**Discussion Papers No. 465, July 2006**
**Statistics Norway, Research Department**

*John K. Dagsvik*

# Axiomatization of Stochastic Models for Choice under Uncertainty

**Abstract:**
This paper develops a theory of probabilistic models for risky choices. Part of this theory can be viewed as an extension of the expected utility theory to account for bounded rationality. One probabilistic version of the *Archimedean Axiom* and two versions of the *Independence Axiom* are proposed. In addition, additional axioms are proposed of which one is Luce's *Independence from Irrelevant Alternatives*. It is demonstrated that different combinations of the axioms yield different characterizations of the probabilities for choosing the respective risky prospects. An interesting feature of the models developed is that they allow for violations of the expected utility theory known as the *common consequence effect* and the *common ratio effect*.

**Keywords:** Random tastes, bounded rationality, independence from irrelevant alternatives, probabilistic choice among lotteries, Allais paradox.

**JEL classification:** C25, D11, D81

**Acknowledgement:** I have benefited from comments by Jørgen Weibull, Steinar Strøm, Marina Rybalka, Terje Skjerpen, Weizhen Zhu and Anne Skoglund.

**Address:** John K. Dagsvik, Statistics Norway, Research Department. E-mail: john.dagsvik@ssb.no

**Discussion Papers** comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

# 1. Introduction

In the standard theory of decision making under uncertainty, it is assumed that the agent's preference functional is deterministic. This assumption is maintained in most of the recent theoretical and empirical literature. It has been recognized for some time, however, that even in seemingly identical repetitions (replications) of the same choice setting, the decision maker often makes different choices, cf. Tversky (1969). This means that the deterministic theory cannot be applied directly in an empirical context unless some additional stochastic "error" is introduced. As Fishburn (1976, 1978), Hey (1995), Carbone (1997), Loomes and Sugden (1995, 1998) and Starmer (2000) discuss, this raises the question of how axiomatization of theories for choice under uncertainty should be extended to accommodate stochastic error.

This paper proposes an axiomatic foundation of probabilistic models for risky choice experiments that may be viewed, in part, as a generalization of the von Neumann–Morgenstern expected utility theory. This setting means that the agent's choice behavior in replications of choice settings (with uncertain outcomes) is assumed to be governed by a probability mechanism. The motivation for this generalization is twofold. First, it is of interest to establish a probabilistic framework that is justified on theoretical grounds and that can be used in microeconometric empirical analysis of choice behavior under uncertainty. Apart from a few rather particular cases, no such framework seems to be available. Second, it is of independent theoretical interest to extend the von Neumann–Morgenstern theory to allow for errors in the decision process of the agents. There is a huge literature on stochastic choice models with certain outcomes; see, for example, chapter 2 in Anderson, Palma and Thisse (1992) and Fishburn (1998) for reviews of discrete choice models. In fact, it was empirical observations of inconsistencies, dating back to Thurstone (1927a,b), that led to the study of probabilistic theories in the first place. Thurstone argued that one reason for observed inconsistent choice behavior is bounded rationality in the sense that the agent is viewed as having difficulties with assessing the precise value (to him or her) of the choice objects. Whereas probabilistic models for certain outcomes have been studied and applied extensively in psychology and economics, it seems that there has been less interest in developing corresponding models for choice with uncertain outcomes. (For a summary of models with uncertain outcomes, see Fishburn (1998) and Starmer (2000, Section 6.2).) This is somewhat curious, as one would expect that if an agent has problems with rank ordering alternatives with certain outcomes, he or she would most certainly find it difficult to choose among lotteries.

The importance of developing theoretically justified stochastic choice models for uncertain outcomes has been articulated by Harless and Camerer (1994) and Hey and Orme (1994). For example, Hey and Orme summarize their view as follows:

> "... we are tempted to conclude by saying that our study indicates that behavior can be reasonably well modeled (to what might be termed a 'reasonable approximation') as 'Expected utility plus noise'. Perhaps we should now spend some time thinking about the noise, rather than about even more alternatives to expected utility?" (pp. 1321–1322)

In this paper, we first consider models for choice among lotteries with "minimal structure" that are *not* analogous to von Neumann–Morgenstern's expected utility theory.[1] We restate axioms known from the theoretical literature on probabilistic choice, which are known as *Order-independence*, *Balance conditions*, the *Quadruple condition* and the *Independence from Irrelevant alternative condition* (IIA). From Order-independence and Balance conditions, it follows that the binary choice probabilities are *simply scalable* (Tversky, 1972), and that the scale is a functional of the lottery outcome probabilities.

From the Quadruple and Balance conditions, combined with a particular regularity condition, it follows by an extension of Debreu's result (cf. Debreu, 1958), that the binary choice probabilities can be represented as a *Fechnerian* model. That is, there exists a suitable utility functional of lottery outcome probabilities such that the binary choice probabilities can be expressed as a symmetric cumulative distribution functional of the respective utility differences. Subsequently, we consider the multinomial case where the agent faces a choice set of several lotteries. In this case, we apply different combinations of the axioms mentioned above together with IIA. In this context, IIA yields a Luce model where the utility of a lottery is a general functional of the lottery outcome probabilities associated with this lottery. The IIA axiom implies that there exists a scale functional (utility) of the lottery outcome probabilities such that the choice probabilities can be expressed as a multinomial logit model in terms of the utilities (Luce, 1959).

Next, we consider a generalization of the Expected Utility theory. We propose axioms that can be viewed as a probabilistic version of the so-called *Archimedean Axiom*, and two of the axioms can be viewed as probabilistic versions of the *Independence Axiom* in the von Neumann–Morgenstern theory of expected utility. These probabilistic versions extend the basic von Neumann–Morgenstern axioms in the following sense: whereas the *Archimedean* and *Independence* Axioms may not necessarily hold in a single-choice experiment, the probabilistic versions state that they will hold in an aggregate sense (to be made precise below) when the agent participates in a large number of replications of a choice experiment. The intuition is that the agent may be boundedly rational and make errors when he or she evaluates the value to him or her of the respective choice alternatives (strategies) in each single replication of the experiment, but on average (across replications of the experiment), the agent shows no systematic departure from the von Neumann–Morgenstern type of axioms. Alternatively, the probabilistic axioms may be conveniently interpreted in the context of an

[1] Most of the results presented in this paper have appeared previously in Dagsvik (2005). However, some results are new, the presentation of results have been reorganized and errors have been corrected.

observationally homogeneous *population* of agents that face the same choice experiment. Whereas each agent's behavior is allowed to deviate from the von Neumann–Morgenstern axioms, the "aggregate" behavior in the population is assumed to be consistent with these axioms. The latter type of interpretation is analogous to the most common one within the theory of discrete choice (see, for example, McFadden, 1981, 1984).

Next, we demonstrate that different combinations of the probabilistic Archimedean and Independence Axioms, combined with other additional axioms, imply particular characterizations of the probabilities for choice among risky prospects as a function of the lottery outcome probabilities.

As a particular case within our generalized Expected Utility theory, we study settings with monetary rewards. What distinguishes this case from the general situation is that the outcomes (money) are realizations of an *ordered* variable. Accordingly, it is possible to use this property to obtain additional characterization of the model. The (additional) axiom that yields this characterization states the following: if the probability of preferring lottery one over lottery two is less than the probability of preferring lottery three over lottery four, this inequality remains true when all outcomes are rescaled by the same factor while the lottery outcome probabilities remain unchanged.

In an empirical context, the choice probabilities implied by the proposed axioms in this paper are essential for establishing the link between theory and the corresponding empirical model. More precisely, the agents' choices among lotteries are, from a statistical point of view, outcomes of a multinomial experiment with probabilities equal to the respective choice probabilities mentioned above. Accordingly, when the structure of the choice probabilities has been obtained, one can, in the context of empirical analysis, apply standard inference methods such as maximum likelihood estimation procedures and likelihood ratio tests.

The first work on stochastic models for choice among lotteries occurred in the 1960s. Becker et al. (1963) proposed a probabilistic model for choice among lotteries, which they called a *Luce Model for Wagers*. Luce and Suppes (1965) considered a special case of the Luce model for wagers, which they called the *Strict Expected Utility Model*. However, neither these authors nor more recent contributions discuss the issue of deriving a stochastic model from axiomatization. To the best of our knowledge, the only contribution that provides a model founded on an axiomatic basis is by Fishburn (1978), who develops the *incremental expected utility advantage* model. However, his model does not contain the expected utility model as a special case, although the expected utility model can be approximated by an incremental expected utility advantage model. As pointed out by Fishburn (1978, pp. 635–636), the incremental advantage model seems extreme as it implies that there is a positive probability of choosing $1 in a choice between $1 for certain or a gamble that yields $10 000 with probability .999 or $0 with probability .001.

Allais (1953) argued that individuals may systematically violate the expected utility theory. His examples have later been viewed as special cases of phenomena called the *common consequence effect* and the *common ratio effect*. It is interesting that the stochastic version of the expected utility theory developed here in fact allows for the common consequence and the common ratio effects.

The paper is organized as follows. In the next section, we present some basic concepts and notation. In Section 3, we discuss different types of axiomatizations and characterizations for binary choice models with "minimal structure". In Section 4, we consider the generalization of the Expected Utility theory. In Section 5, we specialize to the case with monetary rewards, and in Section 6, we discuss a random utility representation. In Section 7, we show that the models developed are able to rationalize the common consequence effect and the common ratio effect. Finally, we discuss an example in Section 8.

## 2. Preliminaries

The aim of this section is to introduce axioms that enable us to characterize choice among lotteries when some randomness is present in the agent's choice behavior in the sense that if he or she faces several replications of a specific choice experiment, then he or she may choose different lotteries each time. The reason for this type of inconsistent behavior may be, as mentioned above, that the agent has difficulties with evaluating the proper value (to him or her) of the respective lotteries.

Let X denote the set that indexes the set of outcomes, which is assumed to be finite and to contain m outcomes, i.e., $X \equiv \{1,2,...,m\}$. In the following, we shall assume, as is customary, that the agent's information about the chances of the different realizations of lottery s can be represented by lottery outcome probabilities

$$g_s := (g_s(1), g_s(2), ..., g_s(m)),$$

where $g_s(k)$ is the probability of outcome k, $k \in X$, if lottery s is chosen. Let S denote the set of simple probability measures on the algebra of all subsets of the set of outcomes. Recall that a *preference relation* refers to a binary relation, $\succsim$, on S that is: (i) *complete*, i.e., for all $g_r$, $g_s \in S$ either $g_r \succsim g_s$ or $g_s \succsim g_r$; and (ii) *transitive*, i.e., for all $g_r$, $g_s$, $g_t$, in S, $g_r \succsim g_s$ and $g_s \succsim g_t$ implies $g_r \succsim g_t$. A real-valued function $L(g_s)$ on S *represents* $\succsim$ if for all $g_r$, $g_s \in S$, $g_r \succsim g_s$, if and only if $L(g_r) \geq L(g_s)$. Let $\mathsf{B}$ be the family of *finite* subsets of S that contain at least two elements.

Consider now the following choice setting. The agent faces n replications of a choice experiment in which a set B of lotteries, $B \in \mathsf{B}$, is presented in each replication. We assume that there is no learning. As there is an element of randomness in the agent's choice behavior, he or she may choose different lotteries in different replications. We assume that the agent's choices in different replications are stochastically independent. Let $P_B(g_s)$, $g_s \in B$, be the probability that $g_s$ is the most preferred vector of lottery outcome probabilities in B. Let $P(g_r, g_s)$ be the probability that lottery $g_r$ is chosen over $g_s$, i.e., $P(g_r, g_s) \equiv P_{\{g_r, g_s\}}(g_r)$. It then follows that $P(g_r, g_s) > P(g_s, g_r)$ if and only if $P(g_r, g_s) > 0.5$. The argument above provides a motivation for the following definition.

**Definition 1**

*For $g_r, g_s \in S$, lottery $g_r$ is said to be strictly preferred to $g_s$ in the aggregate sense, if and only if $P(g_r, g_s) > 0.5$. If $P(g_r, g_s) = 0.5$, then $g_r$ is, in the aggregate sense, indifferent to $g_s$.*

Thus, Definition 1 introduces a binary relation, $\succsim$, where $g_r \succ g_s$ means that $g_r$ is strictly preferred to $g_s$ (in the aggregate sense), whereas $g_r \sim g_s$ means that $g_r$ is indifferent to $g_s$. Note, however, that the relation is not necessarily a *preference relation*. The reason for this is that the binary relation $\succsim$ is *not* necessarily transitive. That is, for $g_1, g_2, g_3 \in S$, the statement that $P(g_1, g_2) \geq 0.5$ and $P(g_2, g_3) \geq 0.5$ imply $P(g_1, g_3) \geq 0.5$ is not necessarily true.

Let $g_1, g_2 \in S$. The mixed lottery, $\alpha g_1 + (1-\alpha) g_2, \alpha \in [0,1]$, is a lottery in S yielding the probability $\alpha g_1(k) + (1-\alpha) g_2(k)$ of outcome k, $k \in X$. Here, we assume that the agents perceive the lotteries $\alpha\beta g_1 + (1-\alpha\beta) g_2$ and $\beta[\alpha g_1 + (1-\alpha) g_2] + (1-\beta) g_2$, $\alpha, \beta \in [0,1]$ as equivalent. This property is known as the axiom of reduction of compound lotteries, cf. Luce and Raiffa (1957).

For sets, $A, B \in \mathcal{B}$ such that $A \subseteq B$, let

$$P_B(A) \equiv \sum_{g_s \in A} P_B(g_s).$$

The interpretation is that $P_B(A)$ is the probability that the agent will choose a lottery within A when B is the choice set.

# 3. Probabilistic models with minimal structure

In this section, we shall consider models that are not necessarily extensions of the von Neumann–Morgenstern expected utility theory. We start by introducing behavioral axioms, which will lead to different types of characterizations of the choice probabilities.

