Skewed Noise*

David Dillenberger† Uzi Segal‡

October 4, 2015

Abstract

We study the attitude of decision makers to skewed noise. For a binary lottery that yields the better outcome with probability p, we identify noise around p with a compound lottery that induces a distribution over the exact value of the probability and has an average value p. We characterize a new notion of skewed distributions, and use a recursive non-expected utility to provide conditions under which rejection of symmetric noise implies rejection of negatively skewed noise, yet does not preclude acceptance of some positively skewed noise, in agreement with recent experimental evidence. In the context of decision making under uncertainty, our model permits the co-existence of aversion to symmetric ambiguity (as in Ellsberg's paradox) and ambiguity seeking for low likelihood "good" events. We also use the model to the study of random allocation problems and show that it can predict systematic preference for one allocation mechanism over the other, even though the two agree on the overall probability distribution over assignments.

Keywords: Skewed distributions, recursive non-expected utility, ambiguity seeking, one-sided matching.

---

*We thank Faruk Gul, Yoram Halevy, Efe Ok, Collin Raymond, Joel Sobel, Tayfun Sönmez, Utku Ünver, and Rakesh Vohra for their comments and help. We thank Francesca Toscano and Zhu Zhu of Boston College for helpful research assistance.

†Department of Economics, University of Pennsylvania (ddill@sas.upenn.edu).

‡Department of Economics, Boston College (segalu@bc.edu) and Warwick Business School.

1 Introduction

Standard models of decision making under risk assume that individuals obey the reduction of compound lotteries axiom, according to which a decision maker is indifferent between any multi-stage lottery and the simple lottery that induces the same probability distribution over final outcomes. Experimental and empirical evidence suggest, however, that this axiom is often violated (see, among others, Kahneman and Tversky [26], Bernasconi and Loomes [6], Conlisk [12], Harrison, Martinez-Correa, and Swarthout [23], and Abdellaoui, Klibanoff, and Placido [1]). Individuals may have preferences over gambles with identical probability distributions over final outcomes that differ in the timing of resolution of uncertainty. Alternatively, if individuals distinguish between the source of risk in each stage and thus perceive risk as a multi-stage prospect, independently of time, they may care about the number of lotteries they participate in or about their order.

Halevy [22] and Miao and Zhong [33], for example, consider preferences over two-stage lotteries and demonstrate that individuals are averse to the introduction of symmetric noise, that is, symmetric mean-preserving spread into the first-stage lottery. One rationale for this kind of behavior is that the realizations of a symmetric noise cancel out each other but create an undesired confusion in evaluation. On the other hand, there is some evidence that asymmetric, positively skewed noise may be desirable. Boiney [7] conducted an experiment in which subjects had to choose one of three investment plans. In all three prospects, the overall probability of success (which results in a prize $\overline{x} = \$200$) is $p = 0.2$, and with the remaining probability the investment fails and $\underline{x} = \$0$ is received. Option A represents an investment plan in which the probability of success is given. In B and C, on the other hand, the probability of success is uncertain. Prospect B (resp., C) represents a negatively (positively) skewed distribution around p in which it is very likely that the true probability slightly exceeds (falls below) p but it is also possible, albeit unlikely, that the true probability is much lower (higher).[1] Boiney's main find-

[1]Specifically, option B was such that with probability 0.9 the true probability of success

ing is that subjects are not indifferent between the three prospects and that most prefer C to A and A to B. Moreover, these preferences are robust to different values of $\overline{x} > \underline{x}$ and p.

In Boiney's experiment, the underlying probability of success p was the same in all three options. In recent experiments, which we discuss in details in Section 3, Abdellaoui, Klibanoff, and Placido [1] and Abdellaoui, l'Haridon, and Nebout [2] found strong evidence that aversion to compound risk (i.e., noise) is an increasing function of p. In particular, their results are consistent with a greater aversion to negatively skewed noise around high probabilities than to positively skewed noise around small probabilities.

In this paper we propose a model that can accommodate the behavioral patterns discussed above. For a binary lottery $(\overline{x}, p; \underline{x}, 1-p)$ with $\overline{x} > \underline{x}$, we identify noise around p with a two-stage lottery that induces a distribution over the exact value of the probability and has an average value p. We introduce and characterize a new notion of skewness, and use a version of Segal's [39] recursive non-expected utility model to outline conditions under which a decision maker who always rejects symmetric noise will also reject any negatively skewed noise (for instance, will prefer option A to B in the example above) but may seek some positively skewed noise.

We suggest two applications. First, our model can be used to address the recently documented phenomenon of some ambiguity seeking in the context of decision making under uncertainty. The recursive model we study here was first suggested by Segal [38] as a way to analyze attitudes towards ambiguity. Under this interpretation, ambiguity is identified as a two-stage lottery, where the first stage captures the decision makers subjective uncertainty about the true probability distribution over the states of the world, and the second stage determines the probability of each outcome, conditional on the probability distribution that has been realized. Our model permits the co-existence of aversion to symmetric ambiguity (as in Ellsberg's [18] famous paradox) and

is 0.22 and with probability 0.1 it is only 0.02. In option C, on the other hand, with probability 0.9 the true probability of success is 0.18 and with probability 0.1 it is 0.38. The average probability of success in all three prospects is 0.2.

ambiguity seeking in situations where the decision maker anticipates a bad outcome, yet believes that there is a small chance that things are not as bad as they seem. In this case, he might not want to know the exact values of the probabilities.

We also apply our model to study a simple variant of the house allocation problem (or one-sided matching), where the goal is to look for a systematic way of assigning a set of indivisible objects to a group of individuals having preferences over these objects. We demonstrate that different mechanisms, which agree on the overall probability distribution over assignments and hence are being treated equivalently in the standard model, induce different compound lotteries. Comparing two familiar mechanisms, variants of the random serial dictator and of the random top cycle, our model suggests that the former will be preferred for a large set of parameters, yet permits the opposite preferences when one type of the goods is scarce, but almost everyone prefers it over the alternative type (see Section 5 for details).

The fact that the recursive evaluation of two-stage lotteries in Segal's model is done using non-expected utility functionals is key to our analysis. It is easy to see that if the decision maker uses the same expected utility functional in each stage he will be indifferent to noise. In Section 6 we further show that a version of the recursive model in which the two stages are evaluated using different expected utility functionals (Kreps and Porteus [27], Klibanoff, Marinacci, and Mukerji [25]) cannot accommodate the co-existence of rejecting all symmetric noise while still accepting some positively skewed noise.

This paper confines attention to the analysis of attitudes to noise related to the probability p in the binary prospect which pays $\overline{x}$ with probability p and $\underline{x} < \overline{x}$ otherwise. In reality the decision maker may face lotteries with many outcomes and the probabilities of receiving each of them may be uncertain. We prefer to deal only with binary lotteries since when there are many outcomes their probabilities depend on each other and therefore skewed noise over the probability of one event may affect noises over other probabilities in too many ways. This complication is avoided when there are only two outcomes — whatever the decision maker believes about the probability of receiving

$\overline{x}$ completely determines his beliefs regarding the probability of receiving $\underline{x}$. Note that while the underlying lottery is binary, the noise itself (that is, the distribution over the value of p) may have many possible values or may even be continuous.

The rest of the paper is organized as follows: Section 2 describes the analytical framework and introduces notations and definitions that will be used in our main analysis. Section 3 studies attitudes towards asymmetric noises and states our main behavioral result. Section 4 and Section 5 are devoted to applications. Section 6 comments on the relationship of our paper to other models. All proofs are relegated to an appendix.

2 The model

Fix two monetary outcomes $\overline{x} > \underline{x}$. The underlying lottery we consider is the binary prospect $(\overline{x}, p; \underline{x}, 1-p)$, which pays $\overline{x}$ with probability p and $\underline{x}$ otherwise. We identify this lottery with the number $p \in [0,1]$ and analyze noise around p as a two-stage lottery, denoted by $\langle p_1, q_1; ...; p_n, q_n \rangle$, that yields with probability q_i the lottery $(\overline{x}, p_i; \underline{x}, 1-p_i)$, $i = 1, 2, ..., n$, and satisfies $\sum_i p_i q_i = p$. Let

$$\mathcal{L}_2 = \left\{ \langle p_1, q_1; ...; p_n, q_n \rangle : p_i, q_i \in [0,1], i = 1, 2, ..., n, \text{ and } \sum_i q_i = 1 \right\}.$$

Let $\succeq$ be a complete and transitive preference relation over $\mathcal{L}_2$, which is represented by $U : \mathcal{L}_2 \to \Re$. Throughout the paper we confine our attention to preferences that admit the following representation:

$$U\left(\langle p_1, q_1; ...; p_n, q_n \rangle\right) = V\left(c_{p_1}, q_1; ...; c_{p_n}, q_n\right) \tag{1}$$

where V is a functional over simple (finite support) one-stage lotteries over the interval $[\underline{x}, \overline{x}]$ and c is a certainty equivalent function (not necessarily the one obtained from V).[2] According to this model, the decision maker evaluates

[2]The function c: $[0,1] \to \Re$ is a certainty equivalent function if for some W over one-stage lotteries, $W(c_p, 1) = W(\overline{x}, p; \underline{x}, 1-p)$.

a two-stage lottery $\langle p_1, q_1; ...; p_n, q_n \rangle$ recursively. He first replaces each of the second-stage lotteries with its certainty equivalent, c_{p_i}. This results in a simple, one-stage lottery over the certainty equivalents, $(c_{p_1}, q_1; ...; c_{p_n}, q_n)$, which he then evaluates using the functional V.[3] We assume throughout that V is monotonic with respect to first-order stochastic dominance and continuous with respect to the weak topology.

There are several reasons that lead us to study this special case of U. First, it explicitly captures the sequentiality aspect of two-stage lotteries, by distinguishing between the evaluations made in each stage (V and c in the first and second stage, respectively). Second, it allows us to state our results using familiar and easy to interpret conditions that are imposed on the functional V, which do not necessarily carry over to a general U. Finally, the model is a special case of the recursive non-expected utility model of Segal [39]. This will facilitate the comparison of our results with other models (see, for example, Section 6).

We identify simple lotteries with their cumulative distribution functions, denoted by capital letters (F, G, and H). Denote by $\mathcal{F}$ the set of all cumulative distribution functions of simple lotteries over $[\underline{x}, \overline{x}]$. We assume that V satisfies the assumptions below (specific conditions on c will be discussed only in the relevant section). These assumptions are common in the literature on decision making under risk.

Definition 1 *The function V is quasi concave if for any $F, G \in \mathcal{F}$ and $\lambda \in [0, 1]$,*

$$V(F) \geqslant V(G) \Longrightarrow V(\lambda F + (1-\lambda) G) \geqslant V(G).$$

Quasi concavity implies preference for randomization among equally valued prospects. Together with risk aversion ($V(F) \geqslant V(G)$ whenever G is a

[3]The functional V thus represents some underlying complete and transitive binary relation over simple lotteries, which is used in the first stage to evaluate lotteries over the certainty equivalents of the second stage. To avoid confusion with the main preferences over $\mathcal{L}_2$, we will impose all the assumptions in the text directly on V.

mean preserving spread of F), quasi concavity implies preference for portfolio diversification (Dekel [14]), which is an important feature when modeling markets of risky assets.[4]

Following Machina [29], we assume that V is smooth, in the sense that it is Fréchet differentiable, defined as follows.

Definition 2 *The function $V : \mathcal{F} \to \Re$ is Fréchet differentiable if for every $F \in \mathcal{F}$ there exists a local utility function $u_F : [\underline{x}, \overline{x}] \to \Re$, such that for every $G \in \mathcal{F}$,*

$$V(G) - V(F) = \int u_F(x) d[G(x) - F(x)] + o(\| G - F \|)$$

where $\| \cdot \|$ is the L_1-norm.

To accommodate the two Allais paradoxes — the common ratio and common consequence effects — and the mutual purchase of insurance policies and lottery tickets, Machina [29] suggested the following assumption on the behavior of the local utility function, which he labeled *Hypothesis II*: If G first-order stochastically dominates F, then at every point x, the Arrow-Pratt measure of absolute risk aversion of the local utility u_G is higher than that of u_F.[5] For the purpose of our analysis, we only need a weaker notion of Hypothesis II, which requires the property to hold just for degenerate lotteries (i.e., Dirac measures), denoted by δ_y. Formally,

Definition 3 *The Fréchet differentiable functional V satisfies Weak Hypoth-*

[4]The evidence regarding the validity of quasi concavity is supportive yet inconclusive: while the experimental literature that documents violations of linear indifference curves (see, for example, Coombs and Huang [13]) found deviations in both directions, that is, either preference for or aversion to randomization, both Sopher and Narramore [42] and Dwenger, Kubler, and Weizsacker [17] found explicit evidence in support of quasi concavity.

