Observability and "Second-Order Acts"*

Nabil I. Al-Najjar† and Luciano De Castro‡

First draft: December 2010;
This version: June 2013

Abstract

This note questions the behavioral content of second-order acts and their use in decision theoretic models. We show that there can be no *verification mechanism* to determine what the decision maker receives under a second-order act. This impossibility applies even in idealized repeated experiments where infinite data can be observed.

*We thank Paolo Ghirardato, Ben Polak, Kyoungwon Seo, and Marciano Siniscalchi for discussions and comments on an earlier version of the paper. We also thank Simone Galperti for his research assistance.

† Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston IL 60208.

‡ Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston IL 60208.

Contents

1 Introduction

This note questions the behavioral content of second-order acts and their use in decision theoretic models. Although our argument extends to any model that uses these acts in its foundation, we focus for concreteness on the smooth model by Klibanoff, Marinacci, and Mukerji (2005) where they play a fundamental role. Preferences in this model have the representation:

$$\mathcal{W}(f) = \int_{\Delta(\Omega)} \phi \left(\int_{\Omega} u(f(\omega))\, dP(\omega) \right) d\nu(P). \tag{1}$$

Here Ω is a state space, $\Delta(\Omega)$ is the set of probability measures on this space, f is an act with values in $\Delta(C)$, where C is a finite set of "pure" consequences, and u and ϕ are real-valued functions with the appropriate domains.

Second-order acts are used in (1) to separate "risk" from "uncertainty:" A probability measure P in the inner integral represents the true governing law, while the probability measure ν reflects the decision maker's uncertainty about what this law might be. If the decision maker knew P (ν puts unit mass on a single measure), then he has expected utility preference with von Neumann-Morgenstern utility u. A non-degenerate ν reflects the decision maker's uncertainty about the true P. A non-linear ϕ captures a non-neutral attitude to uncertainty about P, while a linear ϕ reduces $\mathcal{W}$ to standard expected utility.

A second-order act is a function of the form:

$$\mathbf{f} : \Delta(\Omega) \to \Delta(C)$$

and represent bets on probability measures. They should be distinguished from the (usual) "state-based acts," which are functions of the form $f : \Omega \to \Delta(C)$ and represent bets on events. Klibanoff, Marinacci, and Mukerji (2005) who use them to derive the representation (1) argue that "second-order acts are not as strange or unfamiliar as they might first appear. Consider any parametric setting, i.e., a finite dimensional parameter space Θ [...] Second order acts would simply be bets on the value of the parameter. [...] Similarly, in model uncertainty applications, second-order acts are bets about the values of the relevant parameters in the underlying model."

The use of acts and contracts implicitly assumes a mechanism to verify what the parties should receive. Contracts contingent on probabilities of events are difficult to interpret as objects of economic or decision theoretic analysis. So while it makes sense to think of a bet on a football team wining a game, a contract that stipulates "you receive 1 dollar if this team wins with probability greater than 0.75 and zero otherwise" seems rather implausible. Klibanoff, Marinacci, and Mukerji's answer to this difficulty is that "[v]erification [of probabilities] may also be possible if one has the opportunity to wait and observe a sufficiently long run of data generated by repeated realizations from the [underlying distribution] that obtains." (p. 1856)

An appeal to repeated trials is unhelpful in unique situations, such as a presidential election or a market crash, where repetition is implausible or impossible. In this note we show that second-order acts remain problematic even in settings where repeated experiments might make sense.

Section 2 illustrates how second-order acts might be used to separate risk from uncertainty, relying on the smooth model for context. We use the example to informally argue that these acts are difficult to interpret because there can be no verification mechanism to determine what the decision maker receives once the state of the world is known. In Section 3 we provide a formal definition of what parameters mean in a preference model. This definition closely follows the typical interpretation of the notion of parameters in statistical inference and its applications in empirical sciences.