**Axiom 1** (*Order-independence*)

*For all $g_1, g_2, g_3 \in S$*

*(i)* $P(g_1, g_2) \geq \frac{1}{2}$ *if and only if* $P(g_1, g_3) \geq P(g_2, g_3)$;

*(ii)* *the Balance condition:* $P(g_1, g_2) + P(g_2, g_1) = 1$.

Axiom 1 is a special case of what is called the *order-independence* condition (see Suppes et al., 1989, pp. 411–412). The intuition is that if $g_1$ is chosen more frequently than $g_2$, then the fraction of times that $g_1$ is preferred over $g_3$ is higher than the fraction of times that $g_2$ is preferred over $g_3$. The Balance condition implies that the relation $\succsim$ is *complete*. Note that Axiom 1 implies that the relation $\succsim$ is transitive, because if

$$P(g_1,g_2) \geq \frac{1}{2} \text{ and } P(g_2,g_3) \geq \frac{1}{2},$$

then, by (i)

$$P(g_1,g_3) \geq P(g_2,g_3) \geq \frac{1}{2}.$$

Hence, Axiom 1 implies that the relation given in Definition 1 is a preference relation.

The next result is due to Tversky (1972).

**Theorem 1**

*For all $g_1, g_2 \in S$, Axiom 1 holds if and only if*

$$P(g_1,g_2) = F\left(f(g_1), f(g_2)\right), \tag{3.1}$$

*for some suitable scale functional f, $f : S \rightarrow R$, and $F : R^2 \rightarrow (0,1)$, where F is a function that is strictly increasing in its first argument and strictly decreasing in the second.*

The proof of Theorem 1 is found in Tversky (1972) and Suppes et al. (1989, p. 419).

When $P(g_1,g_2)$ can be represented as in Theorem 1, the choice probabilities are said to be *simply scalable*, cf. Suppes et al. (1989, p. 410). The representation (3.1) seems to be the weakest possible representation for choice under uncertainty that one can think of. It would include any kind of probabilistic binary nonexpected utility model because the *"preference functional"* f is allowed to be completely general. Despite its generality, simple scalability is violated in some choice contexts; see, for example, Problem 2 in Suppes et al. (1989, p. 413).

Although the model in (3.1) involves a scale representation, it is not fully satisfactory from an empirical standpoint because the functional f is specified and the composition rule F is very general.

**Axiom 2**

*Let* $g_1, g_2, g_3, g_4 \in S$ *. The binary choice probabilities satisfy*

*(i)* *the Quadruple condition:* $P(g_1,g_2) \geq P(g_3,g_4)$ *if and only if* $P(g_1,g_3) \geq P(g_2,g_4)$ *; moreover, if either antecedent inequality is strict, so is the conclusion;*

*(ii)* *Solvability: for any* $y \in (0,1)$ *and any* $g_1, g_2, g_3 \in S$ *satisfying* $P(g_1,g_2) \geq y \geq P(g_1,g_3)$*, there exists a* $g \in S$ *such that* $P(g_1,g) = y$ *;*

*(iii)* *the Balance condition:* $P(g_1,g_2) + P(g_2,g_1) = 1$*.*

Axiom 2 is due to Debreu (1958). The intuition of the Quadruple condition is related to the following example, where the binary choice probabilities have the form of the representation:

$$\mathrm{P}(\mathrm{g}_1,\mathrm{g}_2) = \mathrm{K}(\mathrm{f}(\mathrm{g}_1) - \mathrm{f}(\mathrm{g}_2)),$$

where K is a strictly increasing cumulative distribution function on R, and f is a suitable mapping from S to R. Clearly, the choice model in this example satisfies the Quadruple condition. The example shows that when the average value of some lottery outcome probabilities g is represented by a scale functional, f(g), in such a way that the propensity to prefer lottery outcome probabilities $g_1$ over $g_2$ is a function of the "distance", $\mathrm{f}(\mathrm{g}_1) - \mathrm{f}(\mathrm{g}_2)$, then the Quadruple condition must hold. The Solvability condition is fairly intuitive. If K is continuous, the Balance condition will also be fulfilled in the example above.

**Theorem 2**

*Axiom 2 holds if and only if there exists a continuous and strictly increasing cumulative distribution function K with* $K(x) + K(-x) = 1$*, and a mapping f from S to some interval I such that the binary choice probabilities can be represented as*

$$P(g_1,g_2) = K\{f(g_1) - f(g_2)\}, \tag{3.2}$$

*for* $g_1, g_2 \in S$*, where*

$$I = \{x : x = f(g), g \in S\}.$$

*The mapping f is unique up to a linear transformation. The cumulative distribution function K is unique in the sense that if* $K_1$ *and* $K_0$ *are two representations, then* $K_0(x) = K_1(ax)$*, where* $a > 0$ *is a constant.*

The proof of Theorem 2 is given in the Appendix. In the special case with scalar representation of alternatives, proofs are given in Falmagne (1985) and Suppes et al. (1989). However, their proofs do not apply when the lotteries are represented by vectors of lottery outcome probabilities. In the special case with binary outcomes, the lottery outcome probabilities of lottery j can be represented by a scalar, namely by $g_j(1)$, $g_j(1) \in [0,1]$.

The model in (3.2) is known in the literature as the Fechnerian representation (Fechner, 1860/1966), and it clearly implies that the relation given in Definition 1 is a preference relation. It has the advantage of providing a cardinal utility functional representation. This model is, however, still too general to be directly useful in empirical applications because Theorem 2 is silent about the structure of the functional f and the c.d.f. K. If f is linear in the lottery outcome probabilities, we get a binary probabilistic version of the expected utility theory as a special case. Different versions of probabilistic nonexpected utility models follow as special cases when values of f are specified as suitable nonlinear functionals (see, for example, Starmer, 2000).

A crucial building block for corresponding choice probabilities in the multinomial case is the following axiom.

**Axiom 3 (IIA)**

*For a given* $g_s \in S$, $P(g_s, g_r) \in (0,1)$ *for all* $g_r \in S$. *Furthermore, for* $g_s \in A \subseteq B$, $A, B \in \mathcal{B}$,

$$P_B(g_s) = P_A(g_s) P_B(A).$$

Axiom 3 was first proposed by Luce (1959) in the context of probabilistic choice with certain outcomes, and it is called *"Independence from Irrelevant Alternatives"* (IIA). As is well known, it represents a probabilistic version of rationality in the following sense: Suppose the agent faces a set B of feasible lotteries. One may view the agent's choice as if it takes place in two stages. In stage one, the agent selects a subset from B, which contains the most attractive alternatives. In the second stage, he or she chooses the most preferred alternative from this subset. In the second stage, the alternatives outside the subset selected in stage one are *irrelevant*. Thus, rationality is associated with the property that the agent only takes into consideration the lotteries within the *presented* choice set. The probability that a particular set A (say) will be chosen in the first stage is $P_B(A)$, and the probability that $g_s$ is chosen (when alternatives in B\A are irrelevant) is $P_A(g_s)$. Thus, $P_B(A)P_A(g_s)$ is the final probability of choosing $g_s$. As indicated above, the crucial point here is that $P_A(g_s)$ is *independent* of alternatives outside A. For the sake of interpretation, let J(B) denote the agent's chosen lottery from B. With this notation, we can express IIA as:

$$\mathrm{P_B}(\mathrm{g_s}) = \mathrm{P}(\mathrm{J(B)} = \mathrm{g_s}) = \mathrm{P}(\mathrm{J(B)} \in \mathrm{A})\mathrm{P}(\mathrm{J(A)} = \mathrm{g_s}).$$

The conditional probability of choosing $g_s$ given that the choice belongs to A, equals

$$\mathrm{P}(\mathrm{J(B)} = \mathrm{g_s} \,|\, \mathrm{J(B)} \in \mathrm{A}) = \frac{\mathrm{P}(\mathrm{J(B)} = \mathrm{g_s})}{\mathrm{P}(\mathrm{J(B)} \in \mathrm{A})},$$

so that IIA can be rewritten as

$$\mathrm{P}(\mathrm{J(B)} = \mathrm{g_s} \,|\, \mathrm{J(B)} \in \mathrm{A}) = \mathrm{P}(\mathrm{J(A)} = \mathrm{g_s}).$$

Whereas $\mathrm{P}(\mathrm{J(A)} = \mathrm{g_s})$ is the probability of choosing $g_s$ from a given choice set A, the conditional probability

$$\mathrm{P}(\mathrm{J(B)} = \mathrm{g_s} \,|\, \mathrm{J(B)} \in \mathrm{A})$$

expresses the conditional probability of choosing $g_s$ from a given choice set B, given that the choice from B belongs to A. Clearly,

$$\mathrm{P}(\mathrm{J(B)} = \mathrm{g_s} \,|\, \mathrm{J(B)} \in \mathrm{A})$$

will in general be different from

$$\mathrm{P}(\mathrm{J(A)} = \mathrm{g_s}).$$

They only coincide when IIA holds.

As Axiom 3 is a probabilistic statement, it means that it represents probabilistic rationality in the sense that lotteries outside the second-stage choice set A may matter in single-choice experiments but will not affect average behavior. The following result has been obtained by Luce (1959):

**Theorem 3**

*Axiom 3 holds if and only if there exist representative scale values, $f(g_s)$, for some functional f, such that*

*(3.3)*
$$P_B(g_s) = \frac{exp(f(g_s))}{\sum_{g_r \in B} exp(f(g_r))},$$

*for all $g_r \in B, B \in \mathrm{B}$.*

Thus, Axiom 3 implies that the relation given in Definition 1 is a preference relation. We realize that Axioms 1 and 2 are implied by Axiom 3. Under IIA, the representation (3.3) is the weakest possible representation that one can think of. It would include any kind of probabilistic nonexpected utility model because the functional f is allowed to be completely general.

Similarly to the models discussed above, the weakness of the representation (3.3) is that our theory is silent about the structure of the functional f.

# 4. Probabilistic extensions of the expected utility theory

As mentioned above, the theory developed so far has no implication for the structure of the functional f. In this section, we shall introduce axioms that allow further characterization.

The purpose of the first axiom is to impose necessary and sufficient conditions to insure that the binary relation given in Definition 1 is a preference relation.

**Axiom 4** (*Weak Stochastic Transitivity and Completeness*)

*Let* $g_1, g_2, g_3 \in S$. *The binary choice probabilities satisfy*

*(i)* *weak Stochastic Transitivity: if* $P(g_1,g_2) \geq \frac{1}{2}$ *and* $P(g_2,g_3) \geq \frac{1}{2}$, *then* $P(g_1,g_3) \geq \frac{1}{2}$;

*(ii)* *the Balance condition:* $P(g_1,g_2) + P(g_2,g_1) = 1$.

Recall that the *Balance condition* is equivalent to *completeness*. It follows immediately that the binary relation given in Definition 1 is a *preference relation*, provided it satisfies Axiom 4.

Next, we introduce axioms that are intuitive probabilistic versions of the so-called *Archimedean* and *Independence* Axioms of von Neumann and Morgenstern.

**Axiom 5** (*Archimedean*)

*For all* $g_1, g_2, g_3 \in S$, *if*

$$P(g_1,g_2) > \frac{1}{2} \quad and \quad P(g_2,g_3) > \frac{1}{2},$$

*then there exist* $\alpha, \beta \in (0,1)$ *such that*

$$P(\alpha g_1 + (1-\alpha) g_3, g_2) > \frac{1}{2} \quad and \quad P(g_2, \beta g_1 + (1-\beta) g_3) > \frac{1}{2}.$$

Axiom 5 is a probabilistic version of the *Archimedean Axiom* in the von Neumann-Morgenstern expected utility theory because, by Definition 1, it is equivalent to the following statement: if $g_1 \succ g_2$ and $g_2 \succ g_3$, then there exist $\alpha, \beta \in (0,1)$ such that $\alpha g_1 + (1-\alpha) g_3 \succ g_2$ and $g_2 \succ \beta g_1 + (1-\beta) g_3$, cf. Karni and Schmeidler (1991, p. 1769). Note that Axiom 5 is weaker than the assumption that $P(g_r, g_s)$ is continuous. This is because if $P(g_r, g_s)$ is continuous in $(g_r, g_s)$, then whenever $P(g_1, g_2) > 1/2$ and $P(g_2, g_3) > 1/2$, continuity implies that $P(\alpha g_1 + (1-\alpha) g_3, g_2) > 1/2$ and $P(g_2, \beta g_1 + (1-\beta) g_3) > 1/2$ for a suitable $\alpha, \beta \in (0,1)$.

**Axiom 6** (*Independence*)

*For all $g_1, g_2, g_3 \in S$, and all $\alpha \in [0,1]$, if*

$$P(g_1, g_2) \geq \frac{1}{2},$$

*then*

$$P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3) \geq \frac{1}{2}.$$

Axiom 6 is a probabilistic version of the *Independence Axiom* in the von Neumann–Morgenstern expected utility theory because it is equivalent to the statement that if $g_1 \succ g_2$, then $\alpha g_1 + (1-\alpha) g_3 \succ \alpha g_2 + (1-\alpha) g_3$, cf. Karni and Schmeidler (1991, p. 1769).

**Theorem 4** (*von Neumann-Morgenstern*)

*Let $\succsim$ be a binary relation. The following two conditions are equivalent.*

*(i)* $\succsim$ *is a preference relation satisfying Axioms 5 and 6.*

*(ii)* *There exists a function, $u : X \to R$, that is unique up to a positive affine transformation such that the functional $V : S \to R$ defined by*

$$V(g) = \sum_{k \in X} u(k)\, g(k), \tag{4.1}$$

*represents the preference relation.*

Theorem 4 is the von Neumann–Morgenstern expected utility theorem, cf. Karni and Schmeidler (1991, pp. 1769–1770).

Recall that we cannot apply the result of Theorem 4 in our context without additional assumptions because the binary relation of Definition 1 is not necessarily a preference relation.

As the binary relation given in Definition 1 is a preference relation when it satisfies Axiom 4, the next corollary follows.

