Non-EU Risk Preferences

The Allais Paradox:

- Choose A or B.

  A: $1 million for sure $:= \delta_{1M}$

  B: $.1\delta_{5M} + .89\delta_{1M} + .01\delta_0$

- Choose C or D

  C: $.11\delta_{1M} + .89\delta_0$

  D $.1\delta_{5M} + .9\delta_0$

- Non-trivial % of people choose A and D.

- This isn't consistent with expected utility:

$$U(A) = u(1) > .1u(5) + .89u(1) + .01u(0) = U(B)$$

$$\leftrightarrow .11u(1) > .1u(5) + .01u(0)$$

$$\leftrightarrow U(C) = .11u(1) + .89u(0) > .1u(5) + .9u(0) = U(D)$$

Reason: the independence axiom says that if

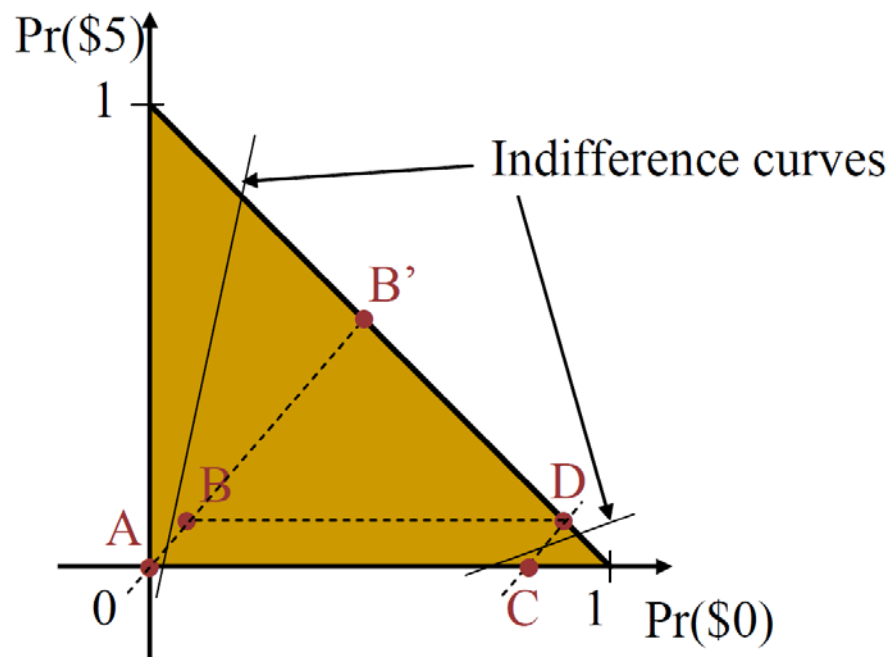- $.11\delta_1 + .89\delta_1 \succ .1\delta_5 + .01\delta_0 + .89\delta_1,$

  then

  $\delta_1 \succ (10/11)\delta_5 + (1/11)\delta_0 \ ,$

  so

  $.11\delta_1 + .89\delta_0 \succ .11\big((10/11)\delta_5 + (1/11)\delta_0\big) + .89\delta_0 = .1\delta_5 + .9\delta_0$

  On next slide, horizontal axis is prob (0) and vertical is prob (5) so with EU
  slope of indifference curve should be $\big[u(1) - u(0)\big] / \big[u(5) - u(1)\big]$

- Horizontal axis measures Pr (0), vertical Pr (5).
- With the "Allais" choices, A preferred to B so the indifference curve through A is steep, and D preferred to C so the indifference curve there is flatter.
- But independence axiom says the slope should be constant.

- This is a "common-consequence" Allais paradox: In both A and B there is probability .89 of 1M, in C and D this common probability is removed.

- There are also "common ratio" Allais paradoxes such as

$$Z = \{0, 3000, 4000\},$$

$$A = .8\delta_{4000} + .2\delta_0, B = \delta_{3000},$$

$$C = .2\delta_{4000} + .8\delta_0, D = .25\delta_{3000} + .75\delta_0$$

- Some people say $B \succ A$ and $C \succ D$ but $C = .25A + .75\delta_0$ and

$$D = .25B + .75\delta_0.$$

- Original Allais and Kahmenan-Tversky experiments for large hypothetical stakes, but Allais-type choice pairs are also observed with small stake real money lab experiments.