In Section 4, we show that there exists no *verification mechanisms* for second-order acts in any parametric model. This impossibility result applies to all verification mechanisms; it is irrelevant how simple or complex they are, or whether they use finite data or refer to an idealized setting where infinite data can be observed. This impossibility indicates that the smooth model is founded on choices whose consequences cannot be determined because they are contingent on unobservables. The repeatability of the experiment is irrelevant.

There is a number of models in the literature that have a "double-integral" representation that appears, superficially at least, similar to $\mathcal{W}$. These include Neilson((1993), (2009)), Nau (2001), Nau (2006), Ergin and Gul (2009), Chew and Sagi (2008), Strzalecki (2010), Grant, Polak, and Strzalecki (2009), among others. In Section 4 we show that these models profoundly differ from the smooth model because they do not appeal to second-order acts and thus do not suffer from this model's observability problems. We then discuss Seo (2009)'s use of lotteries over acts by decision makers who do not reduce objective compound lotteries. We view his contribution as further illustration of the problematic nature of second-order acts. Finally, Epstein (2010) pointed to other paradoxical implications of the smooth model. The points he raises are orthogonal to ours, so our works are complementary.

2 Interpreting Mixtures

We provide a simple example to illustrate why bets on unobservable probabilities (second-order acts) are needed in the smooth model.

To work with a state space where the notion of "parameter" is not trivial, we assume that $\Omega = \{H, T\}^\infty$, the set of outcomes of infinite coin tosses. Assume further that consequences are real numbers and that the decision maker is risk-neutral. Under these assumptions, the smooth model representation ranks acts according to:

$$\mathcal{W}(f) = \int_{\Delta(\Omega)} \phi \left(\int_\Omega f(\omega)\, dP(\omega) \right) d\nu(P).$$

Interpret ν as uncertainty about the true probability law P governing the state ω. The model proposes to separate risk from uncertainty by identifying a different attitude to ν-uncertainty captured by ϕ. For simplicity, assume that ϕ is strictly concave so the decision maker is ambiguity averse.

Interpret a real number $\theta \in [0,1]$ as the probability of *Heads* in a given toss and let P^θ be the corresponding i.i.d. distribution on Ω with marginal θ. Given $0 < \theta_0 < \theta_1 < 1$ and $\alpha \in (0,1)$ define the probability measure $P_\alpha \equiv \alpha P^{\theta_0} + (1-\alpha) P^{\theta_1}$ on Ω. A distribution like P_α has many distinct representations as a two-stage process. Two polar examples of such representations are:

1. *Representation 1:* In the first stage the parameter $\theta_i \in \{\theta_0, \theta_1\}$ is drawn according to the probability measure ν_α that puts mass α on θ_0 and $1-\alpha$ on θ_1. In the second stage, ω is generated according to P^{θ_i}.

2. *Representation 2:* The first stage is trivial, with δ_{P_α} denoting the distribution that puts unit mass on P_α. In the second stage ω is chosen according to $\alpha P^{\theta_0} + (1-\alpha)P^{\theta_1}$.

All such representations are equivalent from the perspective of an expected utility maximizer, yet correspond to different behavior in the smooth model. To illustrate this, suppose that f is the act that pays \$100 if the first coin turns *Heads* and \$0 if it turns *Tails*. A decision maker with *Representation 1* evaluates f according to:

$$\int_{\Delta(\Omega)} \phi\left(\int_\Omega f\, dP\right) d\nu_\alpha = \alpha\phi\left(100\theta_0\right) + (1-\alpha)\phi\left(100\theta_1\right). \tag{2}$$

But under *Representation 2*, he evaluates f according to:

$$\int_{\Delta(\Omega)} \phi\left(\int_\Omega f\, dP\right) d\delta_{P_\alpha} = \phi\left(\int_\Omega f\, dP_\alpha\right) = \phi\big(100[\alpha\theta_0 + (1-\alpha)\theta_1]\big). \tag{3}$$

These two expressions are different, unless ϕ is linear. In (2), the decision maker has subjective uncertainty whether the 'true' parameter is θ_0 or θ_1. This is treated as uncertainty by placing the mixture over θ_1 and θ_2 in the outer integral in (2). In (3), by contrast, the decision maker is confident that the true distribution is P_α. The representation treats the mixture over θ_1 and θ_2 as objective risk by placing it in the inner integral in (3), and is therefore treated by ϕ on equal footing as f.

