



Columbia University

*Department of Economics
Discussion Paper Series*

Estimating Ambiguity

Marc Henry

Discussion Paper #:0102-21

***Department of Economics
Columbia University
New York, NY 10027***

March 2002

Estimating ambiguity

Marc Henry¹

Columbia University

Abstract

We propose a measure of the degree of ambiguity associated with a belief function and a nonparametric method to estimate it. The degree of ambiguity associated with a belief function is measured by the Kullback-Leibler diameter of the set of probability measures compatible with it. It is shown that an estimator based on the empirical version of the unambiguous measure generating the belief function is consistent for the true value of the ambiguity measure. Applications to policy decision making under Knightian uncertainty are discussed.

KEYWORDS: Knightian uncertainty; belief functions; entropy; nonparametric methods, Prohorov metric, core diameter.

1 Introduction

The immense success of the von Neumann-Morgenstern-Savage decision theory paradigm (see Neumann and Morgenstern (1947) and Savage (1954)) can be partly attributed to the essential reduction of all situations of uncertainty to situations of risk with known probabilities on events. Henceforth, we shall refer to “Knightian uncertainty” (Knight (1921)) or “ambiguity” to describe situations in which no objective probability measure on the event space is available to the decision maker. Savage posits that subjective beliefs agents base their decisions upon can be represented by subjective additive probability measures. However, since Ellsberg’s famous experiment (Ellsberg (1961)) uncovered preferences which cannot be supported through expected utility maximization by a single probability measure on events, the framework was modified along several directions, including the representation of beliefs by Choquet capacities (see Cohen and Tallon (2000) for a survey of the use of non-additive belief representations in “non-expected utility” models of choice under uncertainty).

Let \mathcal{A} be a σ -algebra of events on the set Ω . To set the framework of the present paper in context, it is useful to consider the following hierarchy of representations of beliefs on (Ω, \mathcal{A}) . The set of Choquet *capacities* $\nu : \mathcal{A} \rightarrow [0, 1]$ such that $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$ and $\forall A, B \in \mathcal{A}, A \subset B \implies \nu(A) \leq$

¹This research was carried out while the author was visiting IRES, UCL, and financial support from the latter institution in the form of an FSR grant is gratefully acknowledged.

$\nu(B)$, contains the set of *lower probabilities*, which are characterized by super-additivity. *Lower envelopes* of some non-empty class of probability measures on (Ω, \mathcal{A}) are lower probabilities; special cases of lower envelopes are *2-monotone* or *convex* capacities characterized by $\nu(A) + \nu(B) \leq \nu(A \cap B) + \nu(A \cup B)$, and finally, *belief functions*, described in section 2 below, are convex capacities themselves. We therefore have the hierarchy: ν is a probability measure $\implies \nu$ is a belief function $\implies \nu$ is a convex capacity $\implies \nu$ is a lower envelope $\implies \nu$ is a lower probability $\implies \nu$ is a Choquet capacity (see Walley and Fine (1982)). In all cases, the set of dominating probability measures (which may be empty) will be called the *core* of the belief representation.

The main axiomatic extensions of the expected utility model to take into account ambiguity averse behaviour of agents are *Choquet expected utility* (Schmeidler (1989)) in which beliefs are represented by a Choquet capacity, and the multi prior model of Gilboa and Schmeidler (1989) where beliefs are represented by the lower envelope of a closed and convex set of probability measures. In the latter model, as in the non axiomatic dynamic extension of Epstein and Wang (1994), decision is based on a minmax criterion and therefore ambiguity aversion is indistinguishable from the degree of ambiguity underlying the belief representation.

It is the concept of ambiguity as departure from epistemic determinacy in the form of an objective probability measure on events that we are concerned with in this work. It is useful to consider a situation of ontological determinacy and to think of the epistemic indeterminacy as *scientific uncertainty*. A defense of interval valued probabilities in this framework can be based on the practice of reporting scientific predictions on events (particularly earth-science related) as intervals of values produced by different calibrations of deterministic physical models.

