote by $S(B) = \arg\max_{R \subseteq B} V(R)$ the subset of B that determines $\tilde{V}(B)$.[25] That is, $S(B) \subseteq B$. Let $\beta : 2^\Omega \to \mathbb{R}_+$ be the following system of weights,

$$\beta(R) = \sum_{\{B \in 2^\Omega | S(B) = R\}} \tilde{\alpha}^\star_Y(B) + \sum_{\{B \in 2^\Omega | S(B) = B \setminus R\}} \tilde{\alpha}^\star_Y(B)$$

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

$$W^\beta_i = \sum_{\{R \in 2^\Omega | \omega_i \in R\}} \beta(R) = \sum_{\{R \in 2^\Omega | \omega_i \in R\}} \sum_{\{B \in 2^\Omega | S(B) = R\}} \tilde{\alpha}^\star_Y(B) + \sum_{\{R \in 2^\Omega | \omega_i \in R\}} \sum_{\{B \in 2^\Omega | S(B) = B \setminus R\}} \tilde{\alpha}^\star_Y(B)$$

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

$$W^\beta_i = \sum_{\{B \in 2^\Omega | \omega_i \in S(B)\}} \tilde{\alpha}^\star_Y(B) + \sum_{\{B \in 2^\Omega | \omega_i \in B \setminus S(B)\}} \tilde{\alpha}^\star_Y(B) = \sum_{\{B \in 2^\Omega | \omega_i \in B\}} \tilde{\alpha}^\star_Y(B) = Y_i$$

Thus, β is a decomposition of Y.

Note that,

$$\int^{cav} Y d\tilde{V} = \sum_{B \in 2^\Omega} \tilde{\alpha}^\star_Y(B)\tilde{V}(B) = \sum_{R \in 2^\Omega} \sum_{\{B \in 2^\Omega | S(B) = R\}} \tilde{\alpha}^\star_Y(B)\tilde{V}(B) = \sum_{R \in 2^\Omega} \sum_{\{B \in 2^\Omega | S(B) = R\}} \tilde{\alpha}^\star_Y(B)V(R)$$

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,

$$\int^{cav} Y dV \geq \sum_{R \in 2^\Omega} \beta(R)V(R) = \sum_{R \in 2^\Omega} \sum_{\{B \in 2^\Omega | S(B) = R\}} \tilde{\alpha}^\star_Y(B)V(R) + \sum_{R \in 2^\Omega} \sum_{\{B \in 2^\Omega | S(B) = B \setminus R\}} \tilde{\alpha}^\star_Y(B)V(R)$$

$$= \int^{cav} Y d\tilde{V} + \sum_{R \in 2^\Omega} \sum_{\{B \in 2^\Omega | S(B) = B \setminus R\}} \tilde{\alpha}^\star_Y(B)V(R) \geq \int^{cav} Y d\tilde{V}$$

[24] 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.

[25] 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.

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 $\int^{cav} Y d\tilde{V} = \int^{cav} Y dV$. $\square$

A.8 Proof of Proposition 2

Proof. The proof uses an intermediate result from the proof of Proposition 1 by which there exist m vectors $X^j \in C(\bar{G}_j)$ such that $X = \sum_{j=1}^m X^j$ if and only if every $Y \in \mathbb{R}^n_+$ satisfies $\sum_{j=1}^{m} \int^{cav} Y dv_j \leq X^T \cdot Y$. Since the monotonic covers of the raw data sets in a characteristic function form are capacities, by Proposition 1, $X \in C(\bar{G}^{\tilde{\mathscr{V}}})$ if and only if every random variable $Y \in R^n_+$ satisfies $\sum_{\tilde{V}_j \in \tilde{\mathscr{V}}} \int^{cav} Y d\tilde{V}_j \leq Y \cdot X$. By Lemma 5, $\tilde{V}(\Omega) = V(\Omega)$, $C(G) = C(\tilde{G})$ and for every $Y \in \mathbb{R}^n_+ : \int^{cav} Y dV = \int^{cav} Y d\tilde{V}$. Therefore, $X \in C(\bar{G}^{\mathscr{V}})$ if and only if every random variable $Y \in R^n_+$ satisfies $\sum_{V_j \in \mathscr{V}} \int^{cav} Y dV_j \leq Y \cdot X$. $\square$

B Uniform Optimal Decomposition across Practitioners

We say that a set of raw data sets $\mathscr{V}$ satisfies the property of **Uniform Optimal Decomposition across Practitioners** if for every random variable $Y \in R^n_+$, for every event $B \subseteq \Omega$ and for every pair of data sets $v, v' \in \mathscr{V}$: $\alpha^{\star}_{Y,v}(B) = \alpha^{\star}_{Y,v'}(B)$.