[5]Graphically, Hypothesis II implies that for given $x > y > z$, indifference curves in the probability triangle $\{(z, p; y, 1 - p - q; x, q)\colon (p, q) \in \Re^2_+ \text{ and } p + q \leqslant 1\}$ are "fanning out", that is, they become steeper as the probability of the good outcome x rises and the probability of the bad outcome z falls.

esis II if for every x and for every $y > z$,

$$-\frac{u''_{\delta_y}(x)}{u'_{\delta_y}(x)} \geqslant -\frac{u''_{\delta_z}(x)}{u'_{\delta_z}(x)}.$$

3 Asymmetric noise

Our aim in this paper is to analyze attitudes to skewed noise, that is, to noise that is not symmetric around its mean. For that we need first to formally define the notion of a skewed distribution.

For a distribution F on $[a, b] \subset \Re$ with expected value μ and for $\delta \geqslant 0$, let $\eta_1(F, \delta) = \int_a^{\mu-\delta} F(x)dx$ be the area below F between a and $\mu - \delta$ and $\eta_2(F, \delta) = \int_{\mu+\delta}^{b}[1 - F(x)]dx$ be the area above F between $\mu + \delta$ and b (see Figure 1). Note that $\eta_1(F, 0) = \eta_2(F, 0)$. If F is symmetric around its mean, then for every δ these two values are the same. The following definition is based on the case where the left area is systematically larger than the right area.

Definition 4 *The lottery X with the distribution F on $[a, b]$ and expected value μ is skewed to the left (or negatively skewed) if for every $\delta > 0$, $\eta_1(F, \delta) \geqslant \eta_2(F, \delta)$.*

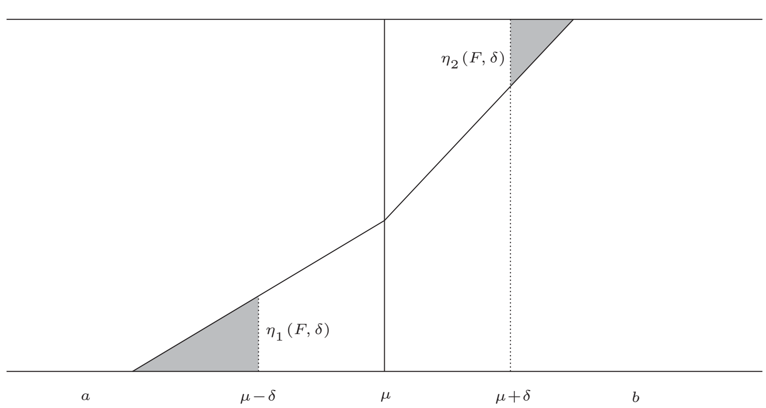

Figure 1: Definition 4, $\eta_1(F, \delta) \geqslant \eta_2(F, \delta)$

Similarly, right-skewness requires that $\eta_2(F,\delta) \geqslant \eta_1(F,\delta)$ for every $\delta > 0$. The usefulness of this new notion of skewness will become clear in Section 3.1, where we discuss the proof of our main behavioral result. (We defer a discussion of related statistical notions to Section 6.)

Recall our notation for two-stage lotteries of the form $\langle p_1, q_1; \ldots; p_m, q_m\rangle$, where p_i stands for the simple lottery $(\overline{x}, p_i; \underline{x}, 1-p_i)$ and $\overline{x} > \underline{x}$. The following definitions of rejection of symmetric and skewed noise are natural.

Definition 5 *The relation $\succeq$ rejects symmetric noise if for all p, α, and ε,*

$$\langle p, 1\rangle \succeq \langle p-\alpha, \varepsilon; p, 1-2\varepsilon; p+\alpha, \varepsilon\rangle.$$

Definition 6 *The relation $\succeq$ rejects negatively (resp., positively) skewed noise if for all $p \in (0,1)$, $\langle p, 1\rangle \succeq \langle p_1, q_1; ...; p_n, q_n\rangle$ whenever $\sum_i p_i q_i = p$ and the distribution of $(p_1, q_1; ...; p_n, q_n)$ is skewed to the left (resp., right).*

As before, we assume that the preference relation $\succeq$ over $\mathcal{L}_2$ can be represented as in eq. (1) by $U(\langle p_1, q_1; ...; p_n, q_n\rangle) = V((c_{p_1}, q_1; ...; c_{p_n}, q_n))$, where V is a functional over simple lotteries and c is a certainty equivalent function. The following theorem is the main result of the paper, showing the connection between the rejection of symmetric and skewed noises.

Theorem 1 *Suppose (i) V is quasi concave, Fréchet differentiable, and satisfies Weak Hypothesis II, and (ii) $\succeq$ rejects symmetric noise. Then $\succeq$ rejects negatively skewed noise, but not necessarily positively skewed noise.*

The first part of Theorem 1 provides conditions under which the decision maker rejects negatively-skewed noise. The conditions on V are familiar in the literature and, as we have pointed out in the introduction and will further discuss below and in Section 4, rejection of symmetric noise is empirically supported. The theoretical link between attitudes toward symmetric and negatively skewed noise will be useful in the applications we consider in subsequent sections. The second part suggests that such behavior is consistent with preference for some positively-skewed noise. The distinction between positive and negative skewness is the basis for our analysis, and as we argue, is also supported by empirical evidence. It is this part of the theorem that distinguishes our model from other known preferences over compound lotteries that cannot accommodate rejections of all symmetric noise with acceptance of some positively-skewed noise (Section 6).

In Example 1 below we go further and introduce a family of functionals for which we can provide sufficient conditions for acceptance of some positively skewed noise. In particular, for every $p > 0$, if the probability q of receiving $(\overline{x}, p; \underline{x}, 1-p)$ is sufficiently small, then the decision maker will prefer the noise $\langle p, q; 0, 1-q\rangle$ over receiving the lottery $(\overline{x}, pq; \underline{x}, 1-pq)$ for sure. To guarantee this property, we show that for the functional form of this example, the first non-zero derivative of $V(c_{pq}), 1) - V(c_p, q; 0, 1-q)$ with respect to q at $q = 0$ is negative (see Appendix B). Note that while Theorem 1 is independent of the function c, the specification of c is crucial for this result.

Example 1 Let $V(c_{p_1}, q_1; ...; c_{p_n}, q_n) = \mathrm{E}[w(c_p)] \times \mathrm{E}[c_p]$, where $w(x) = \frac{\zeta x - x^\zeta}{\zeta - 1}$ and $c_p = \beta p + (1-\beta)p^\kappa$.[6] These functions satisfy all the assumptions of Theorem 1, and there is an open neighborhood of $(\beta, \zeta, \kappa) \in \Re^3$ for which for every $p > 0$ there exists a sufficiently small $q > 0$ such that $\langle p, q; 0, 1-q\rangle \succeq \langle pq, 1\rangle$. We prove these claims in Appendix B.

Theorem 1 does not restrict the location of the skewed distribution, but it is reasonable to find skewed to the left distributions over the value of the probability p when p is high, and skewed to the right distributions when p is low. The theorem is thus consistent with the empirical observation that decision makers reject skewed to the left distributions concerning high probability of a good event, but seek such distributions when the probability of the good event is low.

Our results can explain some of the findings in a recent paper by Abdellaoui, Klibanoff, and Placido [1]. For three different compound lotteries, subjects were asked for their compound-risk premium (as in Dillenberger [15]), that is, the maximal amount they are willing to pay to replace a compound

[6]The function V belongs to the quadratic utility model of Chew, Epstein, and Segal [9].

lottery with its binary, single-stage counterpart. The underlying binary lottery yields €50 with probability p and 0 otherwise. The three two-stage lotteries were $\langle 0.5, \frac{1}{6}; 0, \frac{5}{6}\rangle$, $\langle 1, \frac{5}{22}; 0.5, \frac{12}{22}; 0, \frac{5}{22}\rangle$, and $\langle 1, \frac{5}{6}; 0.5, \frac{1}{6}\rangle$, with base probabilities of winning $p = \frac{1}{12}$, $p = \frac{1}{2}$, and $p = \frac{11}{12}$, respectively. They found that the compound-risk premium is an increasing function of p. Other studies too provide evidence for the pattern of more compound risk aversion for high probabilities than for low probabilities, and even for compound risk seeking for low probabilities (see, for example, Kahn and Sarin [25] and Viscusi and Chesson [45]). Following Theorem 1, we argue that it is not only the magnitude of the probabilities that drive their results, but the fact that in the three lotteries above, noise is positively skewed, symmetric, and negatively skewed, respectively. This is further supported by the results of two recent papers: Masatlioglu, Orhun, and Raymond [31] found that individuals exhibit a strong preference for positively skewed noise over negatively skewed ones, while Abdellaoui, l'Haridon, and Nebout [2] found that for many values of r, subjects prefer the positively skewed noise $\left\langle r, \frac{1}{3}; 0, \frac{2}{3}\right\rangle$ over its reduced version, $\left\langle \frac{r}{3}, 1\right\rangle$.

3.1 Outline of the Proof of Theorem 1

We discuss only the first part of the theorem, according to which rejection of symmetric noise implies rejection of negatively skewed noise (the second part, claiming that rejection of symmetric noise does not imply rejection of positively skewed noise, is proved by Example 1). In the recursive model, rejection of symmetric noise implies that for any p, the local utility of V at δ_{c_p} prefers $\langle p, 1\rangle$ to $\langle p-a, \frac{1}{2}; p+a, \frac{1}{2}\rangle$.[7] By Weak Hypothesis II, this ranking prevails when evaluated using the local utility at $\delta_{c_{p^*}}$, for $p^* > p$. We then use the following characterization of skewed to the left distributions.

Definition 7 *Let μ be the expected value of a lottery X. Lottery Y is obtained from X by a left symmetric split if Y is the same as X, except for that one of*

[7] More precisely, $u_{\delta_{c_p}}(c_p) \geqslant \frac{1}{2}u_{\delta_{c_p}}(c_{p-\alpha}) + \frac{1}{2}u_{\delta_{c_p}}(c_{p+\alpha})$, for any p.

the outcomes $x \leqslant \mu$ *of* X *is split into* $x + \alpha$ *and* $x - \alpha$*, each with half of the probability of* x*.*

Theorem 2 *If the lottery* $Y = (y_1, p_1; \ldots; y_n, p_n)$ *with expected value* μ *is skewed to the left, then there is a sequence of lotteries* X_i*, each with expected value* μ*, such that* $X_1 = (\mu, 1)$*,* $X_i \to Y$*, and* X_{i+1} *is obtained from* X_i *by a left symmetric split. Conversely, any such sequence converges to a skewed to the left distribution.*

Pick a lottery $\langle p, 1\rangle$. By Theorem 2, any skewed to the left noise Q around p can be obtained as the limit of left symmetric splits. The key observation is that by repeatedly applying Weak Hypothesis II, each such split will be rejected when evaluated using the local utility at δ_{c_p}. By Fréchet differentiability, it will also be rejected by $\succeq$. Quasi concavity then implies that $\langle p, 1\rangle \succeq Q$.

4 Ambiguity aversion and seeking

Ambiguity aversion is one of the most investigated phenomena in decision theory. Consider the classic Ellsberg [18] thought experiment: subjects are presented with two urns. Urn 1 contains 100 red and black balls, but the exact color composition is unknown. Urn 2 has exactly 50 red and 50 black balls in it. Subjects are asked to choose an urn from which a ball will be drawn, and to bet on the color of this ball. If a bet on a specific urn is correct the subject wins \$100, zero otherwise. Let C_i be the bet on a color (Red or Black) draw from Urn i. Ellsberg predicted that most subjects will be indifferent between R_1 and B_1 as well as between R_2 and B_2, but will strictly prefer R_2 to R_1 and B_2 to B_1. While, based on symmetry arguments, it seems plausible that the number of red balls in urn 1 equals the number of black balls, Urn 1 is ambiguous in the sense that the exact distribution is unknown whereas urn 2 is risky, as the probabilities are known. An ambiguity averse decision maker will prefer to bet on the risky urn to bet on the ambiguous one. Ellsberg's predictions were confirmed in many experiments.

