Working Paper No. 311

Stochastic Utility Theorem

Pavlo R. Blavatskyy

Januar 2007

Stochastic Utility Theorem

Pavlo R. Blavatskyy[†]

Institute for Empirical Research in Economics
University of Zurich
Winterthurerstrasse 30
CH-8006 Zurich
Switzerland
Phone: +41(1)6343586
Fax: +41(1)6344978
e-mail: pavlo.blavatskyy@iew.unizh.ch

January 2007

Abstract

This paper analyzes individual decision making under risk. It is assumed that an individual does not have a preference relation on the set of risky lotteries. Instead, an individual possesses a probability measure that captures the likelihood of one lottery being chosen over the other. Choice probabilities have a stochastic utility representation if they can be written as a non-decreasing function of the difference in expected utilities of the lotteries. Choice probabilities admit a stochastic utility representation if and only if they are complete, strongly transitive, continuous, independent of common consequences and interchangeable. Axioms of stochastic utility are consistent with systematic violations of betweenness and a common ratio effect but not with a common consequence effect. Special cases of stochastic utility include the Fechner model of random errors, Luce choice model and a tremble model of Harless and Camerer (1994).

Keywords: expected utility theory, stochastic utility, Fechner model, Luce choice model, tremble

JEL Classification codes: C91, D81

Stochastic Utility Theorem

I. Introduction

Experimental studies of repeated decision making under risk demonstrate that individual choices are often contradictory. For example, Camerer (1989) reports that 31.6% of subjects reversed their choices when presented with the same binary choice problem for the second time. Starmer and Sugden (1989) find that 26.5% of all choices are reversed on the second repetition of a decision problem. Hey and Orme (1994) report an inconsistency rate of 25% even when individuals are allowed to declare indifference. Wu (1994) finds that 5% to 45% of choice decisions are reversed (depending on a lottery pair) when decision problem is repeated. Ballinger and Wilcox (1997) report a median switching rate of 20.8%

Although experimental data convincingly show that choice under risk is generally stochastic, this finding remains largely ignored in the theoretical work (e.g. Loomes and Sugden, 1998). The majority of decision theories assume that individuals are born with a unique preference relation on the set of risky lotteries, which typically does not allow for a possibility of stochastic choice (unless individuals happen to be exactly indifferent). As a notable exception, Machina (1985) and Chew et al. (1991) develop a model of stochastic choice as a result of deliberate randomization by individuals with quasi-concave preferences. However, Hey and Carbone (1995) find that randomness in the observed choices cannot be attributed to conscious randomization.

Deterministic decision theories derive representation of unique preference relation and predict that repeated choice is always consistent (except for decision problems where an individual is exactly indifferent). Strictly speaking, all deterministic theories are falsified by experimental data. However, a common approach is to embed a deterministic decision

theory into a model of stochastic choice and fit this compound structure to empirical data. Three models of stochastic choice have been suggested in the literature.

The simplest model is a tremble model of Harless and Camerer (1994). Individuals have a unique preference relation on the set of risky lotteries but they do not choose according to their preferences all the time. With probability $p > 0$ a tremble occurs and individuals choose an alternative which is not their preferred option. With probability $1 - p$ individuals act in accordance with their preferences. Carbone (1997) and Loomes et al. (2002) find that this constant error model fails to explain the experimental data and it is essentially "inadequate as a *general* theory of stochastic choice".

The Fechner model of random errors was originally proposed by Fechner (1860). Individuals possess a unique preference relation on the set of risky lotteries Λ but they reveal their preferences with a random error as a result of carelessness, slips, insufficient motivation etc. (e.g. Hey and Orme, 1994). Preference relation is assumed to have utility representation $U : \Lambda \to \mathbf{R}$ so that it is possible to evaluate a relative advantage $d = U(L_1) - U(L_2)$ of a lottery L_1 over another lottery L_2. If there were no random errors, an individual would choose lottery L_1 whenever $d > 0$. Since actual choices are affected by mistakes, an individual chooses lottery L_1 if $d + \varepsilon > 0$, where ε is a random error.

In a classical Fechner model, random error ε is drawn from a normal distribution with zero mean and constant standard deviation (e.g. Hey and Orme, 1994). Hey (1995) and Buschena and Zilberman (2000) assume that error term is heteroscedastic *i.e.* the standard deviation of ε is higher in certain decision problems, for example, when lotteries have many possible outcomes. Blavatskyy (2005) develops a model where random error ε is drawn from a truncated normal distribution.

Finally, the third model of stochastic choice is a random utility model. Individuals are endowed with several preference relations over the set of risky lotteries and a probability measure over those preference relations. When faced with a decision problem, individuals first draw a preference relation and then choose an alternative which they prefer according to the selected preference relation (e.g. Loomes and Sugden, 1995).