**Corollary 1**

*Assume that Axioms 4, 5 and 6 hold. Then for* $g_1, g_2 \in S$ *,*

$$P(g_1, g_2) \geq \frac{1}{2} \Leftrightarrow V(g_1) \geq V(g_2).$$

*Moreover, if either antecedent inequality is strict, so is the conclusion.*

Even if the binary relation given in Definition 1 satisfies Axioms 4, 5 and 6, we would still not be able to specify choice probabilities that represent a generalization of the expected utility theory. We would at most be able to ascertain whether $g_r$ is preferred to $g_s$ (say) in the aggregate sense. Consequently, similarly to the approach in Section 3, we need to provide additional theoretical building blocks to ascertain precisely how the choice probabilities $\{P(g_r, g_s)\}$ can be represented by the lottery outcome probabilities $g_r$ and $g_s$. This is crucial for establishing a link between the theoretical concepts introduced above and a model that is applicable for empirical modeling and analysis.

The next result is analogous to Theorem 1.

**Theorem 5**

*For all* $g_1$ *and* $g_2 \in S$*, Axioms 1, 5 and 6 hold if and only if*

$$P(g_1, g_2) = F(V(g_1), V(g_2)),$$

*where*

$$V(g) = \sum_{k \in X} u(k) g(k),$$

*and* $F : R^2 \to (0,1)$ *is a function that is strictly increasing in its first argument and strictly decreasing in the second, and u:* $X \to R$ *is a function that is unique up to a positive linear transformation.*

The proof of Theorem 5 is given in the Appendix.

**Theorem 6**

*For all* $g_1, g_2 \in S$ *, Axioms 2, 5 and 6 hold if and only if*

$$(4.2) \qquad P(g_1, g_2) = K\left\{h\left(V(g_1)\right) - h\left(V(g_2)\right)\right\},$$

*where*

$$(4.3) \qquad V(g_s) = \sum_{k \in X} u(k) g_s(k),$$

*and K is a continuous and strictly increasing cumulative distribution function defined on R with* $K(x) + K(-x) = 1$*,* $h : R \to R$ *is strictly increasing and* $u : X \to R$*. The mappings K, h and V are unique in the sense that if* $K_0$ *and* $K_1$*,* $h_0(V_0)$ *and* $h_1(V_1)$ *are two representations, then* $K_0(x) = K_1(ax)$*, where* $a > 0$ *is a constant,* $V_1(g_s) = b_1 V_0(g_s) + c_1$ *and* $h_1(b_1 x + c_1) = b_2 h_0(x) + c_2$*, where* $b_1 > 0$*,* $b_2 > 0$*,* $c_1$ *and* $c_2$ *are constants.*

The proof of Theorem 6 is given in the Appendix.

**Remark**

Note that the formulation in (4.2) is equivalent to

$$\mathrm{P}(\mathrm{g}_1, \mathrm{g}_2) = \tilde{\mathrm{K}}\left(\tilde{\mathrm{h}}\left(\mathrm{V}(\mathrm{g}_1)\right) / \tilde{\mathrm{h}}\left(\mathrm{V}(\mathrm{g}_2)\right)\right),$$

where $\tilde{\mathrm{K}}$ is a continuous and strictly increasing c.d.f. on $R_+$ and $\tilde{\mathrm{h}}$ is positive and strictly increasing. This follows immediately from (4.2), by defining $\tilde{\mathrm{K}}(\mathrm{x}) = \mathrm{K}\left(\mathrm{e}^{\mathrm{x}}\right)$ and $\tilde{\mathrm{h}}(\mathrm{x}) = \log \mathrm{h}(\mathrm{x})$.

**Theorem 7**

*Assume that* $P(g_r, g_s) \in (0,1)$ *for all* $g_r, g_s \in S$*. Then, for* $B \in \mathcal{B}$ *, Axioms 3, 5 and 6 hold if and only if*

$$(4.4) \qquad P_B(g_s) = \frac{exp\left(h\left(V(g_s)\right)\right)}{\sum_{g_r \in B} exp\left(h\left(V(g_r)\right)\right)},$$

*where*

$$V(g_s) = \sum_{k \in X} u(k) g_s(k),$$

*and* $h: R \to R$ *is strictly increasing and* $u: X \to R$. *The mappings h and V are unique in the sense that if* $h_0(V_0)$ *and* $h_1(V_1)$ *are two representations, then* $V_1(g_r) = bV_0(g_r) + c$ *and* $h_1(bx + c) = h_0(x) + d$ *, where* $b > 0$ *, c and d are constants.*

The proof of Theorem 7 is given in the Appendix.

The choice model obtained in Theorem 7 is a special case of the *Luce model for wagers*, proposed by Becker et al. (1963). They postulated that

(4.5) $$P_B(g_s) = \frac{\psi(V(g_s))}{\sum_{g_r \in B} \psi(V(g_r))},$$

where $\psi: R \to R_+$ is a strictly increasing mapping that is unique up to a multiplicative constant. By letting $\log \psi(x) = h(x)$, we find that (4.5) is equivalent to (4.4).

The model in (4.4) characterizes the choice probabilities in terms of a linear preference functional V up to an unknown monotone mapping h. However, the underlying axioms imply no additional restrictions on h.

**Axiom 7** *(Strong independence)*

*For all* $g_1, g_2, g_1^*, g_2^*, g_3 \in S$ *and all* $\alpha \in [0,1]$, *if*

$$P(g_1, g_2) \geq P(g_1^*, g_2^*),$$

*then*

$$P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3) \geq P(\alpha g_1^* + (1-\alpha) g_3, \alpha g_2^* + (1-\alpha) g_3).$$

Axiom 7 states that if the fraction of replications where $g_1^*$ is chosen over $g_2^*$ is less than or equal to the fraction of replications where $g_1$ is chosen over $g_2$, this inequality still holds when $g_j$ is replaced by $\alpha g_j + (1-\alpha) g_3$ and $g_j^*$ is replaced by $\alpha g_j^* + (1-\alpha) g_3$, for $j = 1,2$. Note that in Axiom 7, it is *not* claimed that $P(g_1, g_2)$ is equal to $P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3)$.

It follows that Axiom 7 implies Axiom 6. To realize this, note that when $g_1^* = g_2^*$, then $P(g_1^*, g_2^*) = 1/2$, and

$$P\left(\alpha g_1^* + (1-\alpha) g_3, \alpha g_2^* + (1-\alpha) g_3\right) = 1/2 .$$

Thus, it follows from this and Axiom 7 that when

$$P(g_1, g_2) \geq 1/2 ,$$

then:

$$P\left(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3\right) \geq 1/2 ,$$

which we recognize as Axiom 6.

The intuition why Axiom 7 is stronger than Axiom 6 is related to the fact that it represents a statement that involves comparisons between the degree to which one lottery is chosen over a second and the degree to which a third lottery is chosen over a fourth. It is this strengthening that enables us to derive strong functional form restrictions on the choice probabilities, to be considered next.

**Theorem 8**

*Axioms 2, 5 and 7 hold if and only if the choice probabilities have the form as in (4.2) with*

*(4.6)* $$h(x) = \beta x + \kappa ,$$

*where* $\beta > 0$ *and* $\kappa$ *are constants.*

The proof of Theorem 8 is given in the Appendix.

**Corollary 2**

*For all* $g_r, g_s \in S$, *and* $B \in \mathcal{B}$ , *Axioms 3, 5 and 7 hold if and only if*

$$P_B(g_s) = \frac{exp\left(V(g_s)\right)}{\sum_{g_r \in B} exp\left(V(g_r)\right)} .$$

The proof of Corollary 2 is given in the Appendix.

There are two alternative interpretations of the Axioms above, which represent extensions of the corresponding von Neumann–Morgenstern axioms. The first interpretation is as follows. Consider an agent who participates in a large number of replications of a choice experiment. The agent may be boundedly rational in the sense that he or she has difficulties assessing the precise value (to him or her) of the strategies in each single replication. This may be so even if the agent has no problem with assessing the values of the outcomes, simply because the evaluations of the respective lottery strategies do not follow immediately from the values of the outcomes and the outcome probabilities. The axioms state that whereas the agent is allowed to make "errors" when selecting strategies in each replication of the experiment (in the sense that his or her behavior is not consistent with the von Neumann–Morgenstern theory), the agent will still—in the aggregate sense specified in the axioms—behave according to the respective versions of the probabilistic extension of the von Neumann–Morgenstern theory.

In the alternative interpretation, we consider a large observationally homogeneous population. In this setting, each agent in the population faces the same choice experiment. Although the behavior of each individual agent may be inconsistent with the von Neumann–Morgenstern theory, the axioms above state that aggregate behavior in the population will be consistent with the probabilistic version of the theory.

# 5. Monetary rewards

The setting we shall discuss here is somewhat different from the previous one in that we focus on applications where money is involved. Specifically, we now assume that the set of outcomes is a set of *money amounts*. Thus, the lottery outcomes of the choice experiment consist of pairs $\{(k, w_k) \in X \times W\}$, where W is a subset of $[0, \infty)$ or equal to $[0, \infty)$ and $w_k \in W$ is a given amount associated with outcome k. The corresponding probability of outcome $(k, w_k)$, given lottery s is denoted by $g_s(k, w_k)$. The utilities are now given as $\{u(k, w_k)\}$. Let S and $\mathcal{B}$ be defined as in Section 2. What distinguishes the present setting from the previous one is that one component (money) of the outcome is an *ordered* variable. The purpose of this section is to utilize this property to characterize the functional form of the utility function $u(k, w)$.

Let $\delta_w$ denote the Dirac measure, that is:

$$\delta_w(x) = \begin{cases} 1 & \text{if } w = x \\ 0 & \text{otherwise}, \end{cases}$$

and define $g_s^\lambda$ by:

$$g_s^\lambda(k, \tilde{w}_k) = g_s(k, w_k)\delta_{\lambda w_k}(\tilde{w}_k),$$

where $(k, \tilde{w}_k) \in X \times W$ and $\lambda$ is a positive real number. Although $g_s^\lambda$ also depends on $\{w_k\}$, this is suppressed in the notation.

**Axiom 8**

*Let* $g_s \in S$ *for* $s = 1,2,3,4$. *Then*

$$P\left(g_1^1, g_2^1\right) \le P\left(g_3^1, g_4^1\right),$$

*if and only if*

$$P\left(g_1^\lambda, g_2^\lambda\right) \le P\left(g_3^\lambda, g_4^\lambda\right),$$

*for any* $\lambda > 0$ *and* $w_k \in W$, $k \in X$.

Axiom 8 means that if the fraction of individuals that prefer $g_1$ over $g_2$ is less than the fraction that prefers $g_3$ over $g_4$, then this inequality does not change if all the incomes (potential outcomes) are rescaled by the same factor while the lottery outcome probabilities remain unchanged.

Before we state the next result, we adopt the definition:

$$\frac{x^\theta - 1}{\theta} = \log x,$$

when $\theta = 0$. The function $\left(x^\theta - 1\right)/\theta$ will then be differentiable and strictly increasing for all θ.

The intuition is, loosely speaking, that the agent is "to some degree" viewed as being indifferent with respect to scale transformations of the potential money rewards when comparing lotteries. By the above notion of "some degree", it is meant that the respective binary choice probabilities may change as a result of the rescaling of rewards, but only in such a way that the original inequality of Axiom 8 is preserved.

**Theorem 9**

*Axioms 2, 5, 6 and 8 hold if and only if the choice probabilities have the form as in (4.2) with either*

*(i)* $$h(x) = \frac{\beta\left(e^{\theta x} - 1\right)}{\theta} + \kappa \quad and \quad u(k,w) = b \log w + c,$$

*for* $w > 0$, *and* $\theta \neq 0$,

*(ii)* $$h(x) = \beta x + \kappa \quad and \quad u(k,w) = b_k\left(\frac{w^\rho - 1}{\rho}\right) + c,$$

*for* $w > 0$, *and* $\theta = 0$, *or*

*(iii)* $$h(x)=\frac{\beta(x^{\theta}-1)}{\theta}+\kappa \quad \text{and} \quad u(k,w)=b_k w^{\rho},$$

*for $w \ge 0$, where $\beta > 0$, $b > 0$, $b_k > 0$, $\rho > 0$, $\kappa$ and c are constants.*

The proof of Theorem 9 is given in the Appendix.

**Axiom 9**

*Let $g_1, g_2 \in S$. Then*

$$P(g_1^1, g_2^1) = P(g_1^{\lambda}, g_2^{\lambda}),$$

*for any real number $\lambda > 0$ and $w_k \in W$, $k \in X$.*

Axiom 9 is stronger than Axiom 8 because it postulates that the choice probabilities are invariant under scale transformations of the rewards. This means that the agent is viewed as being indifferent with respect to rescaling of the potential money rewards.

**Corollary 3**

*Axioms 2, 5, 6 and 9 hold if and only if the choice probabilities have the form as in (4.2) with $\theta = 0$, i.e., either*

*(i)* $$h(x)=\beta x+\kappa \quad \text{and} \quad u(k,w)=b_k \log w + c,$$

*for $w > 0$, or*

*(ii)* $$h(x)=\beta \log x+\kappa \quad \text{and} \quad u(k,w)=b_k w^{\rho},$$

*for $w \ge 0$, where $\beta > 0$, $b_k > 0$, $\rho > 0$, $\kappa$ and c are constants.*

The proof of Corollary 3 is given in the appendix. Note that when $\beta = 1$, the choice model in Corollary 3 (ii) reduces to the so-called *Strict Expected Utility* model for uncertain outcomes proposed by Luce and Suppes (1965).

**Corollary 4**

*Axioms 3, 5, 6 and 8 hold if and only if the choice model (4.4) holds with h and u as in Theorem 9.*

**Proof**

The “if” part of the corollary is evident. Consider the “only if” part. Recall that Axioms 3, 5 and 6 imply Theorem 7. As Axiom 2 is implied by Axiom 3, the conditions of Theorem 9 are fulfilled, and thus the structure of h and u must be as in (i), (ii) or (iii) of Theorem 9.

Q.E.D.