The recursive model was suggested by Segal [38] as a way to capture ambiguity aversion.[8] Under this interpretation, ambiguity is identified as two-stage lotteries. The first stage captures the decision maker's uncertainty about the true probability distribution over the states of the world (the true composition of the urn in Ellsberg's example), and the second stage determines the probability of each outcome, conditional on the probability distribution that has been realized. Holding the prior probability distribution over states fixed, an ambiguity averse decision maker prefers the objective (unambiguous) simple lottery to any (ambiguous) compound one. Note that according to Segal's model, preferences over ambiguous prospects are induced from preferences over the compound lotteries that reflect the decision maker's beliefs. That is, the first stage is imaginary and corresponds to the decision maker's subjective beliefs over the values of the true probabilities.

While Ellsberg-type behavior seems intuitive and is widely documented, there are situations where decision makers actually prefer not to know the probabilities with much preciseness. Suppose a person suspects that there is a high probability that he will face a bad outcome (severe loss of money, serious illness, criminal conviction, etc.). Yet he believes that there is some (small) chance things are not as bad as they seem (Federal regulations will prevent the bank from taking possession of his home, it is really nothing, they won't be able to prove it). These beliefs might emerge, for example, from consulting with a number of experts (such as accountants, doctors, lawyers) who disagree in their opinions; the vast majority of which are negative but some believe the risk is much less likely. Does the decision maker really want to know the exact probabilities of these events? The main distinction between the sort of ambiguity in Ellsberg's experiment and the ambiguity in the last examples is that the latter is asymmetric and, in particular, positively skewed. On the other hand, if the decision maker expects a good outcome with high

[8] There are many other ways to model ambiguity aversion. Prominent examples include Choquet expected utility (Schmeidler [37]), maximin expected utility (Gilboa and Schmeidler [21]), variational preferences (Maccheroni, Marinacci, and Rustichini [30]), α-maxmin (Ghirardato, Maccheroni, and Marinacci [20])), and the smooth model of ambiguity aversion (Klibanoff, Marinacci, and Mukerji [28]).

probability, he would probably prefer to know this probability for sure, rather than knowing that there is actually a small chance that things are not that good. In other words, asymmetric but negatively skewed ambiguity may well be undesired.

There is indeed a growing experimental literature that challenges the assumption that people are globally ambiguity averse (see a recent survey by Trautmann and van de Kuilen [43]). A typical finding is that individuals are ambiguity averse for moderate and high likelihood events, but ambiguity seeking for unlikely events. This idea was suggested by Ellsberg himself (see Becker and Bronwson [5, fnt. 4]). In a recent large-scale experiment, Kocher, Lahno, and Trautmann [19] used Ellsberg's two-color design (symmetric ambiguity) and found ambiguity aversion on the domain of moderate likelihood gains. However, they documented either ambiguity neutrality or ambiguity seeking for low likelihood gains, where ambiguity events are implemented as bets on the color of a colored chip drawn from a bag with an unknown distribution of ten different colors. Similar results are reported in Dimmock, Kouwenberg, Mitchell, and Peijnenburg [16]. In their study, individuals were on average indifferent between betting on the event that one out of ten colors is drawn from the ambiguous bag, and a risky lottery with 20 percent known chance of success, showing a significant amount of ambiguity seeking for unlikely events. Camerer and Webber [8] pointed out that such pattern may be due to perceived skewness, which distorts the mean of the ambiguous distributions of high and low probabilities.

Our model is consistent with the co-existence of aversion to both symmetric ambiguity (as in Ellsbergs paradox) and ambiguity seeking for low-probability events. To illustrate, consider (i) a risky urn containing $n > 2$ balls numbered 1 to n, and (ii) an ambiguous urn also containing n balls, each marked by a number from the set $\{1, 2, \ldots, n\}$, but in an unknown composition. Betting that a specific number will *not* be drawn from the risky urn corresponds to the simple lottery with probability of success $\frac{n-1}{n}$. While we dont know what distribution over the composition of the ambiguous urn does the decision maker hold, it is reasonable to invoke symmetry arguments. Let $(m_1, \ldots, m_n)$ be a possible distribution of of the numbers in the ambiguous urn, indicating that number i appears m_i times (of course, $\sum_i m_i = n$). Symmetry arguments require that the decision maker believes that this composition is as likely as any one of its permutations. Unless the decision maker believes that there are *at most* two balls marked with the same number, the same bet over the ambiguous urn corresponds to a compound lottery that induces a negatively skewed distribution around $\frac{n-1}{n}$.[9] The hypotheses of Theorem 1 imply that the bet from the risky urn is preferred.

Consider now the same two urns, but the bet is on a specific number drawn from each of them. The new bet from the risky urn corresponds to the simple lottery with probability of success $\frac{1}{n}$, while the new bet over the ambiguous urn corresponds to to compound lottery that induces a positively skewed distribution around $\frac{1}{n}$. Theorem 1 permits preferences for the ambiguous bet, especially where n is large.

5 Allocation Mechanisms

In this section we apply our results to the comparison of two known allocation mechanisms of indivisible goods. We demonstrate that agents with preferences as studied in this paper may systematically prefer one mechanism to the other, even though both mechanisms are considered to be the same in standard models, in the sense that they induce the same probability distribution over successful matchings.

Consider the following variant of the house allocation problem (Hylland and Zeckhauser [24]). Let $N = \{1, 2, ..., n\}$ be a group of individuals and assume that there are n goods to be allocated among them. The goods are of two types, g_1 and g_2, and we denote by p and $1-p$ the proportion of each type. Some individuals prefer g_1 to g_2 and the rest have the opposite preferences. We assume that the utility from the desired outcome is 1 and the utility from

[9] To see this, let F be the distribution of the decision maker's beliefs. Note that F is non-decreasing and constant on $[\frac{i}{n}, \frac{i+1}{n})$ for $i \leqslant n-1$. Since $\eta_1(F, 0) = \eta_2(F, 0)$ (see Section 3) and $Pr(\frac{n-k}{n}) > 0$ for some k with $n \geqslant k > 2$, it must be that $1 - Pr(1) - Pr(\frac{n-1}{n}) < Pr(1)$, from which the result readily follows.

the other outcome is 0.

Many important goods are allocated using randomizing devices. These include, among others, the allocation of public schools, course schedule, or dormitory rooms to students, and shifts, offices, or tasks to workers. We consider two familiar mechanisms, each consists of two stages.

Random Top Cycle (TC): In the first stage, the allocation of the goods among the agents is randomly determined, so that the probability of person i to hold a unit of type g_1 is p (and $1 - p$ for type g_2). Those who got their desired outcome keep it. In the second stage, the rest will trade according to the following schedule: If k people who hold one type of good and $\ell \leqslant k$ people who hold the other type are unhappy with their holdings, then the latter ℓ will trade and get their desired outcome, while ℓ out of the former k will be selected at random and get their preferred option. The other $k - \ell$ will keep their undesired outcome.[10]

Random Serial Dictatorship (SD): The agents are randomly ordered, so that the probability of person i to be in place $j = 1, \ldots, n$ is $\frac{1}{n}$. Agents then choose the goods according to this order. A person will get his desired outcome if when his turn arrives such a unit is still available.

In practice, the second mechanism is a lot more popular than the first one (see, for example, Sönmez and Ünver [41]). One possible reason is that the SD mechanis is simpler and easier to implement. Another reason is that mechanisms are used when markets fail or are undesired, and the TC mechanism is too close in its spirit to a market environment. We offer here an alternative explanation, based on the preferences analyzed by Theorem 1.

The literature on one-sided and two-sided matching (for a recent survey, see Abdulkadiroğlu and Sönmez [4]) typically maintains the assumption that agents are only interested in the overall probability they'll receive their desired outcome. This leads to equivalence results of different randomized mechanisms

[10]This is a variant of the classic top cycle mechanism. It can, equivalently, be formulated more closely to the familiar top cycle, as a problem of matching with indifferences and using a specific tie-breaking rule. Since the environment we consider is simple, we maintain our formulation and slightly abuse the title "cycle."

(Abdulkadiroğlu and Sönmez [3]; see also Pathak and Sethuraman [34]). In particular, Abdulkadiroğlu and Sönmez's [3] results imply that both TC and SD lead each of the two types of individuals to the same overall probability of success. But as we now show, indifference between the two mechanisms is not preserved when TC is viewed as a two-stage lottery and the analysis of this paper is applied.

In the formal analysis to follow, we confine attention to the case of large (continuum) economies. We assume without loss of generality that p, the proportion of good 1, satisfies $p \geqslant \frac{1}{2}$. Let q be the proportion of individuals who prefers g_1 to g_2. In the TC mechanism we can therefore identify four groups:

1. qp will get g_1 and like it.
2. $(1-q)p$ will get g_1 and will prefer to trade it for g_2.
3. $q(1-p)$ will get g_2 and will prefer to trade it for g_1.
4. $(1-q)(1-p)$ will get g_2 and like it.

Consider first the case $p < q$. The second group is smaller than the third one, and therefore all members of the second group will be able to trade. In other words, all those who prefer g_2 (the second and the fourth group) are guaranteed to receive it. Those who prefer g_1 face a lottery. With probability p they will get their desired outcome, and with probability $1-p$ they will get their desired outcome if they'll be able to trade, the probability of this event is $\frac{(1-q)p}{q(1-p)}$. These people will face the two-stage lottery $\langle 1, p; \frac{(1-q)p}{q(1-p)}, 1-p\rangle$.

In the SD mechanism, those who prefer g_2 are guaranteed to receive it. A person who prefers g_1 will get it only if he is in the top α of the list where $q\alpha \leqslant p$, that is, if his rank is less than $\frac{p}{q}$. If the lottery of this mechanism is performed at the same stage as the trading lottery of the TC mechanism, then this mechanism can be viewed as the (degenerate) two-stage lottery $\langle \frac{p}{q}, 1\rangle$.

We obtain that all those who prefer g_2 to g_1 ($1-q$ of the group) know in advance that since there is an excess supply of their desired good, they will eventually get it under both procedures, and are therefore indifferent between the two mechanisms. Those who prefer g_1 to g_2 have to compare $X_1 = \langle \frac{p}{q}, 1 \rangle$ with $Y_1 = \langle 1, p; \frac{(1-q)p}{q(1-p)}, 1-p \rangle$. Since $p \geqslant \frac{1}{2}$, Y_1 is skewed to the left. If preferences satisfy the assumptions of Theorem 1, then $X_1 \succ Y_1$, and both groups will prefer the SD mechanism to the TC one. suppose for example that $p = \frac{2}{3}$ of the units are of type g_1 and $q = \frac{4}{5}$ of the group prefer it to g_2. It seems indeed plausible that decision makers will prefer to know that their probability of success is $\frac{5}{6}$ rather than to participate in a lottery where with probability $\frac{2}{3}$ they are guaranteed success, but with probability $\frac{1}{3}$ their probability of receiving a desired outcome is only $\frac{1}{2}$.

Consider now the case $p > q$. In the SD mechanism, the q who prefer g_1 are guaranteed to receive it. A person who prefers g_2 will get it only if he is in the top α of the list where $(1-q)\alpha \leqslant 1-p$, that is, if his rank is less than $\frac{1-p}{1-q}$. In the TC mechanism, there are as before four groups. Since $p > q$, the third group is smaller than the second one, and therefore all members of the third group will be able to trade. In other words, all those who prefer g_1 (the first and the third group) are guaranteed to receive it. Those who prefer g_2 face a lottery. With probability $1-p$ they will get their desired outcome, and with probability p they will get their desired outcome if they'll be able to trade, the probability of this event is $\frac{q(1-p)}{(1-q)p}$.

We obtain that all those who prefer g_1 to g_2 (q of the group) know in advance that since there is an excess supply of their desired good, they will eventually get it under both procedures, and are therefore indifferent between the two mechanisms. Those who prefer g_2 to g_1 have to compare $X_2 = \langle \frac{1-p}{1-q}, 1 \rangle$ with $Y_2 = \langle 1, 1-p; \frac{q(1-p)}{(1-q)p}, p \rangle$. This time, Y_2 is skewed to the right. Theorem 1 does not tell us which of the two is better, but using the preferences of Example 1, we know that for large p and small q, TC is preferred to SD.