We propose here an index of ambiguity for metrizable and separable event spaces (heuristically, this index is equal to the diameter of the core) with the aim of characterizing *scientific progress* on physical systems that impact policy, based on the evolution of the index of ambiguity. Dow and Werlang (1994) also consider the issue of the resolution of epistemic indeterminacy, whereas Walley and Fine (1982) and Marinacci (1999) propose limit laws in a frequentist setup which allows for ontological indeterminacy as well. Ontological determinacy is implicit in the work of Hansen and Sargent (2000) who apply the minimax approach on a set of contiguous stochastic processes in a departure from the rational expectations framework. They use a notion of *fear of misspecification* or *taste for robustness* as their concept of ambiguity aversion (which again is not clearly distinct from ambiguity itself). Ontological determinacy is very explicit in Brock and Durlauf (2000) as they advocate a Bayesian averaging rule in a similar policy decision setting.

The next section discusses belief functions in detail and describes the Kullback-Leibler relative entropy diameter used as an index of ambiguity. The reason for the focus on belief functions is the statistical procedure employed to estimate the relative entropy diameter, which relies on random draws from the probability space generating the belief function. The reasons for the use of relative

entropy instead of a proper metric, such as the Prohorov metric of the bounded Lipschitz metric which both metrize the weak topology on the set of probability measures on (Ω, \mathcal{A}) (when Ω is metrizable and separable) are the invariance properties of relative entropy and feasibility of calculation and estimation. The nonparametric estimation procedure is described in section 3 with a discussion of its implementation and potential applications. Proofs are collected in the appendix.

2 Measuring ambiguity

Let Ω be a Polish space with Borel σ -algebra \mathcal{B} and call \mathcal{M} the space of all probability measures on (Ω, \mathcal{B}) . Consider a compact, convex metrizable subset Y of a locally convex topological vector space with Borel σ -algebra \mathcal{B}_Y , and let p be a probability measure on (Y, \mathcal{B}_Y) . Finally, let F be a strongly measurable multivalued mapping taking points in Y onto closed non-empty subsets of Ω .

For all $S \subset \Omega$, we define the Dempster variate, or belief function (see Dempster (1967)), generated on (Ω, \mathcal{B}) by (Y, \mathcal{B}_Y, p, F) in the following way. Define

$$\begin{aligned} S^* &= \{y \in Y \mid F(y) \cap S \neq \emptyset\} \\ S_* &= \{y \in Y \mid F(y) \subset S\}. \end{aligned}$$

The belief function p_* is defined by $p_*(S) = p(S_*)$, and the plausibility function by $p^*(S) = p(S^*)$. The belief (resp. plausibility) function corresponds to the smallest (resp. largest) reliability that can be attached to an event S . They satisfy $p^*(S) \geq p_*(S)$ for all S , with equality if and only if the belief function is a probability measure. Finally, define the set of probability measures on (Ω, \mathcal{B}) compatible with the belief function p_* as

$$\mathcal{C} = \{\nu \in \mathcal{M} \mid \forall S \in \mathcal{B}, p_*(S) \leq \nu(S) \leq p^*(S)\}.$$

We make the following assumptions:

Assumption (i): p_* is almost positive (i.e. $p_*(S) = 0$ implies $p^*(S) = 0$).

Under assumption (i), the all measures in the set \mathcal{C} are absolutely continuous with respect to each other. For two elements ν and ν' in \mathcal{C} , denote by $d\nu/d\nu'$ or f , the Radon-Nikodym derivative of ν with respect to ν' .

We make the further assumption below:

Assumption (ii): $\forall(\nu, \nu') \in \mathcal{C}^2$, $d\nu/d\nu'$ is continuous on $F(Y)$,

and define the relative entropy of measure ν with respect to ν' on S as

$$I_{\nu, \nu'}(S) \equiv \frac{1}{\nu'(S)} \int_S \log \frac{d\nu}{d\nu'}(\omega) \nu'(d\omega),$$

when $\nu'(S) \neq 0$ and zero otherwise. $I_{\nu,\nu'}(\Omega) \geq 0$ with equality if and only if $\nu = \nu'$, and it can be symmetrized by $I_{\nu,\nu'}(\Omega) + I_{\nu',\nu}(\Omega)$. However, it does not satisfy the triangular inequality, and is therefore not a metric. It is used as a measure of information for discriminating between competing hypotheses, as it is invariant in the sense that it is decreased through a measurable transformation of the probability spaces $(\Omega, \mathcal{B}, \nu)$ and $(\Omega, \mathcal{B}, \nu')$ with equality if and only if the transformation is a sufficient statistic (see Kullback and Leibler (1951)).

