Expected Utility Theory with Bounded Probability Nets

by M. Kaneko, 21 July 2016, Stony Brook

- Introduce restrictions on available probabilities into the classical EU theory.

  Bounded probability nets $\Pi_k = \{\frac{t}{10^k} : t = 0,1, \dots, 10^k\}$.

- Two types of use of probability:
  - Measurement of "utility" from a pure alternative.
  - Extension of measured utilities into events involving more risks.
- In the classical EU theory, these are not well separated;
  - Here, they are separated, and both are restricted with bounded probability nets.

- The classical EU theory can be regarded as the case with no restrictions.

- The new theory is applied to the Alleis-Kahneman-Tversky anomaly.
- We connect measurement of "utility" with Simon's satisficing/aspiration.

My three projects.

- **Inductive game theory** - - a player learns the structure of a society from his experiences.

- **Mathematical (epistemic) logic** - - a player has some logical inference ability and makes a decision using his logical ability

- **Social justice** - - what is the ultimate social justice? The Nash social welfare function and its contractarian foundation:

$$\sum_{i=1}^{n} \log(u_i(x) - u_i(x_0))$$

  - So far, this theory is constructed upon EU theory, but how much does Nash theory rely upon EU?

- Key concept for the three projects is: **“bounded rationality”**

The Alleis-Kahneman-Tversky Anomaly

- You are asked to choose one out of each of the following two choice problems:

a: $\frac{8}{10}4000 * \frac{2}{10}0$ [20%] *or* *b:* $\frac{10}{10}3000 * \frac{0}{10}0$ (=3000) [80%]

and

c: $\frac{2}{10}4000 * \frac{8}{10}0$ [65%] *or* *d:* $\frac{25}{100}3000 * \frac{75}{100}0$ [35%]

Modal case ($a \prec b$) & ($c \succ d$): Contradiction to EU theory: In EU, those preferences are expressed as:

$$\frac{8}{10}u(4000) + \frac{2}{10}u(0) < \frac{10}{10}u(3000) + \frac{0}{10}u(0)$$

$$\frac{2}{10}u(4000) + \frac{8}{10}u(0) > \frac{25}{100}u(3000) + \frac{75}{100}u(0).$$

	c: 65%	*d:* 35%
a: 20%	EU	anomaly
b: 80%	anomaly	EU

The opposite inequality of the second is obtained from the first by dividing by 4 ($u(0) = 0$).

- This example will be used to motivate the development of our theory.

Detailed Derivation of the Anomaly in EU Theory

Three alternatives: $y_o = 0$; $y = 3000\ (= b)$; $y^o = 4000$ - - (y_o, y^o: benchmarks)

- **Measurement**: Find a probability λ so that $y \sim \lambda y^o * (1-\lambda) y_o$,
  - y is measured by the benchmark scale, e.g., $y \sim \frac{85}{100} y^o * \frac{15}{100} y_o$

$$a := \frac{80}{100} y^o * \frac{20}{100} y_o \prec \frac{85}{100} y^o * \frac{15}{100} y_o \sim y =: b.$$

- **Extension**: Measurement is extended to $\lambda y * (1-\lambda) y_o$ such as $d = \frac{25}{100} y * \frac{75}{100} y_o$.
  - EU theory allows to substitute $\frac{85}{100} y^o * \frac{15}{100} y_o$ for y in $d := \frac{25}{100} y * \frac{75}{100} y_o$.
  - Then, $d \sim \frac{2125}{10000} y^o * \frac{7875}{10000} y_o \succ \frac{20}{100} y^o * \frac{80}{100} y_o := c$.
  - The AKT observation gives some inconsistency.
  - Recall that level of significance in statistics is usually, 5% or 1%.

Expected Utility Theory

- $L(X)$: the set of probability distributions (with finite supports) over pure alternative choices;
- $\precsim$: a preference relation over $L(X)$.

Axiom NM1: (1) $\precsim$ is complete; (2) $\precsim$ is transitive.

Axiom NM2 (continuity): if $f \prec h \prec g$, there is an indifferent lottery $\alpha f * (1-\alpha)g \sim h$.

Axiom NM3: (Independence): (1): if $f \sim g$, then $\alpha f * (1-\alpha)h \sim \alpha g * (1-\alpha)h$;

(2): if $f \prec g$, then $\alpha f * (1-\alpha)h \prec \alpha g * (1-\alpha)h$.