**Corollary 5**

*Axioms 2, 5, 7 and 8 hold if and only if the choice model in (4.2) holds with*

$$h(x) = \beta x + \kappa \ \ and \ \ u(k,w) = b_k\left(\frac{w^{\rho} - 1}{\rho}\right) + c\,,$$

*for $w \geq 0$ if $\rho \neq 0$, and $w > 0$ if $\rho = 0$.*

**Proof**

Evidently, the “if” part of the corollary is true. Consider the “only if” part. Recall that Axioms 3, 5 and 7 imply Theorem 8. As Axiom 2 is implied by Axiom 3, the conditions of Theorem 9 are fulfilled, and as h must, by Theorem 8, be linear, the structure of h and u must be as in (ii) of Theorem 9, or as in (iii) of Theorem 9 with $\theta = 1$.

Q.E.D.

The next result is analogous to Corollary 4 and follows from Corollary 5.

**Corollary 6**

*Axioms 3, 5, 7 and 8 hold if and only if the choice model (4.4) holds with h and u as in Corollary 5.*

**Figure 1. Overview of axioms**

**Axiom 1**

(i) $P(g_1,g_2) \geq \frac{1}{2} \Leftrightarrow P(g_1,g_3) \geq P(g_2,g_3)$,

(ii) $P(g_1,g_2) + P(g_2,g_1) = 1$.

**Axiom 2**

(i) $P(g_1,g_2) \geq P(g_3,g_4) \Leftrightarrow P(g_1,g_3) \geq P(g_2,g_4)$,

(ii) For y such that $P(g_1,g_3) \geq y \geq P(g_1,g_2)$, there is a $g \in S$ such that $P(g_1,g) = y$,

(iii) $P(g_1,g_2) + P(g_2,g_1) = 1$.

**Axiom 3 (IIA)**

For given $g_s \in S$, $P(g_s,g_r) \in (0,1)$ for all $g_r \in S$,

$P_B(g_s) = P_A(g_s) P_B(A)$, $g_s \in A \subset B, A, B \in \mathcal{B}$ .

**Axiom 4**

(i) If $P(g_1,g_2) \geq \frac{1}{2}$ and $P(g_2,g_3) \geq \frac{1}{2} \Rightarrow P(g_1,g_3) \geq \frac{1}{2}$,

(ii) $P(g_1,g_2) + P(g_2,g_1) = 1$.

**Axiom 5**

If $P(g_1,g_2) > \frac{1}{2}$ and $P(g_2,g_3) > \frac{1}{2}$,

there exist $\alpha, \beta \in (0,1)$ such that:

$P(\alpha g_1 + (1-\alpha) g_3, g_2) > \frac{1}{2}$ and $P(g_2, \beta g_1 + (1-\beta) g_3) > \frac{1}{2}$ .

**Axiom 6**

$$P(g_1,g_2) > \frac{1}{2}$$

$$\Downarrow$$

$$P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3) > \frac{1}{2} .$$

for all $\alpha \in [0,1]$

**Figure 1 (cont). Overview of axioms**

**Axiom 7**

$$P(g_1, g_2) \geq P(g_1^*, g_2^*)$$

$$\Downarrow$$

$$P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3) \geq P(\alpha g_1^* + (1-\alpha) g_3, \alpha g_2^* + (1-\alpha) g_3)$$

for all $\alpha \in [0,1]$.

**Axiom 8**

$$P(g_1^1, g_2^1) \leq P(g_3^1, g_4^1),$$

$$\Updownarrow$$

$$P(g_1^\lambda, g_2^\lambda) \leq P(g_3^\lambda, g_4^\lambda), \forall \lambda > 0,$$

where $g_s^\lambda(k, \tilde{w}_k) = g_s(k, w_k)\delta_{\lambda w_k}(\tilde{w}_k)$.

**Axiom 9**

$P(g_1^1, g_2^1) = P(g_1^\lambda, g_2^\lambda)$ for all $\forall \lambda > 0$.

| | | |
|---|---|---|
| **Axiom 1** | $\Rightarrow$ | **Axiom 4** |
| **Axiom 2** | $\Rightarrow$ | **Axiom 1**<br>**Axiom 4** |
| **Axiom 3** | $\Rightarrow$ | **Axiom 1**<br>**Axiom 2**<br>**Axiom 4** |
| **Axiom 7** | $\Rightarrow$ | **Axiom 6** |
| **Axiom 9** | $\Rightarrow$ | **Axiom 8** |

**Figure 2. Relationship between axioms and binary choice probabilities**

| | | |
|---|---|---|
| **Axiom 1** | $\Leftrightarrow$ | **Theorem 1.** $P(g_1,g_2)=F(f(g_1),f(g_2))$, for F strictly increasing in its first argument and strictly decreasing in the second and f is some mapping from S to R. |
| **Axiom 2** | $\Leftrightarrow$ | **Theorem 2.** $P(g_1,g_2)=K(f(g_1)-f(g_2))$, for some function f that is unique up to a positive linear transformation and a c.d.f. K that is strictly increasing, symmetric and continuous. |
| **Axioms 4, 5, 6** | $\Leftrightarrow$ | **Corollary 1.** $P(g_1,g_2)\geq\frac{1}{2}\Leftrightarrow V(g_1)\geq V(g_2)$, where $V(g_s)=\sum_{k\in X}u(k)g_s(k)$. |
| **Axioms 1, 5, 6** | $\Leftrightarrow$ | **Theorem 5.** $P(g_1,g_2)=F(V(g_1),V(g_2))$, for F strictly increasing in its first argument and strictly decreasing in the second. |
| **Axioms 2, 5, 6** | $\Leftrightarrow$ | **Theorem 6.** $P(g_1,g_2)=K\{h(V(g_1))-h(V(g_2))\}$, with h and K strictly increasing, symmetric and continuous. |
| **Axioms 2, 5, 7** | $\Leftrightarrow$ | **Theorem 8.** $P(g_1,g_2)=K\{V(g_1)-V(g_2)\}$, with K strictly increasing, symmetric and continuous. |

**Figure 3. Relationship between axioms and multinomial choice probabilities**

| | | |
|---|---|---|
| **Axiom 3** | $\Leftrightarrow$ | **Theorem 3.** $P_B(g_s)=\frac{\exp(f(g_s))}{\sum_{g_r\in B}\exp(f(g_r))}$, for some f. |
| **Axioms 3, 5, 6** | $\Leftrightarrow$ | **Theorem 7.** $P_B(g_s)=\frac{\exp(h(V(g_s)))}{\sum_{g_r\in B}\exp(h(V(g_r)))}$, for some strictly increasing h. |
| **Axioms 3, 5, 7** | $\Leftrightarrow$ | **Corollary 2.** $P_B(g_s)=\frac{\exp(V(g_s))}{\sum_{g_r\in B}\exp(V(g_r))}$. |

**Figure 4. Relationship between axioms and choice probabilities for the case with monetary rewards**

| | | |
|---|---|---|
| **Axioms 2, 5, 6, 8** | $\Leftrightarrow$ | **Theorem 9.** Choice probabilities are as in Theorem 6 with $h(x)=\frac{\beta\left(e^{\theta x}-1\right)}{\theta}+\kappa$ and $u(k,w)=b\log w+c,\ w>0,\theta\neq 0,$ $h(x)=\beta x+\kappa$ and $u(k,w)=\frac{b_k\left(w^{\rho}-1\right)}{\rho}+c, w>0,$ or $h(x)=\frac{\beta\left(x^{\theta}-1\right)}{\theta}+\kappa$ and $u(k,w)=b_k w^{\rho},\ w\geq 0$, $\beta>0, b>0, b_k>0$. |
| **Axioms 2, 5, 6, 9** | $\Leftrightarrow$ | **Corollary 3.** Choice probabilities are as in Theorem 6 with $h(x)=\beta x+\kappa$ and $u(k,w)=b_k\log w+c,\ w>0$, or $h(x)=\beta\log x+\kappa$ and $u(k,w)=b_k w^{\rho},\ w\geq 0$, $\beta>0, b>0, b_k>0$. |
| **Axioms 3, 5, 6, 8** | $\Leftrightarrow$ | **Corollary 4.** Choice probabilities are as in Theorem 7 with $h(x)=\frac{\beta\left(e^{\theta x}-1\right)}{\theta}+\kappa$ and $u\left(k,w\right)=b\log w+c,\ w>0,\theta\neq 0,$ $h(x)=\beta x+\kappa$ and $u(k,w)=b_k\log w+c, w>0,$ or $h(x)=\frac{\beta\left(x^{\theta}-1\right)}{\theta}+\kappa$ and $u\left(k,w\right)=b_k w^{\rho},\ w\geq 0$, $\beta>0, b>0, b_k>0$. |
| **Axioms 2, 5, 7, 8** | $\Leftrightarrow$ | **Corollary 5.** Choice probabilities are as in Theorem 8 with $h(x)=\beta x+\kappa$ and $u\left(k,w\right)=\frac{b_k(w^{\rho}-1)}{\rho}+c_k, w\geq 0,\ \rho\neq 0,$ $\beta>0, b_k>0$, and $w>0$ when $\rho=0$. |
| **Axioms 3, 5, 7, 8** | $\Leftrightarrow$ | **Corollary 6.** Choice probabilities are as in Theorem 7 with $h(x)=\beta x+\kappa$ and $u\left(k,w\right)=\frac{b_k(w^{\rho}-1)}{\rho}+c_k, w\geq 0,\ \rho\neq 0,$ $\beta>0, b_k>0$, and $w>0$ when $\rho=0$. |

Figures 1–4 display a convenient overview and summary of the results obtained in the paper. It is an important feature of the axioms that they have direct empirical counterparts. Figures 2 and 3 emphasize the equivalences between sets of axioms and the structure of the respective choice probabilities. However, some of these choice probabilities depend on unknown functional forms (f, K and h). For example, all the binary choice probabilities depend on an unknown c.d.f. K. Only

Corollaries 2, 3, 4 and 5 and Theorem 9 yield fully specified functional forms for the choice probabilities. As all the axioms have explicit empirical counterparts, they can be used to test these models without relying on ad hoc functional form specifications. To carry out rigorous nonparametric tests of these axioms is in itself a complicated task. In fact, it seems that the general case with ordinal restrictions on choice probabilities of the type displayed in Figure 1 lies outside the scope of a large body of literature devoted to statistical hypotheses testing under ordinal constraints. As far as we know, only Iverson and Falmagne (1985) and Dagsvik and Røine (2006) have explicitly addressed the challenge of developing test procedures for this type of setting. In particular, Iverson and Falmagne (1985) discuss how one can test property (i) of Axiom 1 and property (i) of Axiom 2 within a maximum likelihood setting.

# 6. A random utility representation

In this section, we shall consider the problem of a random utility representation of the agent's preferences over lotteries that yield choice probabilities satisfying Axioms 2, 3 and 7. From the theory of discrete choice, we know that the Luce choice model is consistent with an additive random utility representation in which the error terms are independent (across alternatives) with extreme value c.d.f., $\exp\left(-e^{-x}\right)$. Here, the setting is not as simple as in the standard discrete choice case because S is *not* countable and g is a vector. Therefore, if a random utility representation $\{U(g), g \in S\}$ exists, it must be a multiparameter stochastic process, i.e., a *random field*.

**Theorem 10** *(Random utility representation)*

*There exist a probability space and random variables $\{\varepsilon(g), g \in S\}$ defined on it, such that $\varepsilon(g_s), s = 1,2,..., g_s \in S$, are independent for distinct $g_1, g_2, ...,$ and*

$$(6.1) \qquad P\left(\varepsilon(g_s) \le y\right) = exp\left(-e^{-y}\right),$$

*for $y \in R$. The random utility representation*

$$(6.2) \qquad U(g) = h\left(V(g)\right) + \varepsilon(g),$$

*for $g \in S$, is consistent with Axioms 3, 5 and 6, i.e., for $B \in \mathcal{B}$.*

$$(6.3) \qquad P_B(g_s) = P\left(U(g_s) = \max_{g_r \in B} U(g_r)\right) = \frac{exp\left(h\left(V(g_s)\right)\right)}{\sum_{g_r \in B} exp\left(h\left(V(g_r)\right)\right)}.$$

The proof of Theorem 10 is given in the appendix. The next result is immediate.

**Corollary 6**

*Axioms 3, 6 and 7 are consistent with the random utility representation*

$$U(g) = V(g) + \varepsilon(g),$$

*for* $g \in S$ *, where the c.d.f. of* $\varepsilon(g)$ *is given in (6.1).*

**Remark**

The distribution function given in (6.1) is a so-called type III *extreme value* distribution.[2] In statistics, the extreme value distributions arise as the asymptotic distributions of the maximum of i.i.d. random variables. Many authors have studied this distribution in the context of the theory of discrete choice and random utility models; see, for example, McFadden (1973), Yellott (1977) and Strauss (1979). Under different regularity conditions, they have demonstrated that (6.1) is the *only* distribution that implies a random utility representation that is consistent with the Luce model (IIA).

# 7. The Allais paradox

Starting with Allais (1953), it has long been known that people's behavior under uncertainty may systematically violate the Independent axiom in the expected utility theory. The examples that Allais (1953) discussed have played an important role in the development of nonexpected utility theory. The examples discussed by Allais are special cases of more general phenomena called the *common consequence effect* and the *common ratio effect*. To explain what these phenomena mean, let $g_1$ and $g_2$ be two lotteries with binary outcomes such that lottery one has payoff y with probability g and payoff c with probability $1-g$ . Lottery two has payoff q with probability g and payoff c with probability $1-g$ , where q is also a lottery that has payoff x with probability $\mu$ and payoff $x_0$ with probability $1-\mu$, $0<\mu<1$. The expected utilities of the first and second lotteries, $V_1, V_2$, are

$$V_1 = g x_2 + (1-g) c,$$

and

$$V_2 = \mu g x_1 + (1-\mu) g x_0 + (1-g) c .$$

[2] There seems to be some confusion in the literature about the terminology. Some authors call (5.1) the type III extreme value distribution, whereas other authors call it the type I extreme value distribution. Some authors also call it the Double Exponential Distribution.