- However, the choices have to be "carefully designed."

- In particular, it matters if one of the alternatives has a sure payoff.   Conlisk *AER* [1989] reports far fewer Allais-type violations of independence when none of the four options has a certain payoff, and the Harless and Camerer *Ema* [1994] meta-analysis/survey concludes "*in the triangle interior, however, EU manages a miraculous recovery.*"

- But Allais-type choices do definitely occur in the lab, and there are related anomalies in field data.

- For example Kahneman and Tversky *Ema* [1979] say that while people typically underweight outcomes that are probable vs. ones that are certain, they overweight very low probabilities:

Risk Tolerant at Low Probabilities

PROBLEM 7:

A: (6,000, .45), B: (3,000, .90).

$N = 66$ [14] [86]*

PROBLEM 8:

C: (6,000, .001), D: (3,000, .002).

$N = 66$ [73]* [27]

Here the lotteries are all *p* of stated prize, *1-p* of 0 so

$$A = .45\delta_{6000} + .55\delta_0, B = .9\delta_{3000} + .1\delta_0,$$

$$C = .001\delta_{6000} + .999\delta_0, D = .002\delta_{3000} + .998\delta_0$$

Prospect Theory

- Reference-dependent preferences for changes to current wealth (to accommodate "loss aversion."

- Let *r* denote a fixed reference point. ($r = 0$ *is the most common choice.*) *(If we let r depend on the menu, the model gets much more flexible.)*

- Utility function $u(z \mid r)$ typically neither concave nor convex but "S-shaped" about *r* so can be risk-loving for losses and risk-averse for gains.

- Preference over lotteries <u>non-linear in probabilities</u> via *probability weighting function* $\pi : [0,1] \to [0,1]$. (*pre-editing combines equal outcomes*.)

- Assume π is weakly increasing, continuous, $\pi(0) = 0, \pi(1) = 1$

- Evaluate lotteries by $U(p) = \sum_z \pi(p_z) u(z \mid r)$.

<u>Problem:</u> As KT realized, prospect theory allows violations of FOSD.

- For example let $\pi(x) = x^2, u(z) = z$. Let $A := \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2$, $B = \delta_1$.

- Most people prefer A to B but $u(B) = 1 > .25*1 + .25*2 = u(A)$.

- In general $\pi(p) = p$ is the only weighting function guaranteed to respect FOSD . (HW)

- This led to the development of *Rank Dependent Expected Utility* (*RDEU*) and *Cumulative Prospect Theory* (*CPT*).

- Assume *r=0,* and to simplify consider lotteries ℓ with only 3 outcomes, all positive: $0 < z_1 < z_2 < z_3$ with probabilities (p_1, p_2, p_3)

- Suppress *r* in the following.

- CPT says to set
$$U(\ell) = u(z_1) + [\pi(p_2 + p_3)](u(z_2) - u(z_1)) + \pi(p_3)(u(z_3) - u(z_2))$$

.