For the smooth model to be non-vacuous, one needs to find behavioral implications to identify ν and ϕ (thus deciding which of (2) or (3) to use). This is accomplished by the use of the second-order acts, which is the set $\mathbf{F}$ of all functions of the form:

$$\mathbf{f} : \Delta(\Omega) \to \Delta(C).$$

The smooth model postulates two classes of acts (state-based and second-order acts, $\mathcal{F}$ and $\mathbf{F}$ respectively), two preference relations ($\succcurlyeq$ on $\mathcal{F}$ and another preference $\breve{\succcurlyeq}$ on $\mathbf{F}$), and axioms relating the two. Once these ingredients are in place, deriving the representation $\mathcal{W}$ is straightforward and follows the argument outlined by Neilson((1993), (2009)): Apply Savage's theorem with state space $\Delta(\Omega)$ and consequences $\Delta(C)$ to identify the probability measure ν and function ϕ appearing in $\mathcal{W}$.

Second-order acts are bets on probabilities. Think of them as a contract that stipulates, for instance, that: "you receive $2 if the probability of rain is higher than .63, $1 if the probability is between 0.27 and 0.63, and zero otherwise." Klibanoff, Marinacci, and Mukerji recognize the problem that bets on unobserved *probabilities* raise, but note that "[v]erification [of probabilities] may also be possible if one has the opportunity to wait and observe a sufficiently long run of data generated by repeated realizations."

Our point is that repeated trials confound rather than clarify the issues. A model of choice should point to, at a minimum, a thought experiment where the decision maker can be presented with the consequences of his choice. Proposition 4.1 will formally show that the smooth model cannot explain how bets on probabilities can be implemented, even in repeated trials and with infinite amount of data.

The intuition is simple: consider the setting with two i.i.d. parameters θ_0, θ_1 introduced earlier and the subset of second-order acts of the form $\mathbf{f} : [0,1] \to \mathbb{R}$. Here, $\mathbf{f}(0)$ and $\mathbf{f}(1)$ denote the consequence obtained if θ_0 or θ_1 occur, respectively, while $\mathbf{f}(\alpha)$ is the consequence if the 'true' distribution consists of first flipping a coin that picks θ_0 with probability α, and θ_1 with probability $1-\alpha$. One can easily conceive of a mechanism to verify whether θ_0 or θ_1 occurred (*e.g.,* by examining the limiting frequencies). Proposition 4.1 formally defines the concept of *verification mechanism* and shows that none exists to verify statements of the form "P_α occurred" and to determine what the decision maker should receive under $\mathbf{f}$.

3 Parameters, Risk, and Uncertainty

Klibanoff, Marinacci, and Mukerji suggest that the problem of observability and, ultimately, the meaning of the smooth model can be settled by considering "any parametric setting, i.e., a finite dimensional parameter space Θ [...] Second order acts would simply be bets on the value of the parameter."

To evaluate this claim and justify the interpretation of parameters as risk, we need a formal notion of what a "parameter" means. In this section, assume for simplicity a finite set of consequences C and that the decision maker has a von Neumann-Morgenstern utility u over $\Delta(C)$.[1]

Definition 1 (Parametrizations) *Let $\mathcal{P} \subset \Delta(\Omega)$ be a compact convex set of probabilities, with extreme points $\{P^\theta\}_{\theta\in\Theta}$, where Θ is an index set.*[2] *We say that Θ is a parametrization if there exists a map $\vartheta : \Omega \to \Theta$ and $\mathcal{E} \subset \Sigma$ such that the following is true:*

(i) ϑ is measurable with respect to $\mathcal{E}$;

(ii) $P^\theta(\vartheta^{-1}(\theta)) = 1$ for all $\theta \in \Theta$;

(iii) For every $A \in \Sigma$ and probability measure $P \in \mathcal{P}$, $P^{\vartheta(\omega)}(A)$ is a version of the conditional probability of A given $\mathcal{E}$.