We shall therefore use this Kullback-Leibler contrast to define an index of ambiguity on (Ω, \mathcal{B}) in view of the following lemma.

Lemma 1: Under assumptions (i) and (ii),

$$\mathcal{A}(F, p) \equiv \sup_{(\nu, \nu') \in \mathcal{C}^2} I_{\nu, \nu'}(\Omega) < +\infty.$$

In view of lemma 1, we define an index of ambiguity implicit in the pair (F, p) as the ‘‘Kullback-Leibler diameter’’ of \mathcal{C} , i.e. $\mathcal{A}(F, p)$. As noted in the introduction, in a minmax decision framework, this diameter serves also as a measure of ambiguity aversion on the part of the decision maker.

3 Estimating ambiguity

Consider a sample $\{y_i\}_{i=1}^n$ of i.i.d. uniform random variables on (Y, \mathcal{B}_Y, p) as the result of an experiment with known link F to the measurable space of interest (Ω, \mathcal{B}) . The problem considered here is the estimation of the index of ambiguity induced on (Ω, \mathcal{B}) by (Y, \mathcal{B}_Y, p, F) .

A rationale for setting the problem in this way, and particularly for assuming prior knowledge of the mapping F can be constructed from the following example (presented in schematic form): suppose the ‘‘experiment’’ consists in n independent small scale introductions of a genetically modified corn seed in n similar controlled ecological environments. Such an experiment would make little sense if it weren’t carried out with clear prior knowledge of the link between small scale introduction in controlled environments and large scale implementation for agricultural purposes. An elementary event in the the controlled environment, say the appearance of a genetic modification in an insect, would naturally be linked to a composite event, such as the appearance of a collection of related genetic modifications in a family of insects ‘‘genetically close’’ to the former.

Consider first the problem of estimating relative entropy of ν with respect to ν' , where ν and ν' are two elements of \mathcal{C} , from a sample of hypothetical i.i.d. random variables $\{X_i\}_{i=1}^n$ distributed according to ν on the probability space $(\Omega, \mathcal{B}, \nu')$. Ahmad and Lin (1976) propose a nonparametric estimator of entropy for absolutely continuous density functions with respect to Lebesgue measure on

the real line, and Robinson (1991) extends it to relative entropy in a Euclidian space. However, the topological vector spacial structure is not necessary, nor is a particular metric, and we can construct an entropy estimator from $\{X_i\}_{i=1}^n$ as follows:

Denote by $\hat{f}_{\nu, \nu'}$ a histogram estimator of $d\nu/d\nu'$ (described below) based on $\{X_i\}_{i=1}^n$, and construct an estimator of $I_{\nu, \nu'}(\Omega)$ in the form:

$$\hat{I}_{\nu, \nu'} \equiv \int_{\Omega} \ln \hat{f}_{\nu, \nu'}(\omega) \nu_n(d\omega) = \frac{1}{n} \sum_{i=1}^n \ln \hat{f}_{\nu, \nu'}(X_i),$$

where ν_n is the empirical measure $(1/n) \sum_{i=1}^n \delta_{X_i}$ (δ denoting dirac measure).

For the construction of $\hat{f}_{\nu, \nu'}$ (denoted \hat{f} in what follows), we make the following assumption:

Assumption (iii): $F(Y)$ is compact.

Remark: Hall (1990) shows that properties of kernel estimators of entropy of the type proposed by Ahmad and Lin (1976), Robinson (1991) and others, depends crucially on the tail behaviour of the density. In particular, they show that in Euclidian spaces, \sqrt{n} -consistency requires drastic conditions on tail behaviour and/or excess smoothness (to apply bias-reducing techniques such as higher-order kernels as in Robinson (1991)) for any dimension higher than 1 for histogram density estimates (which we use here) and 3 for more general kernel estimates (where large kernel tails are needed to offset the effect of large tails in the density).

Of course, we do not try to achieve \sqrt{n} -consistency here, as it would require moment and smoothness conditions on f , which we are trying to avoid as they are difficult to relate to the mapping F . However, we assume compactness of $F(Y)$ and continuity of f to avoid clouding the central issue with tail behaviour considerations.

