CIRJE-F-113

A Note on Learning under the Knightian Uncertainty

Kiyohiko G. Nishimura
University of Tokyo

Hiroyuki Ozaki
Tohoku University

April 2001

"A Note on Learning under the Knightian Uncertainty"

K. G. Nishimura and H. Ozaki

Abstract

In contrast to the traditional model of uncertainty, where the uncertainty is characterized by a single distribution function that a decision maker faces, the Knightian-uncertainty approach characterizes it as a set of distributions rather than a single one. Hence, learning in the context of Knightian uncertainty is characterized by an update process of the set of distributions after each of random sampling.

This note presents two examples in which the Dempster-Shafer update rule, the one which attracts much attention since it seems intuitive, does not at all reduce the Knightian uncertainty (Example 1) and it actually increases the Knightian uncertainty (Example 2). Thus, what is a sensible update process is still an open question under the Knightian uncertainty.

A NOTE ON LEARNING UNDER THE KNIGHTIAN UNCERTAINTY*

by
Kiyohiko G. Nishimura

Faculty of Economics
The University of Tokyo

and

Hiroyuki Ozaki

Faculty of Economics
Tohoku University

April 15, 2001

Abstract

In contrast to the traditional model of uncertainty, where the uncertainty is characterized by a single distribution function that a decision maker faces, the Knightian-uncertainty approach characterizes it as a set of distributions rather than a single one. Hence, learning in the context of Knightian uncertainty is characterized by an update process of the set of distributions after each of random sampling. This note presents two examples in which the Dempster-Shafer update rule, the one which attracts much attention since it seems intuitive, does not at all reduce the Knightian uncertainty (Example 1) and it actually increases the Knightian uncertainty (Example 2). Thus, what is a sensible update process is still an open question under the Knightian uncertainty.

*The first author's work is partially supported by a grant to the Project of Information Technology and the Market Economy, from the Japan Society for Promotion of Sciences.

1. Introduction

In a traditional framework, uncertainty is characterized by a single distribution function that a decision maker faces. Many economic applications assume that the distribution function is known to the decision maker. In a more general framework, the distribution function is not known but the decision maker tries to learn from observations taken from the distribution. As in Rothschild (1974) and others, uncertainty is typically characterized by a parametric family of distributions with unknown parameters, say, a normal distribution with an unknown mean and a known variance. Then, learning is characterized by an update process of a prior distribution over the space of unknown parameters after each of random sampling. The conjugate distribution is usually assumed for a prior to guarantee the updated distribution to be in the same parametric family, which reduces mathematical complexity greatly.

In contrast to the traditional approach, the Knightian-uncertainty approach characterizes uncertainty as a set of distributions, instead of a single distribution. The distributions are not necessarily in the same parametric family. Not only the parameter value of the distribution is unknown, but also the shape of distribution itself is unknown. Hence, learning is characterized by an update process of the set of distributions after each of random sampling. Among update rules under the Knightian uncertainty, the maximum-likelihood update rule, which is often called the Dempster-Shafer rule, attracts much attention since it seems intuitive and sensible. This rule chooses, among all distributions in the set, those that put the highest probability on the occurrence of an actual observation, and updates the chosen distributions by using the Bayes rule. By using the Dempster-Shafer rule, learning is expected to reduce the Knightian uncertainty, in the sense that the set of distributions that the decision maker faces shrinks after each observation.

This note presents two counter-examples of this popular conviction. In particular, we show that the Dempster-Shafer update rule does not at all reduce the Knightian uncertainty (Example 1) and it actually increases the Knightian uncertainty (Example 2). Thus, what is a sensible update process is still an open question under the Knightian uncertainty.