Theorem 1 (Representation): A binary relation $\precsim$ satisfies Axiom NM1-NM3 if and only if

$\exists u: L(X) \to R$ s.t. $\forall f, g \in L(X)$ and $\alpha \in [0,1]$,

(1-1) $f \precsim g \Leftrightarrow u(f) \leq u(g)$;

(1-2) $u(\alpha f * (1-\alpha)g) = \alpha u(f) + (1-\alpha)u(g)$;

Theorem 2 (Uniqueness): If $u: L(X) \to R$ and $v: L(X) \to R$ satisfy (1-1) and (1-2) with the same $\precsim$,

then $\exists \alpha > 0$ *and* β such that $v(f) = \alpha u(f) + \beta \ \forall f \in L(X)$.

Bounded probability nets

- For any integer $k \geq 0$,

  $\Pi_k := \{\frac{t}{10^k} : t = 0,1, \dots, 10^k\}$ - - available probability values.

- **Extension Lemma 1.**

  $\Pi_{k+1} = \{\boldsymbol{e} \cdot \boldsymbol{\lambda} = \sum_{t=1}^{10} \frac{1}{10} \lambda_t : \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{10}) \in (\Pi_k)^{10}\}$ for $k \geq 0$.

  where $\boldsymbol{e} = (\frac{1}{10}, \dots, \frac{1}{10})$.

- **Depth of a probability**: $\delta(\lambda) = k \Leftrightarrow_{def} \lambda \in \Pi_k - \Pi_{k-1}$.

- Let ρ be an **upper bound** (nonnegative integer or $+\infty$) for available probabilities.

  $\Pi_\rho = \{\frac{t}{10^\rho} : t = 0,1, \dots, 10^\rho\}$ for $\rho < +\infty$; or $\Pi_\infty = \cup_{k>0} \Pi_k$ if $\rho = \infty$.

  - **Our main concern** is the case of small ρ, e.g., $\rho = 2$ in the AKT example.
  - $\rho = \infty$ is the reference case - - this is the case of the classical EU theory.

Pure Alternatives and Lotteries:

- Y is the set of pure alternatives - - in the AKT example, $Y = \{y^o,\ y,\ y_o\}$.
- $L_k(Y) := \{f: Y \to \Pi_k \mid \sum_{y \in Y} f(y) = 1\}$ - - the set of lotteries of depth k.

Extension Lemma 2.

$$L_{k+1}(Y) = \left\{ \boldsymbol{e} * \boldsymbol{f} = \sum_{t=1}^{10} \frac{1}{10} f_t : \boldsymbol{f} = (f_1, \dots, f_{10}) \in L_k(Y)^{10} \right\} \text{ for } k \geq 0.$$

- This connects $L_{k+1}(Y)$ with $L_k(Y)$ and enables us to use **mathematical induction**:

$$L_0(Y) = Y \to L_1(Y) \to L_2(Y) \to L_3(Y) \to L_4(Y) \to \dots$$

$$d = \frac{25}{100} y * \frac{75}{100} y_o \sim \frac{2125}{10000} y^o * \frac{7875}{10000} y_o$$

In AKT example, bound ρ can be 2.