These results seem plausible. Generally speaking, for those who are on the long side of the market (and are therefore not guaranteed ex ante to receive their desired outcome), the serial dictatorship mechanism is more attractive. But when almost all the available unites are of type g_1 (that is, p is close to 1) but not too many people like it (q is low), the SD mechanism, with its one stage resolution of uncertainty, is a very unattractive lottery. The TC procedure will probably give a person who prefers g_2 a unit of g_1, which he'll find very hard to trade. However, there is also some chance that it will give him his desired outcome right away. From an ex ante perspective, this slim chance for a very good outcome may compensate him for his otherwise aversion to unknown situations.

6 Relations to the literature

In this section we highlight the connection and differences between our definitions and analysis and some other related models.

Recursive Utility: Consider a recursive model in which the decision maker is an expected utility maximizer in each of the two stages, with Bernoulli utility functions u and v. The decision maker evaluates a two-stage lottery $\langle p_1, q_1; ...; p_n, q_n\rangle$ recursively by

$$U\left(\langle p_1, q_1; ...; p_n, q_n\rangle\right) = \sum_i q_i u\left(v^{-1}(\mathrm{E}_v[p_i])\right) \tag{2}$$

where $v^{-1}\left(\mathrm{E}_v[p]\right)$ is the certainty equivalent of lottery p calculated using the function v. In the context of temporal lotteries, this model is a special case of the one studied by Kreps and Porteus [27]. In the context of ambiguity, this is the model of Klibanoff, Marinacci, Mukerji [28]. We now show that this model cannot accommodate rejection of symmetric noise with acceptance of some positively skewed noise.

Proposition 1 *Suppose $\succeq$ admits a representation as in eq. (2). If $\succeq$ rejects symmetric noise, then it rejects all noise.*

Another special form of recursive utility is discussed in Dillenberger [15], who studied a property called preferences for one-shot resolution of uncertainty. In the language of this paper, this property means that the decision maker rejects *all* noise. Dillenberger confined his attention to recursive preferences over two-stage lotteries in which the certainty equivalent functions in the second stage are calculated using the same V that applied in the first stage (this is known as the time neutrality axiom (Segal [39]).

Orders of Risk Aversion: Our analysis in this paper is based on functionals V that are Fréchet differentiable. Segal and Spivak [40] defined a preference relation as exhibiting second-order (resp., first-order) risk aversion if the derivative of the implied risk premium on a small, actuarially fair gamble vanishes (resp., does not vanish) as the size of the gamble converges to zero.[11] If V is Fréchet differentiable, then it satisfies second-order risk aversion. There are interesting preference relations that are not Fréchet differentiable (and that do not satisfy second-order risk aversion), e.g., the rank-dependent utility model of Quiggin [35]. The analysis of attitudes towards skewed noise in the recursive model for general first-order risk aversion preferences is not vacuous, but is significantly different than the one presented in this paper and will be developed in future work.

Skewed Distributions: Our definition of skewed distributions (Def. 4) is stronger than a possible alternative definition, according to which the lottery X with the distribution F and expected value μ is skewed to the left if $\int_{\underline{x}}^{\overline{x}}(y-\mu)^3 dF(y) \leqslant 0$.

Proposition 2 *If X with distribution F and expected value μ is skewed to the left as in Definition 4, then for all odd n, $\int_{\underline{x}}^{\overline{x}}(y-\mu)^n dF(y) \leqslant 0$.*[12]

In addition, our notion of skewed to the left distribution implies the following relationship between the mean and the median, which does not necessarily hold under the assumption of negative third moment above.

[11] Formally, if $\pi(t)$ is the amount of money that an agent would pay to avoid the non-degenerate gamble $x+t\widetilde{\varepsilon}$, where $\mathrm{E}(\widetilde{\varepsilon})=0$, then $\pi(t)$ is $O(t)$ and $o(t)$ for first and second-order risk averse preferences, respectively.

[12] The converse of Proposition 2 is false. For example, let F be the distribution of the lottery $(-10, \frac{1}{10}; -2, \frac{1}{2}; 0, \frac{4}{35}; 7, \frac{2}{7})$. Note that its expected value μ is zero. Moreover, $\mathrm{E}\left[(X-\mu)^3\right] = -6 < 0$ and $\mathrm{E}\left[(X-\mu)^{2n+1}\right]$ is decreasing with n, which means that all odd moments of F are negative. Nevertheless, the area below the distribution from -10 to -5 is $\frac{1}{2}$, but the area above the distribution from 5 to 10 is $\frac{4}{7} > \frac{1}{2}$, which means that F is not skewed to the left according to Definition 4.

Proposition 3 *If X with distribution F and expected value μ is skewed to the left, then the highest median $\bar{m}(X)$ of X satisfies $\bar{m}(X) \geqslant \mu$.*

Increasing Downside Risk: Menezes, Geiss, and Tressler [32] characterize a notion of increasing downside risk by combining a mean-preserving spread of an outcome below the mean followed by a mean-preserving contraction of an outcome above the mean, in a way that the overall result is a transfer of risk from the right to the left of a distribution, keeping the variance intact. Distribution F has more downsize risk than distribution G if one can move from G to F in a sequence of such mean-variance-preserving transformations. Menezes et al. [32] do not provide a definition (and a characterization as in our Theorem 2) of a skewed to the left distribution. Observe that our characterization involves a sequence of only symmetric left splits, starting in the degenerate lottery that puts all the mass on the mean. In particular, our splits are not mean-variance-preserving and occur only in one side of the mean.

Appendix A: Proofs

Proof of Theorem 1: Let c_p be the certainty equivalent of the lottery $(\overline{x}, p; \underline{x}, 1-p)$. The two-stage lottery $\langle p-\alpha, \varepsilon; p, 1-2\varepsilon; p+\alpha, \varepsilon\rangle$ translates in the recursive model into the lottery $(c_{p-\alpha}, \varepsilon; c_p, 1-2\varepsilon; c_{p+\alpha}, \varepsilon)$. Since the decision maker always rejects symmetric noise, it follows that the local utility $u_{\delta_{c_p}}$ satisfies

$$u_{\delta_{c_p}}(c_p) \geqslant \tfrac{1}{2}u_{\delta_{c_p}}(c_{p-\alpha}) + \tfrac{1}{2}u_{\delta_{c_p}}(c_{p+\alpha}).$$

By Weak Hypothesis II, for every $r \geqslant p$,

$$u_{\delta_{c_r}}(c_p) \geqslant \tfrac{1}{2}u_{\delta_{c_r}}(c_{p-\alpha}) + \tfrac{1}{2}u_{\delta_{c_r}}(c_{p+\alpha}). \tag{3}$$

Consider first the lottery over the probabilities given by $Q = \langle p_1, q_1; \ldots; p_m, q_m\rangle$ where $\sum q_i p_i = p$ (we deal with the distributions with non-finite support at the end of the proof) . If Q is skewed to the left, then it follows by Theorem 2 that there is a sequence of lotteries $Q_i = \langle p_{i,1}, q_{i,1}, \ldots, p_{i,n_i}, q_{i,n_i}\rangle \to Q$ such that $Q_1 = \langle p, 1\rangle$ and for all i, Q_{i+1} is obtained from Q_i by a left symmetric split. For each i, let $\tilde{Q}_i = (c_{p_{i,1}}, q_{i,1}; \ldots; c_{p_{i,n_i}}, q_{i,n_i})$. Suppose $p_{i,j}$ is split into $p_{i,j} - \alpha$ and $p_{i,j} + \alpha$. By eq. (3), as $p > p_{i,j}$,

$$\begin{aligned}
&\mathrm{E}[u_{\delta_p}(\tilde{Q}_i)] = \\
&q_{i,j}u_{\delta_{c_p}}(c_{p_{i,j}}) + \sum_{m \neq j} q_{i,m}u_{\delta_{c_p}}(c_{p_{i,m}}) \geqslant \\
&\tfrac{1}{2}q_{i,j}u_{\delta_{c_p}}(c_{p_{i,j}-\alpha}) + \tfrac{1}{2}q_{i,j}u_{\delta_{c_p}}(c_{p_{i,j}+\alpha}) + \sum_{m \neq j} q_{i,m}u_{\delta_{c_p}}(c_{p_{i,m}}) = \\
&\mathrm{E}[u_{\delta_{c_p}}(\tilde{Q}_{i+1})].
\end{aligned}$$

As $Q_i \to Q$, and as for all i, $u_{\delta_{c_p}}(c_p) \geqslant \mathrm{E}[u_{\delta_p}(\tilde{Q}_i)]$, it follows by continuity that $u_{\delta_{c_p}}(c_p) \geqslant \mathrm{E}[u_{\delta_{c_p}}(\tilde{Q})]$. By Fréchet Differentiability

$$\frac{\partial}{\partial\varepsilon} V\left(\varepsilon\tilde{Q} + (1-\varepsilon)\delta_{c_p}\right)\bigg|_{\varepsilon=0} \leqslant 0.$$

Quasi-concavity now implies that $V(\delta_{c_p}) \geqslant V(\tilde{Q})$, or $\langle p, 1\rangle \succeq$ Q. Finally, as preferences are continuous, it follows by that the theorem holds for all Q, even if its support is not finite (see remark 1 at the end of the proof of Theorem 2).

For the second part of the theorem, we show in Appendix B that the functional form in Example 1 satisfies all the assumptions of the theorem and always accepts some positively skewed noise. ■

Proof of Theorem 2: The main difficulty in proving the first part of this theorem is the fact that whereas outcomes to the left of μ can be manipulated, any split that lands an outcome to the right of μ must hit its exact place according to Y, as we will not be able to touch it later again.

Lemma 1 proves part *1* of the theorem for binary lotteries Y. After a preparatory claim (Lemma 2), the general case of this part is proved in Lemma 3 for lotteries Y with $F_Y(\mu) \geqslant \frac{1}{2}$, and for all lotteries in Lemma 4. That this can be done with bounded shifts is proved in Lemma 5. Part *2* of the theorem is proved in Lemma 6.

Lemma 1 *Let $Y = (x, r; z, 1-r)$ with mean $\mathrm{E}[Y] = \mu$, $x < z$, and $r \leqslant \frac{1}{2}$. Then there is a sequence of lotteries X_i with expected value μ such that $X_1 = (\mu, 1)$, $X_i \to Y$, and X_{i+1} is obtained from X_i by a left symmetric split. Moreover, if r_i and r'_i are the probabilities of x and z in X_i, then $r_i \uparrow r$ and $r'_i \uparrow 1-r$.*

Proof: The main idea of the proof is to have at each step at most five outcomes: x, μ, z, and up to two outcomes between x and μ. In a typical move either μ or one of the outcomes between x and μ, denote it w, is split "as far as possible," which means:

1. If $w \in (x, \frac{x+\mu}{2}]$, then split its probability between x and $w + (w - x) = 2w - x$. Observe that $x < 2w - x \leqslant \mu$.

2. If $w \in [\frac{x+z}{2}, \mu]$, then split its probability between z and $w - (z - w) = 2w - z$. Observe that $x \leqslant 2w - z < \mu$.

3. If $w \in (\frac{x+\mu}{2}, \frac{x+z}{2})$, then split its probability between μ and $w - (\mu - w) = 2w - \mu$. Observe that $x < 2w - \mu < \mu$.