This paper uses alternative framework for analyzing decision making under risk. Since repeated decisions are often inconsistent, a natural interpretation of this fact is that individuals do not have a unique preference relation on the set of risky lotteries. It is assumed that individuals possess a probability measure that captures the likelihood of one lottery being chosen over another lottery. These choice probabilities serve as primitives of choice and they admit stochastic utility representation if they satisfy five intuitive axioms. A preference relation can be easily translated into choice probabilities. Thus, there is no need for a stochastic choice model as a mediator between a deterministic preference relation and an empirical stochastic choice pattern. A related axiomatization of choice probabilities for riskless alternatives is given in Debreu (1958) and for risky lotteries—in Fishburn (1978). In particular, Fishburn (1978) represents choice probabilities by an increasing function of the *ratio* of incremental expected utility advantages of two lotteries.

The remainder of this paper is organized as follows. Section II introduces a new framework, where primitives of choice are choice probabilities rather than a deterministic preference relation. Section III presents five intuitively appealing properties (axioms) of individual choice probabilities. Section IV contains the main result of the paper—a representation theorem (stochastic utility theorem) for choice probabilities and discusses its main implications and possible extensions. Section V concludes.

II. Framework

Let $X = \{x_1,...,x_n\}$ denote a finite set of all possible outcomes (consequences). Outcomes are not necessarily monetary payoffs. They can be, for example, consumption bundles or portfolios of assets. Objects of choice are risky lotteries. A risky lottery $L(p_1,...,p_n)$ is a probability distribution on X i.e. it delivers outcome x_i with probability $p_i \in [0,1]$, $i \in \{1,...,n\}$, and $\sum_{i=1}^{n} p_i = 1$. A compound lottery $\alpha L_1 + (1-\alpha)L_2$ yields a risky lottery L_1 with probability $\alpha \in [0,1]$ and a risky lottery L_2 with probability $1-\alpha$. The set of all risky lotteries is denoted by Λ.

An individual (decision maker) does not have a preference relation on Λ. Instead, an individual possesses a probability measure on $\Lambda \times \Lambda$, which serves as a primitive of choice. For two risky lotteries $L_1, L_2 \in \Lambda$ a choice probability $\Pr(L_1, L_2) \in [0,1]$ denotes a likelihood that an individual chooses L_1 over L_2 in a repeated binary choice. Choice probability $\Pr(L_1, L_2)$ is clearly observable from a (relative) frequency with which an individual chooses L_1 when he or she is asked to choose repeatedly between L_1 and L_2.

A deterministic preference relation can be easily converted into a choice probability. If an individual strictly prefers L_1 over L_2, then $\Pr(L_1, L_2) = 1$. If an individual strictly prefers L_2 over L_1, then $\Pr(L_1, L_2) = 0$. Finally, if an individual is exactly indifferent between L_1 and L_2, then $\Pr(L_1, L_2) = 1/2$. Thus, a deterministic binary preference relation on Λ, which is a starting building block of nearly all decision theories, can be considered as a special case of a more general framework where individual decisions are governed by choice probabilities.

III. Axioms

Axiom 1 (Completeness) For any two lotteries $L_1, L_2 \in \Lambda$ there exist a choice probability $\Pr(L_1, L_2) \in [0,1]$ and a choice probability $\Pr(L_2, L_1) = 1 - \Pr(L_1, L_2)$.

Axiom 1 immediately implies that $\Pr(L, L) = 1/2$ for any $L \in \Lambda$. In a binary choice between L_1 and L_2 only two events are possible—either an individual chooses L_1 over L_2 or an individual chooses L_2 over L_1 (but not both at the same time). This distinguishes Axiom 1 from completeness imposed on a binary preference relation. Axiom 1 can be extended by allowing for a "neutral" event when an individual does not care about a choice problem and delegates choice decision to an arbitrary third party (but bears consequences of the third party decision). However, empirical evidence suggests that such option, when available, is seldom used (Camerer, 1989) and it is not necessary for a theoretical analysis.

Axiom 2 (Strong Stochastic Transitivity) For any three lotteries $L_1, L_2, L_3 \in \Lambda$ if $\Pr(L_1, L_2) \geq 1/2$ and $\Pr(L_2, L_3) \geq 1/2$ then $\Pr(L_1, L_3) \geq \max\{\Pr(L_1, L_2), \Pr(L_2, L_3)\}$.

Axiom 3 (Continuity) For any three lotteries $L_1, L_2, L_3 \in \Lambda$ the sets $\{\alpha \in [0,1] \mid \Pr(\alpha L_1 + (1-\alpha)L_2, L_3) \geq 1/2\}$ and $\{\alpha \in [0,1] \mid \Pr(\alpha L_1 + (1-\alpha)L_2, L_3) \leq 1/2\}$ are closed.

Intuitively, continuity insures that a small change in the probability distribution over outcomes does not result in a significant change in the choice probabilities.