2. Update Rules under Multi-Period Knightian Uncertainty

2.1. Preliminaries: Multi-Period Knightian Uncertainty

Let $(W, \mathcal{B}_W)$ be a measurable space, where W is a Borel subset of $\mathbb{R}_+$ and $\mathcal{B}_W$ is the Borel σ-algebra on W. Denote the set of all probability measures on $(W, \mathcal{B}_W)$ by $\mathcal{M}(W)$. Given $\mathcal{Q}$, $\mathcal{R} \subseteq \mathcal{M}(W)$, the *product* of $\mathcal{Q}$ and $\mathcal{R}$, denoted $\mathcal{Q} \times \mathcal{R}$, is defined as

$$\mathcal{Q} \times \mathcal{R} = \{ \mu \times \nu \mid \mu \in \mathcal{Q} \text{ and } \nu \in \mathcal{R} \}$$

where $\mu \times \nu$ denotes the product probability measure constructed from μ and ν.

We consider a two-period model. As is apparent in the following analysis, to extend to a n-period model with $n > 2$ is straightforward (though notationally cumbersome). We thus consider $(W \times W, \mathcal{B}_{W \times W})$, which is the (self-)product measurable space constructed from $(W, \mathcal{B}_W)$. As in the one-period model, we denote the set of all probability measures on $(W \times W, \mathcal{B}_{W \times W})$ by $\mathcal{M}(W \times W)$. Obviously, for $\mathcal{Q}$ and $\mathcal{R} \subseteq \mathcal{M}(W)$, we have $\mathcal{Q} \times \mathcal{R} \subseteq \mathcal{M}(W \times W)$.

In characterizing the Knightian uncertainty, we follow in this note the multiple prior approach (see Gilboa and Schmeidler (1993) for this and an alternative non-additive probability approach). In this framework, the decision maker's state of the *Knightian uncertainty* is represented by a subset of $\mathcal{M}(W)$ in the case of one period, and a subset of $\mathcal{M}(W \times W)$ in the case of two periods.

One technical definition is useful in the following analysis. Given any $\pi \in \mathcal{M}(W \times W)$ and $X \in \mathcal{B}_W$, a probability measure $\pi_X \in \mathcal{M}(W)$ is defined as

$$\pi_X(Y) = \pi(X \times Y)$$

for all $Y \in \mathcal{B}_W$.

Let us now define the *marginal Knightian uncertainty* of two-period Knightian uncertainty. Then, given two-period Knightian uncertainty $\mathcal{P} \subseteq \mathcal{M}(W \times W)$, its *second-period marginal Knightian uncertainty*, denoted by $\mathcal{P}|_2$, is defined as

$$\mathcal{P}|_2 = \{ \nu \in \mathcal{M}(W) \mid (\exists \pi \in \mathcal{P})\ \nu = \pi_W \}.$$

If $\mathcal{P} = \mathcal{Q} \times \mathcal{R}$ for some $\mathcal{Q}$ and $\mathcal{R}$, then $\mathcal{P}|_2 = \mathcal{R}$. Also we have that $\mathcal{M}(W \times W)|_2 = \mathcal{M}(W)$. This relation is proved in the following way. First, the inclusion $\mathcal{M}(W \times W)|_2 \subseteq \mathcal{M}(W)$ is obvious. Second, to see that the opposite inclusion also holds, let $\nu \in \mathcal{M}(W)$ and $\mu \in \mathcal{M}(W)$ and define $\pi \in \mathcal{M}(W \times W)$ by $\pi = \mu \times \nu$. Then, $(\forall X \in \mathcal{B}_W)\ \nu(X) = \mu(W)\nu(X) = \pi(W \times X) = \pi_W(X)$, and hence, $\nu \in \mathcal{M}(W \times W)|_2$.

2.2. Update and Dempster-Shafer Rule

In general, an *updating rule* in the two-period framework is a function ϕ such that

$$\phi : 2^{\mathcal{M}(W \times W)} \times \mathcal{B}_{W \times W} \to 2^{\mathcal{M}(W \times W)}$$

where $2^{\mathcal{M}(W \times W)}$ denotes a set consisting of all subsets of $\mathcal{M}(W \times W)$. Given two-period Knightian uncertainty $\mathcal{P} \subseteq \mathcal{M}(W \times W)$ and an observation $A \in \mathcal{B}_{W \times W}$, $\phi(\mathcal{P}, A)$ $(\subseteq \mathcal{M}(W \times W))$ represents the state of Knightian uncertainty which is updated from $\mathcal{P}$ upon the observation of the event A according to the updating rule ϕ.