If $r = \frac{1}{2}$, that is, if $\mu = \frac{x+z}{2}$ then the sequence terminates after the first split. We will therefore assume that $r < \frac{1}{2}$. Observe that the this procedures never split the probabilities of x and z hence these probabilities form increasing sequences. We identify and analyze three cases: *a*. For every i the support of X_i is $\{x, y_i, z\}$. *b*. There is $k > 1$ such that the support of X_k is $\{x, \mu, z\}$. *c*. Case *b* does not happen, but there is $k > 1$ such that the support of X_k is $\{x, w_k, \mu, z\}$. We also show that if for all $i > 1$, μ is not in the support of X_i, then case *a* prevails.

a. The simplest case is when for every i the support of X_i has three outcomes at most, $x < y_i < z$. By construction, the probability of y_i is $\frac{1}{2^i}$, hence X_i puts $1 - \frac{1}{2^i}$ probability on x and z. In the limit these converge to a lottery over x and z only, and since for every i, $\mathrm{E}[X_i] = \mu$, this limit must be Y. For the former, let $X = (3, 1)$ and $Y = (0, \frac{1}{4}; 4, \frac{3}{4})$ and obtain

$$X = (3, 1) \to (2, \tfrac{1}{2}; 4, \tfrac{1}{2}) \to (0, \tfrac{1}{4}; 4, \tfrac{1}{4} + \tfrac{1}{2}) = Y.$$

For a sequence that does not terminate, let $X = (5, 1)$ and $Y = (0, \frac{1}{6}; 6, \frac{5}{6})$. Here we obtain

$$\begin{aligned} X &= (5, 1) \to (4, \tfrac{1}{2}; 6, \tfrac{1}{2}) \to (2, \tfrac{1}{4}; 6, \tfrac{3}{4}) \to (0, \tfrac{1}{8}; 4, \tfrac{1}{8}; 6, \tfrac{3}{4}) \to \dots \\ & (0, \tfrac{1}{2} \textstyle\sum_1^n \tfrac{1}{4^i}; 4, \tfrac{1}{2 \cdot 4^n}; 6, \tfrac{1}{2} + \textstyle\sum_1^n \tfrac{1}{4^i}) \to \dots (0, \tfrac{1}{6}; 6, \tfrac{5}{6}) = Y. \end{aligned}$$

b. Suppose now that even though at a certain step the obtained lottery has more than three outcomes, it is nevertheless the case that after k splits we reach a lottery of the form $X_k = (x, p_k; \mu, q_k;\ z, 1 - p_k - q_k)$. For example, let $X = (17, 1)$ and $Y = (24, \frac{17}{24}; 0, \frac{7}{24})$. The first five splits are

$$\begin{aligned} X &= (17, 1) \to (10, \tfrac{1}{2}; 24, \tfrac{1}{2}) \to (3, \tfrac{1}{4}; 17, \tfrac{1}{4}; 24, \tfrac{1}{2}) \to \\ & (0, \tfrac{1}{8}; 6, \tfrac{1}{8}; 17, \tfrac{1}{4}; 24, \tfrac{1}{2}) \to (0, \tfrac{3}{16}; 12, \tfrac{1}{16}; 17, \tfrac{1}{4}; 24, \tfrac{1}{2}) \to \\ & (0, \tfrac{7}{32}; 17, \tfrac{1}{4}; 24, \tfrac{17}{32}) \end{aligned} \quad (4)$$

By construction $k \geqslant 2$ and $q_k \leqslant \frac{1}{4}$. Repeating these k steps j times will yield the lottery $X_{jk} = (x, p_{jk}; \mu, q_{jk}; z, 1 - p_{jk} - q_{jk}) \to Y$ as $q_{jk} \to 0$ and as the expected value of all lotteries is μ, $p_{jk} \uparrow r$ and $1 - p_{jk} - q_{jk} \uparrow 1 - r$.

c. If at each stage X_i puts no probability on μ then we are in case *a*. The reason is that as splits of type 3 do not happen, in each stage the probability of the outcome between x and z is split between a new such outcome and either x or z, and the number of different outcomes is still no more than three. Suppose therefore that at each stage X_i puts positive probability on at least one outcome w strictly between x and μ (although these outcomes w may change from one lottery X_i to another) and at some stage X_i puts (again) positive probability on μ. Let $k \geqslant 2$ be the first split that puts positive probability on μ. We consider two cases.

c_1. $k = 2$: In the first step, the probability of μ is divided between z and $2\mu - z$ and in the second step the probability of $2\mu - z$ is split and half of it is shifted back to μ (see for example the second split in eq. (4) above). In other words, the first split is of type 2 while the second is of type 3. By the description of the latter,

$$\frac{x+\mu}{2} < 2\mu - z < \frac{x+z}{2} \Longleftrightarrow \frac{2}{3} < \frac{\mu - x}{z - x} < \frac{3}{4} \tag{5}$$

The other one quarter of the original probability of μ is shifted from $2\mu - z$ to

$$2\mu - z - (\mu - [2\mu - z]) = 3\mu - 2z \leqslant \frac{x+\mu}{2} \Longleftrightarrow 4(z-x) \geqslant 5(\mu - x)$$

Which is satisfied by eq. (5). Therefore, in the next step a split of type 1 will be used, and one eighth of the original probability of μ will be shifted away from $2\mu - z$ to x. In other words, in three steps $\frac{5}{8}$ of the original probability of μ is shifted to x and z, one quarter of it is back at μ, and one eighth of it is now on an outcome $w_1 < \mu$. ◇

c_2. $k \geqslant 3$: For example, $X = (29, 1)$ and $Y = (48, \frac{29}{48}; 0, \frac{19}{48})$. Then

$$\begin{aligned} X &= (29,1) \to (10, \tfrac{1}{2}; 48, \tfrac{1}{2}) \to (0, \tfrac{1}{4}; 20, \tfrac{1}{4}; 48, \tfrac{1}{2}) \to \\ &\quad (0, \tfrac{1}{4}; 11, \tfrac{1}{8}; 29, \tfrac{1}{8}; 48, \tfrac{1}{2}) \to \ldots \end{aligned} \tag{6}$$

After k splits $\frac{1}{2^k}$ of the original probability of μ is shifted back to μ and $\frac{1}{2^k}$ is shifted to another outcome $w_1 < \mu$. The rest of the original probability is split (not necessarily equally) between x and z. $\diamond$

Let $\ell = \max\{k, 3\}$. We now construct inductively a sequence of cycles, where the length of cycle j is $\ell + j - 1$. Such a cycle will end with the probability distributed over $x < w_j < \mu < z$. Denote the probability of μ by p_j and that of w_j by q_j. We show that $p_j + q_j \to 0$. The probabilities of x and z are such that the expected value is kept at μ, and as $p_j + q_j \to 0$, it will follow that the probabilities of x and z go up to r and $1 - r$, respectively. In the example of eq. (6), $\ell = 3$, the length of the first cycle (where $j = 1$) is 3, and $w_1 = 11$.

Suppose that we've finished the first j cycles. Cycle $j + 1$ starts with splitting the p_j probability of μ to $\{x, w_1, \mu, z\}$ as in the first cycle. One of the outcomes along this sequence may be w_j, but we will continue to split only the "new" probability of this outcome (and will not yet touch the probability q_j of w_j). At the end of this part of the new cycle, the probability is distributed over x, w_1, w_j, μ, and z. At least half of p_j, the earlier probability of μ, is shifted to $\{x, z\}$, and the probabilities of both these outcomes did not decrease. Continuing the example of eq. (6), the first part of the second cycle (where $j = 1$) is

$$(0, \tfrac{1}{4}; 11, \tfrac{1}{8}; 29, \tfrac{1}{8}; 48, \tfrac{1}{2}) \to (0, \tfrac{1}{4}; 10, \tfrac{1}{16}; 11, \tfrac{1}{8}; 48, \tfrac{9}{16}) \to$$
$$(0, \tfrac{9}{32}; 11, \tfrac{1}{8}; 20, \tfrac{1}{32}; 48, \tfrac{9}{16}) \to (0, \tfrac{9}{32}; 11, \tfrac{9}{64}; 29, \tfrac{1}{64}; 48, \tfrac{9}{16})$$

The second part of cycle $j + 1$ begins with $j - 1$ splits starting with w_1. At the end of these steps, the probability is spread over x, w_j, μ, and z. Split the probability of w_j between an element of $\{x, \mu, z\}$ and w_{j+1} which is not in this set to get p_{j+1} and q_{j+1}. In the above example, as $j = 1$ there is only one split at this stage to

$$(0, \tfrac{45}{128}; 22, \tfrac{9}{128}; 29, \tfrac{1}{64}; 48, \tfrac{9}{16})$$

And $w_2 = 22$. The first part of the third cycle ($j = 2$) leads to

$$(0, \tfrac{91}{256}; 11, \tfrac{1}{512}; 22, \tfrac{9}{128}; 29, \tfrac{1}{512}; 48, \tfrac{73}{128})$$

The second part of this cycle has two splits. Of $w_1 = 11$ into 0 and 22, and then of $w_2 = 22$ into $\mu = 29$ and $w_3 = 15$.

$$\to (0, \tfrac{365}{1024}; 22, \tfrac{73}{1024}; 29, \tfrac{1}{512}; 48, \tfrac{73}{128}) \to (0, \tfrac{365}{1024}; 15, \tfrac{73}{2048}; 29, \tfrac{77}{2048}; 48, \tfrac{73}{128})$$

We now show that for every j,

$$p_{j+2} + q_{j+2} \leqslant \tfrac{3}{4}(p_j + q_j) \tag{7}$$

We first observe that for every j, $p_{j+1} + q_{j+1} < p_j + q_j$. This is due to the fact that the rest of the probability is spread over x and z, the probability of z must increase (because of the initial split in the probability of μ), and the probabilities of x and z cannot go down.

When moving from (p_j, q_j) to (p_{j+2}, q_{j+2}), half of p_j is switched to z. Later on, half of q_j is switched either to x or z, or to μ, in which case half of it (that is, one quarter of q_j) will be switched to z on the move from p_{j+1} to p_{j+2}. This proves inequality (7), hence the lemma. □

Lemma 2 *Let $X = (x_1, p_1; \ldots; x_n, p_n)$ and $Y = (y_1, q_1; \ldots; y_m, q_m)$ where $x_1 \leqslant \ldots \leqslant x_n$ and $y_1 \leqslant \ldots \leqslant y_m$ be two lotteries such that X dominates Y by second-order stochastic dominance. Then there is a sequence of lotteries X_i such that $X_1 = X$, $X_i \to Y$, X_{i+1} is obtained from X_i by a symmetric (not necessarily always left or always right) split of one of the outcomes of X_i, all the outcomes of X_i are between y_1 and y_m, and the probabilities the lotteries X_i put on y_1 and y_m go up to q_1 and q_m, respectively.*

Proof: From Rothschild and Stiglitz [36, p. 236] we know that we can present Y as $(y_{11}, q_{11}; \ldots; y_{nn}, q_{nn})$ such that $\sum_j q_{kj} = p_k$ and $\sum_j q_{kj} y_{kj}/p_k = x_k$, $k = 1, \ldots, n$.

Let $Z = (z_1, r_1; \ldots; z_\ell, r_\ell)$ such that $z_1 < \ldots < z_\ell$ and $\mathrm{E}[Z] = z$. Let $Z_0 = (z, 1)$. One can move from Z_0 to Z in at most ℓ steps, where at each step some of the probability of z is split into two outcomes of Z without affecting the expected value of the lottery, in the following way. If

$$\frac{r_1 z_1 + r_\ell z_\ell}{r_1 + r_\ell} \geqslant z \tag{8}$$

then move r_1 probability to z_1 and $r'_\ell \leqslant r_\ell$ to z_ℓ such that $r_1 z_1 + r'_\ell z_\ell = z(r_1 + r'_\ell)$. However, if the sign of the inequality in (8) is reversed, then move r_ℓ probability to z_ℓ and $r'_1 \leqslant r_1$ probability to z_ℓ such that $r'_1 z_1 + r_\ell z_\ell = z(r'_1 + r_\ell)$. Either way the move shifted all the required probability from z to one of the outcomes of Z without changing the expected value of the lottery.

Consequently, one can move from X to Y in ℓ^2 steps, where at each step some probability of an outcomes of X is split between two outcomes of Y. By Lemma 1, each such split can be obtained as the limit of symmetric splits (recall that we do not require in the current lemma that the symmetric splits will be left or right splits). That all the outcomes of the obtained lotteries are between y_1 and y_m, and that the probabilities these put on y_1 and y_m go up to q_1 and q_m follow by Lemma 1. □

Lemma 3 *Let $Y = (y_1, p_1; \ldots; y_n, p_n)$, $y_1 \leqslant \ldots \leqslant y_n$, with expected value μ be skewed to the left such that $F_Y(\mu) \geqslant \frac{1}{2}$. Then there is a sequence of lotteries X_i with expected value μ such that $X_1 = (\mu, 1)$, $X_i \to Y$, and X_{i+1} is obtained from X_i by a left symmetric split. Moreover, if r_i and r'_i are the probabilities of y_1 and y_n in X_i, then $r_i \uparrow p_1$ and $r'_i \uparrow p_n$.*

Proof: Suppose wlg that $y_{j^*} = \mu$ (of course, it may be that $p_{j^*} = 0$). Since $F_Y(y_{j^*}) \geqslant \frac{1}{2}$, it follows that $t := \sum_{j=j^*+1}^{n} p_j \leqslant \frac{1}{2}$. As Y is skewed to the left, $y_n - \mu \leqslant \mu - y_1$, hence $2\mu - y_n \geqslant y_1$. Let $m = n - j^*$ be the number of outcomes of Y that are strictly above the expected value μ. Move from X_1 to $X_m = (2\mu - y_n, p_n; \ldots; 2\mu - y_{j^*+1}, p_{j^*+1}; y_{j^*}, 1 - 2t; y_{j^*+1}, p_{j^*+1}; \ldots; y_n, p_n)$ by repeatedly splitting probabilities away from μ. All these splits are symmetric, hence left symmetric splits.