Every probability measure μ on Θ defines a unique probability on Ω by:

$$M_\mu(A) = \int_\Theta P^\theta(A)\, d\mu(\theta).$$

The notion of parametrization above is closely related with the notion of sufficient statistics as extreme points of sets of measures. (See, for instance, Dynkin (1978) and Lauritzen (1984)). Al-Najjar and De Castro (2010) derive sufficient parameterizations from primitive invariance conditions on preferences and study them in great details.

Parameterizations are linked to a preference through the concept of sufficiency:

Definition 2 (Sufficiency) *A parametrization Θ is* sufficient *for a preference $\succcurlyeq$ if ϑ is the essentially unique function*[3] *such that for every $f, g \in \mathcal{F}$:*

$$f \succcurlyeq g \iff \int_\Omega f\, dP^{\vartheta(\cdot)} \succcurlyeq \int_\Omega g\, dP^{\vartheta(\cdot)}. \tag{4}$$

[1] Assume further that Ω is a Polish space with Borel σ-algebra Σ and, therefore, $\Delta(\Omega)$ is also Polish.

[2] Θ can be identified with a subset of $\Delta(\Omega)$; thus Θ inherits the topology and Borel structure of $\Delta(\Omega)$.

[3] That is, if ϑ' is another function satisfying (4), the set $\{\omega \in \Omega : \vartheta(\omega) \neq \vartheta'(\omega)\}$ is $\succcurlyeq$-null.

A preference is *parametric* if it has a sufficient parameterization. A parametric preference can be decomposed into:

- *An expected utility preference that captures risk conditional on parameters:* The expression $\int f \, dP^{\vartheta(\omega)} \succcurlyeq \int g \, dP^{\vartheta(\omega)}$ compares two lotteries in $\Delta(C)$ and so incorporates the decision maker's risk attitude. If he evaluated gambles according to expected utility with von Neumann-Morgenstern utility u, then this comparison reduces to comparing the "certainty equivalent" of f conditional on each θ, $\int_\Omega u(f) \, dP^\theta$, with that of g, $\int_\Omega u(g) \, dP^\theta$.
- *An aggregator that captures how the decision maker deals with parameter uncertainty.*[4]

The notion of parameterization provides a formal justification for interpreting parameters as risk. It is, however, silent on what aggregator the decision maker uses.

We conclude this section by providing two examples of parameterizations:

1. *The i.i.d. parametrization:* $\Theta = [0, 1]$, P^θ is the i.i.d. distribution with probability of heads equal to θ, $\mathcal{E}$ is the set of events that are invariant with respect to finite permutations, and ϑ is defined by setting $\vartheta(\omega)$ equal to the limiting frequency of heads in ω if this limit exists, and define it arbitrarily otherwise.

   This parametrization is the one that appears in de Finetti's representation of exchangeable processes as mixtures of i.i.d. distributions.

2. *The trivial parametrization:* $\Theta = \Omega$, $P^\omega = \delta_\omega$ is the probability measure that puts mass 1 on the state ω, $\mathcal{E} = \Sigma$, and ϑ is the identity. Given P on Ω, we have:
$$M_P = \int \delta_\omega \, dP = P.$$
   This parameterization is trivial in that parameters are states, and so they provide no useful compression of information.

[4]Formally, the aggregator is the restriction of the preference to parameter-based acts. See Al-Najjar and De Castro (2010).