A natural extension of Bayesian learning in the framework of unique prior to the multiple prior case is so-called *Dempster-Shafer rule* or the maximum-likelihood update rule.[1] Given $\mathcal{P} \subseteq \mathcal{M}(W \times W)$ and $A \in \mathcal{B}_{W \times W}$, the *Dempster-Shafer rule*, $\phi_{DS}(\mathcal{P}, A)$, is defined as

$$\begin{aligned} & \pi' \in \phi_{DS}(\mathcal{P}, A) \\ \Leftrightarrow \quad & (\exists \pi^* \in \mathcal{P})\ \ \pi^* \in \arg\max\{\pi(A) \,|\, \pi \in \mathcal{P}\} \text{ and } (\forall B \in \mathcal{B}_{W \times W}) \ \ \pi'(B) = \frac{\pi^*(A \cap B)}{\pi^*(A)}\,. \end{aligned}$$

Here $\mathcal{P}$ and A are assumed to be such that $\max\{\pi(A) \,|\, \pi \in \mathcal{P}\}$ exists and is strictly positive.

When $\mathcal{P} = \{\pi\}$, that is, when there is no Knightian uncertainty, we immediately have $\phi_{DS}(\mathcal{P}, A) = \{\pi'\}$ where $\pi'(\cdot) = \pi(A \cap \cdot)/\pi(A)$ for any A such that $\pi(A) \neq 0$, which is the familiar Bayesian update rule. Hence, within the current setup of a two-period model where the event E is observed in the first period, if there is no Knightian uncertainty, the Dempster-Shafer rule corresponds to updating $\pi(W \times \cdot) \in \mathcal{M}(W)$ into $\pi(E \times \cdot)/\pi(E \times W) \in \mathcal{M}(W)$.

---

[1] Originally, the Dempster-Shafer rule is proposed in the non-additive measure approach. Gilboa and Schmeidler (1993) shows the equivalence of the Dempster-Shafer rule and the maximum-likelihood update rule for preferences which can be simultaneously represented by a non-additive measure and by multiple priors. Taking this property in mind, we use the term, the Dempster-Shafer rule, for the maximum-likelihood update rule in the text.

2.3. First-Period Observation and *Ex-Post* Knightian Uncertainty

We now embed the Dempster-Shafer update rule in our sequential two-period framework. Let $\mathcal{P} \in \mathcal{M}(W \times W)$ be the initial Knightian uncertainty over two periods. If there were no observation of an event in the first period, the second-period marginal Knightian uncertainty, $\mathcal{P}|_2$ would be the state of the Knightian uncertainty in the second period. However, we observed $E \in \mathcal{B}_W$ in the first period. The event of having only the first-period observation E can be characterized as $E \times W$ in the two-period framework of the previous section.. We then update our state of the Knightian uncertainty using the Dempster-Shafer update rule ϕ_{DS}. The updated second-period Knightian uncertainty, denoted by $\mathcal{P}|_2^{\phi_{DS}}(E)$, is defined as

$$\mathcal{P}|_2^{\phi_{DS}}(E) = \{\, \nu \in \mathcal{M}(W) \,|\, (\exists \pi \in \phi_{DS}(\mathcal{P}, E \times W))\ \nu = \pi_E \,\} \,.$$

Our main focus is whether learning (updating) reduces the degree of the Knightian uncertainty. If the updated Knightian uncertainty is "smaller" than the original Knightian uncertainty, that is, $\mathcal{P}|_2^{\phi_{DS}}(E) \subset \mathcal{P}|_2$, it is natural to say that the degree of the Knightian uncertainty is reduced. Thus, we are interested in whether $\mathcal{P}|_2^{\phi_{DS}}(E) \subset \mathcal{P}|_2$ holds or not.

3. Learning May Not Reduce Knightian Uncertainty: Two Examples

This section offers two examples where learning in the form of the Dempster-Shafer updating does not reduce the Knightian uncertainty. In the first example, learning does not change the degree of the Knightian uncertainty, while in the second example, learning actually *increases* the degree of the Knightian uncertainty.