Axiom 4 (Common Consequence Independence) For any four lotteries $L_1, L_2, L_3, L_4 \in \Lambda$ and any probability $\alpha \in [0,1]$: $\Pr(\alpha L_1 + (1-\alpha)L_3, \alpha L_2 + (1-\alpha)L_3) =$
$$= \Pr(\alpha L_1 + (1-\alpha)L_4, \alpha L_2 + (1-\alpha)L_4).$$

The intuition behind Axiom 4 is the following. In a choice between two lotteries a choice probability is independent of the consequences that are common to both lotteries. When risk is resolved, an individual receives only one realized outcome of a risky lottery. Thus, if two risky lotteries yield identical chances of the same outcome (or, more generally, if two compound lotteries yield identical chances of the same risky lottery) this *common consequence* does not affect the choice probability. Axiom 4 is weaker than condition $\Pr(L_1, L_2) = \Pr(\alpha L_1 + (1-\alpha)L_3, \alpha L_2 + (1-\alpha)L_3)$ for any three lotteries $L_1, L_2, L_3 \in \Lambda$ and any probability $\alpha \in [0,1]$. The latter condition can be interpreted as a stochastic version of the independence axiom of expected utility theory.

Axiom 5 (Interchangeability) For any three lotteries $L_1, L_2, L_3 \in \Lambda$ if $\Pr(L_1, L_2) = \Pr(L_2, L_1) = 1/2$, then $\Pr(L_1, L_3) = \Pr(L_2, L_3)$.

Intuitively, if an individual chooses between two lotteries at random then he or she does not mind which of the two lotteries is involved in another decision problem. If in a direct binary choice an individual is equally likely to choose either of the two lotteries then he or she does not favor either lottery. Thus, these two lotteries can be interchanged in any other decision problem without affecting choice probabilities. Axiom 5 holds trivially for a transitive binary preference relation.

IV. Stochastic Utility Theorem

Theorem 1 (Stochastic Utility Theorem) Probability measure on $\Lambda \times \Lambda$ satisfies Axioms 1-5 if and only if there exist an assignment of real numbers u_i to every outcome x_i, $i \in \{1,...,n\}$, and there exist a non-decreasing function $\Psi : \mathbf{R} \to [0,1]$ such that for any two risky lotteries $L_1(p_1,...,p_n), L_2(q_1,...,q_n) \in \Lambda$: $\Pr(L_1, L_2) = \Psi\left(\sum_{i=1}^{n} u_i p_i - \sum_{i=1}^{n} u_i q_i\right)$.

Proof is presented in the Appendix.

Function $\Psi(.)$ has to satisfy a restriction $\Psi(x) = 1 - \Psi(-x)$ for every $x \in \mathbf{R}$, which immediately implies that $\Psi(0) = 1/2$. If a vector $U = \{u_1,...,u_n\}$ and function $\Psi(.)$ represent a probability measure on $\Lambda \times \Lambda$ then a vector $U' = aU + b$ and a function $\Psi'(.) = \Psi(./a)$ represent the same probability measure for any two real numbers a and b, $a \neq 0$. Vector $U = \{u_1,...,u_n\}$ can be regarded as a vector of von Neumann-Morgenstern utilities that capture the "goodness" of outcomes $\{x_1,...,x_n\}$. Function $\Psi(.)$ can be regarded as a measure of relative advantage of one lottery over another.

Several popular models of stochastic choice emerge from a general stochastic utility representation in Theorem 1 as special cases. For example, when function $\Psi(.)$ is a cumulative distribution function of the normal distribution with mean zero and constant standard deviation $\sigma > 0$, stochastic utility representation becomes the Fechner model of random errors (e.g. Fechner, 1860; Hey and Orme, 1994). When function $\Psi(.)$ is a cumulative distribution function of the logistic distribution, i.e. $\Psi(x) = 1/(1 + \exp(-\lambda x))$, where $\lambda > 0$ is constant, stochastic utility representation becomes Luce choice model (e.g. Luce and Suppes, 1965; Camerer and Ho, 1994; Wu and Gonzalez, 1996). Finally, if

function $\Psi(.)$ is the step function (piecewise constant function) of the form $\Psi(x) = p$ if $x < 0$, $\Psi(x) = 1/2$ if $x = 0$ and $\Psi(x) = 1 - p$ if $x > 0$, where $p > 0$ is constant, stochastic utility representation becomes a tremble model of Harless and Camerer (1994).

Axiomatization of stochastic utility allows characterizing the above mentioned models in terms of the properties of individual choice probabilities that are observable from empirical data and that can be directly tested in a controlled laboratory experiment. In particular, axioms allow us to relate stochastic utility theory to the well known stylized empirical facts about individual choice behavior under risk. Unlike expected utility theory, stochastic utility theory is consistent with systematic violations of betweenness and a common ratio effect. However, a subjective probability measure that admits stochastic utility representation cannot exhibit a common consequence effect.