Under assumption (iii), let $\{C_j^n\}_{j=1}^{k(n)}$ be a measurable partition of $F(Y)$ such that the following two conditions are satisfied:

Assumption (iv): There exists positive constants c_1 and c_2 such that

$$\frac{c_1}{k(n)} \leq \min_{j \leq k(n)} \nu'(C_j^n) \leq \max_{j \leq k(n)} \nu'(C_j^n) \leq \frac{c_2}{k(n)}$$

Assumption (v): $k(n)^{-1} + n^{-1}k(n) = o(1)$.

Now let \hat{f} be defined on $F(Y)$ as

$$\hat{f}(\omega) \equiv \frac{1}{n} \sum_{\ell=1}^n \sum_{j=1}^{k(n)} \delta_{X_\ell}(C_j^n) \delta_\omega(C_j^n) / \nu'(C_j^n),$$

and we now state

Lemma 2: Under assumptions (i) to (v),

$$\hat{I}_{\nu, \nu'} \rightarrow_{\nu'} I_{\nu, \nu'}(\Omega).$$

The link between the probability space $(\Omega, \mathcal{B}_Y, p)$ and the elements of \mathcal{C} is provided intuitively by Dempster's characterization of $\nu \in \mathcal{C}$ by the existence of a set of probability kernels indexed by $y \in Y$ and with support $F(y)$ on (Ω, \mathcal{B}) .

This prompts the construction of analogues of the unachievable random variables $\{X_i\}_{i=1}^n$ in the pair of n -uples of elements of Ω , $\{\omega_{i*}\}_{i=1}^n$ and $\{\omega_i^*\}_{i=1}^n$, with $(\omega_{i*}, \omega_i^*) \in F(y_i)^2$. For each n -uple, an empirical density is constructed with the slightly modified assumption:

Assumption (iv'): There exists positive constants c_1 and c_2 such that

$$\frac{c_1}{k(n)} \leq \min_{j \leq k(n)} p_*(C_j^n) \leq \max_{j \leq k(n)} p^*(C_j^n) \leq \frac{c_2}{k(n)}.$$

Remark: Note that assumption (iv') implies assumption (iii) for all ν' in \mathcal{C} .

We denote by \hat{f}_* and \hat{f}^* the empirical density estimators constructed from $\{\omega_{i*}\}_{i=1}^n$ and $\{\omega_i^*\}_{i=1}^n$ respectively. More precisely:

$$\hat{f}_*(\omega) \equiv \frac{1}{n} \sum_{\ell=1}^n \sum_{j=1}^{k(n)} \delta_{\omega_\ell}(C_j^n) \delta_\omega(C_j^n) / p_*(C_j^n),$$

$$\hat{f}^*(\omega) \equiv \frac{1}{n} \sum_{\ell=1}^n \sum_{j=1}^{k(n)} \delta_{\omega_\ell}(C_j^n) \delta_\omega(C_j^n) / p^*(C_j^n).$$

Finally, call $\hat{\mathcal{A}}_n(F, p)$ the proposed estimator for the Kullback-Leibler diameter $\mathcal{A}_n(F, p)$ of \mathcal{C} in \mathcal{M} , defined as:

$$\hat{\mathcal{A}}_n(F, p) = \max_{\substack{(\omega_{i*}, \omega_i^*) \in F(y_i) \\ i=1, \dots, n}} \sum_{i=1}^n \log \frac{f^*(\omega_i^*)}{f_*(\omega_{i*})}.$$

We can now state the main result which is an immediate consequence of Theorem 2.1 of Wasserman (1990) and Lemmata 1 and 2 above:

Theorem 1: $\hat{\mathcal{A}}_n(F, p) \rightarrow_p \mathcal{A}(F, p)$.

Theorem 1 is a weak result, due mostly to the degree generality of the setting, and more precise asymptotic results (on the rate of convergence and

possible limiting distribution) would be needed to infer comparisons on the informativeness of different experiments. However, such results will necessarily entail smoothness assumptions on the densities of measures within the core of the belief function, and therefore will be highly “ F -specific”. Naturally, implementation of the estimator will rely on algorithms which are also F -specific, so that the present note should be mostly regarded as a blueprint for the modeling of scientific uncertainty in policy decisions, the modeling of its evolution over time (or “learning”) and the definition of a precautionary approach in decision making.