$$= [1 - \pi(p_2 + p_3)]u(z_1) + [\pi(p_2 + p_3) - \pi(p_3)]u(z_2) + \pi(p_3)u(z_3)$$
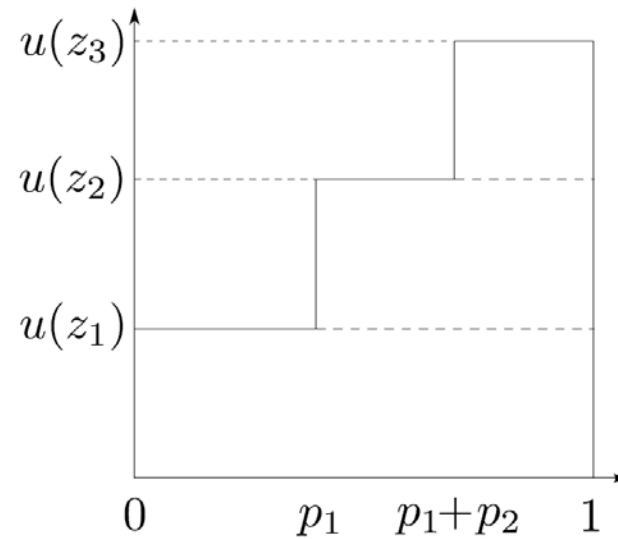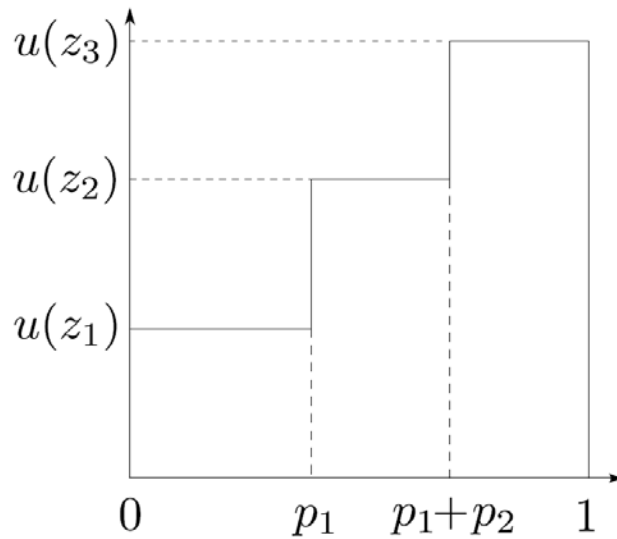
- Under EU, one way to compute expected utility is using "vertical stripes" as in Riemann integrals:

  EU= $p_1 u(z_1) + \big((p_2 + p_1) - p_1\big) u_2(z_2) + \big(1 - p_1 - p_2\big) u_3(z_3)$

- Another way is using "horizontal stripes" as in Lebesgue integration:

  $$EU = u(z_1) + \big(1 - p_1\big)\big(u(z_2) - u(z_1)\big) + \big(1 - p_1 - p_2\big)\big(u(z_3) - u(z_2)\big).$$

- CPT is like integrating the horizontal stripes, replacing *p* with $\pi(p)$; this gives preferences that respect FOSD. (*Remember that if $p \geq^{fosd} q$, p and q are the marginals of a pair of perfectly correlated lotteries with p always at least as large as q.*)

- CPT has been used to explain a wide range of lab and field data, e.g. why some people both buy insurance and lottery tickets.

- Most calibrations show an "S-shaped" weighting function, as in this figure from Tversky and Kahneman *J Risk Uncertainty* [1992]. (*real payoffs, n=25)*
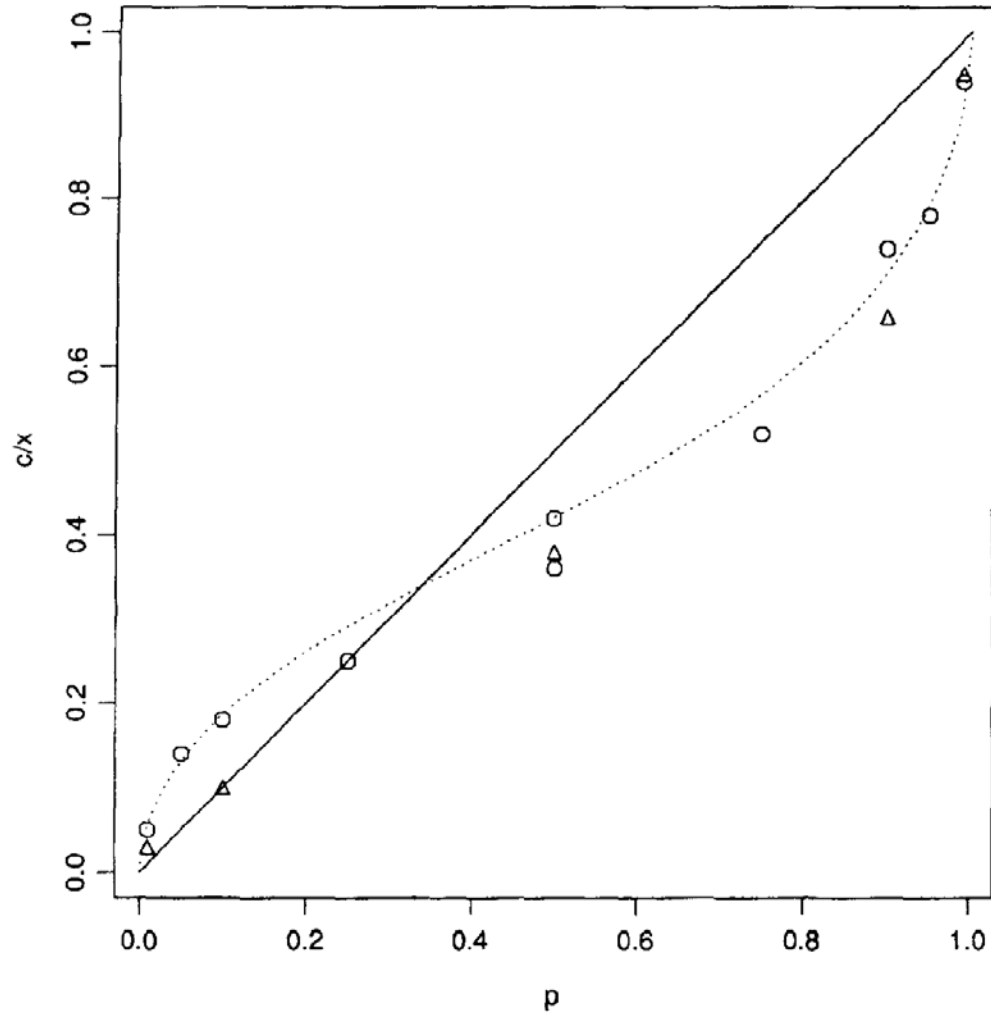
- Let's apply CPT to Allais' common-consequence choices Choose A or B, then C or D.