For two lotteries $L_1, L_2 \in \Lambda$ and $\alpha \in [0,1]$ let event A denote a choice of L_1 over L_2 <u>and</u> a choice of $\alpha L_1 + (1 - \alpha)L_2$ over L_1 and let event B denote a choice of L_2 over L_1 <u>and</u> a choice of L_1 over $\alpha L_1 + (1 - \alpha)L_2$. Systematic violations of betweenness are observed if A happens more frequently than B or vice versa (e.g. Camerer and Ho, 1994). If binary choices are independent then $A(B)$ is observed more frequently than $B(A)$ when $\Pr(L_1, L_2)$ is greater (smaller) than $\Pr(L_1, \alpha L_1 + (1 - \alpha)L_2)$. If a probability measure on $\Lambda \times \Lambda$ satisfies Axioms 2 and 4 then $\Pr(L_1, L_2)$ is greater (smaller) than $\Pr(L_1, \alpha L_1 + (1 - \alpha)L_2)$ when $\Pr(L_1, \alpha L_1 + (1 - \alpha)L_2)$ is greater (smaller) than one half. Thus, stochastic utility can exhibit a pattern known as systematic violations of betweenness (e.g. Blavatskyy, 2006).

Let $n = 3$ and lotteries have monetary outcomes $x_3 > x_2 > x_1 = 0$. A common ratio effect is observed when $\Pr(L_1 \succ L_2) > \Pr(L_3 \succ L_4)$ for lotteries $L_1(0,1,0)$, $L_2(1 - p, 0, p)$,

$L_3(1-\theta,\theta,0)$, $L_4(1-\theta p,0,\theta p)$ and probabilities $p,\theta \in (0,1)$ (e.g. Allais, 1953). Choice probabilities that have stochastic utility representation can exhibit a common ratio effect. If choice probabilities satisfy Axiom 4 then $\Pr(L_3,L_4) = \Pr(L_1,\theta L_2 + (1-\theta)L_1)$. When they additionally satisfy Axiom 2 then $\Pr(L_1,L_2) \geq \Pr(L_1,\theta L_2 + (1-\theta)L_1)$ if $\Pr(L_1,\theta L_2 + (1-\theta)L_1) \geq 1/2$. Thus, stochastic utility is consistent with a common ratio effect $\Pr(L_1,L_2) \geq \Pr(L_3,L_4) \geq 1/2$ but it cannot explain a common ratio effect $\Pr(L_1,L_2) \geq 1/2 \geq \Pr(L_3,L_4)$ (e.g. Loomes, 2005).

A common consequence effect is observed when $\Pr(L_1,L_2) > \Pr(L_3,L_4)$ for lotteries $L_1(0,1,0)$, $L_2(p-q,1-p,q)$, $L_3(1-p,p,0)$, $L_4(1-q,0,q)$ and probabilities $p,q \in (0,1)$, $p > q$ (e.g. Allais, 1953). Choice probabilities that admit stochastic utility representation cannot exhibit common consequence effect because Axiom 4 implies $\Pr(L_1,L_2) = \Pr(L_3,L_4)$. To accommodate a common consequence effect within stochastic utility framework Axiom 4 can be replaced with a stochastic analogue of one of the axioms of non-expected utility theories. For example, Axiom 4 can be weakened into Axiom 4a.

Axiom 4a (Betweenness) For any $L_1,L_2 \in \Lambda$ and any $\alpha,\beta,\gamma,\delta \in [0,1]$ such that $\alpha - \beta = \gamma - \delta$: $\Pr(\alpha L_1 + (1-\alpha)L_2, \beta L_1 + (1-\beta)L_2) = \Pr(\gamma L_1 + (1-\gamma)L_2, \delta L_1 + (1-\delta)L_2)$.

Probability measure on $\Lambda \times \Lambda$ that satisfies Axioms 1-3, 4a and 5 can be represented by implicit stochastic utility $\Pr(L_1,L_2) = \Psi(U(L_1) - U(L_2))$ for any $L_1, L_2 \in \Lambda$, where $U : \Lambda \to \mathbf{R}$ is implicit expected utility of a lottery (e.g. Dekel, 1986). Thus, stochastic utility theorem can be extended to represent choice probabilities that do not necessarily exhibit common consequence independence.

V. Conclusion

One of the robust findings from experimental research on repeated decision making under risk is that individuals often make contradictory choices when they face the same binary choice problem within a short period of time. This evidence suggests that individuals do not possess a unique preference relation on the space of risky lotteries Λ. Either individuals have multiple preference relations on Λ that can be represented by a random utility model (recently axiomatized by Gul and Pesendorfer, 2006) or individuals have a probability measure on $\Lambda \times \Lambda$ that can be represented by stochastic utility model.