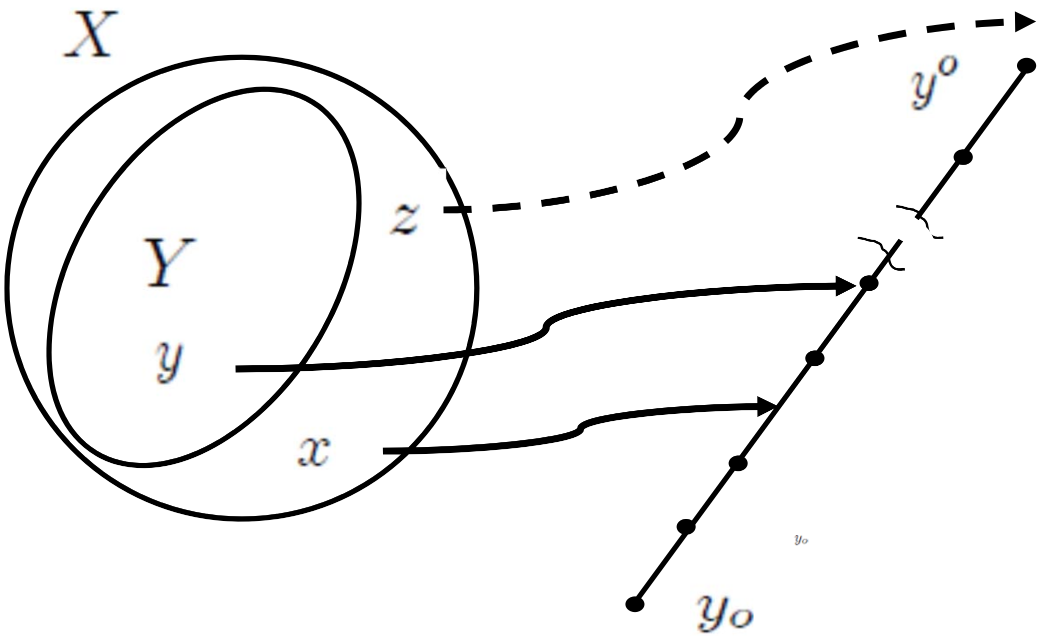
- Let $L_\rho(Y) = L_k(Y)$ if $\rho = k < \infty$; and $L_\rho(Y) = \cup_{k<\infty} L_k(Y)$ if $\rho = \infty$.

Analogy: Mathematical Induction

- **Induction Base** - - Measurement: Base Facet
- **Induction Step** - - Extension

Base Facet $F = \left\langle Y; y^o, y_o; \{\lambda_y\}_{y \in Y} \right\rangle$

- Y is a subset of X: each $y \in Y$ is measured with y^o, y_o and probability λ_y.
- y^o, y_o in Y are the *upper, lower benchmarks*.
- λ_y is a probability in Π_ρ.



Benchmark scale:

$\mathbf{B}_\rho(y^o, y_o) = \{\lambda y^o * (1-\lambda) y_o : \lambda \in \Pi_\rho\} \subseteq L_\rho(Y)$.

- **Preference relation $\precsim$ over $L_\rho(Y)$.**

Axiom B0.(1)(Reflexivity): for any $f \in L_\rho(Y)$, $f \precsim f$;

(2): (Transitivity): for any $f, g, h \in L_\rho(Y)$, $f \precsim g$ and $g \precsim h$ imply $f \precsim h$.

- **We do not require completeness.**

Axiom B1(Benchmarks): $y_o \prec y \prec y^o$ for all $y \in Y - \{y^o, y_o\}$.

Axiom $\mathbf{B2}^o$(Monotonicity along the Benchmark scale): for any $\lambda, \lambda' \in \Pi_\rho$,

if $\lambda < \lambda'$, then $\lambda y^o * (1-\lambda) y_o \prec \lambda' y^o * (1-\lambda') y_o$.

Axiom B3(Measurement with the benchmark scale): $y \sim \lambda_y y^o * (1-\lambda_y) y_o$ for all $y \in Y$.

- B1, $\mathrm{B2}^o$, B3 determine the meaning of a base facet $F = \left\langle Y; y^o, y_o; \{\lambda_y\}_{y \in Y} \right\rangle$.

Axiom B4$^{\infty}$ $(\textbf{Extension})$: For any $\boldsymbol{f} = (f_1, \dots, f_{10}) \in L_{\infty}(Y)^{10}$ and $\boldsymbol{g} = (g_1, \dots, g_{10}) \in \mathbf{B}_{\infty}(y^o, y_o)^{10}$,

if $\boldsymbol{f} = (f_1, \dots, f_{10}) \sim \boldsymbol{g} = (g_1, \dots, g_{10})$, then $\boldsymbol{e} * \boldsymbol{f} \sim \boldsymbol{e} * \boldsymbol{g}$ (recall $\boldsymbol{e} = \left(\frac{\mathbf{1}}{\mathbf{10}}, \dots, \frac{\mathbf{1}}{\mathbf{10}}\right)$).

Theorem: Let $\rho = \infty$, and B0,B1, B2^o, B3, B4$^{\infty}$. Then, for all $f, g \in L_{\rho}(Y)$,

$$f \precsim g \Leftrightarrow u_{eu}(f) \leq u_{eu}(g),$$

where the EU function u_{eu} is defined as $u_{eu}(f) := \sum_{y \in Y} f(y)\lambda_y$ for all $f \in L_{\rho}(Y)$.