A: $1 million for sure $:= \delta_{1M}$, B: $.1\delta_{5M} + .89\delta_{1M} + .01\delta_0$

C: $.11\delta_{1M} + .89\delta_0$, D $.1\delta_{5M} + .9\delta_0$

$A \succ B$ implies $u(1) > \pi(.1)u(5) + [\pi(.99) - \pi(.1)]u(1) + [1 - \pi(.99)]u(0)$

$$\leftrightarrow \frac{1 - \pi(.99)}{\pi(.1)} > \frac{u(5) - u(1)}{u(1) - u(0)}.$$

$D \succ C$ implies $\pi(.1)u(5) + [1 - \pi(.1)]u(0) > \pi(.11)u(1) + [1 - \pi(.11)]u(0)$

$$\leftrightarrow \frac{\pi(.11) - \pi(.1)}{\pi(.1)} < \frac{u(5) - u(1)}{u(1) - u(0)}$$

So these choices occur if $\dfrac{\pi(.11) - \pi(.1)}{\pi(.1)} < \dfrac{u(5) - u(1)}{u(1) - u(0)} < \dfrac{1 - \pi(.99)}{\pi(.1)}.$

Now consider the KT lotteries

$$A = .45\delta_{6000} + .55\delta_0, B = .9\delta_{3000} + .1\delta_0,$$

$$C = .001\delta_{6000} + .999\delta_0, D = .002\delta_{3000} + .998\delta_0$$

Here a number of subjects said $B \succ A$ and $C \succ D$.