This paper shows that choice probabilities admit a stochastic utility representation if and only if they are complete, strongly transitive, continuous, independent of common consequences and interchangeable. Axioms of stochastic utility are consistent with several choice patterns (such as systematic violations of betweenness and a common ratio effect) that contradict to the axioms of expected utility framework. Special cases of stochastic utility representation include the Fechner model of random errors, Luce choice model and a tremble model of Harless and Camerer (1994).

At least one axiom of stochastic utility—common consequence independence—has been extensively tested in controlled experiments and it is known to be frequently violated. However, interchangeability axiom is likely to be problematic too. When lottery outcomes are monetary, choice among outcomes for certain is deterministic. In fact, when one lottery transparently first-order stochastically dominates the other lottery, subjects seldom choose a dominated alternative (e.g. Loomes and Sugden, 1998). Thus, reality appears to be somewhere between a deterministic decision theory derived from a preference relation on Λ and a stochastic decision theory derived from a probability measure on $\Lambda \times \Lambda$.

Apparently, an individual possesses a preference relation on a non-empty subset of Λ (that includes, for example, the set of lottery outcomes) and a probability measure on $\Lambda \times \Lambda$. The preference relation allows to make cognitively undemanding decisions, e.g. when one alternative clearly dominates the other alternative, and the probability measure governs the remaining decisions. Such a hybrid model of decision making can be characterized within the framework of choice probabilities introduced in this paper because a deterministic preference relation can be easily translated into choice probabilities. A restriction on the interchangeability axiom so that it holds only when lotteries do not transparently dominate each other appears to be a necessary first step for extending stochastic utility theorem along these lines.

References

Allais, M. (1953) "Le comportement de l'homme rationnel devant le risque: critique des postulates et axiomes de l'école Américaine" *Econometrica* **21**, 503-546

Ballinger, P. and N. Wilcox (1997) "Decisions, error and heterogeneity" *Economic Journal* **107**, 1090-1105

Blavatskyy, P. (2005) "A Stochastic Expected Utility Theory" IEW WP 231

Blavatskyy, P. (2006) "Violations of Betweenness or Random Errors?" *Economics Letters* **91(1)**, 34-38

Buschena, D. and Zilberman, D. (2000) "Generalized expected utility, heteroscedastic error, and path dependence in risky choice" *Journal of Risk and Uncertainty* **20**, 67-88

Camerer, C. (1989) "An experimental test of several generalized utility theories." *Journal of Risk and Uncertainty* **2**, 61-104

Camerer, C. and Ho, T. (1994) "Violations of the Betweenness Axiom and Nonlinearity in Probability" *Journal of Risk and Uncertainty* **8**, 167-196

Carbone, E. (1997) "Investigation of stochastic preference theory using experimental data" *Economics Letters* **57**, 305-311

Chew, S., L. Epstein and U. Segal (1991) "Mixture symmetry and quadratic utility" *Econometrica* **59**, 139-163

Debreu, G. (1958) "Stochastic choice and cardinal utility" *Econometrica* **26**, 440-444

Dekel, E. (1986) "An axiomatic characterization of preferences under uncertainty" *Journal of Economic Theory* **40**, 304-318

Gul, F. and Pesendorfer, W. (2006) "Random Expected Utility" *Econometrica* **71(1)**, 121-146

Fechner, G. (1860) "Elements of Psychophysics" NewYork: Holt, Rinehart and Winston

Fishburn, P. (1978) A probabilistic expected utility theory of risky binary choices, *International Economic Review* **19**, 633-646

Harless, D. and C. Camerer (1994) The predictive utility of generalized expected utility theories, *Econometrica* **62**, 1251-1289

Hey, J. (1995) "Experimental investigations of errors in decision making under risk" *European Economic Review* **39**, 633-640

Hey, J.D. and C. Orme (1994) Investigating generalisations of expected utility theory using experimental data, *Econometrica* **62**, 1291-1326

Hey, J. and Carbone, E. (1995) "Stochastic Choice with Deterministic Preferences: An Experimental Investigation" *Economics Letters* **47**, 161-167

Loomes, G. (2005) "Modelling the stochastic component of behaviour in experiments: some issues for the interpretation of data" *Experimental Economics* **8** 301-323

Loomes, G. and Sugden, R. (1995) "Incorporating a stochastic element into decision theories" *European Economic Review* **39**, 641-648

Loomes, G. and Sugden, R. (1998) "Testing different stochastic specifications of risky choice" *Economica* **65**, 581-598

Loomes, G., Moffatt, P. and Sugden, R. (2002) "A microeconomic test of alternative stochastic theories of risky choice" *Journal of Risk and Uncertainty* **24**, 103-130