Next we show that Y is a mean preserving spread of X_m. Obviously, $\mathrm{E}[X_m] = \mathrm{E}[Y] = \mu$. Integrating by parts, we have for $x \geqslant \mu$

$$y_n - \mu = \int_{y_1}^{y_n} F_Y(z)dz = \int_{y_1}^{x} F_Y(z)dz + \int_{x}^{y_n} F_Y(z)dz$$
$$y_n - \mu = \int_{y_1}^{y_n} F_{X_m}(z)dz = \int_{y_1}^{x} F_{X_m}(z)dz + \int_{x}^{y_n} F_{X_m}(z)dz$$

Since F_Y and F_{X_m} coincide for $z \geqslant \mu$, we have, for $x \geqslant \mu$, $\int_{y_1}^{x} F_{X_m}(z)dz = \int_{y_1}^{x} F_Y(z)dz$ and in particular, $\int_{y_1}^{x} F_{X_m}(z)dz \leqslant \int_{y_1}^{x} F_Y(z)dz$.

For $x < \mu$ it follows by the assumption that Y is skewed to the left and by the construction of X_m as a symmetric lottery around μ that

$$\begin{aligned}\int_{y_1}^{x} F_{X_m}(z)dz &= \int_{2\mu-x}^{2\mu-y_1} [1 - F_{X_m}(z)]dz = \\ & \int_{2\mu-x}^{2\mu-y_1} [1 - F_Y(z)]dz \leqslant \int_{y_1}^{x} F_Y(z)dz\end{aligned}$$

Since to the right of μ, X_m and Y coincide, we can view the left side of Y as a mean preserving spread of the left side of X_m. By Lemma 2 the left side of Y is the limit of symmetric mean preserving spreads of the left side of X_m. Moreover, all these splits take place between y_1 and μ and are therefore left symmetric splits. By Lemma 2 it also follows that $r_i \uparrow p_1$ and $r'_i \uparrow p_n$. □

We now show that Lemma 3 holds without the restriction $F_Y(\mu) \geqslant \frac{1}{2}$.

Lemma 4 *Let Y with expected value μ be skewed to the left. Then there is a sequence of lotteries X_i with expected value μ such that $X_1 = (\mu, 1)$, $X_i \to Y$, and X_{i+1} is obtained from X_i by a left symmetric split.*

Proof: The first step in the proof of Lemma 3 was to create a symmetric distribution around μ such that its upper tail (above μ) agrees with F_Y. Obviously this can be done only if $F_Y(\mu) \geqslant \frac{1}{2}$, which is no longer assumed. Instead, we apply the proof of Lemma 3 successively to mixtures of F_Y and δ_μ, the distribution that yields μ with probability one.

Suppose that $F_Y(\mu) = \lambda < \frac{1}{2}$. Let $\gamma = 1/2(1-\lambda)$ and define Z to be the lottery obtained from the distribution $\gamma F_Y + (1-\gamma)\delta_\mu$. Observe that

$$F_Z(\mu) = \gamma F_Y(\mu) + (1-\gamma)\delta_\mu(\mu) = \frac{\lambda}{2(1-\lambda)} + \frac{1-2\lambda}{2(1-\lambda)} = \frac{1}{2}$$

It follows that the lotteries Z and $(\mu, 1)$ satisfy the conditions of Lemma 3, and therefore there is sequence of lotteries X_i with expected value μ such that $X_1 = (\mu, 1)$, $X_i \to Z$, and X_{i+1} is obtained from X_i by a left symmetric split. This is done in two stages. First we create a symmetric distribution around μ that agrees with Z above μ (denote the number of splits needed in this stage by t), and then we manipulate the part of the distribution which is weakly to the left of μ by taking successive symmetric splits (which are all left symmetric splits when related to μ) to get nearer and nearer to the second-order stochastically dominated left side of Z as in Lemma 2. Observe that the highest outcome of this part of the distribution Z is μ, and its probability is $1-\gamma$. By Lemma 2, for every $k \geqslant 1$ there is ℓ_k such that after ℓ_k splits of this second phase the probability of μ will be at least $r_k = (1-\gamma)(1-\frac{1}{k+1})$ and $\| X_{t+\ell_k} - Z \| < \frac{1}{k}$.

The first cycle will end after $t+\ell_1$ splits with the distribution F_{Z_1}. Observe that the probabilities of the outcomes to the right of μ in Z_1 are those of the lottery Y multiplied by γ. The first part of second cycle will be the same as the first cycle, applied to the r_k conditional probability of μ. At the end of this part we'll get the lottery Z_1' which is the same as Z_1, conditional on the probability of μ. We now continue the second cycle by splitting the combination of Z_1 and Z_1' for the total of $t + \ell_1 + \ell_2$ steps. As we continue to add such cycles inductively we get closer and closer to Y, hence the lemma. □

Next we show that part *1* of the theorem can be achieved by using bounded spreads. The first steps in the proof of Lemma 3 involve shifting probabilities from μ to all the outcomes of Y to the right of μ, and these outcomes are not more than $\max y_i - \mu$ away from μ. All other shifts are symmetric shifts involving only outcomes to the left of μ. The next lemma shows that such shifts can be achieved as the limit of symmetric bounded shifts.

Lemma 5 *Let $Z = (z-\alpha, \frac{1}{2}; z+\alpha, \frac{1}{2})$ and let $\varepsilon > 0$. Then there is a sequence of lotteries Z_i such that $Z_0 = (z, 1)$, $Z_i \to Z$, and Z_{i+1} is obtained from Z_i by a symmetric (not necessarily left or right) split of size smaller than ε.*

Proof: The claim is interesting only when $\varepsilon < \alpha$. Fix n such that $\varepsilon > \frac{\alpha}{n}$. We show that the lemma can be proved by choosing the size of the splits to be $\frac{\alpha}{n}$. Consider the $2n+1$ points $z_k = z + \frac{k}{n}$, $k = -n, \ldots, n$ and construct the sequence $\{Z_i\}$ where $Z_i = (z-\alpha, p_{i,-n}; z - \frac{n-1}{n}\alpha, p_{i,-n+1}; \ldots; z+\alpha, p_{i,n})$ as follows.

THE INDEX i IS ODD: Let z_j be the highest outcome in $\{z, \ldots, z + \frac{n-1}{n}\alpha\}$ with the highest probability in Z_{i-1}. Formally, j satisfies:

- $0 \leqslant j \leqslant n-1$
- $p_{i-1,j} \geqslant p_{i-1,k}$ for all k
- If for some $j' \in \{0, \ldots, n-1\}$, $p_{i-1,j'} \geqslant p_{i-1,k}$ for all k, then $j \geqslant j'$.

Split the probability of z_j between $z_j - \frac{\alpha}{n}$ and $z_j + \frac{\alpha}{n}$ (i.e., between z_{j-1} and z_{j+1}). That is, $p_{i,j-1} = p_{i-1,j-1} + \frac{1}{2}p_{i-1,j}$, $p_{i,j+1} = p_{i-1,j+1} + \frac{1}{2}p_{i-1,j}$, $p_{i,j} = 0$, and for all $k \neq j-1, j, j+1$, $p_{i,k} = p_{i-1,k}$.

THE INDEX i IS EVEN: In this step we create the mirror split of the one done in the previous step. Formally, If j of the previous stage is zero, do nothing. Otherwise, split the probability of z_{-j} between $z_{-j} - \frac{\alpha}{n}$ and $z_{-j} + \frac{\alpha}{n}$. That is, $p_{i,-j-1} = p_{i-1,-j-1} + \frac{1}{2}p_{i-1,-j}$, $p_{i,-j+1} = p_{i-1,-j+1} + \frac{1}{2}p_{i-1,-j}$, $p_{i,-j} = 0$, and for all $k \neq -j-1, -j, -j+1$, $p_{i,k} = p_{i-1,k}$.

After each pair of these steps, the probability distribution is symmetric around z. Also, the sequences $\{p_{i,-n}\}_i$ and $\{p_{i,n}\}$ are non decreasing. Being bounded by $\frac{1}{2}$, they converge to a limit L. Our aim is to show that $L = \frac{1}{2}$. Suppose not. Then at each step the highest probability of $\{p_{i-1,-n+1}, \ldots, p_{i-1,n-1}$ must be at least $\ell := (1-2L)/(2n-1) > 0$. The variance of Z_i is bounded from above by the variance of $(\mu - \alpha, \frac{1}{2}; \mu + \alpha, \frac{1}{2})$, which is α^2. Splitting p probability from z to $z - \frac{\alpha}{n}$ and $z + \frac{\alpha}{n}$ will increase the variance by $p(\frac{\alpha}{n})^2$. Likewise, for $k \neq -n, 0, n$, splitting p probability from $z + \frac{k\alpha}{n}$ to $z + \frac{(k+1)\alpha}{n}$ and

$z - \frac{(k-1)\alpha}{n}$ will increase the variance by $\frac{p}{2}(\frac{\alpha}{n})^2$. Therefore, for positive even i we have

$$\sigma^2(Z_i) - \sigma^2(Z_{i-2}) \geqslant \frac{1-2L}{2n-1}\left(\frac{\alpha}{n}\right)^2$$

If $L < \frac{1}{2}$, then after enough steps the variance of Z_i will exceed α^2, a contradiction. □

That we can do the theorem for all lotteries Y follows by the fact that a countable set of countable sequences is countable. To finish the proof of the theorem, we need the following result:

Lemma 6 *Consider the sequence $\{X_i\}$ of lotteries where $X_1 = (\mu, 1)$ and X_{i+1} is obtained from X_i by a left symmetric split. Then the distributions F_i of X_i converge (in the L^1 topology) to a skewed to the left distribution with expected value μ.*

Proof: That such sequences converge follows from the fact that a symmetric split will increase the variance of the distribution, but as all distributions are over the bounded $[\underline{x}, \overline{x}]$ segment of $\Re$, the variances of the distributions increase to a limit. Replacing (x, p) with $(x - \alpha, \frac{p}{2}; x + \alpha, \frac{p}{2})$ increases the variance of the distribution by

$$\frac{p}{2}(x - \alpha - \mu)^2 + \frac{p}{2}(x + \alpha - \mu)^2 - p(x - \mu)^2 = p\alpha^2$$

and therefore the distance between two successive distributions in the sequences in bounded by $\overline{x} - \underline{x}$ times the change in the variance. The sum of the changes in the variances is bounded, as is therefore the sum of distances between successive distributions, hence Cauchy criterion is satisfied and the sequence converges.

Next we prove that the limit is a skewed to the left distribution with expected value μ. Let F be the distribution of $X = (x_1, p_1; \ldots; x_n, p_n)$ with expected value μ be skewed to the left. Suppose wlg that $x_1 \leqslant \mu$, and break it symmetrically to obtain $X' = (x_1 - \alpha, \frac{p_1}{2}; x_1 + \alpha, \frac{p_1}{2}; x_2, p_2; \ldots; x_n, p_n)$ with the distribution F'. Note that $\mathrm{E}[X'] = \mu$. Consider the following two cases.

Case 1: $x_1+\alpha \leqslant \mu$. Then for all δ, $\eta_2(F,\delta) = \eta_2(F',\delta)$. For δ such that $\mu-\delta \leqslant x_1-\alpha$ or such that $x_1+\alpha \leqslant \mu-\delta$, $\eta_1(F',\delta) = \eta_1(F,\delta) \geqslant \eta_2(F,\delta) = \eta_2(F',\delta)$. For δ such that $x_1-\alpha < \mu-\delta \leqslant x_1$, $\eta_1(F',\delta) = \eta_1(F,\delta)+[(\mu-\delta)-(x_1-\alpha)]\frac{p_1}{2} > \eta_1(F,\delta) \geqslant \eta_2(F,\delta) = \eta_2(F,\delta)$. Finally, for δ such that $x_1 < \mu-\delta < x_1+\alpha\,(\leqslant \mu)$, $\eta_1(F',\delta) = \eta_1(F,\delta)+[(x_1+\alpha)-(\mu-\delta)]\frac{p_1}{2} > \eta_1(F,\delta) \geqslant \eta_2(F,\delta) = \eta_2(F',\delta)$.