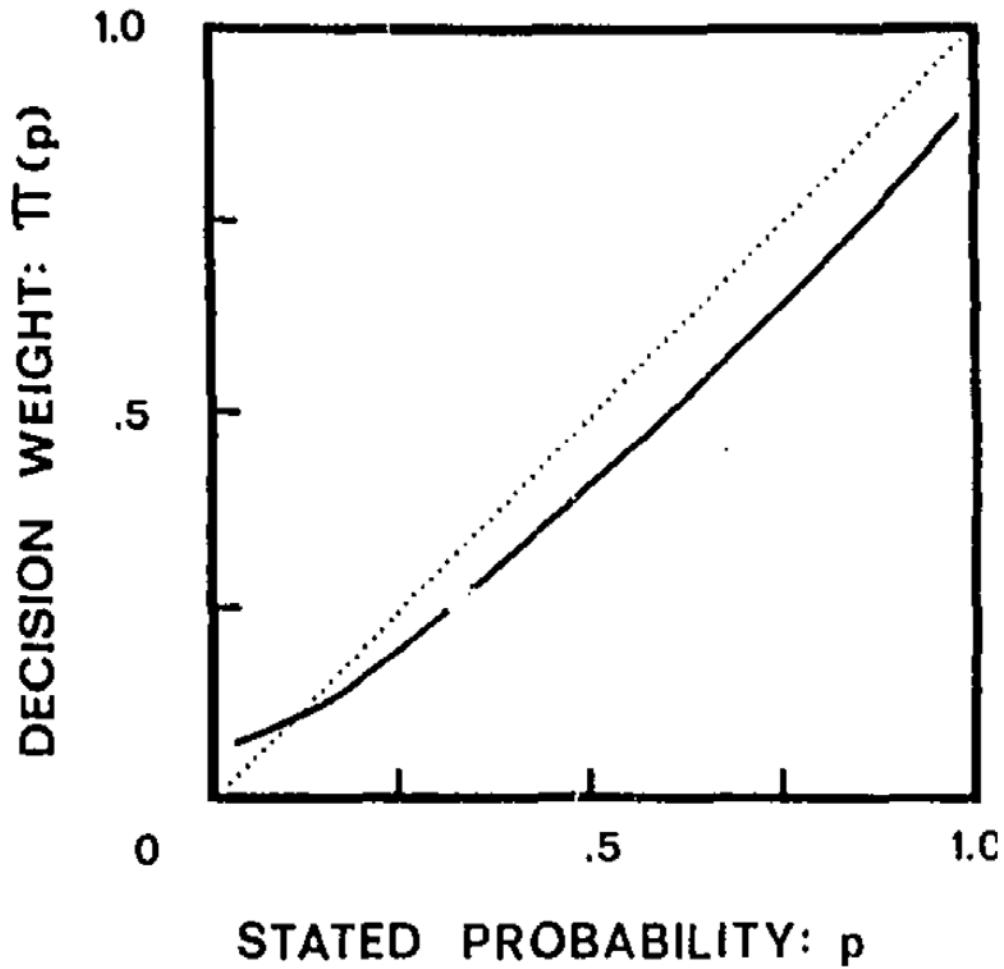This doesn't fit EU. KT explained it with this figure

FIGURE 4.—A hypothetical weighting function.

"Typical" CPT weighting functions also predict Allais-type violations in Conlisk's off-boundary version:

Lottery A	Lottery B
1/100 Chance of Nothing 89/100 Chance of $1 Million 10/100 Chance of $5 Million	71/100 Chance of Nothing 19/100 Chance Of $1 Million 10/100 Chance of $5 Million

Lottery A*	Lottery B*
2/100 Chance of Nothing 78/100 Chance of $1 Million 20/100 Chance of $5 Million	72/100 Chance of Nothing 8/100 Chance of $1 Million 20/100 Chance of $5 Million

CPT predicts the "paradoxical" choices $A \succ A^{*}, B^{*} \succ B$ if π is flatter around .28-.29 than around .98-.99, but this pair of choices isn't common.

Calibrating Prospect Theory

KT used (ad-hoc) functional forms

$$\pi(p) = \frac{p^{\gamma}}{\left(p^{\gamma} + (1-p)^{\gamma}\right)^{1/\gamma}}, \quad 0 < \gamma \leq 1$$

$$u(z) = z^{\alpha} \text{ (for z>0)}, \quad u(z) = -\lambda(-z)^{\alpha} \quad z < 0 \ .$$

For $0 < \alpha < 1$ the utility function is risk averse over gains and risk seeking over losses; $\lambda > 1$ says the decision maker is more sensitive to losses than gains.

The probability weighting function overweights extreme outcomes when their probabilities are low and underweights them when their probabilities are high- "S-shaped." Decreasing γ causes the weighting function to become more curved and to cross the diagonal farther to the right.

Overweighting rare gains can make an "otherwise risk averse" ($\alpha < 1$) agent risk seeking.

Prelec (*Ema* [1998]) shows that *if* preferences have a CPT representation, and if preferences satisfy some other "natural for CPT" assumptions, the probability weighting function must instead be $\pi(p) = \exp(-(-\ln p)^{\gamma})$.

Similar shape as KT's, but always satisfies $\pi(1/e) = 1/e \approx .37$. And is "subproportional" so fits some (KT!) data the KT specification doesn't.

