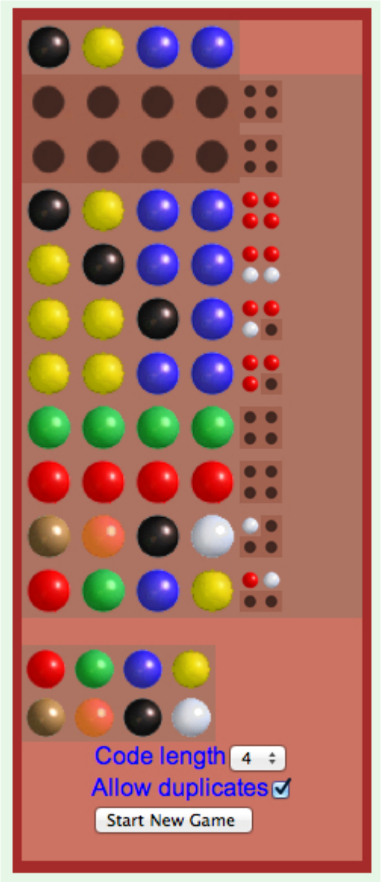
Critical Tests Among Models of Risky Decision Making

Michael H. Birnbaum

Konstanz, September, 2019

Critical Test



Critical Tests are Theorems of One Model that are Violated by Another Model

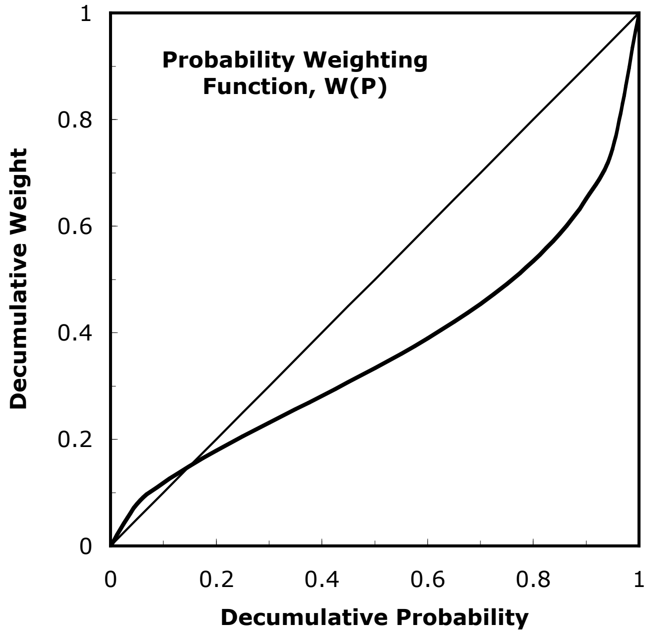
- This approach has advantages over tests or comparisons of fit.
- It is not the same as "axiom testing."
- Use model-fitting to rival model to predict where to find violations of theorems deduced from model tested.

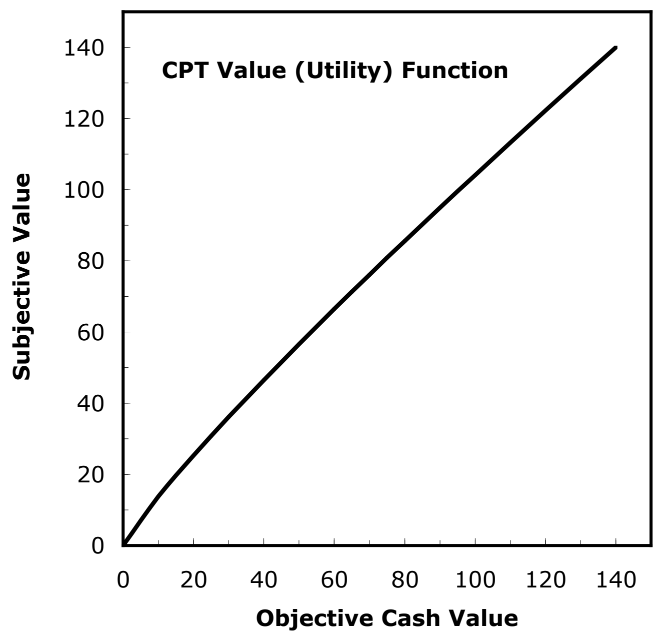
Outline

- I will discuss critical properties that test between nonnested theories: CPT and TAX.
- Lexicographic Semiorders vs. family of transitive, integrative models (including CPT and TAX).
- Integrative Contrast Models (e.g., Regret, Majority Rule) vs. transitive, integrative models.