3.1. Knightian Uncertainty Represented by a Product of Probability-Measure Sets

The first example is taken from a special case of a general search model under the Knightian uncertainty examined by Nishimura and Ozaki (2001). Consider an unemployed worker searching for a job. The worker has some idea about the distribution of job offers in the current period and the next, so that she is certain that the true distribution is in a small subset of all distributions. However, she neither have confidence about which distribution in the set of candidate distributions is the true one, nor that the true distribution is the same for two periods. This worker's Knightian uncertainty in two periods can be characterized as a product of a set of distribution functions in the current period and that in the next period. In this case, the following theorem and corollary show that the Dempster-Shafer update rule does not change the degree of the Knightian uncertainty.

Theorem 1. *Let* $\mathcal{Q}, \mathcal{R} \in \mathcal{M}(W)$ *and the observation* $E \in \mathcal{B}_W$ *be such that* $\max\{\mu(E) \,|\, \mu \in \mathcal{Q}\}$ *exists and is strictly positive. Then, if* $\mathcal{P} = \mathcal{Q} \times \mathcal{R}$*, we have* $\mathcal{P}|_2^{\phi_{DS}}(E) = \mathcal{R}$.

Proof. First note that

$$
\begin{aligned}
&\pi' \in \phi_{DS}(\mathcal{Q} \times \mathcal{R}, E \times W) \\
\Leftrightarrow\ & (\exists\, (\mu \times \nu)^* \in \arg\max\{\mu \times \nu(E \times W) \,|\, \mu \times \nu \in \mathcal{Q} \times \mathcal{R}\})\ \ \pi'(\cdot) = \frac{(\mu \times \nu)^*\,((E \times W) \cap \cdot)}{(\mu \times \nu)^*\,(E \times W)}
\end{aligned}
$$

Since $\mu \times \nu(E \times W) = \mu(E)\nu(W)$ and $\nu(W) = 1$ for all ν, we have $(\mu \times \nu)^* = \mu^* \times \nu$ where $\mu^* \in \arg\max\{\mu(E) \,|\, \mu \in \mathcal{Q}\})$ and $\nu \in \mathcal{R}$, and

$$
\frac{(\mu \times \nu)^*\,((E \times W) \cap \cdot)}{(\mu \times \nu)^*\,(E \times W)} = \frac{\mu^* \times \nu((E \times W) \cap \cdot)}{\mu^* \times \nu(E \times W)} = \frac{\mu^* \times \nu((E \times W) \cap \cdot)}{\mu^*(E)\nu(W)} = \frac{\mu^* \times \nu((E \times W) \cap \cdot)}{\mu^*(E)}
$$

because $\max\{\mu(E) \,|\, \mu \in \mathcal{Q}\}$ exists and is strictly positive by the assumption. Consequently, we obtain

$$
\begin{aligned}
&\pi' \in \phi_{DS}(\mathcal{Q} \times \mathcal{R}, E \times W) \\
\Leftrightarrow\ \ & (\exists \mu^* \in \arg\max\{\mu(E) \,|\, \mu \in \mathcal{Q}\})(\exists \nu \in \mathcal{R})\ \ \pi'(\cdot) = \frac{\mu^* \times \nu((E \times W) \cap \cdot)}{\mu^*(E)} \qquad (1)
\end{aligned}
$$

To show $\mathcal{P}|_2^{\phi_{DS}}(E) \subseteq \mathcal{R}$, let $\nu \in \mathcal{P}|_2^{\phi_{DS}}(E)$. Then, there exists $\pi' \in \phi_{DS}(\mathcal{Q} \times \mathcal{R}, E \times W)$ such that $\nu = \pi'_E$. Hence, $\nu \in \mathcal{R}$ because (1) implies that $(\forall F)\ \pi'_E(F) = \pi'(E \times F) = \mu^* \times \nu'(E \times F)/\mu^*(E) = \nu'(F)$ for some $\nu' \in \mathcal{R}$.