Case 2: $x_1+\alpha > \mu$. Then for all δ such that $\mu+\delta \geqslant x_1+\alpha$, $\eta_2(F,\delta) = \eta_2(F',\delta)$. For δ such that $\mu-\delta \leqslant x_1-\alpha$, $\eta_1(F',\delta) = \eta_1(F,\delta) \geqslant \eta_2(F,\delta) = \eta_2(F',\delta)$. For δ such that $x_1-\alpha < \mu-\delta \leqslant x_1$, $\eta_1(F',\delta) = \eta_1(F,\delta)+[(\mu-\delta)-(x_1-\alpha)]\frac{p_1}{2} \geqslant \eta_2(F,\delta)+[(\mu-\delta)-(x_1-\alpha)]\frac{p_1}{2} \geqslant \eta_2(F,\delta)+\max\{0,(x_1+\alpha)-(\mu+\delta)\}\frac{p_1}{2} = \eta_2(F',\delta)$. Finally, for δ such that $\mu-\delta > x_1$, $\eta_1(F',\delta) = \eta_1(F,\delta)+[(x_1+\alpha)-(\mu-\delta)]\frac{p_1}{2} \geqslant \eta_2(F,\delta)+\max\{0,(x_1+\alpha)-(\mu+\delta)\}\frac{p_1}{2} = \eta_2(F',\delta)$.

If $X_n \to Y$, all have the same expected value and for all n, X_n is skewed to the left, then so is Y. ■

Remark 1 The two parts of Theorem 2 do not create a simple if and only if statement, because the support of the limit distribution F in part *2* need not be finite. On the other hand, part *1* of the theorem does not hold for continuous distributions. By the definition of left symmetric splits, if the probability of $x > \mu$ in X_i is p, then for all $j > i$, the probability of x in X_j must be at least p. It thus follows that the distribution F cannot be continuous above μ. However, it can be shown that if F with expected value μ is skewed to the left, then there is a sequence of finite skewed to the left distributions F_n, each with expected value μ, such that $F_n \to F$. This enables us to use Theorem 2 even for continuous distributions.

Proof of Proposition 1: Note that by eq. (2), for any p, the value of the noise $\langle p+\alpha, \frac{1}{2}; p-\alpha, \frac{1}{2}\rangle$ is

$$\begin{aligned}&\tfrac{1}{2}u\left(v^{-1}[(p+\alpha)v(x)+(1-p-\alpha)v(y)]\right)+\\&\qquad\tfrac{1}{2}u\left(v^{-1}[(p-\alpha)v(x)+(1-p+\alpha)v(y)]\right).\end{aligned}$$

The value of the simple lottery $(\overline{x}, p; \underline{x}, 1-p)$ is

$$u\left(v^{-1}[pv(\overline{x}) + (1-p)v(\underline{x})]\right).$$

Rejection of symmetric noise implies that for any p and α in the relevant range,

$$\begin{aligned} & u\left(v^{-1}[pv(\overline{x}) + (1-p)v(\underline{x})]\right) \geqslant \\ & \quad \tfrac{1}{2}u\left(v^{-1}[(p+\alpha)v(\overline{x}) + (1-p-\alpha)v(\underline{x})]\right) + \\ & \qquad \tfrac{1}{2}u\left(v^{-1}[(p-\alpha)v(\overline{x}) + (1-p+\alpha)v(\underline{x})]\right). \end{aligned} \tag{9}$$

Pick any two numbers $a > b$ in $[0, 1]$ and note that by setting $p = \frac{1}{2}(a+b)$ and $\alpha = \frac{1}{2}(a-b)$, inequality (9) is equivalent to the requirement that

$$\begin{aligned} & u\left(v^{-1}[\tfrac{1}{2}(a+b)v(\overline{x}) + (1-\tfrac{1}{2}(a+b))v(\underline{x})]\right) \geqslant \\ & \quad \tfrac{1}{2}u\left(v^{-1}[av(\overline{x}) + (1-a)v(\underline{x})]\right) + \tfrac{1}{2}u\left(v^{-1}[bv(\overline{x}) + (1-b)v(\underline{x})]\right). \end{aligned}$$

Since a and b are arbitrary, this inequality should hold for all such pairs. This is the case if and only if the function $u \circ v^{-1}$ is mid-point concave, which by continuity implies that $u \circ v^{-1}$ is concave. But then the decision maker would reject any noise. ■

Proof of Proposition 2: Let the lottery Y be obtained from the lottery Z by a left symmetric split and denote by x their common mean. For example, the outcome $z_i \leqslant x$ with probability p_i of Z is split into $z_i - \alpha$ and $z_i + \alpha$, each with probability $\frac{p_i}{2}$. Denote the distributions of Y and Z by F and G. Since for $t < 0$ and odd n, t^n is a concave function, it follows that if $z_i + \alpha \leqslant x$, then

$$\begin{aligned} & \int_{\underline{x}}^{\overline{x}} (t-x)^n dF(t) - \int_{\underline{x}}^{\overline{x}} (t-x)^n dG(t) = \\ & \frac{p_i}{2}[(z_i - \alpha - x)^n + (z_i + \alpha - x)^n] - p_i(z_i - x)^n \leqslant 0. \end{aligned} \tag{10}$$

If $z_i + \alpha > x$ we need to manipulate eq. (10) a little further. Let $\xi = z_i - x$

and obtain

$$\frac{p_i}{2}[(z_i-\alpha-x)^n+(z_i+\alpha-x)^n]-p_i(z_i-x)^n=$$
$$\frac{p_i}{2}\left[(\xi-\alpha)^n+(\xi+\alpha)^n\right]-p_i\xi^n=$$
$$\frac{p_i}{2}\xi^n+\frac{p_i}{2}\sum_{j=1}^{\frac{n-1}{2}}\binom{n}{2j-1}\xi^{2j-1}\alpha^{n-2j+1}-\frac{p_i}{2}\sum_{j=0}^{\frac{n-1}{2}}\binom{n}{2j}\xi^{2j}\alpha^{n-2j}+$$
$$\frac{p_i}{2}\xi^n+\frac{p_i}{2}\sum_{j=1}^{\frac{n-1}{2}}\binom{n}{2j-1}\xi^{2j-1}\alpha^{n-2j+1}+\frac{p_i}{2}\sum_{j=0}^{\frac{n-1}{2}}\binom{n}{2j}\xi^{2j}\alpha^{n-2j}-p_i\xi^n=$$
$$p_i\sum_{j=1}^{\frac{n-1}{2}}\binom{n}{2j-1}\xi^{2j-1}\alpha^{n-2j+1}\leqslant 0$$

where the last inequality follows by the fact that $\xi \leqslant 0$. Since X with expected value μ is skewed to the left it follows by Theorem 2 that it can be obtain as the limit of a sequence of left symmetric splits. At δ_μ (the distribution of $(\mu, 1)$), $\int_{\underline{x}}^{\overline{x}}(y-\mu)^n d\delta_\mu = 0$. The claim follows by the fact that each left symmetric split reduces the value of the integral. ■

Proof of Proposition 3: Let the lottery Y be obtained from the lottery Z by a left symmetric split. Denote by μ their common mean and assume that $\bar{m}(Z) \geqslant \mu$. As Y is obtained from Z by splitting one of its outcomes $z_i \leqslant \mu \leqslant \bar{m}(Z)$, this split can only increase the mass on the distribution above μ, thus (weakly) increasing its median.

By Theorem 2, X is the limit of a sequence of left symmetric splits starting with $(x, 1)$, hence the claim. ■

Appendix B: Example 1

A quadratic utility (Chew, Epstein, and Segal [9]) functional is given by $V(p) = \sum_x \sum_y p_x p_y \theta(x, y)$, where θ is symmetric. Following [9, Example 5 (p. 145)], If $\theta(x, y) = \frac{v(x)w(y)+v(y)w(x)}{2}$, where v and w are positive functions, then $V(p) = \mathrm{E}[v(p)] \times \mathrm{E}[w(p)]$. This is the form of V we analyze below.

The function V is the product of two positive linear functions of the probabilities, hence quasi concave. To see why, observe that $\ln V(p) = \ln \mathrm{E}[v(p)] + \ln \mathrm{E}[w(p)]$. The sum of concave functions is concave, hence quasi concave, and any monotone nondecreasing transformation of a quasi concave function is quasi concave.

Direct calculations show that the local utility function of any quadratic utility is given by $u_F(x) = 2 \int \theta(x, y) dF(y)$. Since we are only interested in the behavior of the function in lotteries of the form $\delta_y := (y, 1)$, we have

$$u_{\delta_y}(x) = 2\theta(x, y) = v(x)w(y) + v(y)w(x)$$

Take $v(x) = x$ and let w be any increasing, concave, and differential function such that $w(0) = 0$. We now show that V satisfies Weak Hypothesis II. That is, we show that

$$RA := -\frac{u''_{\delta_y}(x)}{u'_{\delta_y}(x)} = -\frac{yw''(x)}{w(y) + yw'(x)}$$

is an increasing function of y. We have

$$\begin{aligned}
&-\frac{\partial}{\partial y}\left(\frac{yw''(x)}{w(y) + yw'(x)}\right) > 0 \iff \\
&w''(x)(w(y) + yw'(x)) < (w'(y) + w'(x))yw''(x) \iff \\
&w(y) > w'(y)y \iff \\
&w(y)/y > w'(y)
\end{aligned}$$

which holds since w is concave.

Next we analyze the functional form $V(\langle p_1, q_1; \ldots; p_n, q_n\rangle) = \mathrm{E}[w(c_p)] \times \mathrm{E}[c_p]$ where $w(x) = \frac{\zeta x - x^\zeta}{\zeta - 1}$, $c_p = \beta p + (1-\beta)p^\kappa$, $\zeta = 1.024$,[13] $\kappa = 1.1$, and $\beta = 0.15$. Since all the inequalities below are strict, there is an open set of parameters for which they are satisfied as well. Observe that

$$w(c_p) = \frac{\zeta\left[\beta p + (1-\beta)p^\kappa\right] - \left[\beta p + (1-\beta)p^\kappa\right]^\zeta}{\zeta - 1}$$

We show first that this functional rejects all symmetric noise. For any $0 < p < 1$ and $\varepsilon \leqslant \min\{p, 1-p\}$, let

$$f(\varepsilon, p) := [w(c_{p+\varepsilon}) + w(c_{p-\varepsilon})] \times [c_{p+\varepsilon} + c_{p-\varepsilon}]$$

Rejection of symmetric noise requires that $f(0,p) - f(\varepsilon, p) > 0$ for all $p \in (0,1)$ and $\varepsilon \in (0, \min\{p, 1-p\})$. Numerical calculations show that this is indeed the case. See graph below.

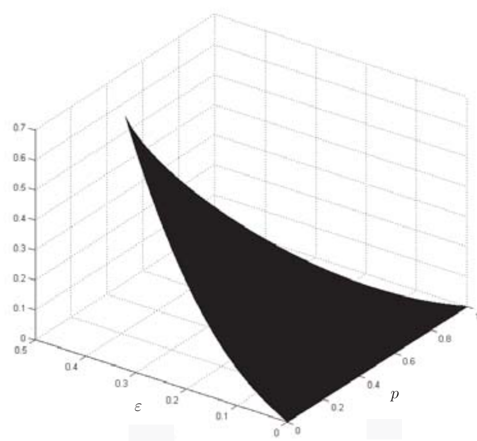

Figure A1: Rejection of symmetric noise

Using the same functional as above, we now show that for every $p > 0$

[13] since $\zeta > 1$, we have that $w'(x) = \frac{\zeta - \zeta x^{\zeta-1}}{\zeta-1} > 0$ and $w''(x) = \frac{(\zeta-1)\zeta x^{\zeta-2}}{\zeta-1} < 0$, hence w is increasing and concave.

there exists a sufficiently small $q > 0$ such that $\langle p, q; 0, 1-q\rangle \succeq \langle pq, 1\rangle$, that is, the decision maker always accepts some positively skewed noise.