Luce, R. D. and Suppes, P. (1965) "Preference, utility, and subjective probability" in R. D. Luce, R. R. Bush & E. Galanter (eds.), *Handbook of mathematical psychology*, Vol. III, 249–410, Wiley, New York NY

Machina, M. (1985) "Stochastic Choice Functions Generated from Deterministic Preferences over Lotteries" *Economic Journal* **95**, 575-594

Starmer, Ch. and Sugden, R. (1989) "Probability and juxtaposition effects: An experimental investigation of the common ratio effect." *Journal of Risk and Uncertainty* **2**, 159-178

Wu, G. (1994) "An Empirical Test of Ordinal Independence" *Journal of Risk and Uncertainty* **9**, 39-60

Wu, G. and Gonzalez, R. (1996) "Curvature of the Probability Weighting Function." *Management Science* **42**, 1676-90

Appendix

Proof of Theorem 1

It is straightforward to verify that if a probability measure on $\Lambda \times \Lambda$ has a stochastic utility representation then it satisfies all Axioms 1-5. We will now prove that if a probability measure on $\Lambda \times \Lambda$ satisfies Axioms 1-5 then it has a stochastic utility representation. Since the set of possible outcomes X is finite, completeness and strong stochastic transitivity imply that there exist "the best" and "the worst" lottery in Λ i.e. $\exists \overline{L}, \underline{L} \in \Lambda$ such that $\Pr(\overline{L}, L) \geq 1/2$ and $\Pr(\underline{L}, L) \leq 1/2$ for any $L \in \Lambda$.

Consider first the case when $\Pr(\overline{L}, \underline{L}) = 1/2$. If $\Pr(\underline{L}, \overline{L}) = 1/2$ and $\Pr(\overline{L}, L) \geq 1/2$ for any $L \in \Lambda$ (by definition of $\overline{L}$) strong stochastic transitivity implies that $\Pr(\underline{L}, L) \geq 1/2$. Since $\Pr(\underline{L}, L) \leq 1/2$ by definition of $\underline{L}$, it must be the case that $\Pr(\underline{L}, L) = 1/2$ for any $L \in \Lambda$. If $\Pr(L_1, \underline{L}) \geq 1/2$ and $\Pr(\underline{L}, L_2) = 1/2$ for any $L_1, L_2 \in \Lambda$, strong stochastic transitivity implies that $\Pr(L_1, L_2) \geq 1/2$. If $\Pr(L_1, L_2) \geq 1/2$ and $\Pr(L_2, L_1) \geq 1/2$ it must be the case that $\Pr(L_1, L_2) = 1/2$ for any $L_1, L_2 \in \Lambda$ and this degenerate probability measure can be represented by assignment $u_i = 0$, $i \in \{1, ..., n\}$, and $\Psi(0) = 1/2$. Therefore, from this point onwards only the case when $\Pr(\overline{L}, \underline{L}) > 1/2$ is considered.

Continuity axiom implies that the sets $\left\{ \alpha \in [0,1] \mid \Pr(\alpha \overline{L} + (1-\alpha)\underline{L}, L) \geq 1/2 \right\}$ and $\left\{ \alpha \in [0,1] \mid \Pr(\alpha \overline{L} + (1-\alpha)\underline{L}, L) \leq 1/2 \right\}$ are closed for any $L \in \Lambda$. Completeness axiom guarantees that every $\alpha \in [0,1]$ belongs to at least one of these two sets. Since both sets are nonempty ($\alpha = 1$ belongs to the first set and $\alpha = 0$ belongs to the second set), there is at

14

least one α that belongs to both sets. Thus, for any $L \in \Lambda$ there exist $\alpha_L \in [0,1]$ such that $\Pr(\alpha_L \overline{L} + (1 - \alpha_L)\underline{L}, L) = 1/2$. We will now prove that α_L is unique for every $L \in \Lambda$.

Suppose there is $\alpha'_L \neq \alpha_L$ such that $\Pr(\alpha'_L \overline{L} + (1 - \alpha'_L)\underline{L}, L) = 1/2$. Assume without a loss of generality that $\alpha'_L > \alpha_L$ and let $\delta = \alpha'_L - \alpha_L$. If $\Pr(\alpha'_L \overline{L} + (1 - \alpha'_L)\underline{L}, L) = 1/2$ and $\Pr(L, \alpha_L \overline{L} + (1 - \alpha_L)\underline{L}) = 1/2$ then $\Pr(\alpha'_L \overline{L} + (1 - \alpha'_L)\underline{L}, \alpha_L \overline{L} + (1 - \alpha_L)\underline{L}) = 1/2$ by strong stochastic transitivity. Common consequence independence implies $\Pr(\alpha'_L \overline{L} + (1 - \alpha'_L)\underline{L}, \alpha_L \overline{L} + (1 - \alpha_L)\underline{L}) = \Pr(\delta \overline{L} + (1 - \delta)\underline{L}, \underline{L}) = 1/2$. If $\Pr(\delta \overline{L} + (1 - \delta)\underline{L}, \underline{L}) = 1/2$ then $\Pr(2\delta \overline{L} + (1 - 2\delta)\underline{L}, \delta \overline{L} + (1 - \delta)\underline{L}) = 1/2$ by common consequence independence and $\Pr(2\delta \overline{L} + (1 - 2\delta)\underline{L}, \underline{L}) = 1/2$ by strong stochastic transitivity. Applying this reasoning $k = \text{int}\{\alpha'_L / \delta\}$ times we obtain that $\Pr(k\delta \overline{L} + (1 - k\delta)\underline{L}, \underline{L}) = 1/2$.