- The complete relation $\precsim$ is uniquely determined by $F = \left\langle Y; y^o, y_o; \{\lambda_y\}_{y \in Y} \right\rangle$, B0,B1, B2^o, B3, B4$^{\infty}$.

Extension: $Y \longrightarrow L_1(Y) \longrightarrow L_2(Y) \longrightarrow L_3(Y) \longrightarrow L_4(Y) \longrightarrow \dots$

($\longrightarrow$ One-way)

$$y \sim \frac{85}{100} y^o * \frac{15}{100} y_o \qquad \frac{25}{100} y * \frac{75}{100} y_o \sim \frac{2125}{10000} y^o * \frac{7875}{10000} y_o.$$

- However, no bounds for probability nets.

- When $\rho < \infty$, the one-way extension is not available.
- The "switchback" method may be applied.

Switchback Extension $\rho = 2$:

$$L_2(Y) \dashrightarrow L_1(Y) \rightarrow L_2(Y)$$

$$y \sim \frac{85}{100} y^o * \frac{15}{100} y_o \dashrightarrow \frac{8}{10} y^o * \frac{2}{10} y_o$$

$$\frac{25}{100} y * \frac{75}{100} y_o \dashrightarrow \frac{2}{10} y * \frac{8}{10} y_o$$

$$\frac{16}{100} y^o * \frac{84}{100} y_o \prec \frac{2}{10} y * \frac{8}{10} y_o \prec \frac{25}{100} y * \frac{75}{100} y_o$$

Axiom B2(Monotonicity): Let $y, z \in Y$ and $\lambda_y > \lambda_z$. If $f(y) > g(y)$ and $f(x) = g(x)$ for all $x \in Y - \{y, z\}$, then $g \prec f$.

- We write $\boldsymbol{f} = (f_1, \ldots, f_{10}) \precsim \boldsymbol{g} = (g_1, \ldots, g_{10})$ and recall $\boldsymbol{e} * \boldsymbol{f} = \sum_{t=1}^{10} \frac{1}{10} f_t$.

Axiom B4 (Extension): for any $\boldsymbol{f} \in L_{\rho-1}(Y)^{10}$ and $\boldsymbol{g} \in \mathbf{B}_{\rho-1}(y^o; y_o)^{10}$,

(1): if $\boldsymbol{f} \precsim \boldsymbol{g}$, then , $\boldsymbol{e} * \boldsymbol{f} \precsim \boldsymbol{e} * \boldsymbol{g}$; and (2) if $\boldsymbol{g} \precsim \boldsymbol{f}$, then , $\boldsymbol{e} * \boldsymbol{g} \precsim \boldsymbol{e} * \boldsymbol{f}$.

Each has the strict part.

- Constructive: $\precsim$ over $Y \rightarrow \precsim$ over $L_1(Y) \rightarrow \ldots \rightarrow \precsim$ over $L_{\rho-1}(Y) \rightarrow \precsim$ over $L_\rho(Y)$

Axiom B0: $\precsim$ is reflexive and transitive over $L_\rho(Y)$.

Axiom B1: $y^o \prec y \prec y_o$ for all $y \in Y - \{y^o, y_o\}$.

Axiom B2: : Let $y, z \in Y$ and $\lambda_y > \lambda_z$. If $f(y) > g(y)$ and $f(x) = g(x)$ for all $x \in Y - \{y, z\}$, then $g \prec f$.

Axiom B3: $y \sim \lambda_y y^o * (1 - \lambda_y) y_o$ for all $y \in Y$.

Axiom B4 : For any $\boldsymbol{f} \in L_{\rho-1}(Y)^{10}$ and , $\boldsymbol{g} \in \mathbf{B}_{\rho-1}(y^o, y_o)^{10}$,

(1): if $\boldsymbol{f} \precsim \boldsymbol{g}$, then , $\boldsymbol{e} * \boldsymbol{f} \precsim \boldsymbol{e} * \boldsymbol{g}$; and (2) if $\boldsymbol{g} \precsim \boldsymbol{f}$, then , $\boldsymbol{e} * \boldsymbol{g} \precsim \boldsymbol{e} * \boldsymbol{f}$,

with strict part.

- Axioms B0-B4 are consistent for any ρ.