For $q = 0$, $V(c_{pq}, 1) - V(c_p, q; 0, 1-q) = 0$. We show that for every $p < 1$, the first non-zero derivative of this expression with respect to q at $q = 0$ is negative. We get

$$(\zeta - 1)V(c_{pq}, 1) = (\zeta - 1)w(c_{pq})c_{pq} =$$
$$\left(\zeta\,[\beta pq + (1-\beta)p^\kappa q^\kappa] - [\beta pq + (1-\beta)p^\kappa q^\kappa]^\zeta\right) \times [\beta pq + (1-\beta)p^\kappa q^\kappa]$$

Differentiate with respect to q to obtain

$$\left(\zeta\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right] - \zeta\,[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-1}\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]\right) \times$$
$$[\beta pq + (1-\beta)p^\kappa q^\kappa] +$$
$$\left(\zeta\,[\beta pq + (1-\beta)p^\kappa q^\kappa] - [\beta pq + (1-\beta)p^\kappa q^\kappa]^\zeta\right) \times \left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]$$

At $q = 0$, this expression equals 0. Differentiate again with respect to q to obtain

$$\zeta\left(\kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-2} - (\zeta-1)\,[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-2}\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]^2 -\right.$$
$$\left.[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-1}\,\kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-2}\right) \times [\beta pq + (1-\beta)p^\kappa q^\kappa] +$$
$$2\zeta\left(\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right] - [\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-1}\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]\right) \times$$
$$\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right] +$$
$$\left(\zeta\,[\beta pq + (1-\beta)p^\kappa q^\kappa] - [\beta pq + (1-\beta)p^\kappa q^\kappa]^\zeta\right) \times \kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-2}$$

Observe that

$$\zeta\left(\kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-2} - (\zeta-1)\,[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-2}\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]^2 -\right.$$
$$\left.[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-1}\,\kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-2}\right) \times [\beta pq + (1-\beta)p^\kappa q^\kappa] =$$
$$\zeta\left(\kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-1} - (\zeta-1)q\,[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-2}\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]^2 -\right.$$
$$\left.[\beta pq + (1-\beta)p^\kappa q^\kappa]^{\zeta-1}\,\kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-1}\right) \times \left[\beta p + (1-\beta)p^\kappa q^{\kappa-1}\right]$$

This expression converges to zero with q. This is obvious for $\zeta \geqslant 2$. If $2 > \zeta > 1$, then notice that by l'Hospital's rule

$$\lim_{q\to 0} \frac{q}{\left[\beta pq + (1-\beta)p^\kappa q^\kappa\right]^{2-\zeta}} = \lim_{q\to 0} \frac{\left[\beta pq + (1-\beta)p^\kappa q^\kappa\right]^{\zeta-1}}{(2-\zeta)\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]} = 0$$

Also, as $q \to 0$, the limit of the expression

$$2\zeta\left(\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right] - \left[\beta pq + (1-\beta)p^\kappa q^\kappa\right]^{\zeta-1}\left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]\right) \times \left[\beta p + \kappa(1-\beta)p^\kappa q^{\kappa-1}\right]$$

is $2\zeta\beta^2p^2$. Finaly,

$$\left(\zeta\left[\beta pq + (1-\beta)p^\kappa q^\kappa\right] - \left[\beta pq + (1-\beta)p^\kappa q^\kappa\right]^\zeta\right) \times \kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-2} =$$
$$\left(\zeta\left[\beta p + (1-\beta)p^\kappa q^{\kappa-1}\right] - \left[\beta pq^{1-\frac{1}{\zeta}} + (1-\beta)p^\kappa q^{\kappa-\frac{1}{\zeta}}\right]^\zeta\right) \times \kappa(\kappa-1)(1-\beta)p^\kappa q^{\kappa-1}$$

As $\zeta, \kappa > 1$, this expression goes to zero with q.

On the other hand, $(\zeta-1)V(c_p, q; 0, 1-q)$ equals

$$q^2\left(\zeta\left[\beta p + (1-\beta)p^\kappa\right] - \left[\beta p + (1-\beta)p^\kappa\right]^\zeta\right) \times \left[\beta p + (1-\beta)p^\kappa\right]$$

Its first order derivative with respect to q at $q = 0$ is zero, while the second derivative at this point equals

$$2\left(\zeta\left[\beta p + (1-\beta)p^\kappa\right] - \left[\beta p + (1-\beta)p^\kappa\right]^\zeta\right) \times \left[\beta p + (1-\beta)p^\kappa\right]$$

We therefore get that the first order derivative of $V(c_{pq}), 1) - V(c_p, q; 0, 1-q)$ at $q = 0$ is zero, and that

$$(\zeta-1)\lim_{q\to 0}\frac{\partial^2}{\partial q^2}\left[V(c_{pq}), 1) - V(c_p, q; 0, 1-q)\right] = g(p;\beta,\zeta,\kappa) :=$$
$$2\zeta\beta^2p^2 - 2\left(\zeta\left[\beta p + (1-\beta)p^\kappa\right] - \left[\beta p + (1-\beta)p^\kappa\right]^\zeta\right) \times \left[\beta p + (1-\beta)p^\kappa\right]$$

The graph below shows $g(p;\beta,\zeta,\kappa)$ for $\beta = 0.15$, $\kappa = 1.1$, and $\zeta = 1.024$.

Note that for these values $g(p;\beta,\zeta,\kappa) < 0$ for all $p \in (0,1)$, which means that for $q > 0$ small enough, the positively skewed noise $\langle p, q; 0, 1-q\rangle$ is accepted.

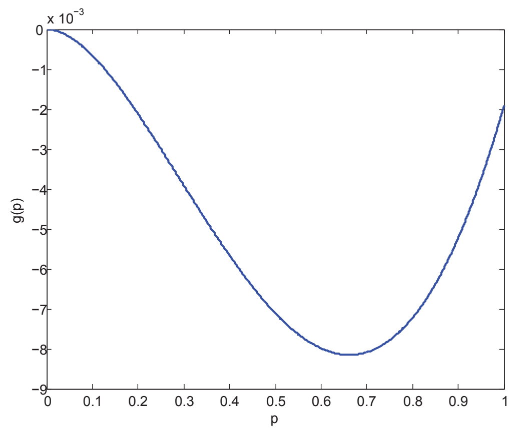

Figure A2: $g(p; 0.15, 1.024, 1.1)$

References

[1] Abdellaoui, M., P. Klibanoff, and L. Placido, 2013. “Experiments on compound risk in relation to simple risk and to ambiguity,” mimeo.

[2] Abdellaoui, M., O. lHaridon, and A. Nebout, 2015. “Probability Weighting in Recursive Evaluation of Two-stage Prospects,” mimeo.

[3] Abdulkadiroğlu, A. and T. Sönmez, 1998. “Random serial dictatorship and the core from random endowments in house allocation problems,” *Econometrica* 66:689–701.

[4] Abdulkadiroğlu, A. and T. Sönmez, 2013. “Matching markets: Theory and practice,” *Advances in Economics and Econometrics*, D. Acemoglu, M. Arello, and E. Dekel (eds), Cambridge: CUP.

[5] Becker, S.W. and F.O. Brownson, 1964: “What price ambiguity? Or the role of ambiguity in decision making,” *Journal of Political Economy* 72:62–73.

[6] Bernasconi, M. and G. Loomes, 1992. “Failures of the reduction principle in an Ellsberg-type problem,” *Theory and Decision* 32:77–100.

[7] Boiney, L.G., 1993. “The effects of skewed probability on decision making under ambiguity,” *Organizional Behavior and Human Decisions Processes* 56:134–148.

[8] Camerer, C. F. and M. Weber, 1992. “Recent Developments in Modeling Preferences: Uncertainty and Ambiguity,” *Journal of Risk and Uncertainty*, 5: 325-370.

[9] Chew, S.H., L.G. Epstein, and U. Segal, 1991. “Mixture symmetry and quadratic utility,” *Econometrica* 59:139–163.

[10] Chew S.H., B. Miao, and S. Zhong, 2013. “Partial ambiguity,” working paper.

[11] Chew, S.H. and W.S. Waller, 1986. "Empirical tests of weighted utility theory," *Journal of Mathematical Psychology* 30:55–72.

[12] Conlisk, J., 1989. "Three variants on the Allais example," *American Economic Review* 79:392–407.

[13] Coombs, C. and L. Huang, 1976. "Tests of the betweenness property of expected utility," *Journal of Mathematical Psychology* 13:323-337.

[14] Dekel, E., 1989. "Asset demands without the independence Axiom," *Econometrica* 57:163–69.

[15] Dillenberger, D., 2010. "Preferences for one-shot resolution of uncertainty and Allais-type behavior," *Econometrica* 78:1973–2004.

[16] Dimmock, S. G., R. Kouwenberg, O. S. Mitchell, and K. Peijnenburg, 2013. "Ambiguity attitudes and economic behavior," Working paper, Bocconi.

[17] Dwenger, N., D. Kubler, and G. Weizsacker, 2013. "Flipping a coin: Theory and evidence," Mimeo, Humboldt-Universität zu Berlin.

[18] Ellsberg, D., 1961. "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics* 75:643–669.

[19] Kocher, M. G, A. Lahno, and S.T. Trautmann, 2015. "Ambiguity aversion is the exception," Mimeo, University of Munich, Department of Economics.

[20] Ghirardato, P., Maccheroni, F., and Marinacci, M., 2004. "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory* 118:133–73.

[21] Gilboa, I. and D. Schmeidler, 1989. "Maxmin expected utility with a non-unique prior," *Journal of Mathematical Economics* 18:141–153.

[22] Halevy, Y., 2007. "Ellsberg revisited: An experimental study," *Econometrica* 75:503–536.

[23] Harrison, G. W., J. Martinez-Correa, and J.T. Swarthout, 2012. "Reduction of compound lotteries with objective probabilities: Theory and evidence," working paper.

[24] Hylland, A., and Zeckhauser, R., 1979. "The efficient allocation of individuals to positions," *Journal of Political Economy* 87:293–314.

[25] Kahn, B.E. and R.K. Sarin, 1988. "Modeling ambiguity in decisions under uncertainty," *Journal of Consumer Research* 15:265–272.

[26] Kahneman, D. and A. Tversky, 1979. "Prospect theory: An analysis of decision under risk," *Econometrica* 47:263–291.

[27] Kreps, D.M. and E.L. Porteus, 1978. "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica* 46:185–200.

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

[29] Machina, M.J., 1982. " 'Expected utility' analysis without the independence axiom," *Econometrica* 50:277–323.

[30] Maccheroni, F., M. Marinacci, and A. Rustichini, 2006. " Ambiguity aversion, robustness, and the variational representation of preferences," *Econometrica* 74:1447–1498

[31] Masatlioglu, Y., Y. Orhun, and C. Raymond, 2015. "Skewness and intrinsic preferences for information, Mimeo.

[32] Menezes, C.F., C. Geiss, and J. Tressler, 1980. "Increasing downside risk," *American Economic Review* 70:921–932.

[33] Miao, B., and S. Zhong, 2012. "An experimental study of attitude towards two-stage lottery," working paper.

[34] Pathak, P.A. and J. Sethuraman, 2011. "Lotteries in student assignment: An equivalence result," *Theoretical Economics* 6:1–17.

[35] Quiggin, J., 1982. "A theory of anticipated utility," *Journal of Economic Behavior and Organization* 3:323–343.

[36] Rothschild, M. and J.E. Stiglitz, 1970. "Increasing risk: I. A definition," *Journal of Economic Theory* 2:225–243.

[37] Schmeidler, D., 1989. "Subjective probability and expected utility without additivity," *Econometrica* 57:571–587.

[38] Segal, U., 1987. "The Ellsberg paradox and risk aversion: An anticipated utility approach," *International Economic Review* 28:175–202.

[39] Segal, U., 1990. "Two-stage lotteries without the reduction axiom," *Econometrica* 58:349–377.

[40] Segal, U. and A. Spivak, 1990. "first-order versus second-order risk aversion," *Journal of Economic Theory* 51:111–125.

[41] Sönmez, T. and M. U. Ünver, 2011. "Matching, Allocation, and Exchange of Discrete Resources," in J. Benhabib, A. Bisin, and M. Jackson (eds.), Handbook of Social Economics, Vol. 1A. The Netherlands: North-Holland, 781-852.

[42] Sopher, B. and J. M. Narramore, 2000. "Stochastic choice and consistency in decision making under risk: An experimental study," *Theory and Decision* 48:323-350.

[43] Trautmann, S. T. and G. van de Kuilen, 2014. "Ambiguity Attitudes," Forthcoming. In: Keren, G., Wu, G. (Eds.), Blackwell Handbook of Judgment and Decision Making (Chapter 3), Blackwell.

[44] van de Kuilen, G. and P.P. Wakker, 2011. "The midweight method to measure attitudes toward risk and ambiguity," *Management Science* 57:582–598.

[45] Viscusi, W.K. and H. Chesson, 1999: "Hopes and fears: The confliction effects of risk ambiguity," *Theory and Decision* 47:153–178.