4 Parameters and Observability

Given a parametrization Θ, let $\mathbf{F}_\Theta$ (or just $\mathbf{F}$ when the context is clear) denote the set of second-order acts with respect to Θ, *i.e.,* the set of all functions of the form $\mathbf{f} : \Delta(\Theta) \to \Delta(C)$. In the case of the trivial parameterization $\Theta = \Omega$, this is just the set of second-order acts in the smooth model.[5] The next result shows that neither the smooth model nor any model involving second-order acts can be parametric. To make this formal, define:

Definition 3 *Given a parameterization* Θ*, a collection of events* $\{E_\mu\}_{\mu\in\Delta(\Theta)}$ *is a* verification mechanism *for the set of second-order acts* $\mathbf{F}$ *if for every* $\nu, \nu' \in \Delta(\Theta), \nu \neq \nu'$*,*

$$E_\nu \cap E_{\nu'} = \emptyset \qquad \textit{and} \qquad M_\nu(E_\nu) = 1.$$

We interpret $\omega \in E_\nu$ to mean that "observing ω verifies that ν occurred." The first requirement says that if ω verifies ν, then it cannot also verify $\nu' \neq \nu$. The second requirement says that if ν is 'true,' then it will be verified with probability 1 under M_ν.

Proposition 4.1 *Fix any parametrization* Θ*. Then there exists no verification mechanism for its set of second-order acts* $\mathbf{F}$*.*

Proof: Assume, by way of contradiction, that there is such verification mechanism $\{E_\nu\}_{\nu\in\Delta(\Theta)}$. Let $\nu \in \Delta(\Theta) - \Theta$ be of the form $\alpha\theta_0 + (1-\alpha)\theta_1$ for some $\theta_0, \theta_1 \in \Theta, \theta_0 \neq \theta_1,$ and $\alpha \in (0, 1)$. Then

$$M_\nu(E_\nu) = \alpha P^{\theta_0}(E_\nu) + (1-\alpha)P^{\theta_1}(E_\nu) = 1.$$

This implies that $P^{\theta_0}(E_\nu) = 1$. Since $P^{\theta_0}(E_0) = 1$, we also have $P^{\theta_0}(E_0 \cap E_\nu) = 1$. But since $E_\nu \cap E_0 = \emptyset$, we have $P^{\theta_0}(E_0 \cap E_\nu) = 0$, a contradiction. ∎

[5] The trivial parameterization, where $\Theta = \Omega$, is always sufficient for any preference. The question is whether a preference has an "interesting" parameterization (*e.g.,* exchangeable, Markovian, ... etc), not whether some parametrization exists. Al-Najjar and De Castro (2010) discuss parametrizations in greater details.

Some bets on probabilities make sense, namely those made on parameters: any individual parameter θ is associated with a unique event $\vartheta^{-1}(\theta) \subset \Omega$ and thus can, at least in principle, be verified. For instance, one can in principle verify whether the data is governed by the i.i.d. parameter θ_0 or θ_1 by inspecting the limiting frequencies. What does not make sense is bets on mixtures of parameters which no amount of data can verify.

This sheds light on how the smooth model differs from others in the literature, such as Neilson((1993), (2009)), Nau (2001), Nau (2006), Ergin and Gul (2009), Chew and Sagi (2008), Strzalecki (2010), Grant, Polak, and Strzalecki (2009), Klibanoff, Mukerji, and Seo (2010), among others. All these models use versions of the "double-integral" representation, making it tempting to lump them together, or to view them as special cases of the smooth model. Holding the states, consequences, and acts fixed, the smooth model relates two sets of objects:

$$\left\{ \begin{array}{c} \text{Preference } \succcurlyeq \text{ on state-based acts } \mathcal{F} \\ + \\ \text{Preference } \breve{\succcurlyeq} \text{ on unobservable acts } \mathbf{F} \\ + \\ \text{Axioms on } \succcurlyeq \text{ and } \breve{\succcurlyeq} \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{c} \text{Preference} \\ \text{represented by} \end{array} \quad \mathcal{W} \right\}.$$

The models cited above fundamentally differ from the smooth model in that they do not involve bets on unobservable probabilities.[6] As an illustration, consider preferences with *second-order expected utility representation*:

$$\mathcal{V}(f) = \int_\Theta \phi \left(\int_\Omega u(f(\omega))\, dP^\theta \right) d\mu. \tag{5}$$