$\rho = \infty$:

1. $\precsim$ is represented by expected utility over $L_\infty(Y)$;
2. $\precsim$ is determined uniquely and satisfies completeness;
3. B2 and B4 are derived from Axioms B0,B1,$\mathrm{B2}^o$,B3,$\mathrm{B4}^\infty$ (Indifference part of B4 for $\rho = \infty$)

→

$\rho < \infty$:

1. Only some part
2. Multiplicity and incomparability
3. B2 and B4 are crucial

Canonical Preferences $\precsim_c$ for $\rho < \infty$

Intersection Lemma: Let $\precsim$ and $\precsim'$ be preference relations over $L_\rho(Y)$ with Axioms B0-B4.

Then, the intersection $\precsim^* = (\precsim \cap \precsim')$ satisfies B0-B4.

- We define the canonical preference relation $\precsim_c$ by

  $\precsim_c = \cap \{\precsim$: it is a relation over $L_\rho(Y)$ with B0−B4$\}$.

- The central domain $\mathbf{C}(F;\rho) = \{f \in L_\rho(Y): f \sim_c g \text{ for some } g \in \mathbf{B}_\rho(y^o, y_o)\}$.

  - - This is the constructive domain by B4.

Theorem : For any relation $\precsim$ with B0-B4, and for any $f, g \in \mathbf{C}(F;\rho)$,

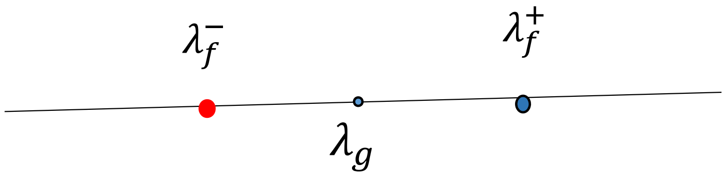
- $f \precsim g$ if and only if $f \precsim_c g$;
- $u_{eu}(f) \leq u_{eu}(g)$ if and only if $f \precsim_c g$.

- Thus, the eu-preference relation is the unique once over the central domain $\mathbf{C}(F;\rho)$.
- What happens outside $\mathbf{C}(F;\rho)$? Incomparability ? Multiplicity of $\precsim$?

Greatest Lower Bound λ_f^- and Least Upper Bound λ_f^+

- For each $f \in L_\rho(Y)$, define
  - $\lambda_f^- = max\{\lambda \in \Pi_\rho: \ \lambda y^o * (1-\lambda)y_o \precsim_c f\}$
  - $\lambda_f^+ = min\{\lambda \in \Pi_\rho: \ f \precsim_c \lambda y^o * (1-\lambda)y_o\}$.

- $\lambda_f^- = \lambda_f = \lambda_f^+ \Leftrightarrow f \in \mathbf{C}(F;\rho)$.
- $\lambda_f^- < \lambda_f^+ \Leftrightarrow f \notin \mathbf{C}(F;\rho)$.



Relevant case: $f \notin \mathbf{C}(F;\rho)\ and\ g \in \mathbf{C}(F;\rho)$:

Theorem A: f, g are incomparable w.r.t. $\precsim_c$ if and only if $\lambda_f^- < \lambda_g < \lambda_f^+$.

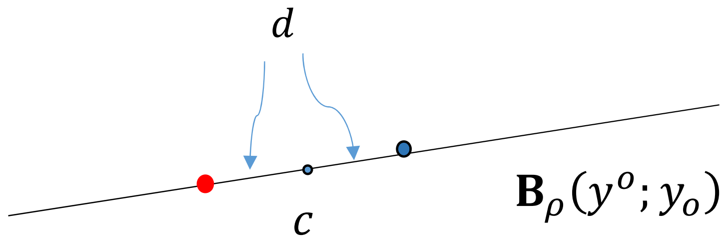
The AKT example with $\rho = 2$; Greatest lower bound /Least upper bound

- For $\rho = \infty$, $\quad d = \frac{25}{100} y * \frac{75}{100} y_o \sim \frac{2125}{10000} y^o * \frac{7875}{10000} y_o \succ \frac{20}{100} y^o * \frac{80}{100} y_o = c$.

- Let $\rho = 2$.
  - d is a non-benchmark lottery and c is a benchmark lottery.
  - It can be shown that d and c are incomparable

- GLB and LUB $d = \frac{25}{100} y * \frac{75}{100} y_o$ in the benchmark scale $\mathbf{B}_\rho(y^o; y_o)$ are:

$$\frac{16}{100} y^o * \frac{84}{100} y_o \prec_c d \prec_c \frac{25}{100} y^o * \frac{75}{100} y_o$$



Interpretations:

1: d is in this interval including $c = \frac{20}{100} y^o * \frac{80}{100} y_o$;

and, either d or c could be chosen (if he is asked).