Various representation theorems for CPT and related forms of RDEU, e.g. Yaari *Ema* [1987], Wakker and Tversky *J Risk Uncertainty* [1993].

Many studies have been done to estimate CPT parameters, usually by eliciting certain equivalents for a range of 2-outcome lotteries, and fitting a representative-agent model.

Typically find "S-shaped" probability weighting.

Neilson and Stowe *J Risk Uncertainty* [2002] take a critical view of the empirical support for CPT, working both with KT and Prelec probability weighting functions.

A risk-neutral EU agent is always indifferent between $p\delta_z + (1-p)\delta_0$ and its expected value δ_{pz} .

When $u(z) = z^\alpha$ then under CPT, if the agent is indifferent between a lottery and its expected value,

$$\pi(p)z^\alpha + (1-\pi(p))*0 = p^\alpha z^\alpha \text{ so } \pi(p) = p^\alpha .$$

If $\pi(p) = \dfrac{p^\gamma}{\left(p^\gamma + (1-p)^\gamma\right)^{1/\gamma}}$ as in KT, the indifference curve is

$$p^\alpha = \dfrac{p^\gamma}{\left(p^\gamma + (1-p)^\gamma\right)^{1/\gamma}} .$$

When α is higher than this the agent is less risk averse and prefers the gamble.

The next figure (Neilson Stowe Figure 3) plots these indifference curves for various parameter values, and notes the region where "Allais choices" occur for the original Allais problem, with hypothetical payoffs of $5million and $1 million, along with and various estimates of the parameters. (Their Figure 5 reports a similar exercise for some small-stakes real-payoff experiments.)

Points above the indifference curves are parameters where the agent likes to bet (is risk seeking) on unlikely gains.
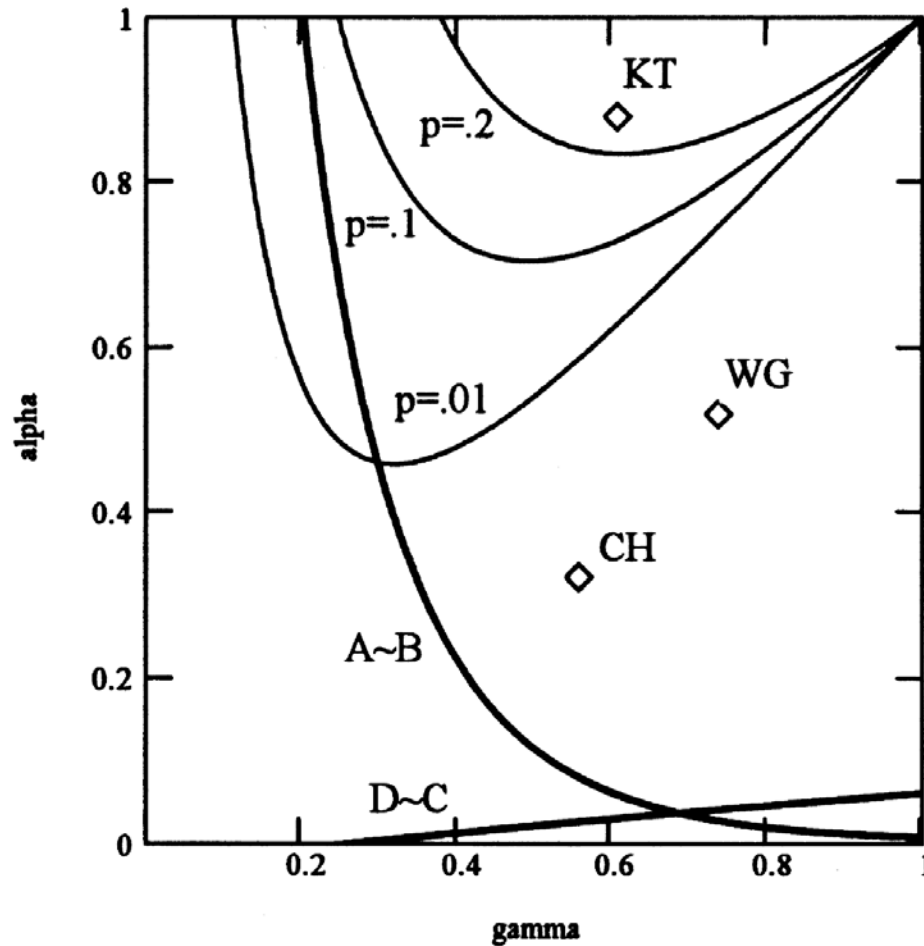
CH= Camerer and Ho *J. R. Uncert.* [1994], WG= Wu and Gonzales *Man. Sci.* [1996]. Allais paradox occurs in the region above $D \sim C$, below $A \sim B$; to also get risk-seeking on rare gains need *p* small and α not too low.