By a similar pattern, common consequence independence implies $\Pr(\alpha'_L \overline{L} + (1 - \alpha'_L)\underline{L}, \alpha_L \overline{L} + (1 - \alpha_L)\underline{L}) = \Pr(\overline{L}, (1 - \delta)\overline{L} + \delta\underline{L}) = 1/2$. If $\Pr(\overline{L}, (1 - \delta)\overline{L} + \delta\underline{L}) = 1/2$ then $\Pr((1 - \delta)\overline{L} + \delta\underline{L}, (1 - 2\delta)\overline{L} + 2\delta\underline{L}) = 1/2$ by common consequence independence and $\Pr(\overline{L}, (1 - 2\delta)\overline{L} + 2\delta\underline{L}) = 1/2$ by strong stochastic transitivity. Applying this reasoning $m = \text{int}\{(1 - \alpha_L) / \delta\}$ times we obtain that $\Pr(\overline{L}, (1 - m\delta)\overline{L} + m\delta\underline{L}) = 1/2$.

We prove next that $\Pr(\eta \overline{L} + (1 - \eta)\underline{L}, v\overline{L} + (1 - v)\underline{L}) = 1/2$ for any $\eta, v \in [\alpha_L, \alpha'_L]$. $\Pr(\eta \overline{L} + (1 - \eta)\underline{L}, \alpha_L \overline{L} + (1 - \alpha_L)\underline{L}) = \Pr((\eta - \alpha_L)\overline{L} + (1 - \eta + \alpha_L)\underline{L}, \underline{L}) \geq 1/2$ with equality due to common consequence independence and inequality due to definition of $\underline{L}$. Similarly, $\Pr(\alpha'_L \overline{L} + (1 - \alpha'_L)\underline{L}, \eta \overline{L} + (1 - \eta)\underline{L}) = \Pr((\alpha'_L - \eta)\overline{L} + (1 - \alpha'_L + \eta)\underline{L}, \underline{L}) \geq 1/2$. Strong stochastic

transitivity then implies $\Pr\left(\alpha'_L\overline{L}+(1-\alpha'_L)\underline{L},\alpha_L\overline{L}+(1-\alpha_L)\underline{L}\right)=1/2\geq$

$\geq\Pr\left(\eta\overline{L}+(1-\eta)\underline{L},\alpha_L\overline{L}+(1-\alpha_L)\underline{L}\right)$. Therefore it must be the case that

$\Pr\left(\eta\overline{L}+(1-\eta)\underline{L},\alpha_L\overline{L}+(1-\alpha_L)\underline{L}\right)=1/2$ for any $\eta\in[\alpha_L,\alpha'_L]$. If

$\Pr\left(\eta\overline{L}+(1-\eta)\underline{L},\alpha_L\overline{L}+(1-\alpha_L)\underline{L}\right)=1/2$ and $\Pr\left(\nu\overline{L}+(1-\nu)\underline{L},\alpha_L\overline{L}+(1-\alpha_L)\underline{L}\right)=1/2$ strong

stochastic transitivity then implies $\Pr\left(\eta\overline{L}+(1-\eta)\underline{L},\nu\overline{L}+(1-\nu)\underline{L}\right)=1/2$ for any

$\eta,\nu\in[\alpha_L,\alpha'_L]$

Since $k\delta\in[\alpha_L,\alpha'_L]$ and $1-m\delta\in[\alpha_L,\alpha'_L]$ by our choice of integers k and m it

follows that $\Pr\left((1-m\delta)\overline{L}+m\delta\underline{L},k\delta\overline{L}+(1-k\delta)\underline{L}\right)=1/2$. Strong stochastic transitivity then

implies $\Pr\left(\overline{L},k\delta\overline{L}+(1-k\delta)\underline{L}\right)=1/2$ and $\Pr\left(\overline{L},\underline{L}\right)=1/2$. However, we already considered

this case at the beginning of this proof. Therefore it must be the case that $\alpha'_L=\alpha_L$. In other

words, for any $L\in\Lambda$ there exist unique $\alpha_L\in[0,1]$ such that $\Pr\left(\alpha_L\overline{L}+(1-\alpha_L)\underline{L},L\right)=1/2$.