2: He may have a preference of c over d (d over c); either preference is possible. (This is proved).

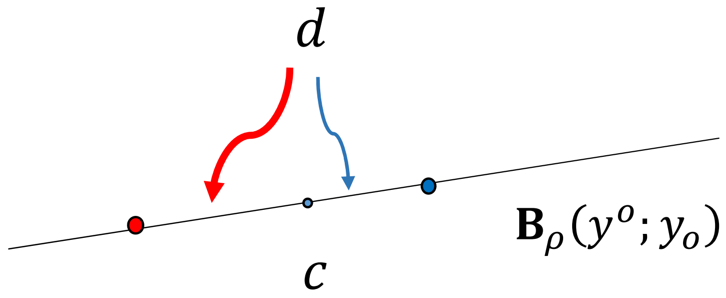
AKT anomaly

- AKT anomaly cases: $b \wedge c$ and $a \wedge d$
- EU cases: $a \wedge c$ and $b \wedge d$

	c: 65%	d: 35%
a:20%	$a \wedge c$: 0 // 13 //20 EU	$a \wedge d$: 20 // 7 //0 AKT-anomaly
b: 80%	$b \wedge c$: 65 // 52 //45 ATK-anomaly	$b \wedge d$: 15 // 28 //35 EU

Case b: our theory is compatible with either choice c or d.

- c - - 65% and d - - 35%



Case a: similar.

Measurement by bounded rationality (**Satisficing/aspiration**) due to Simon

Step: Decision Maker asks himself (thought experiment) about:

(Π_1): for each $\lambda \in \Pi_1$, is y indifferent to $\lambda y^o * (1-\lambda)y_o$?

- If he finds a satisfactory one, he accepts λ_y as his decision.
  - if he find multiple ones, he takes the lowest one λ_y (or the highest one μ_y).
- Otherwise, he goes to the step 2:
- Step 2 is essentially the same but he considers Π_2 instead of Π_1.
- Goes on until he gets tired.

- Extension step is highly demanding from the viewpoint of "bounded" rationality,
- However, substitutability is a property different from measurement in term of probability.

Conclusions

We developed an EU theory with bounded probability nets.

- It covers the range from the classical EU theory to bounded case.
- In bounded cases, it allows multiplicity of preference relations and incomparable lotteries.
- We applied the theory to the AKT anomaly.

Remaining Problems

- Even for bounded ρ, I conjecture some utility representation theorem; for incomparability, different orderings may be given.
- The Anscombe-Aumann approach to “subjective probability” may be in the scope of our theory.
- Experimental tests?

- Alleis, M., (1953), Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine, Econometrica 21, 503--546.
- Hammond, P., (1998), Objective Expected Utility, Handbook of Utility Theory Vol.1: Principles, Ed. S. Barbera, et al. 143-211.
- Fishburn, P., (1982), The Foundations of Expected Utility, D. Reidel Publishing Co. London.
- Hu, T., (2013), Expected Utility Theory from the Frequentist Perspective, Economic Theory 53, 9--25.
- Kahneman, D., and A. Tversky, (1979), Prospect Theory: An Analysis of Decision under Risk, Econometrica 47, 263-292.
- Kaneko, M., and K. Nakamura, (1979), The Nash social welfare function, Econometrica 47 (1979), 423—435.
- Kaneko, M., and J. J. Kline, Inductive Game Theory: A Basic Scenario, Journal of Mathematical Economics 44, (2008), 1332--1363.
- Rubinstein, A., (1988), Similarity and decision making under risk: is there a utility theory resolution to the Allais paradox, Journal of Economic Theory 46, 145-153.
- Rubinstein, A., and Y. Salant, (2006), A Model of Choice from Lists, Theoretical Economics 1, 3-17.
- Simon, H. A., (1983), Reason in Human Affairs, Stanford University Press, Stanford.
- Sugden, R., (2004), Alternatives to Expected Utility Theory: Foundations, Handbook of Utility Theory Vol.2, 685-755.
- von Neumann, J., and O. Morgenstern, (1944), Theory of Games and Economic Behavior, 2nd ed. 1947, Princeton University Press, Princeton, N.J.