To show $\mathcal{P}|_2^{\phi_{DS}}(E) \supseteq \mathcal{R}$, let $\nu \in \mathcal{R}$. Also, let $\mu^* \in \arg\max\{\mu(E) \,|\, \mu \in \mathcal{Q}\}$. Such a μ^* certainly exists and $\mu^*(E) > 0$ by the assumption. Then, if we construct π' such that $\pi'(\cdot) \equiv \mu^* \times \nu((E \times W) \cap \cdot)/\mu^*(E)$, then we have $\pi' \in \phi_{DS}(\mathcal{Q} \times \mathcal{R}, E \times W)$ by (1). Hence, $\nu \in \mathcal{P}|_2^{\phi_{DS}}(E)$ since $(\forall F)\ \nu(F) = \mu^*(E)\nu(F)/\mu^*(E) = \mu^* \times \nu(E \times F)/\mu^*(E) = \pi'(E \times F) = \pi'_E(F)$. ∎

Since $\mathcal{P}|_2 = \mathcal{R}$ holds, the theorem implies the following corollary.

Corollary 1. $\mathcal{P}|_2 = \mathcal{P}|_2^{\phi_{DS}}(E)$.

In this example, the observation in the first period does not reduce the Knightian uncertainty at all. This property is explained heuristically in the following way. One implication of the fact that the unemployed worker's Knightian uncertainty is represented by a product of probability measures is that she does not rule out the possibility that the wage distribution may be very different between the present and future periods. The future wage distribution may be very different from the current distribution from which a particular observation is drawn. Then, today's observation may not convey useful information about the future wage distribution.

3.2. The ε-contamination

In the next example, a decision maker is nearly certain that uncertainty she faces over two periods is characterized by a product of a particular probability measure (that is, i.i.d. random variables.) However, she does not have perfect confidence about this characterization. She thinks that there might be a possibility, though small, that she is wrong and that true probability measure is different from the one she assumes. She does not know what probability measure she faces if her conviction turns to be wrong.

To characterize the decision maker's situation, we use the idea of ε-contamination. Formally, let $\varepsilon > 0$, let P_0 be a probability measure on $(W, \mathcal{B}_W)$, and let $P_0 \times P_0$ be the product probability measure on $(W \times W, \mathcal{B}_{W \times W})$. Here, we assume that $P_0 \times P_0$ is the distribution that the decision maker thinks is the "true" distribution. However, she is not perfectly certain about her conviction. We define the *ε-contamination* of $P_0 \times P_0$ on $(W \times W, \mathcal{B}_{W \times W})$ as

$$\{P_0 \times P_0\}^{\varepsilon} = \{\, (1-\varepsilon)(P_0 \times P_0) + \varepsilon Q \,|\, Q \in \mathcal{M}(W \times W) \,\}.$$

The ε-contamination of $P_0 \times P_0$ formulates the state of the Knightian uncertainty the decision maker faces. She is $(1-\varepsilon) * 100\%$ certain that the true distribution is $P_0 \times P_0$. However, with $\varepsilon * 100\%$ probability the true distribution is different from $P_0 \times P_0$. She is perfectly ignorant about the true distribution when her conviction is wrong, and thus she thinks that any distribution on $(W \times W, \mathcal{B}_{W \times W})$ can be the true distribution.

In the following analysis, the one-period counterpart of the two-period ε-contamination turns out to be important. Applying the same idea to the one-period case, we define the ε-contamination of P_0 on $(W, \mathcal{B}_W)$ as

$$\{P_0\}^{\varepsilon} = \{\,(1-\varepsilon)P_0 + \varepsilon Q \,|\, Q \in \mathcal{M}(W)\,\}\,.$$

The following lemma shows that the second-period marginal Knightian uncertainty of the ε-contamination of $P_0 \times P_0$ is equal to the ε-contamination of P_0.[2]