The payoffs are nonnegative (usually monetary) consequences such that $x_2 > x_1 < x_0$. Note that both lotteries yield payoff c with probability $1-g$. This is the "common consequence". As $x_2 > \mu x_1 + (1-\mu)x_0$, it follows that $V_1 > V_2$, irrespective of the value of c. However, researchers have found that behavior is indeed systematically influenced by c, with a tendency to choose the first lottery when $c = x_2$ and the second when $c = x_0$. This kind of behavior was predicted by Allais and is known as the Allais paradox, cf. Allais (1953).

A second type of phenomenon, also discussed by Allais, is called the *common ratio effect*. To explain what this means, consider lotteries three and four, where lottery three has payoff $x_2$ with probability g and payoff $x_0$ with probability $1-g$. Lottery four has payoff $x_1$ with probability $\mu g$ and payoff $x_0$ with probability $1-\mu g$, where $x_2 > x_1 > x_0$. The corresponding expected utilities are:

$$V_3 = g x_2 + (1-g)x_0 = g(x_2 - x_0) + x_0,$$

and

$$V_4 = \mu g x_1 + (1-\mu g)x_0 = \mu g(x_1 - x_0) + x_0.$$

Evidently, $V_3 > V_4$, irrespective of the value of g. However, experimental evidence indicates that when $\mu$ is fixed, individuals reveal a tendency to switch towards lottery four as g decreases.

Let us now consider these phenomena in the present case with probabilistic choice, and under the Axioms 2, 5 and 6. Then, the choice probability of preferring lottery 1 over lottery 2 is given by:

$$K\big(h(V_1) - h(V_2)\big) = K\big(h(g x_2 + (1-g)c) - h(\mu g x_1 + (1-\mu) g x_0 + (1-g)c)\big).$$

From this expression, we realize that the choice probability will depend on the common consequence c, provided the mapping h is nonlinear. Although

$$K\big(h(g x_2 + (1-g)c) - h(\mu g x_1 + (1-\mu) g x_0 + (1-g)c)\big) > \frac{1}{2},$$

owing to the fact that $g x_2 > \mu g x_1 + (1-\mu) g x_0$, the fraction that prefers lottery two is less than $1/2$ but may be close to $1/2$. Similarly

$$K\big(h(V_3) - h(V_4)\big) = K\big(h(g(x_2 - x_0) + x_0) - h(\mu g(x_1 - x_0) + x_0)\big).$$

In this case, the choice probability will depend on g and $x_0$ even if h is linear. Also, in this case:

$$K\Big(h\big(g(x_2 - x_0) + x_0\big) - h\big(\mu g(x_1 - x_0) + x_0\big)\Big) > \frac{1}{2}.$$

Thus, we realize that with probabilistic models, such as the ones developed in this paper, the common consequence and common ratio effect may occur for less than 50 per cent of the population.

Only under Axioms 2, 5 and 7 does the common consequence effect vanish. The common ratio effect will only vanish when $h(x) = \beta \log x + \kappa$ and $x_0 = 0$.

# 8. An example

For the sake of illustrating the empirical relevance and usefulness of the theory developed above, we provide a discussion of the following example. The agent has the choice of working in either of two wage sectors or in a self-employment sector, denoted by alternatives one, two and three, respectively. In wage work sector j, the agent receives earnings $w_j$, $j = 1, 2$, with perfect certainty. In sector 3, earnings are uncertain. Hours of work in each sector are given. An example of a self-employment activity with fixed hours of work is running a café or a bar with fixed opening hours. We assume that the agent has been running the business—or similar businesses—for many periods and consequently is able to calculate the empirical distribution of returns to his or her business. For simplicity, we approximate this distribution with a discrete distribution. Let u(j,w) be the utility of working in sector j at wage income w. Let $g_3(w)$ be the lottery outcome probability that the agent receives wage w given that he or she chooses to work in the self-employment sector. The expected utilities of working in the wage sector reduce to $u(1, w_1)$ and $u(2, w_2)$, respectively, whereas the expected utility of working in sector 3 equals:

$$\sum_{w \in W} u(3, w) g_3(w).$$

Under the assumptions of Corollary 2, it follows that the probability of working in wage sector j equals:

$$\tilde{P}_B(j) = \frac{\exp\big(u(j, w_j)\big)}{\sum_{s=1}^{2} \exp\big(u(s, w_s)\big) + \exp\left(\sum_{w \in W} u(3, w) g_3(w)\right)}, \tag{8.1}$$

for $j = 1, 2$, where $B = \{1, 2, 3\}$. The probability of working in sector 3 equals:

$$\tilde{P}_B(3) = \frac{\exp\left(\sum_{w \in W} u(3, w) g_3(w)\right)}{\sum_{s=1}^{2} \exp\big(u(s, w_s)\big) + \exp\left(\sum_{w \in W} u(3, w) g_3(w)\right)}. \tag{8.2}$$

With convenient parametric specification of the utility function u(j,w), one can estimate the unknown parameters of the utility function by the method of maximum likelihood, provided data on agents' choices are available.

Alternatively, under the assumptions of Corollary 4, it follows that the probability of working in sector j becomes:

(8.3)
$$P_B(j)=\frac{\exp\left(h\left(u\left(j,w_j\right)\right)\right)}{\sum_{s=1}^{2}\exp\left(h\left(u\left(s,w_s\right)\right)\right)+\exp\left(h\left(\sum_{w\in W}u\left(3,w\right)g_3\left(w\right)\right)\right)},$$

for $j=1,2$, and a similar expression applies to the probability of working in sector 3, where h and u are given as in Corollary 4. Consequently, one can apply likelihood ratio test procedures, for example, to test the hypothesis that specification (i) of Corollary 4 is true, against the alternative specification (ii) in Corollary 4. Recall that the maximum likelihood estimation procedure goes as follows. Let $Y_j$ denote the number of agents that have chosen to work in sector j as observed in the data, and assume for simplicity that the choice probabilities above do not depend on observed individual characteristics. As is well known, the loglikelihood function can be expressed as:

$$\text{LogL}=\sum_j Y_j \log P_B(j),$$

from which the unknown parameters are obtained by maximization of logL. The general case with individual characteristics is completely analogous. We refer to Amemiya (1985), and Ben-Akiva and Lerman (1985) for details about inference methods for discrete choice models.

# 9. Conclusion

In this paper, we have developed a theory of probabilistic choice for risky choices based on different combinations of particular axioms. First, we have considered choice models with "minimal" structure on the choice probabilities. Second, we have generalized the expected utility theory to a probabilistic version. We have explored the relationship between sets of axioms and the structure of the corresponding choice probabilities. In particular, some sets of axioms imply a complete characterization of the functional form of the choice probabilities. The case in which the outcomes are money amounts is given particular attention, and it is demonstrated that particular invariance axioms that may apply in this setting yield an explicit characterization of the functional form of the model.

An interesting property of the models is that they rationalize the so-called *common consequence effect* and the *common ratio effect*.

As most of the axioms proposed are nonparametric, they can be utilized to carry out nonparametric tests of the respective structures of the choice probabilities. Finally, to illustrate the potential for applications we have discussed a concrete example.

# References

Amemiya, T. (1985): *Advanced econometrics*. Basil Blackwell, Ltd., Oxford, UK.

Anderson, S.P., A. de Palma and J.F. Thisse (1992): *Discrete choice theory of product differentiation*. MIT Press, London.

Allais, M. (1953): Le Comportement de l'Homme Rationel devant le Risque: Critique des Postulates et Axiomes de l'Ecole Americaine. *Econometrica*, **21**, 503–546.

Becker, G.M., M.H. De Groot and J. Marschak (1963): An Experimental Study of Some Stochastic Models for Wagers. *Behavioral Sciences*, **8**, 199–202.

Ben-Akiva, M. and S.R. Lerman (1985): *Discrete choice analysis*. MIT Press, Cambridge.

Billingsley, P. (1995): *Probability and measure*. Wiley, New York.

Carbone, E. (1997): Investigation of Stochastic Preference Theory using Experimental Data. *Economic Letters*, **57**, 305–311.

Dagsvik, J.K. (2005): Choice under Uncertainty and Bounded Rationality. Discussion Papers no. 409, Statistics Norway, Oslo.

Dagsvik, J.K. and S. Røine (2006): Testing Dimensional Invariance in a Stochastic Model of the Utility of Income. Discussion Papers, Statistics Norway, Oslo (forthcoming).

Debreu, G. (1958): Stochastic Choice and Cardinal Utility. *Econometrica*, **26**, 440–444.

Falmagne, J.C. (1985): *Elements of psychophysical theory*. Oxford University Press, London.

Fechner, G.T. (1860/1966): *Elements of psychophysics*. Vol. I. Holt, Rinehart and Winston, New York. (Translated from German. Originally published in 1860.)

Fishburn, P.C. (1976): Binary Choice Probabilities between Gambles: Interlocking Expected Utility Models. *Journal of Mathematical Psychology*, **14**, 99–122.

Fishburn, P.C. (1978): A Probabilistic Expected Utility Theory of Risky Binary Choices. *International Economic Review*, **19**, 633–646.

Fishburn, P.C. (1998): Stochastic Utility. In S. Barberà, P.J. Hammond and C. Seidl (eds.), *Handbook of utility theory*, Vol. 1; Principles. Kluwer Academic Publishers, London.

Harless, D.W. and C.F. Camerer (1994): The Predictive Utility of Generalized Expected Utility Theories. *Econometrica*, **62**, 1251–1290.

Hey, J.D. and C. Orme (1994): Investigating Generalizations of Expected Utility Theory Using Experimental Data. *Econometrica*, **62**, 1291–1326.

Hey, J.D. (1995): Experimental Investigations of Errors in Decision Making under Risk. *European Economic Review*, **39**, 633–640.

Iverson, G. and J.C. Falmagne (1985): Statistical Issues of Measurement. *Mathematical Social Sciences,* **10**, 131–153.

Karni, E. and D. Schmeidler (1991): Utility Theory with Uncertainty. In W. Hildenbrand and H. Sonnenschein (eds.), *Handbook of mathematical economics*, Vol. 4, Elsevier Science Publishers B.V., New York.

Loomes, G. and R. Sugden (1995): Incorporating a Stochastic Element into Decision Theories. *European Economic Review*, **39**, 641–648.

Loomes, G. and R. Sugden (1998): Testing Different Stochastic Specifications of Risky Choice. *Economica*, **65**, 581–598.

Luce, R.D. (1959): *Individual choice behavior: A theoretical analysis*. Wiley, New York.

Luce, R.D. and H. Raiffa (1957): *Games and decisions*. Wiley, New York.

Luce, R.D. and P. Suppes (1965): Preference, Utility and Subjective Probability. In R.D. Luce, R.R. Bush, and E. Galanter (eds.), *Handbook of mathematical psychology, III*, 249–410. Wiley, New York.

McFadden, D. (1973): Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (ed.), *Frontiers in econometrics*, Academic Press, New York.

McFadden, D. (1981): Econometric Models for Probabilistic Choice. In C. Manski and D. McFadden (eds.), *Structural analysis of discrete choice.* MIT Press, Cambridge, MA.

McFadden, D. (1984): Econometric Analysis of Qualitative Response Models. In Z. Griliches and M.D. Intriligator (eds.), *Handbook of econometrics*, Vol. II, North Holland, Amsterdam.

Suppes, P., R.D. Luce, D. Krantz and A. Tversky (1989): *Foundations of measurement.* Vol. II. Academic Press, London.

Starmer, C. (2000): Developments in Nonexpected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *Journal of Economic Literature*, **38**, 332–382.

Strauss, D. (1979): Some Results on Random Utility Models. *Journal of Mathematical Psychology*, **17**, 35–52.

Thurstone, L.L. (1927a): A Law of Comparative Judgment. *Psychological Review*, **34**, 272–286.

Thurstone, L.L. (1927b): Psychophysical Analysis. *American Journal of Psychology*, **38**, 368–389.

Tversky, A. (1969): Intransitivity of Preferences. *Psychological Review*, **76**, 31–48.

Tversky, A. (1972): Choice by Elimination. *Journal of Mathematical Psychology*, **9**, 341–367.

Yellott, J.I., Jr. (1977): The Relationship between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment and the Double Exponential Distribution. *Journal of Mathematical Psychology*, **15**, 109–144.

## Appendix

**Proof of Theorem 2**

Debreu (1958) has proved that Axiom 2 implies that there exists a cardinal representation $f(g), g \in S$, such that for $g_1, g_2, g_3, g_4 \in S$

$$P(g_1, g_2) \le P(g_3, g_4) \Leftrightarrow f(g_1) - f(g_2) \le f(g_3) - f(g_4), \tag{A.1}$$

where the inequality on one side is strict if and only if the inequality on the other side is strict. From (A.1), it follows that $g_1$, $g_2$, $g_3$ and $g_4$ satisfy $P(g_1, g_2) = P(g_3, g_4)$, if and only if $f(g_1) - f(g_2) = f(g_3) - f(g_4)$. However, this means that we can write

$$P(g_1, g_2) = K\{f(g_1) - f(g_2)\},$$

for some suitable function K. Evidently, K(x) is strictly increasing and takes values in [0,1]. Without loss of generality, it can be chosen to be a cumulative distribution function. The Balance condition implies that $K(x) + K(-x) = 1$, which means that K is symmetric. Recall that a cumulative distribution function is continuous to the right. As K is symmetric, it must also be continuous to the left. Hence, K is continuous.

Next, we shall prove the uniqueness of K. Suppose that $(f_0, K_0)$ and $(f_1, K_1)$ are two representations of the binary choice probabilities. Then

$$K_0(f_0(g_1) - f_0(g_2)) = K_1(f_1(g_1) - f_1(g_2)),$$

for any $g_1, g_2 \in S$. As $f_0$ and $f_1$ are unique up to a linear transformation, we can write

$$f_1(g) = a f_0(g) + b,$$

for $g \in S$, where a and b are constants and $a > 0$. This yields

$$K_0(f_0(g_1) - f_0(g_2)) = K_1(a(f_0(g_1) - f_0(g_2))),$$

which demonstrates that $K_0(x) = K_1(ax)$.