CPT allows someone to both bet on sufficiently unlikely gains and insure against sufficiently unlikely losses.

Neilson and Stowe do a similar calibration exercise here and find

"there are no parameter combinations that allow for both the desired gambling/insurance behavior and a series of choices made by a strong majority of subjects and reasonable risk premia...parameterizations that work well for choices over likely gains (that is, in the subset of the gains space where preferences are risk averse) do not work well for choices over unlikely gains (... where preferences are risk-seeking), and vice-versa...Taken as a whole, the analysis suggests that the functional forms proposed in the literature are not suitable for generalization to applied settings."

Almost all of these studies fit the same preferences to all subjects.

- Bruhin, Fehr-Duda and Epper *Ema* [2010] allow for heterogeneity.

- Elicited certain equivalents for 30-50 binary lotteries using a "random lottery pairs" procedure: at the end of the experiment 1 choice is implemented. (*a case of the independence axiom*)

- Didn't enforce a single cut off but in the Swiss lab they "nudged" people who entered more than one. Excluded subjects with more than 2 decision sheets with multiple switches (about 10% of them)

- Some lotteries framed as losses, but with show up fees subjects were guaranteed not to take a net loss (*can't impose real net losses in lab experiments even if they have high EV)*

Figure 1: Design of the Decision Sheet

Decision situation: 22							
	Option A		**Your Choice:**				**Option B** **Guaranteed payoff amounting to:**
1		A			o	B	20
2		A			o	B	19
3		A			o	B	18
4		A			o	B	17
5		A			o	B	16
6		A			o	B	15
7	A profit of CHF 20 with	A			o	B	14
8		A	o			B	13
9	probability 75%	A	o			B	12
10		A	o			B	11
11	and a profit of CHF 0 with	A	o			B	10
12		A	o			B	9
13	probability 25%	A	o			B	8
14		A	o			B	7
15		A	o			B	6
16		A	o			B	5
17		A	o			B	4
18		A	o			B	3
19		A	o			B	2
20		A	o			B	1

OK

23

- Using the subject's choices they identify the subject's certain equivalent for each lottery.

  - They estimate a mixture model with a number of CPT types.

  - Functional forms $\pi(p) = \dfrac{\delta p^{\gamma}}{\left(\delta p^{\gamma} + (1-p)^{\gamma}\right)^{1/\gamma}}$ (1 more parameter than KT) .

  - And specify $u(z) = \begin{cases} z^{\alpha} & z \geq 0 \\ -(-z)^{\beta} & z < 0 \end{cases}$

  - Add normally-distributed "mental errors" to get positive likelihood of all certain equivalents under all models, with standard error $\sigma_{ig} = \xi_i \left| range(g) \right|$ for individual i on gamble g. These errors could be from people trying to do CPT and not paying attention/making a computation error, or because they are following a different behavior rule.

Then do MLE: first compute the likelihood a given rule produced the observed choices, then sum over subjects to get overall likelihood.

In their notation, c is a behavior rule, g is a "gamble" or lottery, and ce is certain equivalent; hatted version is theoretical prediction.

Density of type c for individual i :

$$f(ce_i, \mathcal{G}; \theta_c, \xi_i) = \prod_{g=1}^{G} \frac{1}{\sigma_{ig}} \phi \left(\frac{ce_{ig} - \hat{ce}_g}{\sigma_{ig}} \right)$$

Overall log likelihood :

$$\ln \mathcal{L}(\Psi; ce, \mathcal{G}) = \sum_{i=1}^{N} \ln \sum_{c=1}^{C} \pi_c f(ce_i, \mathcal{G}; \theta_c, \xi_i)$$

3 data sets, 2 Swiss and 1 Chinese, run in 2005-06.