Interchangeability implies $\Pr(L_1,L_2)=\Pr\left(\alpha_{L_1}\overline{L}+(1-\alpha_{L_1})\underline{L},\alpha_{L_2}\overline{L}+(1-\alpha_{L_2})\underline{L}\right)$ for

any two risky lotteries $L_1,L_2\in\Lambda$. If $\alpha_{L_1}>\alpha_{L_2}$ then a common consequence independence

implies $\Pr(L_1,L_2)=\Pr\left((\alpha_{L_1}-\alpha_{L_2})\overline{L}+(1-\alpha_{L_1}+\alpha_{L_2})\underline{L},\underline{L}\right)\equiv\Psi\left(\alpha_{L_1}-\alpha_{L_2}\right)$. If $\alpha_{L_1}<\alpha_{L_2}$ then

a common consequence independence implies

$\Pr(L_1,L_2)=\Pr\left(\underline{L},(\alpha_{L_2}-\alpha_{L_1})\overline{L}+(1-\alpha_{L_2}+\alpha_{L_1})\underline{L}\right)$ and completeness additionally implies

$\Pr(L_1,L_2)=1-\Pr\left((\alpha_{L_2}-\alpha_{L_1})\overline{L}+(1-\alpha_{L_2}+\alpha_{L_1})\underline{L},\underline{L}\right)=1-\Psi\left(\alpha_{L_2}-\alpha_{L_1}\right)\equiv\Psi\left(\alpha_{L_1}-\alpha_{L_2}\right)$.

Notice that function $\Psi(.)$ is non-decreasing. For any $\lambda,\mu\in[0,1]$ such that $\lambda>\mu$

common consequence independence implies $\Pr\left(\lambda\overline{L}+(1-\lambda)\underline{L},\mu\overline{L}+(1-\mu)\underline{L}\right)=$

$= \Pr\left((\lambda - \mu)\overline{L} + (1 - \lambda + \mu)\underline{L}, \underline{L}\right) \geq 1/2$ with the last inequality due to definition of $\underline{L}$. Since

$\Pr\left(\lambda\overline{L} + (1 - \lambda)\underline{L}, \mu\overline{L} + (1 - \mu)\underline{L}\right) \geq 1/2$ and $\Pr\left(\mu\overline{L} + (1 - \mu)\underline{L}, \underline{L}\right) \geq 1/2$ (again by definition of

$\underline{L}$), strong stochastic transitivity implies $\Pr\left(\lambda\overline{L} + (1 - \lambda)\underline{L}, \underline{L}\right) \geq \Pr\left(\mu\overline{L} + (1 - \mu)\underline{L}, \underline{L}\right)$.

Therefore, $\Psi(\lambda) \geq \Psi(\mu)$ for any $\lambda, \mu \in [0,1]$ such that $\lambda > \mu$.

Interchangeability axiom implies that for any $\beta \in [0,1]$ and any three lotteries

$L_1, L_2, L_3 \in \Lambda$ a choice probability $\Pr\left(\beta L_1 + (1 - \beta)L_2, L_3\right)$ can be rewritten as

$\Pr\left(\beta\left(\alpha_{L_1}\overline{L} + (1 - \alpha_{L_1})\underline{L}\right) + (1 - \beta)\left(\alpha_{L_2}\overline{L} + (1 - \alpha_{L_2})\underline{L}\right), L_3\right)$. After rearranging terms we obtain

$\Pr\left(\beta L_1 + (1 - \beta)L_2, L_3\right) = \Pr\left(\left[\beta\alpha_{L_1} + (1 - \beta)\alpha_{L_2}\right] \cdot \overline{L} + \left[1 - \beta\alpha_{L_1} - (1 - \beta)\alpha_{L_2}\right] \cdot \underline{L}, L_3\right) =$

$= \Psi\left(\beta\alpha_{L_1} + (1 - \beta)\alpha_{L_2} - \alpha_{L_3}\right)$. Thus, assignment of numbers α_L is linear in probabilities i.e.

$\alpha_{\beta L_1 + (1 - \beta)L_2} = \beta\alpha_{L_1} + (1 - \beta)\alpha_{L_2}$.

To summarize, for every outcome x_i, $i \in \{1,...,n\}$, we can find a number $u_i \in [0,1]$

such that $\Pr\left(u_i\overline{L} + (1 - u_i)\underline{L}, x_i\right) = 1/2$. For any two risky lotteries

$L_1(p_1,...,p_n), L_2(q_1,...,q_n) \in \Lambda$ a choice probability $\Pr(L_1, L_2)$ can be then written as

$\Pr(L_1, L_2) = \Psi\left(\sum_{i=1}^{n} u_i p_i - \sum_{i=1}^{n} u_i q_i\right)$. Q.E.D.