References

- ABOU-JAOUDE, S. (1976): “Conditions nécessaires et suffisantes de convergence L_1 en probabilité de l’histogramme pour une densité,” *Annales de l’Institut Henri Poincaré, Section B*, 12, 213–231.
- AHMAD, I. A., and P. LIN (1976): “A nonparametric estimation of the entropy for absolutely continuous distributions,” *IEEE Transactions in Information Theory*, IT-22, 372–375.
- BROCK, W. A., and S. N. DURLAUF (2000): “Growth economics and reality,” preprint.
- COHEN, M., and J.-M. TALLON (2000): “Décision dans l’incertain: l’apport des modèles non additifs,” *Revue d’économie politique*, forthcoming.
- DEMPSTER, A. P. (1967): “Upper and lower probabilities induced by a multi-valued mapping,” *Annals of Mathematical Statistics*, 38, 325–339.
- DOW, J., and S. WERLANG (1994): “Learning under Knightian uncertainty: the Law of Large Numbers for non-additive probabilities,” unpublished manuscript.
- ELLSBERG, D. (1961): “Risk, ambiguity and the Savage axioms,” *Quarterly Journal of Economics*, 75, 643–669.
- EPSTEIN, L. G., and T. WANG (1994): “Intertemporal asset pricing under Knightian uncertainty,” *Econometrica*, 62, 283–322.
- GILBOA, I., and D. SCHMEIDLER (1989): “Maximin expected utility with non-unique priors,” *Journal of Mathematical Economics*, 18, 141–153.
- HALL, P. (1990): “Akaike’s information criterion and Kullback-Leibler loss for histogram density estimation,” *Probability Theory and Related Fields*, 85, 449–467.
- HANSEN, L. P., and T. SARGENT (2000): “Wanting robustness in macroeconomics,” <http://www.stanford.edu/~sargent>.

- HUBER, P. J. (1981): *Robust statistics*. New York: Wiley.
- HUBER, P. J., and V. STRASSEN (1973): “Minimax tests and the Neyman-Pearson lemma for capacities,” *Annals of Statistics*, 1, 251–263.
- KNIGHT, F. H. (1921): *Risk, Uncertainty and Profit*. New York: Kelley.
- KULLBACK, S., and R. A. LEIBLER (1951): “On information and sufficiency,” *Annals of Mathematical Statistics*, 22, 79–86.
- MARINACCI, M. (1999): “Limit laws for non-additive probabilities and their frequentist interpretation,” *Journal of Economic Theory*, 84, 145–195.
- NEUMANN, J. v., and O. MORGENSTERN (1947): *Theory of games and economic behaviour*. Princeton: Princeton University Press.
- ROBINSON, P. M. (1991): “Consistent nonparametric entropy-based testing,” *Review of Economic Studies*, 58, 437–453.
- SAVAGE, L. (1954): *The foundations of statistics*. New York: Wiley.
- SCHMEIDLER, D. (1989): “Subjective probability and expected utility without additivity,” *Econometrica*, 57, 571–587.
- VARADARAJAN, V. S. (1958): “On the convergence of sample probability distributions,” *Sankhyā, Series A*, 19, 23–26.
- WALLEY, P., and T. L. FINE (1982): “Towards a frequentist theory of upper and lower probability,” *Annals of Statistics*, 10, 741–761.
- WASSERMAN, L. (1990): “Prior envelopes based on belief functions,” *Annals of Statistics*, 18, 454–464.

Appendix

Proof of lemma 1: The belief function p_* is a monotone Choquet capacity of infinite order, and \mathcal{M} is metrizable with the Prohorov metric and is also Polish (see for instance Theorem 3.9 in Huber (1981)); hence, by Lemma 2.2 of Huber and Strassen (1973), \mathcal{C} is a compact subset of \mathcal{M} . The result follows from the continuity of $I_{\nu, \nu'}(\Omega)$ under assumptions (i) and (ii). Q.E.D.

Proof of lemma 2: In the proof, we shall drop the subscript for the Radon-Nikodym derivative and its empirical counterpart. Noting that ν_n is absolutely continuous with respect to ν and ν' , we have

$$\begin{aligned}
 |\hat{I}_{\nu, \nu'} - I_{\nu, \nu'}(\Omega)| &= \left| \int_{\Omega} \ln f(\omega) \nu_n(d\omega) - \int_{\Omega} \ln f(\omega) \nu(d\omega) \right| \\
 &\leq \left| \int_{\Omega} \ln(f(\omega)/f(\omega)) \nu_n(d\omega) \right| \\
 &\quad + \left| \int_{\Omega} \ln f(\omega) \nu_n(d\omega) - \int_{\Omega} \ln f(\omega) \nu(d\omega) \right| \\
 &= A + B.
 \end{aligned}$$

ν_n converges weakly to ν with ν' -probability 1 by Theorem 3 of Varadarajan (1958) and f is continuous on a compact subset of Ω , therefore, $B \rightarrow 0$ with ν' -probability 1. Note that assumption (ii) that Ahmad and Lin (1976) used for convergence of first moment in B is not needed here.