To prove that I is an interval, let $g_0 \in S$ be a fixed point of reference. Let $g_1, g_2 \in S$ be such that $f(g_2) \ge f(g_1)$, and let $x \in [f(g_1), f(g_2)]$ be arbitrary. Hence, $f(g_1) - f(g_0) \le x - f(g_0) \le f(g_2) - f(g_0)$, or equivalently

$$K^{-1}(P(g_1, g_0)) \le x - f(g_0) \le K^{-1}(P(g_2, g_0)),$$

which yields

(A.2) $$P(g_1, g_0) \le K\left(x - f(g_0)\right) \le P(g_2, g_0).$$

By Axiom 2 (ii), there exists a $g^* \in S$ such that $P\left(g_0, g^*\right) = K\left(f(g_0) - x\right)$. Thus, (A.2) implies that

$$K\left(f(g_0) - f\left(g^*\right)\right) = P\left(g_0, g^*\right) = K\left(x - f(g_0)\right),$$

so that $x = f\left(g^*\right)$. Therefore, $x \in I$. Hence, we have proved that I is an interval.

Q.E.D.

**Proof of Theorem 5**

When the choice probabilities given in Theorem 5 hold, then evidently Axioms 1, 5 and 6 hold. Consider next the "only if" part of the proof. Tversky (1972) and Suppes et al. (1989, p. 419) have proved that Axiom 1 is equivalent to the representation

$$P(g_1, g_2) = \tilde{F}\left(f(g_1), f(g_2)\right),$$

for $g_1, g_2 \in S$, where f is an ordinal scale functional defined on S and $\tilde{F}$ is a function that is strictly increasing in its first argument and strictly decreasing in the second. Hence

$$P(g_1, g_2) \ge \frac{1}{2} \Leftrightarrow f(g_1) \ge f(g_2),$$

and $\{f(g), g \in S\}$ therefore represents the binary relation $\succsim$ given in Definition 1. Accordingly, $\succsim$ is a preference relation so that by Theorem 4 and Axioms 5 and 6

$$f(g) = h\left(V(g)\right),$$

for some strictly increasing function h. Hence

$$P(g_1, g_2) = F\left(V(g_1), V(g_2)\right),$$

where

$$F(x, y) = \tilde{F}\left(h(x), h(y)\right).$$

Q.E.D.

**Proof of Theorem 6**

When the choice probabilities in Theorem 6 hold, then Axioms 2, 5 and 6 are satisfied. Consider the "only if" part. Debreu (1958) proved that Axiom 2 implies that there exists a mapping f from S to some interval such that for $g_1, g_2, g_3, g_4 \in S$

$$P(g_1,g_2) \geq P(g_3,g_4),$$

if and only if

$$f(g_1) - f(g_2) \geq f(g_3) - f(g_4).$$

Thus, with $g_3 = g_4$ we get

$$P(g_1,g_2) \geq 0.5 \Leftrightarrow f(g_1) \geq f(g_2),$$

and $\{f(g), g \in S\}$ therefore represents $\succsim$ on S. Consequently, $\succsim$ is a preference relation. Then, Theorem 4 and Axioms 5 and 6 imply that $f(g)$ must be a strictly increasing function h (say) of $V(g)$. That is

(A.3) $$f(g) = h(V(g)).$$

As Axiom 2 implies Theorem 2, we can combine (A.3) and (3.2), from which we get the desired result. Furthermore, by Theorem 2, V(·) is unique up to a linear transformation. As evidently f(·) must also be unique up to a linear transformation, we obtain the restrictions on $h(V(\cdot))$ stated in the theorem.

Q.E.D.

**Proof of Theorem 7**

It follows immediately that the "if" part of the theorem is true. Consider the "only if" part. From the theory of discrete choice (see, for example, McFadden, 1984), it follows that Axiom 3 holds if and only if for any $B \in \mathcal{B}$

$$P_B(g_s) = \frac{a(g_s)}{\sum_{g_r \in B} a(g_r)},$$

where $a(g_s), g_s \in S$, is a positive scalar that depends solely on $g_s$ and is unique apart from a multiplicative positive constant. Let $B = \{g_r, g_s\}$. Then

$$P(g_s, g_r) = \frac{a(g_s)}{a(g_s) + a(g_r)} = \frac{1}{1 + a(g_r)/a(g_s)}.$$

Thus

$$P(g_s, g_r) \geq 0.5 \iff a(g_s) \geq a(g_r),$$

and $\{a(g_s), g_s \in S\}$ therefore represents $\succsim$ on S. Consequently, $\succsim$ is a preference relation. Then, by Theorem 4, $a(g_s)$ must be a strictly increasing function of $V(g_s)$. Hence

$$\log a(g_s) = h(V(g_s)),$$

for some strictly increasing function h.

Q.E.D.

**Proof of Theorem 8**

Note first that when choice probabilities are given as in Theorem 8, it follows readily that Axioms 2, 5 and 7 are satisfied. Note next that when Axiom 7 holds, if

(A.4) $$P(g_1, g_2) = P(g_1^*, g_2^*),$$

then

(A.5) $$P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3) = P(\alpha g_1^* + (1-\alpha) g_3, \alpha g_2^* + (1-\alpha) g_3),$$

for $g_1, g_2, g_1^*, g_2^*, g_3 \in S$ and $\alpha \in [0,1]$.

To realize this, note that

$$P(g_1, g_2) = P(g_1^*, g_2^*)$$

is equivalent to

$$P(g_1, g_2) \geq P(g_1^*, g_2^*) \quad \text{and} \quad P(g_1, g_2) \leq P(g_1^*, g_2^*).$$

When applying Axiom 7 twice, with the inequality sign reversed the second time, we obtain (A.5).

Let $x_j = V(g_j)$, $j = 1,2,3$, where $V(\cdot)$ is given as in Theorem 6. Then, as Axiom 7 implies Axiom 6, it follows that Theorem 6 holds. Accordingly, (4.2) yields

(A.6) $$P(\alpha g_1 + (1-\alpha) g_3, \alpha g_2 + (1-\alpha) g_3) = \tilde{K}\left(\frac{\tilde{h}(\alpha x_1 + (1-\alpha) x_3)}{\tilde{h}(\alpha x_2 + (1-\alpha) x_3)}\right),$$

where $\tilde{K}$ and $\tilde{h}$ are defined by $\tilde{K}(x) = K(\log x)$ and $\log \tilde{h}(x) = h(x)$, where $\tilde{h} > 0$ is a strictly increasing function defined on R.

By (A.4), (A.5) and (A.6), we have that whenever $x_j^*$, given by $x_j^* = V(g_j^*)$, $g_j^* \in S,\ j = 1, 2,$ satisfies

$$\text{(A.7)} \qquad \tilde{K}\left(\frac{\tilde{h}(x_1)}{\tilde{h}(x_2)}\right) = \tilde{K}\left(\frac{\tilde{h}(x_1^*)}{\tilde{h}(x_2^*)}\right),$$

then it follows that

$$\text{(A.8)} \qquad \tilde{K}\left(\frac{\tilde{h}(\alpha x_1 + (1-\alpha)x_3)}{\tilde{h}(\alpha x_2 + (1-\alpha)x_3)}\right) = \tilde{K}\left(\frac{\tilde{h}(\alpha x_1^* + (1-\alpha)x_3)}{\tilde{h}(\alpha x_2^* + (1-\alpha)x_3)}\right),$$

for any $\alpha \in [0,1]$. Without loss of generality, we normalize V such that when $g_0 = (1, 0, 0, ...)$, $V(g_0) = 0$. In particular, when $g_3 = g_0$, then $x_3 = 0$, and it follows from (A.7) and (A.8) that whenever $x_1^*$ and $x_2^*$ are such that:

$$\text{(A.9)} \qquad \frac{\tilde{h}(x_1)}{\tilde{h}(x_2)} = \frac{\tilde{h}(x_1^*)}{\tilde{h}(x_2^*)},$$

then

$$\text{(A.10)} \qquad \frac{\tilde{h}(\alpha x_1)}{\tilde{h}(\alpha x_2)} = \frac{\tilde{h}(\alpha x_1^*)}{\tilde{h}(\alpha x_2^*)},$$

for all $\alpha \in [0,1]$. Next, note that (A.9) and (A.10) imply that we can write:

$$\text{(A.11)} \qquad \frac{\tilde{h}(\alpha x_1)}{\tilde{h}(\alpha x_2)} = f_\alpha\left(\frac{\tilde{h}(x_1)}{\tilde{h}(x_2)}\right),$$

for some strictly increasing continuous function $f_\alpha$ that depends on α. To realize this, observe that $\tilde{h}(\alpha x_1)/\tilde{h}(\alpha x_2)$ depends on $x_1, x_2$ solely through $\tilde{h}(x_1)/\tilde{h}(x_2)$ because (by (A.10)) the value of $\tilde{h}(\alpha x_1)/\tilde{h}(\alpha x_2)$ is unchanged when $(x_1, x_2)$ is replaced by $(x_1^*, x_2^*)$ when (A.9) is satisfied.

Let $u = \tilde{h}(x_1)$, $1/v = \tilde{h}(x_2)$. From (A.11) we then get

(A.12) $$\frac{\tilde{h}\left(\alpha\tilde{h}^{-1}(u)\right)}{\tilde{h}\left(\alpha\tilde{h}^{-1}\left(\frac{1}{v}\right)\right)}=f_\alpha\left(uv\right).$$

From (A.12), it follows that $f_\alpha(z)$ is strictly increasing in z.

Without loss of generality, assume now that $\tilde{h}$ is normalized such that for some $g\in S$, $\tilde{h}\left(V(g)\right)=1$. This implies that u and v can attain the value one. By letting u and v successively be equal to one, (A.12) implies that

(A.13) $$f_\alpha(u)=\frac{1}{f_\alpha\left(\frac{1}{u}\right)}.$$

Hence, by (A.12) and (A.13)

(A.14) $$f_\alpha\left(uv\right)=\frac{f_\alpha(u)}{f_\alpha\left(\frac{1}{v}\right)}=f_\alpha(u)f_\alpha(v).$$

(A.14) is a functional equation of the Cauchy type. As $f_\alpha(u)$ is strictly increasing, the only possible solution of (A.14) is given by

(A.15) $$f_\alpha(u)=u^{c(\alpha)},$$

where $c(\alpha)$ is a function of $\alpha$; see, for example, Falmagne (1985), Theorem 3.4.

Recall that $\tilde{h}(\cdot)$ is unique only up to a multiplicative constant. Therefore, $\tilde{h}(\cdot)$ can be normalized such that $\tilde{h}(1)=1$. From (A.11) and (A.15), with $x_1=x$ and $x_2=1$, we obtain that:

(A.16) $$h\left(\alpha x\right)=c(\alpha)h(x)+h(\alpha),$$

where h is defined on [0,1]. In the following, it will be convenient to organize the rest of the proof into two cases depending on whether or not $c(\alpha)$ is a constant.

*Case (i). $c(\alpha)$ is a constant.*

In this case (A.16) yields, by symmetry

$$h\left(\alpha x\right)=ch(x)+h(\alpha)=h\left(x\alpha\right)=ch(\alpha)+h(x),$$

and hence

$$\left(c-1\right)h(x)=\left(c-1\right)h(\alpha),$$

which must hold for all $x, \alpha \in [0,1]$. This implies that $c = 1$. Thus, (A.16) reduces to a well-known Cauchy type functional equation. Then, necessarily

$$h(x) = \beta \log x + \gamma, \tag{A.17}$$

where $\beta$ and $\gamma$ are constants; see, for example, Falmagne (1985), Theorem 3.4.

*Case (ii). c(α) is not a constant.*

In this case, there is at least one $\alpha$, say $\alpha_0$, such that $c(\alpha_0) \neq 1$. Hence, (A.16) leads to:

$$h(\alpha_0 x) = c(\alpha_0) h(x) + h(\alpha_0) = h(x\alpha_0) = c(x) h(\alpha_0) + h(x). \tag{A.18}$$

The last equation yields

$$h(x) = (c(x) - 1) b_0, \tag{A.19}$$

where

$$b_0 = \frac{h(\alpha_0)}{c(\alpha_0) - 1}.$$

When (A.19) is inserted into (A.16) and the terms are rearranged, we obtain

$$c(\alpha x) = c(\alpha) c(x), \tag{A.20}$$

for $\alpha, x \in [0,1]$. The only strictly increasing solution of (A.20) is given by

$$c(\alpha) = \alpha^{\kappa}, \tag{A.21}$$

for some constant $\kappa$ (see Falmagne, 1985, Theorem 3.4). When (A.19) and (A.21) are combined we get

$$h(x) = b_0 \left( x^{\kappa} - 1 \right), \tag{A.22}$$

for $x \in [0,1]$. Note next that (A.7) and (A.8) imply that

$$\left(\alpha x_1 + (1-\alpha) x_3\right)^{\kappa} - \left(\alpha x_2 + (1-\alpha) x_3\right)^{\kappa} = \left(\alpha x_1^* + (1-\alpha) x_3\right)^{\kappa} - \left(\alpha x_2^* + (1-\alpha) x_3\right)^{\kappa}, \tag{A.23}$$

whenever

$$x_1^{\kappa} - x_2^{\kappa} = \left(x_1^*\right)^{\kappa} - \left(x_2^*\right)^{\kappa}. \tag{A.24}$$

Now, keep $x_1^*, x_2^*$ and $x_3$ fixed and differentiate (A.23) with respect to $x_1$ subject to (A.24). This gives

$$\text{(A.25)} \qquad \left(\alpha x_1 + (1-\alpha)x_3\right)^{\kappa-1} = \left(\alpha x_2 + (1-\alpha)x_3\right)^{\kappa-1}\frac{dx_2}{dx_1} = \left(\alpha x_2 + (1-\alpha)x_3\right)^{\kappa-1}\left(\frac{x_1}{x_2}\right)^{\kappa-1}.$$

Suppose that $\kappa \neq 1$. Then, (A.25) implies that $x_1 = x_2$, which is a contradiction. We therefore conclude that $\kappa = 1$, i.e.,

$$\text{(A.26)} \qquad h(x) = b_0\left(x-1\right).$$

Recall that the normalization $h(1) = 0$ we adopted above was made purely for notational convenience so that the general form of h is $h(x) = b_0 x + \gamma$, where $\gamma$ is an arbitrary constant.