Lemma 1. $\{P_0 \times P_0\}^{\varepsilon}|_2 = \{P_0\}^{\varepsilon}$

Proof. To show $\{P_0 \times P_0\}^{\varepsilon}|_2 \subseteq \{P_0\}^{\varepsilon}$, let $\nu \in \{P_0 \times P_0\}^{\varepsilon}|_2$. Then, there exists $\pi \in \{P_0 \times P_0\}^{\varepsilon}$ such that $\nu = \pi_W$. That $\pi \in \{P_0 \times P_0\}^{\varepsilon}$ in turn implies that there exists $\pi' \in \mathcal{M}(W \times W)$ such that $\pi = (1-\varepsilon)(P_0 \times P_0) + \varepsilon\pi'$. Hence, $(\forall F)\ \nu(F) = \pi_W(F) = \pi(W \times F) = (1-\varepsilon)P_0(F) + \varepsilon\pi'(W \times F) = (1-\varepsilon)P_0(F) + \varepsilon\pi'_W(F)$. This shows that $\nu \in \{P_0\}^{\varepsilon}$ because $\pi'_W \in \mathcal{M}(W \times W)|_2 = \mathcal{M}(W)$.

To show $\{P_0 \times P_0\}^{\varepsilon}|_2 \supseteq \{P_0\}^{\varepsilon}$, let $\nu \in \{P_0\}^{\varepsilon}$. Then, there exists $\nu' \in \mathcal{M}(W)$ such that $\nu = (1-\varepsilon)P_0 + \varepsilon\nu'$. Let $\mu \in \mathcal{M}(W)$ and let $\pi = (1-\varepsilon)(P_0 \times P_0) + \varepsilon(\mu \times \nu')$. Then, $\pi \in \{P_0 \times P_0\}^{\varepsilon}$ and $\pi_W = (1-\varepsilon)P_0 + \varepsilon\nu' = \nu$, and hence, $\nu \in \{P_0 \times P_0\}^{\varepsilon}|_2$. ∎

Theorem 2. *Let E be an observation, which is a nonempty measurable subset of W. Then,*

$$\{P_0 \times P_0\}^{\varepsilon}|_2^{\phi_{DS}}(E) = \{P_0\}^{\varepsilon'} \equiv \{\,(1-\varepsilon')P_0 + \varepsilon' Q \,|\, Q \in \mathcal{M}(W)\,\}\,,$$

where

$$\varepsilon' = \frac{\varepsilon}{(1-\varepsilon)P_0(E) + \varepsilon}\,.$$

Proof. Let $Q^+ \in \mathcal{M}(W \times W)$ be any probability measure such that $Q^+(E \times W) = 1$. Such a Q^+ certainly exists since E is nonempty. Then, it follows that $\pi \in \phi_{DS}(\{P_0 \times P_0\}^{\varepsilon}, E \times W)$ if and only if

$$\pi(\cdot) = \frac{(1-\varepsilon)(P_0 \times P_0)((E \times W) \cap \cdot) + \varepsilon Q^+((E \times W) \cap \cdot)}{(1-\varepsilon)P_0(E) + \varepsilon}$$

[2] The lemma shows that $\{P_0 \times P_0\}^{\varepsilon}|_2 = \{P_0\}^{\varepsilon}$. And we also have $(\{P_0\}^{\varepsilon} \times \{P_0\}^{\varepsilon})|_2 = \{P_0\}^{\varepsilon}$. However, $\{P_0 \times P_0\}^{\varepsilon} \neq \{P_0\}^{\varepsilon} \times \{P_0\}^{\varepsilon}$ in general. To see this, consider the following simple example: Let P_0 be the Bernoulli distribution with the probability of success being a half, and let $\varepsilon = 1/2$. Then, $(1/16, 3/16, 3/16/, 9/16) \in (\{P_0\}^{\varepsilon} \times \{P_0\}^{\varepsilon}) \setminus \{P_0 \times P_0\}^{\varepsilon}$. In contrast to the risk, the marginals do not characterize the uncertainty. For more on this, see Ghirardato (1997).