Now, calling N_j the number of X_i 's in bin C_j^n , and define

$$p_j = \int_{C_j^n} f(\omega) \nu'(d\omega) = \frac{1}{n} E_{\nu'}(N_j),$$

we can write A as

$$A = \frac{1}{n} \sum_{j \in J_n} N_j \ln(N_j/n p_j) + \frac{1}{n} \sum_{i=1}^n \ln(E_{\nu'} f(X_i)/f(X_i)) = A_1 + A_2,$$

where J_n is the set of indexes of non-empty bins. Take $\eta > 0$, and consider the partition of J_n in $J_1 = \{j | n p_j > n^\eta\}$ and $J_2 = J_1^c$. We can write

$$n|A_1| \leq \left| \sum_{j \in J_1} N_j \ln(N_j/n p_j) \right| + \left| \sum_{j \in J_2} N_j \ln(N_j/n p_j) \right| = S_1 + S_2.$$

Now, for η small enough, $S_2 = o_{\nu'}(k(n))$, whereas by differentiability of the logarithm, we have

$$S_1 \leq K \sum_{j \in J_1} |N_j(N_j - n p_j)|/n p_j \leq K n^{-\eta} \sum_{j \in J_1} N_j |N_j - n p_j|$$

where K is some positive constant, and

$$\sum_{j \in J_1} N_j |N_j - n p_j| \leq \sum_i^n |\hat{f}(X_i) - E_{\nu'} f(X_i)| = o_{\nu'}(n)$$

by proposition 6 of Abou-Jaoudé (1976). By a similar argument and proposition 2 of Abou-Jaoudé (1976), the bias term A_2 also converges, whence the theorem. Q.E.D.

Proof of Theorem 1: Consider a pair (ν, ν') of mesures in \mathcal{C} . For p -almost every y in Y , there exists, by Theorem 2.1 of Wasserman (1990), a probability measure π_y on \mathcal{B} with support $F(y)$ such that

$$\forall S \in \mathcal{B} \quad \nu(S) = \int_Y \pi_y(S) dp(y).$$

Call π'_y the probability measures corresponding to ν' and satisfying the same property. Samples $\{\omega_i\}_{i=1}^n$ and $\{\omega'_i\}_{i=1}^n$ can be drawn in Ω according to ν and ν' using this result. First draw y randomly in Y and then draw ω in $F(y)$ according to π_y and ω' in $F(y)$ according to π'_y . By construction, samples of size n so generated are such that,

$$\sum_{i=1}^n \log \frac{f_n(\omega_i)}{f'_n(\omega'_i)} \leq \hat{A}_n(F, p),$$

where $f_n(\omega_i)$ and $f'_n(\omega'_i)$ are defined from the ω_i and ω'_i in the same way as f^* above. Conversely, for any choice ω_i^* and ω_{i*} , $i = 1$ to n , there exists mesures ν and ν' in \mathcal{C} such that the sets

$$\begin{aligned} S &= \left\{ \omega \in \Omega \mid \sum_{i=1}^n \log \frac{f^*(\tilde{\omega}_i)}{f_*(\omega_{i*})} \leq \sum_{i=1}^n \log \frac{f^*(\omega_i^*)}{f_*(\omega_{i*})} \right\} \\ \text{and } S' &= \left\{ \omega \in \Omega \mid \sum_{i=1}^n \log \frac{f^*(\omega_i^*)}{f_*(\tilde{\omega}_i)} \geq \sum_{i=1}^n \log \frac{f^*(\omega_i^*)}{f_*(\omega_{i*})} \right\} \end{aligned}$$

where the $\tilde{\cdot}$ indicates that for any one index i , ω_{i*} and ω_i^* are replaced by ω , satisfy $\nu(S) = \nu'(S') = 0$. The result follows from lemma 2. Q.E.D.