Neilson((1993), (2009)) characterized these preferences in the special case of the trivial parametrization $\Theta = \Omega$. The representation $\mathcal{V}$ is covered by

[6] For example, the Savage and von Neumann-Morgenstern models both use the expected utility representation, but they are of course very different models because, among other things, they assume different choice data.

the framework of Al-Najjar and De Castro (2010) who show that Neilson's argument extends to general parametrizations. Although this superficially appears similar to the functional $\mathcal{W}$, the differences are fundamental. The behavioral characterization of $\mathcal{V}$ is *in terms of the set of state-based acts $\mathcal{F}$ only*. Second-order acts do not appear in these models and bets on mixtures of parameters are meaningless.

In contrast with the smooth model, a mixture like $\alpha P^{\theta_0} + (1-\alpha)P^{\theta_1}$ has a unique interpretation in $\mathcal{V}$ as the decision maker's subjective uncertainty about the parameters. This is because the outer integral in $\mathcal{V}$ is over Θ so

$$\text{``} \int_\Theta \phi\left(\int_\Omega u(f(\omega))\, dP_\alpha\right) d\mu \text{''}$$

is a meaningless expressions because $P_\alpha \notin \Theta$ and thus cannot appear as object of integration. A parametric model separates risk and uncertainty via behavioral assumptions about the sufficiency of a parametrization. The smooth model does so by blurring the difference between bets on observable parameters and bets on unobservable mixtures of parameters.

Along the same lines, the smooth model's problems with observability remain even if it so happens that the support of ν contains only observable parameters. The second-order belief ν and utility ϕ are obtained as a consequence of the assumption that the decision maker can reveal a preference $\breve{\succcurlyeq}$ over unobservable second-order acts. Our critique extends to all models that assume a preference over such acts as their foundation; it is irrelevant whether or not the derived beliefs happen to be on observables only.

5 Lotteries over Acts

Seo (2009) attempts to provide foundations for the smooth model without invoking unobservable second-order acts. He does this by expanding the domain of the preference to $\Delta(\mathcal{F})$, the set of mixtures over acts.[7] Elements of $\Delta(\mathcal{F})$ are probability measures over $\mathcal{F}$, so $\Delta(\mathcal{F})$ is a mixture space with

[7] The set of state-based acts $\mathcal{F}$ is already a mixture space, with convex combinations defined state-wise: Given any two acts f, g and $\alpha \in [0,1]$, we define $\alpha f + (1-\alpha)g$ to mean the act that yields the consequence $\alpha f(\omega) + (1-\alpha)g(\omega)$ at state ω.

the operation $\beta Q \oplus (1-\beta)W$, with $Q, W \in \Delta(\mathcal{F}), \beta \in [0,1]$, as the convex combination of two measures Q, W.

As discussed in Kreps (1988), one may formulate subjective expected utility a la Anscombe and Aumann (1963) by defining the preference over the domain $\Delta(\mathcal{F})$ rather than the usual domain $\mathcal{F}$ provided one assumes (in addition to the other standard assumptions) that for any two acts f, g and $\alpha \in [0,1]$:

$$\alpha\delta_f \oplus (1-\alpha)\delta_g \sim \alpha f + (1-\alpha)g. \tag{6}$$

Kreps (1988) notes that $\Delta(\mathcal{F})$ is strictly larger domain than $\mathcal{F}$ and that assuming the Anscombe and Aumann axioms on the larger domain $\Delta(\mathcal{F})$ amounts to "establishing large indifference classes [..making..] the mixture space axioms much more powerful." (p. 107)

By dropping (6), Seo obtains the representation:

$$\mathcal{Y}(Q) = \int_{\mathcal{F}} \left[\int_{\Delta(\Omega)} \phi \left(\int_{\Omega} u(f)\, dP^{\theta} \right) d\nu \right] dQ(f). \tag{7}$$

Specialized to degenerate distributions δ_f and abusing notation, we obtain the representation:

$$\mathcal{Y}(f) = \int_{\Delta(\Omega)} \phi \left(\int_{\Omega} u(f)\, dP^{\theta} \right) d\nu.$$

The idea is to derive a representation over $\mathcal{F}$ by expanding the domain of the preference, except that now one considers bets over acts, whereas the smooth model introduces bets on unobservable probabilities.