This completes the proof.

Q.E.D.

**Proof of Corollary 2**

As Axiom 7 implies Axiom 6, it follows from Theorem 7 that (4.4) must hold. Consider the special case with $B = \{g_1, g_2\}$. In this case, (4.4) reduces to a special case of (4.2) with

$$K(x) = \frac{1}{1+\exp(-x)}.$$

Hence, Theorem 8 applies and implies (4.6). Without loss of generality, we can set $\kappa = 0$ and $\beta = 1$ because $\kappa$ cancels and $\beta$ is absorbed in the utilities $\{u(k)\}$ in the expression for the choice probability.

Corollary 2 represents the most satisfactory model so far, in the sense that the choice probabilities are characterized completely in terms of a linear preference functional (4.1) of the respective lottery outcome probabilities. This is a rather strong result, and it is achieved at the cost of strong assumptions such as Axioms 3 and 7. In the special case with binary comparisons, i.e., $B = \{g_1, g_2\}$, the Luce model is not particularly restrictive. Thus, in this case, Axiom 7 is the most objectionable assertion because it implies the Independence Axiom (Axiom 6).

Q.E.D.

**Proof of Theorem 9**

Note first that it follows immediately that when (i) or (ii) in Theorem 9 hold, then Axioms 2, 5, 6 and 8 are true. We shall next prove that (i) or (ii) is also necessary. Without loss of generality, we consider lotteries with only two outcomes, that is, lottery j has outcome $\left(1, w_j\right)$ or (2,1) with probabilities $g_j(1,w_j)$ and $g_j(2,1)=1-g_j(1,w_j)$ for $j=1,2,3,4$, with $g_2=g_4$, $w_1=w$ and $w_3=a$, where a is a fixed positive number.

Let

(A.27) $$V(g_j^\lambda)=u(1,w_j\lambda)g_j(1,w_j)+u(2,1)g_j(2,1)=g_j(1,w_j)\big(u(1,w_j\lambda)-u(2,1)\big)+u(2,1),$$

for $j=1,2,3,4$, and $\lambda>0$. Clearly, $V(g_j^\lambda)$ is the expected utility of lottery j when $\{g_j^\lambda\}$ represents the outcome probabilities. From Axioms 2, 5 and 6, Theorem 6 follows, which yields

(A.28) $$P(g_1^\lambda,g_2^\lambda)=K\Big(h\big(V(g_1^\lambda)\big)-h\big(V(g_2^\lambda)\big)\Big),$$

where K is a c.d.f. that is continuous and strictly increasing, and h is strictly increasing. Similarly to the proof of Theorem 8, it follows that Axiom 8 implies that if

(A.29) $$P(g_1^1,g_2^1)=P(g_3^1,g_2^1),$$

then

(A.30) $$P(g_1^\lambda,g_2^\lambda)=P(g_3^\lambda,g_2^\lambda),$$

for $\lambda>0$. By (A.28), and because $g_2=g_4$, this is equivalent to the statement that if

(A.31) $$V(g_1^1)=V(g_3^1),$$

then

(A.32) $$V(g_1^\lambda)=V(g_3^\lambda),$$

for $\lambda>0$. If (A.31) holds, then by (A.27)

(A.33) $$\frac{g_3(1,a)}{g_1(1,w)}=\frac{u(1,w)-u(2,1)}{u(1,a)-u(2,1)}.$$

Let

$$\psi(x)=\frac{u(1,x)-u(2,1)}{u(1,a)-u(2,1)}$$

and

$$k(x)=\frac{u(2,x)-u(2,1)}{u(1,a)-u(2,1)}.$$

When (A.33) is inserted into (A.32), we obtain

(A.34) $$\psi(\lambda w)=k(\lambda)+\psi(w)\big(\psi(\lambda a)-k(\lambda)\big).$$

(A.34) is a functional equation, the solution to which can be found in Falmagne (1985, p. 89) case (iv)). The solution is given by

(A.35) $$\psi(w)=c\left(\frac{w^{\rho}-1}{\rho}\right)+1,$$

and

(A.36) $$k(w)=(a-1)\frac{\left(w^{\rho}-1\right)}{\rho},$$

where c and $\rho$ are constants. Hence, it follows that

(A.37) $$u(k,w)=b_k\frac{w^{\rho}-1}{\rho}+c_k,$$

for suitable constants, $b_k$ and $c_k$.

Next, consider the functional form of h. Let $g_j$, $j=1,2,3,4$, represent four lotteries with binary outcomes $(1,w_j)$ and (1,1) with probabilities $\alpha$ and $(1-\alpha)$, respectively, for $j=1,2,3,4$.

*Case (i):* $\rho\neq 0$.

Write

(A.38) $$V\left(g_j^{\lambda}\right)=\alpha\lambda^{\rho}w_j^{\rho}d_1+\alpha c_1+(1-\alpha)\lambda^{\rho}d_2+(1-\alpha)c_2=\mu z_j+\gamma,$$

where $\mu=\lambda^{\rho}$, $z_j=\alpha d_1 w_j^{\rho}+(1-\alpha)d_2$ and $\gamma=\alpha c_1+(1-\alpha)c_2$, where $d_1$, $d_2$, $c_1$ and $c_2$ are constants. Let $w_1$, $w_2$, $w_3$ and $w_4$ be such that

$$P\left(g_1^1,g_2^1\right)\leq P\left(g_3^1,g_4^1\right).$$

Then, by Axiom 8

$$P\left(g_1^{\lambda},g_2^{\lambda}\right)\leq P\left(g_3^{\lambda},g_4^{\lambda}\right).$$

An equivalent statement of Axiom 8 is that whenever

$$\tilde{K}\left(\tilde{h}(z_1)/\tilde{h}(z_2)\right) \le \tilde{K}\left(\tilde{h}(z_3)/\tilde{h}(z_4)\right)$$

then

$$\tilde{K}\left(\tilde{h}(\mu z_1)/\tilde{h}(\mu z_2)\right) \le \tilde{K}\left(\tilde{h}(\mu z_3)/\tilde{h}(\mu z_4)\right),$$

for $\mu > 0$, where $\tilde{K}(x) = K(\log x)$ and $\log \tilde{h}(x) = h(x+\gamma)$. We can now apply Theorem 14.19, in Falmagne (1985, p. 338), which yields

(A.39) $$K\left(h(z_1+\gamma) - h(z_2+\gamma)\right) = F\left(\frac{a_1\left(z_1^\theta - 1\right) - a_2\left(z_2^\theta - 1\right)}{\theta}\right),$$

for some strictly increasing continuous function, F, where θ, $a_1$ and $a_2$ are independent of $z_1$ and $z_2$. By symmetry, one must have that $a_1 = a_2$. Let $M(x) = K^{-1}F(ax)$. Hence, (A.39) yields

(A.40) $$h(z_1+\gamma) - h(z_2+\gamma) = M\left(\frac{a\left(z_1^\theta - 1\right) - a\left(z_2^\theta - 1\right)}{\theta}\right).$$

With $z_2 = 1$ we get

(A.41) $$h(z_1+\gamma) = h(1+\gamma) + M\left(\frac{\left(z_1^\theta - 1\right)}{\theta}\right),$$

and with $z_1 = 1$, we get

(A.42) $$h(z_2+\gamma) = h(1+\gamma) + M\left(\frac{-\left(z_2^\theta - 1\right)}{\theta}\right).$$

Let $x_j = \left(z_j^\theta - 1\right)/\theta$. By subtracting (A.42) from (A.41), (A.40) follows, and we can express the resulting equation as

$$M(x_1) - M(x_2) = M(x_1 - x_2),$$

or, setting $x = x_2$ and $y = x_1 - x_2$, we obtain the equivalent equation

(A.43) $$M(x+y) = M(x) + M(y),$$

for x and y belonging to a suitable interval. As M is continuous, we must have that $M(x) = bx$ where b is a constant (see, for example, Falmagne, 1985, Theorem 2). Consequently, we obtain that

(A.44)
$$h(z) = \beta\frac{(z+\gamma)^{\theta} - 1}{\theta} + \kappa,$$

where θ, γ, κ and $\beta > 0$ are constants. Furthermore, we realize that both $c_k = \gamma = 0$, because otherwise Axiom 8 will not hold.

It remains to prove that Axiom 8 implies that $\gamma - 1/\rho = c_k = 0$ for all k. To this end, consider four lotteries with outcomes $\{k, w_k\}$ and lottery outcome probabilities $g_j$, $j = 1,2,3,4$, such that

(A.45)
$$\sum_k c_k g_j(k) = \sum_k c_k g_1(k),$$

for $k = 2,3,4$. Let

$$B_j = \sum_k g_j(k) b_k\, w_k^{\rho}/\rho,$$

and

(A.46)
$$d = \gamma - \frac{1}{\rho} + \sum_k c_k g_1(k).$$

Similarly to the proof of Theorem 8, it follows that Axiom 8 implies that if

(A.47)
$$P\left(g_1^1, g_2^1\right) = P\left(g_3^1, g_4^1\right),$$

then

(A.48)
$$P\left(g_1^{\lambda}, g_2^{\lambda}\right) = P\left(g_3^{\lambda}, g_4^{\lambda}\right),$$

for $\lambda > 0$. If (A.44) holds with $\theta \neq 0$ and $\theta \neq 1$, then (A.47) and (A.48) yield

(A.49)
$$\left(B_1 + d/\lambda\right)^{\theta} - \left(B_2 + d/\lambda\right)^{\theta} = \left(B_3 + d/\lambda\right)^{\theta} - \left(B_4 + d/\lambda\right)^{\theta}.$$

By dividing (A.48) by $\left(B_2 + d/\lambda\right)^{\theta}$, we obtain

(A.50)
$$\left(C_1 + x\, r_1\, d\right)^{\theta} - 1 = \left(C_2 + x\, r_2\, d\right)^{\theta} - \left(C_3 + x\, r_3\, d\right)^{\theta},$$

where $C_1 = B_1/B_2$, $C_2 = B_3/B_2$, $C_3 = B_4/B_2$, $r_1 = 1 - B_1/B_2$, $r_2 = 1 - B_3/B_2$, $r_3 = 1 - B_4/B_2$ and $x = 1/\left(d + \lambda B_2\right)$, where $0 < x < 1/d$. When differentiating (A.50) with respect to x, we obtain

(A.51) $$\left(C_1 + x r_1 d\right)^{\theta-1} r_1 d = \left(C_2 + x r_2 d\right)^{\theta-1} r_2 d - \left(C_3 + x r_3 d\right)^{\theta-1} r_3 d .$$

Now, divide (A.51) by $\left(C_2 + x r_2 d\right)^{\theta-1}$, which gives

(A.52) $$\left(D_1 + y s_1 d\right)^{\theta-1} r_1 d = r_2 d - \left(D_2 + y s_2 d\right)^{\theta-1} r_3 d ,$$

where $D_1 = C_1/C_2$, $D_2 = C_3/C_2$, $s_1 = r_1 - r_2 C_1/C_2$, $s_2 = r_3 - r_2 C_3/C_2$ and $y = x/\left(C_2 + x r_2 d\right)$, with $0 < y < 1/\left(C_2 + r_2\right) d$. If we differentiate (A.52) with respect to y, we get

(A.53) $$\left(D_1 + y s_1 d\right)^{\theta-2} r_1 s_1 d^2 = \left(D_2 + y s_2 d\right)^{\theta-2} r_3 s_2 d^2 .$$

Note that

$$r_1 s_1 = \left(1 - \frac{B_1}{B_3}\right)\left(1 - \frac{B_1}{B_2}\right),\ r_3 s_2 = \left(1 - \frac{B_4}{B_3}\right)\left(1 - \frac{B_4}{B_2}\right).$$

Evidently, one can select lotteries such that $B_1$ is different from $B_2$ and $B_3$, or $B_4$ is different from $B_2$ and $B_3$. This implies that either $r_1 s_1 \neq 0$ or $r_3 s_2 \neq 0$. Consequently, (A.53) implies that $d = 0$. As d given in (A.46) must be zero for all choices of lottery outcome probabilities, $\{g_1(k)\}$, and $\{c_k\}$ are independent of $\{g_1(k)\}$, we conclude that $\gamma - 1/\rho = 0$ and $c_k = 0$ for all k.

Next, consider the case with $\theta = 1$ without the restriction in (A.45). Similarly to (A.49), we obtain that

(A.54) $$B_1 - B_2 + \left(d_1 - d_2\right)/\lambda = B_3 - B_4 + \left(d_3 - d_4\right)/\lambda ,$$

for $\lambda > 0$, where

$$d_j = \sum_k c_k\, g_j(k) ,$$

and given that (A.54) holds for $\lambda = 1$. Evidently, this implies that $d_1 = d_2$ and $d_3 = d_4$, which in general can happen only if $c_k = c$.

*Case (ii):* $\rho = 0$.