- Baseline treatments had expected payment of about 31 Swiss Francs/20 Chinese Yuan, plus a show up fee of 10 Swiss Francs/ 20 Chinese Yuan.

- They conclude that in all 3 cases the data can be well modelled using three groups, with 20% expected utility types and 30% +50% for two CPT groups, one with "moderately" S-shaped probability weighting and one with a more pronounced version. (*almost every individual has a clear assignment to one of these groups*)

- Swiss data sets and Chinese data very different in the gains domain: one of Chinese groups has a "very optimistic" probability weighting, with more overweighting of unlikely gains, and one of the Swiss is "pessimistic." (They link this to claims that Chinese are more risk-taking in general.)

In Zurich data the EU types have low "mental errors": median ξ about .06. Zurich CPT types had median error .12 and .13 in loss and gain domain respectively. For a typical lottery range of 20 Euros this is a std. deviation of 2.5.

Estimates for the Beijing subjects show a bit more noise.

In a high-stakes Chinese session (about a student's monthly income) the estimated probability weighting functions become less "optimistic."

Epper and Fehr-Duda *An. Review Econ.* [2012]

"Stake dependence potentially constitutes a serious challenge to rank-dependent models if it were to emerge as a robust feature of behavior. Clearly, more data are needed here.

They expect delayed payments to change things too (as has been shown in some lab experiments- fewer Allais-type choice when the lotteries have delayed payoffs):

"Delay dependence is particularly important because the majority of real-world decisions are both risky and delayed."

Yet still confirm non-linear probability weighting:

"probability dependence is not just a robust feature of laboratory data, but also provides a rationale for a host of puzzling real-world phenomena."

Discuss how to evaluate CPT*, and what it means to "explain" or "provide a rationale."*

Bruhin et al, like most past calibrations of CPT, elicit certain equivalents.

Andreoni-Sprenger *NBER* [2011] elicit both certain equivalents and "uncertain equivalents": Suppose $\ell = p * \delta_x + (1-p)\delta_y, \ \ y > x$. For what q is

$$\ell \sim q * \delta_y + (1-q)\delta_0?$$

They find (in a representative-agent calibration) that the parameters that fit the certain equivalents don't fit the uncertain equivalents.

CPT and gambling: for a wide range of CPT specifications at any wealth level there is always some unfavorable small binomial gamble that the agent likes.

This implies that a "naïve" agent- one who misforecasts his future decisions- who needs to decide when to stop a negative-drift diffusion (and claim the current states as payoff) will never stop, because a simple stopping rule with asymmetric upper and lower boundaries can generate such lotteries. (Ebert and Strack *AER* [2015]).   Conversely a "sophisticated" CPT agent typically declines to start diffusions with high positive drift. (Ebert and Strack [2016]).

And a firm who knows the agent's CPT preferences can design lotteries the consumer will buy that have arbitrarily high expected losses. (Avezedo and Gottlieb *JET* [2012].)

- Take-aways for now:

    - Substantial evidence for some failures of expected utility theory.

    - CPT with an S-shaped probability weighting fits many experiments and is the second-most- used (after EU) theory of risk preferences.

    - But it is far from perfect and should not be adopted uncritically.

Readings for next time: Bernheim-Sprenger [2017] *on testing rank dependence and Strzalecki 7.1, 7.6, MWG 6F on EU with subjective probability*

- **Appendix:** **T**he choices in KT [1979] problems 7 and 8 can't be fit with the one-parameter KT probability weighting for any utility function:

They would need $\dfrac{\pi(.9)}{\pi(.45)} > \dfrac{u(6000)}{u(3000)} > \dfrac{\pi(.002)}{\pi(.001)}$, not possible when in KT.

To accommodate all such examples, we need the property Prelec called "subproportionality":

SUBPROPORTIONALITY: *For any* $0 < \lambda < 1$, $p \neq q$, $(x, p) \sim (y, q)$ *implies*: $(y, \lambda q) > (x, \lambda p)$ *if* $y > x > 0$, *and* $(y, \lambda q) < (x, \lambda p)$ *if* $0 > x > y$.