While observability is not an issue with this approach, new difficulties arise with how objective lotteries are treated. As noted by Seo (2009, Corollary 5.2), the preferences he assumes over $\Delta(\mathcal{F})$ fail to reduce *objective* compound lotteries. Formally, let $p, q \in \Delta(C)$ be two lotteries over consequences. With our interpretation of lotteries as constant acts, and fixing $\alpha \in [0,1]$, we can now talk about two objects:

- $\alpha p + (1-\alpha)q \quad \in \mathcal{F}$;
- $\alpha\delta_p \oplus (1-\alpha)\delta_q \in \Delta(\mathcal{F})$.

The first mixture is an act which, in every state ω, randomizes between p, q with objective probabilities $\alpha, (1-\alpha)$. In the second mixture, one first randomizes $\alpha, (1-\alpha)$ to select between p and q, then implement the result of the selection in each state ω. The representation (7) requires that:

$$\alpha p + (1-\alpha) q \not\sim \alpha \delta_p \oplus (1-\alpha)\delta_q$$

for some choice of p, q, α.

The failure to reduce objective compound lotteries is a common behavioral bias which is shown by Halevy (2007) to be strongly correlated with ambiguity averse preferences among experimental subjects. Incorporating such a bias as part of rational choice theory can be unsettling. To appreciate the striking implications of such behavior, let f, g be two state-contingent contracts for determining the end-of-year bonus of an agent. Assume that the agent is risk-neutral (or that the payments are in utils). The agent knows that: (1) a fair coin is tossed to select one contract from the set $\{f, g\}$; (2) the bonus will be determined as a function of the state ω using the selected contract. At the end of the year, an impartial referee receives two sealed envelopes with the following information:

Envelope A: The result of a fair coin toss to select either f or g.

Envelope B: An observation of the realized state ω.

No observability issues arise here: the information contained in the two envelopes enables the referee to unambiguously determine what bonus should be paid. There are, however, at least two ways for the agent to evaluate the outcome.

- Envelope A then B: The agent believes the referee first opens Envelope A, resolving uncertainty about the coin toss and determining whether f or g is used, then Envelope B is opened, resolving uncertainty about ω and determining whether he receives $f(\omega)$ or $g(\omega)$. In this case, the value he attaches to this random contract is $0.5\mathcal{Y}(f) + 0.5\mathcal{Y}(g)$.

$$\mathcal{Y}(0.5f + 0.5g) = 0.5 \int_{\Theta} \phi \left(\int_{\Omega} f \, dP^{\theta} \right) d\nu + 0.5 \int_{\Theta} \phi \left(\int_{\Omega} g \, dP^{\theta} \right) d\nu.$$

- Envelope B then A: The agent believes the referee first opens Envelope B, resolving uncertainty about ω and determining that the agent receives either $f(\omega)$ or $g(\omega)$. Then the referee opens Envelope A, resolving uncertainty about the coin toss and determining whether f or g is used. In this case, the value the agent attaches to this random contract is:

$$\mathcal{Y}\left(\delta_{0.5f+0.50g}\right) = \int_{\Theta} \phi \left[\int_{\Omega} (0.50f + 0.50g)\, dP^{\theta} \right] d\nu.$$

Since the terms $\mathcal{Y}(0.5f+0.5g)$ and $\mathcal{Y}\left(\delta_{0.5f+0.50g}\right)$ are, in general, not equal (unless ϕ is linear), the agent cares about the order in which the referee opens the envelopes. If one insists that all objective lotteries should be reduced, then Seo proves that (7) collapses to subjective expected utility. This makes Seo's approach based on lotteries over acts one of behavioral biases rather than an account of how rational individuals deal with ambiguity.