In this case (A.35) yields

(A.55) $$\begin{aligned} V\left(g_j^{\lambda}\right) &= \alpha b_1 \log\left(\lambda w_j\right) + \left(1-\alpha\right) b_2 \log\lambda + \alpha c_1 + \left(1-\alpha\right) c_2 \\ &= \left(\alpha b_1 + \left(1-\alpha\right) b_2\right) \log\lambda + \alpha b_1 \log w_j + \alpha c_1 + \left(1-\alpha\right) c_2 , \\ &= \log\mu + \log z_j \end{aligned}$$

where $\log\mu = (\alpha b_1 + (1-\alpha) b_2)\log\lambda$ and $\log z_j = \alpha b_1 \log w_j + \alpha c_1 + (1-\alpha)c_2$. By Axiom 8, it follows that whenever

$$\tilde{K}\left(h^*(z_1)/h^*(z_2)\right) \le \tilde{K}\left(h^*(z_3)/h^*(z_4)\right),$$

then

$$\tilde{K}\left(h^*(\mu z_1)/h^*(\mu z_2)\right) \le \tilde{K}\left(h^*(\mu z_3)/h^*(\mu z_4)\right)$$

for $\mu > 0$, where $h^*(x) = \exp(h(\log x))$. By Theorem 14.19 in Falmagne (1985, p. 338), we obtain that

$$K\left(h(\log z_1) - h(\log z_2)\right) = F\left(\frac{a_1\left(z_1^\theta - 1\right) - a_2\left(z_2^\theta - 1\right)}{\theta}\right), \tag{A.56}$$

for some strictly increasing and continuous function F where $a_1$ and $a_2$ are positive constants. Eq. (A.56) is completely analogous to (A.39), and it therefore follows in the same way as the analysis under Case (i) that

$$h(\log z) = \beta\left(\frac{z^\theta - 1}{\theta}\right) + \kappa,$$

implying that

$$h(x) = \beta\frac{\left(e^{\theta x} - 1\right)}{\theta} + \kappa,$$

where $\theta$, $\kappa$ and $\beta > 0$ are constants. It follows readily that also in this case, $c_k = c$, for all k. This completes the proof.

Q.E.D.

**Proof of Corollary 3**

Clearly, Axiom 9 implies Axiom 8. Therefore, by Theorem 8, it follows that Theorem 9 must hold. Let

$$x_j = \sum_k b g_j(k) \log w_k + c,$$

for j = 1, 2, and consider the functional forms in (i) of Theorem 9. In this case, Axiom 9 implies that:

$$h\left(V\left(g_1^{\lambda}\right)\right)-h\left(V\left(g_2^{\lambda}\right)\right)=\frac{\beta\exp\left(\theta\left(b\log\lambda+x_1\right)\right)-\beta\exp\left(\theta\left(b\log\lambda+x_2\right)\right)}{\theta}$$

$$=h\left(V\left(g_1^{1}\right)\right)-h\left(V\left(g_2^{1}\right)\right)=\frac{\beta\exp\left(\theta x_1\right)-\beta\exp\left(\theta x_2\right)}{\theta},$$

for $\lambda > 0$. The equation above implies that $\exp\left(\theta b\log\lambda\right)=\lambda^{\theta b}=1$, which can only be true if $\theta = 0$. The proof of (ii) is completely analogous.

Q.E.D.

**Proof of Theorem 10**

By applying a special case of Kolmogorov's Theorem on the construction of random variables, the existence of the probability space on which the random field $\{\varepsilon(g), g\in S\}$ is defined follows. See, for example, Billingsley (1995). This corollary establishes the desired results for the case that is relevant in our context, namely when $\varepsilon\left(g_s\right), s=1,2,...,$ are independent and identically distributed (i.i.d). The choice probability in (4.3) follows from a well-known result in discrete choice theory (see, for example, McFadden, 1984). The result now follows from Theorem 7.

Q.E.D.

# Recent publications in the series Discussion Papers

373 G. Liu (2004) Estimating Energy Demand Elasticities for OECD Countries. A Dynamic Panel Data Approach

374 K. Telle and J. Larsson (2004): Do environmental regulations hamper productivity growth? How accounting for improvements of firms' environmental performance can change the conclusion

375 K.R. Wangen (2004): Some Fundamental Problems in Becker, Grossman and Murphy's Implementation of Rational Addiction Theory

376 B.J. Holtsmark and K.H. Alfsen (2004): Implementation of the Kyoto Protocol without Russian participation

377 E. Røed Larsen (2004): Escaping the Resource Curse and the Dutch Disease? When and Why Norway Caught up with and Forged ahead of Its Neughbors

378 L. Andreassen (2004): Mortality, fertility and old age care in a two-sex growth model

379 E. Lund Sagen and F. R. Aune (2004): The Future European Natural Gas Market - are lower gas prices attainable?

380 A. Langørgen and D. Rønningen (2004): Local government preferences, individual needs, and the allocation of social assistance

381 K. Telle (2004): Effects of inspections on plants' regulatory and environmental performance - evidence from Norwegian manufacturing industries

382 T. A. Galloway (2004): To What Extent Is a Transition into Employment Associated with an Exit from Poverty

383 J. F. Bjørnstad and E.Ytterstad (2004): Two-Stage Sampling from a Prediction Point of View

384 A. Bruvoll and T. Fæhn (2004): Transboundary environmental policy effects: Markets and emission leakages

385 P.V. Hansen and L. Lindholt (2004): The market power of OPEC 1973-2001

386 N. Keilman and D. Q. Pham (2004): Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area

387 G. H. Bjertnæs and T. Fæhn (2004): Energy Taxation in a Small, Open Economy: Efficiency Gains under Political Restraints

388 J.K. Dagsvik and S. Strøm (2004): Sectoral Labor Supply, Choice Restrictions and Functional Form

389 B. Halvorsen (2004): Effects of norms, warm-glow and time use on household recycling

390 I. Aslaksen and T. Synnestvedt (2004): Are the Dixit-Pindyck and the Arrow-Fisher-Henry-Hanemann Option Values Equivalent?

391 G. H. Bjønnes, D. Rime and H. O.Aa. Solheim (2004): Liquidity provision in the overnight foreign exchange market

392 T. Åvitsland and J. Aasness (2004): Combining CGE and microsimulation models: Effects on equality of VAT reforms

393 M. Greaker and Eirik. Sagen (2004): Explaining experience curves for LNG liquefaction costs: Competition matter more than learning

394 K. Telle, I. Aslaksen and T. Synnestvedt (2004): "It pays to be green" - a premature conclusion?

395 T. Harding, H. O. Aa. Solheim and A. Benedictow (2004). House ownership and taxes

396 E. Holmøy and B. Strøm (2004): The Social Cost of Government Spending in an Economy with Large Tax Distortions: A CGE Decomposition for Norway

397 T. Hægeland, O. Raaum and K.G. Salvanes (2004): Pupil achievement, school resources and family background

398 I. Aslaksen, B. Natvig and I. Nordal (2004): Environmental risk and the precautionary principle: "Late lessons from early warnings" applied to genetically modified plants

399 J. Møen (2004): When subsidized R&D-firms fail, do they still stimulate growth? Tracing knowledge by following employees across firms

400 B. Halvorsen and Runa Nesbakken (2004): Accounting for differences in choice opportunities in analyses of energy expenditure data

401 T.J. Klette and A. Raknerud (2004): Heterogeneity, productivity and selection: An empirical study of Norwegian manufacturing firms

402 R. Aaberge (2005): Asymptotic Distribution Theory of Empirical Rank-dependent Measures of Inequality

403 F.R. Aune, S. Kverndokk, L. Lindholt and K.E. Rosendahl (2005): Profitability of different instruments in international climate policies

404 Z. Jia (2005): Labor Supply of Retiring Couples and Heterogeneity in Household Decision-Making Structure

405 Z. Jia (2005): Retirement Behavior of Working Couples in Norway. A Dynamic Programming Approch

406 Z. Jia (2005): Spousal Influence on Early Retirement Behavior

407 P. Frenger (2005): The elasticity of substitution of superlative price indices

408 M. Mogstad, A. Langørgen and R. Aaberge (2005): Region-specific versus Country-specific Poverty Lines in Analysis of Poverty

409 J.K. Dagsvik (2005) Choice under Uncertainty and Bounded Rationality

410 T. Fæhn, A.G. Gómez-Plana and S. Kverndokk (2005): Can a carbon permit system reduce Spanish unemployment?

411 J. Larsson and K. Telle (2005): Consequences of the IPPC-directive's BAT requirements for abatement costs and emissions

412 R. Aaberge, S. Bjerve and K. Doksum (2005): Modeling Concentration and Dispersion in Multiple Regression

413 E. Holmøy and K.M. Heide (2005): Is Norway immune to Dutch Disease? CGE Estimates of Sustainable Wage Growth and De-industrialisation

414 K.R. Wangen (2005): An Expenditure Based Estimate of Britain's Black Economy Revisited

415 A. Mathiassen (2005): A Statistical Model for Simple, Fast and Reliable Measurement of Poverty

416 F.R. Aune, S. Glomsrød, L. Lindholt and K.E. Rosendahl: Are high oil prices profitable for OPEC in the long run?

417 D. Fredriksen, K.M. Heide, E. Holmøy and I.F. Solli (2005): Macroeconomic effects of proposed pension reforms in Norway

418 D. Fredriksen and N.M. Stølen (2005): Effects of demographic development, labour supply and pension reforms on the future pension burden

419 A. Alstadsæter, A-S. Kolm and B. Larsen (2005): Tax Effects on Unemployment and the Choice of Educational Type

420 E. Biørn (2005): Constructing Panel Data Estimators by Aggregation: A General Moment Estimator and a Suggested Synthesis

421 J. Bjørnstad (2005): Non-Bayesian Multiple Imputation

422 H. Hungnes (2005): Identifying Structural Breaks in Cointegrated VAR Models

423 H. C. Bjørnland and H. Hungnes (2005): The commodity currency puzzle

424 F. Carlsen, B. Langset and J. Rattsø (2005): The relationship between firm mobility and tax level: Empirical evidence of fiscal competition between local governments

425 T. Harding and J. Rattsø (2005): The barrier model of productivity growth: South Africa

426 E. Holmøy (2005): The Anatomy of Electricity Demand: A CGE Decomposition for Norway

427 T.K.M. Beatty, E. Røed Larsen and D.E. Sommervoll (2005): Measuring the Price of Housing Consumption for Owners in the CPI

428 E. Røed Larsen (2005): Distributional Effects of Environmental Taxes on Transportation: Evidence from Engel Curves in the United States

429 P. Boug, Å. Cappelen and T. Eika (2005): Exchange Rate Rass-through in a Small Open Economy: The Importance of the Distribution Sector

430 K. Gabrielsen, T. Bye and F.R. Aune (2005): Climate change- lower electricity prices and increasing demand. An application to the Nordic Countries

431 J.K. Dagsvik, S. Strøm and Z. Jia: Utility of Income as a Random Function: Behavioral Characterization and Empirical Evidence

432 G.H. Bjertnæs (2005): Avioding Adverse Employment Effects from Energy Taxation: What does it cost?

433. T. Bye and E. Hope (2005): Deregulation of electricity markets—The Norwegian experience

434 P.J. Lambert and T.O. Thoresen (2005): Base independence in the analysis of tax policy effects: with an application to Norway 1992-2004

435 M. Rege, K. Telle and M. Votruba (2005): The Effect of Plant Downsizing on Disability Pension Utilization

436 J. Hovi and B. Holtsmark (2005): Cap-and-Trade or Carbon Taxes? The Effects of Non-Compliance and the Feasibility of Enforcement

437 R. Aaberge, S. Bjerve and K. Doksum (2005): Decomposition of Rank-Dependent Measures of Inequality by Subgroups

438 B. Holtsmark (2005): Global per capita $CO_2$ emissions - stable in the long run?

439 E. Halvorsen and T.O. Thoresen (2005): The relationship between altruism and equal sharing. Evidence from inter vivos transfer behavior

440 L-C. Zhang and I. Thomsen (2005): A prediction approach to sampling design

441 Ø.A. Nilsen, A. Raknerud, M. Rybalka and T. Skjerpen (2005): Lumpy Investments, Factor Adjustments and Productivity

442 R. Golombek and A. Raknerud (2005): Exit Dynamics with Adjustment Costs

443 G. Liu, T. Skjerpen, A. Rygh Swensen and K. Telle (2006): Unit Roots, Polynomial Transformations and the Environmental Kuznets Curve

444 G. Liu (2006): A Behavioral Model of Work-trip Mode Choice in Shanghai

445 E. Lund Sagen and M. Tsygankova (2006): Russian Natural Gas Exports to Europe. Effects of Russian gas market reforms and the rising market power of Gazprom

446 T. Ericson (2006): Households' self-selection of a dynamic electricity tariff

447 G. Liu (2006): A causality analysis on GDP and air emissions in Norway

448 M. Greaker and K.E. Rosendahl (2006): Strategic Climate Policy in Small, Open Economies

449 R. Aaberge, U. Colombino and T. Wennemo (2006): Evaluating Alternative Representation of the Choice Sets in Models of Labour Supply

450 T. Kornstad and T.O. Thoresen (2006): Effects of Family Policy Reforms in Norway. Results from a Joint Labor Supply and Child Care Choice Microsimulation Analysis

451 P. Frenger (2006): The substitution bias of the consumer price index

452 B. Halvorsen (2006): When can micro properties be used to predict aggregate demand?

453 J.K. Dagsvik, T. Korntad and T. Skjerpen (2006): Analysis of the disgouraged worker phenomenon. Evidence from micro data

454 G. Liu (2006): On Nash equilibrium in prices in an oligopolistic market with demand characterized by a nested multinomial logit model and multiproduct firm as nest

455 F. Schroyen and J. Aasness (2006): Marginal indirect tax reform analysis with merit good arguments and environmental concerns: Norway, 1999

456 L-C Zhang (2006): On some common practices of systematic sampling

457 Å. Cappelen (2006): Differences in Learning and Inequality

458 T. Borgersen, D.E. Sommervoll and T. Wennemo (2006): Endogenous Housing Market Cycles

459 G.H. Bjertnæs (2006): Income Taxation, Tuition Subsidies, and Choice of Occupation

460 P. Boug, Å. Cappelen and A.R. Swensen (2006): The New Keynesian Phillips Curve for a Small Open Economy

461 T. Ericson (2006): Time-differentiated pricing and direct load control of residential electricity consumption

462 T. Bye, E. Holmøy and K. M. Heide (2006): Removing policy based comparative advantage for energy intensive production. Necessary adjustments of the real exchange rate and industry structure

463 R. Bjørnstad and R. Nymoen (2006): Will it float? The New Keynesian Phillips curve tested on OECD panel data

464 K.M.Heide, E. Holmøy, I. F. Solli and B. Strøm (2006): A welfare state funded by nature and OPEC. A guided tour on Norway's path from an exceptionally impressive to an exceptionally strained fiscal position

465 J.K. Dagsvik: Axiomatization of Stochastic Models for Choice under Uncertainty