Verifying whether a set of raw data sets satisfies Uniform Optimal Decomposition across Practitioners is not always straightforward. However, in certain cases, this property is known to hold. For example, Lovász (1983) demonstrates that if all raw data sets are convex, their optimal decompositions are identical.

The following proposition shows that if the set $\mathscr{V}$ satisfies Uniform Optimal Decomposition across Practitioners, then raw data set auditors need only verify the condition stated in Proposition 2 for χ, the set of indicator vectors (see Footnote 14), a finite subset of random variables Y, rather than for all random variables $Y \in R^n_+$.

Proposition 3. *Let $X \in \mathbb{R}^n_+$ be an aggregated frequency distribution on $T = \sum_{i=1}^m T_i$ observations for the set of m data sets $\mathscr{D}$. Suppose that $\mathscr{V}$ satisfies Uniform Optimal Decomposition across Practitioners. $X \in C(\bar{G}^{\mathscr{V}})$ if and only if every random variable $Y \in \chi$ satisfies $\sum_{V_j \in \mathscr{V}} \int^{cav} Y dV_j \leq Y \cdot X$*

The proof of Proposition 3 depends on Proposition 2, which asserts that if $X \notin C(\bar{G}^{\mathscr{V}})$, then there exists a random variable Y, whose sum of the optimal decompositions across the respective data sets violates the inequality stated in Proposition 2. In the current case, these optimal decompositions, which are identical for all data sets, due to Uniform Optimal Decomposition across Practitioners, can be expressed as the same weighted sums of indicator vectors for all data sets. It then follows that at least one of these indicator vectors, when examined in isolation, violates the inequality in Proposition 3.

Proof. One direction is trivial: If $X \in C(\bar{G}^{\mathscr{V}})$ then by Proposition 2, every random variable $Y \in \chi$ satisfies $\sum_{V_j \in \mathscr{V}} \int^{cav} Y dV_j \leq Y \cdot X$.

For the other direction we assume that $X \notin C(\bar{G}^{\mathscr{V}})$ and show that there exists $B \subseteq \Omega$ such that $\sum_{V_j \in \mathscr{V}} \int^{cav} \chi^B dV_j > \chi^B \cdot X$.

If $X \notin C(\bar{G}^{\mathscr{V}})$ then by Proposition 2 there exists $Y \in R^n_+$ such that $\sum_{V_j \in \mathscr{V}} \int^{cav} Y dV_j > Y \cdot X$.

That is, there exists $Y \in R^n_+$ such that $\sum_{V_j \in \mathscr{V}} \sum_{B \in 2^\Omega} \alpha^\star_{Y,V_j}(B) V_j(B) > Y \cdot X$. According to Remark 1.1, $V_j(B) = \int^{cav} \chi^B dV_j$ for all $B \in 2^\Omega$ and for all $V_j \in \mathscr{V}$ for which $\alpha^\star_{Y,V}(B) > 0$. Consequently, $\sum_{V_j \in \mathscr{V}} \sum_{B \in 2^\Omega} \alpha^\star_{Y,V_j}(B) \int^{cav} \chi^B dV_j > Y \cdot X$. Since the raw data sets satisfy Uniform Optimal Decomposition across Practitioners we have $\sum_{B \in 2^\Omega} \alpha^\star_Y(B) \sum_{V_j \in \mathscr{V}} \int^{cav} \chi^B dV_j > Y \cdot X$, and since $\sum_{B \in 2^\Omega} \alpha^\star_Y(B) \chi^B = Y$ we obtain $\sum_{B \in 2^\Omega} \alpha^\star_Y(B) \sum_{V_j \in \mathscr{V}} \int^{cav} \chi^B dV_j > \sum_{B \in 2^\Omega} \alpha^\star_Y(B) \chi^B \cdot X$. For this inequality to hold true there must be a $B \subseteq \Omega$ such that $\alpha^\star_{Y,V}(B) > 0$ for which $\sum_{V_j \in \mathscr{V}} \int^{cav} \chi^B dV_j > \chi^B \cdot X$ which concludes the proof. $\square$