The second-order expected utility model $\mathcal{V}$ (with non-linear ϕ) is consistent with the reduction of objective compound lotteries: expand the domain of $\mathcal{V}$ to $\Delta(\mathcal{F})$ and assume the indifference (6). In this case, compound objective lotteries always reduce: for any $Q \in \Delta(\mathcal{F})$ we have

$$\int_{\mathcal{F}} \left[\int_{\Theta} \phi \left(\int_{\Omega} u(f)\, dP^{\theta} \right) d\mu \right] dQ \sim \int_{\Theta} \phi \left[\int_{\Omega} u \left(\int_{\mathcal{F}} f\, dQ \right) dP^{\theta} \right] d\mu,$$

which reduces to $\mathcal{V}$, since $\int_{\mathcal{F}} f\, dQ$ is just an ordinary act obtained by integrating the consequences state by state using the distribution Q.

To summarize, Seo (2009)'s contribution is to show that second-order acts' may be avoided, provided one is willing to accept failures to reduce compound lotteries. Such failures have disturbing implications for a theory of rational choice under uncertainty.

6 Concluding Remarks

In discussing the issue of implementing second-order acts, Klibanoff, Marinacci, and Mukerji note that "in Ellsberg urn experiments, all you would need to do is dump the urn and verify the proportion of balls in it. Verification may also be possible if one has the opportunity to wait and observe a sufficiently long run of data." Both of these fixes is problematic in situations involving ambiguity. For one-off events, like a U.S. presidential election or a disruptive technological innovation, there is no urn to dump or balls to count. For repeated events, we argued in this paper that no matter how long one waits, we will never "observe a sufficiently long run of data" to make second-order acts behaviorally meaningful.

References

Al-Najjar, N. I., and L. De Castro (2010): "Parametric Representation of Preferences," Northwestern University.

Anscombe, F., and R. Aumann (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34(1), 199–205.

Chew, S., and J. Sagi (2008): "Small worlds: Modeling attitudes toward sources of uncertainty," *Journal of Economic Theory*, 139(1), 1–24.

Dynkin, E. (1978): "Sufficient statistics and extreme points," *The Annals of Probability*, 6(5), 705–730.

Epstein, L. (2010): "A Paradox for the 'Smooth Ambiguity' Model of Preference," *Econometrica (forthcoming).*

Ergin, H., and F. Gul (2009): "A Subjective Theory of Compound Lotteries," *Journal of Economic Theory*, pp. 899–929.

Grant, S., B. Polak, and T. Strzalecki (2009): "Second-Order Expected Utility," Discussion paper, Harvard University.

Halevy, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75(2), 503–536.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73(6), 1849–1892.

Klibanoff, P., S. Mukerji, and K. Seo (2010): "Relevance and Symmetry," Northwestern University.

Kreps, D. (1988): *Notes on the Theory of Choice*. Westview Press.

Lauritzen, S. (1984): "Extreme Point Models in Statistics," *Scandinavian Journal of Statistics*, pp. 65–91.

Nau, R. (2001): "Uncertainty Aversion with Second-Order Utilities and Probabilities," in *2nd International Symposium on Imprecise Probabilities and Their Applications, Ithaca, New York, June 2001.*

——— (2006): "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science*, 52(1), 136.

Neilson, W. (1993): "Ambiguity aversion: An axiomatic approach using second order probabilities," Discussion paper, mimeo.

——— (2009): "A simplified axiomatic approach to ambiguity aversion," *Journal of Risk and Uncertainty*, forthcoming.

Seo, K. (2009): "Ambiguity and Second-Order Belief," *Econometrica*, 77, 1575–1606.

Strzalecki, T. (2010): "Axiomatic Foundations of Multiplier Preferences," *Econometrica (forthcoming).*