for such a Q^+, where π is well-defined even when $P_0(E) = 0$ since $\varepsilon > 0$. Therefore, $\nu \in \{P_0 \times P_0\}^{\varepsilon}|_2^{\phi_{DS}}(E)$ if and only if

$$\begin{aligned}
\nu(\cdot) &= \pi_E(\cdot) \\
&= \frac{(1-\varepsilon)P_0(E)P_0(\cdot) + \varepsilon Q_E^+(\cdot)}{(1-\varepsilon)P_0(E) + \varepsilon} \\
&= \left(1 - \frac{\varepsilon}{(1-\varepsilon)P_0(E) + \varepsilon}\right) P_0(\cdot) + \frac{\varepsilon}{(1-\varepsilon)P_0(E) + \varepsilon} Q_E^+(\cdot) \\
&= (1-\varepsilon')P_0(\cdot) + \varepsilon' Q_E^+(\cdot)
\end{aligned}$$

for such a Q^+. Clearly, $Q_E^+ \in \mathcal{M}(W)$, and hence, $\{P_0 \times P_0\}^{\varepsilon}|_2^{\phi_{DS}}(E) \subseteq \{P_0\}^{\varepsilon'}$. For the opposite inclusion, it suffices to show that any $Q \in \mathcal{M}(W)$ is represented as Q_E^+ for some Q^+ such that $Q^+(E \times W) = 1$. This can be done by setting Q^+ equal to $\mu^+ \times Q$ for some $\mu^+ \in \mathcal{M}(W)$ such that $\mu^+(E) = 1$. ∎

It follows that $\{P_0\}^{\varepsilon'}$ is a proper *super*set of $\{P_0\}^{\varepsilon}$ if $\varepsilon' > \varepsilon$, which holds whenever $P_0(E) < 1$. Hence, Lemma 1 and Theorem 2 immediately prove the next corollary.

Corollary 2. *Suppose that an observation E be a nonempty subset of W such that $P_0(E) < 1$. Then,*

$$\{P_0 \times P_0\}^{\varepsilon}|_2 \subset \{P_0 \times P_0\}^{\varepsilon}|_2^{\phi_{DS}}(E)$$

where the inclusion is strict.

This corollary shows a striking result. In the example of ε-contamination, the Dempster-Shafer update rule, which is considered to be sensible in the multiple prior framework, actually *increases,* rather than decreases, the degree of the Knightian uncertainty. In a sense, learning (in the form of the Dempster-Shafer update rule) increases the decision maker's confused state of confidence, rather than resolve it.

Heuristically, this fact is explained in the following way. The Dempster-Shafer rule, or in other words, the maximum-likelihood update rule, chooses probability measures among possible priors that put the highest probability on the occurrence of the first-period observation

E, and updates the chosen probability measures using the Bayes rule. This procedure is supposed to eliminate implausible probability measures from the original set of priors and to slant remaining priors to put more probability on the occurrence of the event E. The basic idea is that plausible probability measures should have the highest probability on the actual occurrence. However, in the ε-contamination case $\{P_0 \times P_0\}^{\varepsilon}$, the set of original priors is a weighted average of a *known* distribution $P_0 \times P_0$ and a distribution Q taken from *complete ignorance* (that is, all possible distributions). Thus, the presumption that plausible probability measures should put the highest probability on the actual occurrence is not very much sensible in the case of ε-contamination. Moreover, since we allow all possible distributions for Q, the maximum-likelihood procedure chooses probability measure Q^{+} putting the highest probability $(= 1)$ on the first-period observation E, and dictates chosen probability measures to put more weight on Q^{+}. This produces the counter-intuitive result of Corollary 2, in which learning increases the Knightian uncertainty.

REFERENCES

Ghirardato, Paolo (1997): "On Independence for Non-Additive Measures, with a Fubini Theorem," *Journal of Economic Theory* 73, 261-291.

Gilboa, I. and D. Schmeidler (1993): "Updating Ambiguous Beliefs," *Journal of Economic Theory* 59, 33-49.

Nishimura, K. G. and H. Ozaki (2001): "Search and Knightian Uncertainty," Discussion Paper, University of Tokyo.

Rothschild, M. (1974): "Searching for the Lowest Price When the Distribution of Prices is Unknown," *Journal of Political Economy* 82, 689-711.