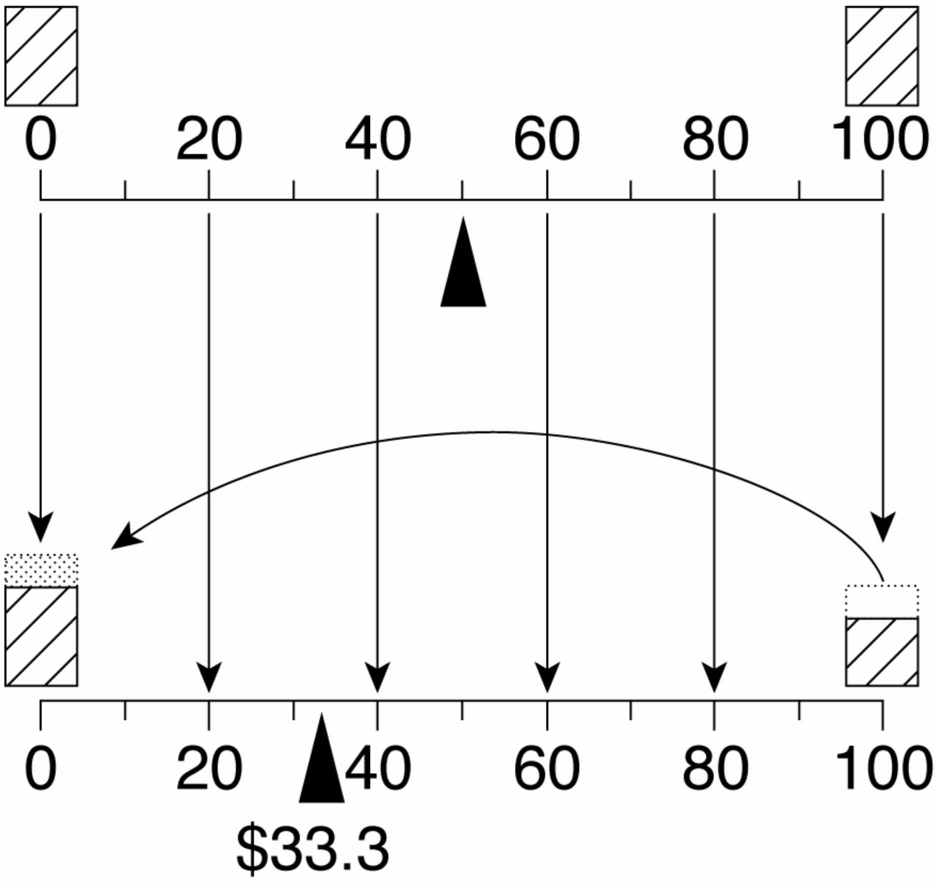
Cumulative Prospect Theory/ Rank-Dependent Utility (RDU)

$$CPU(G) = \sum_{i=1}^{n} [W(\sum_{j=1}^{i} p_j) - W(\sum_{j=1}^{i-1} p_j)] u(x_i)$$





TAX Model



“Prior” TAX Model

Assumptions:

$$G = (x, p; y, q; z, 1 - p - q)$$

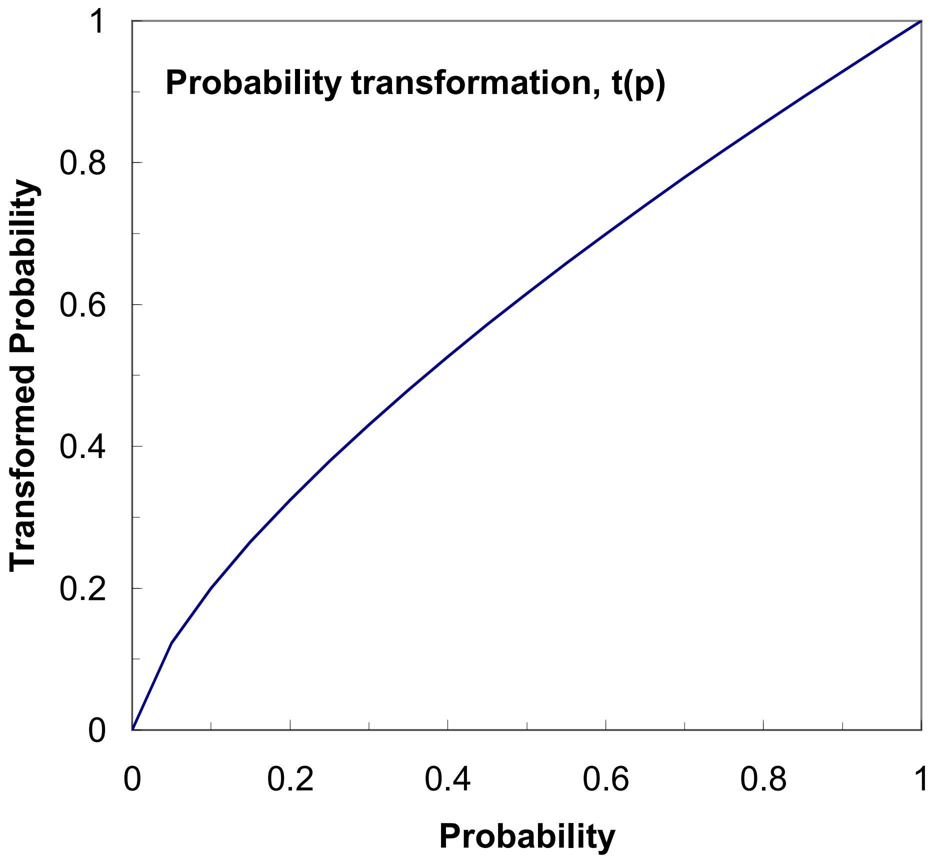
$$U(G) = \frac{Au(x) + Bu(y) + Cu(z)}{A + B + C}$$

$$A = t(p) - \delta t(p)/4 - \delta t(p)/4$$

$$B = t(q) - \delta t(q)/4 + \delta t(p)/4$$

$$C = t(1 - p - q) + \delta t(p)/4 + \delta t(q)/4$$

TAX Parameters

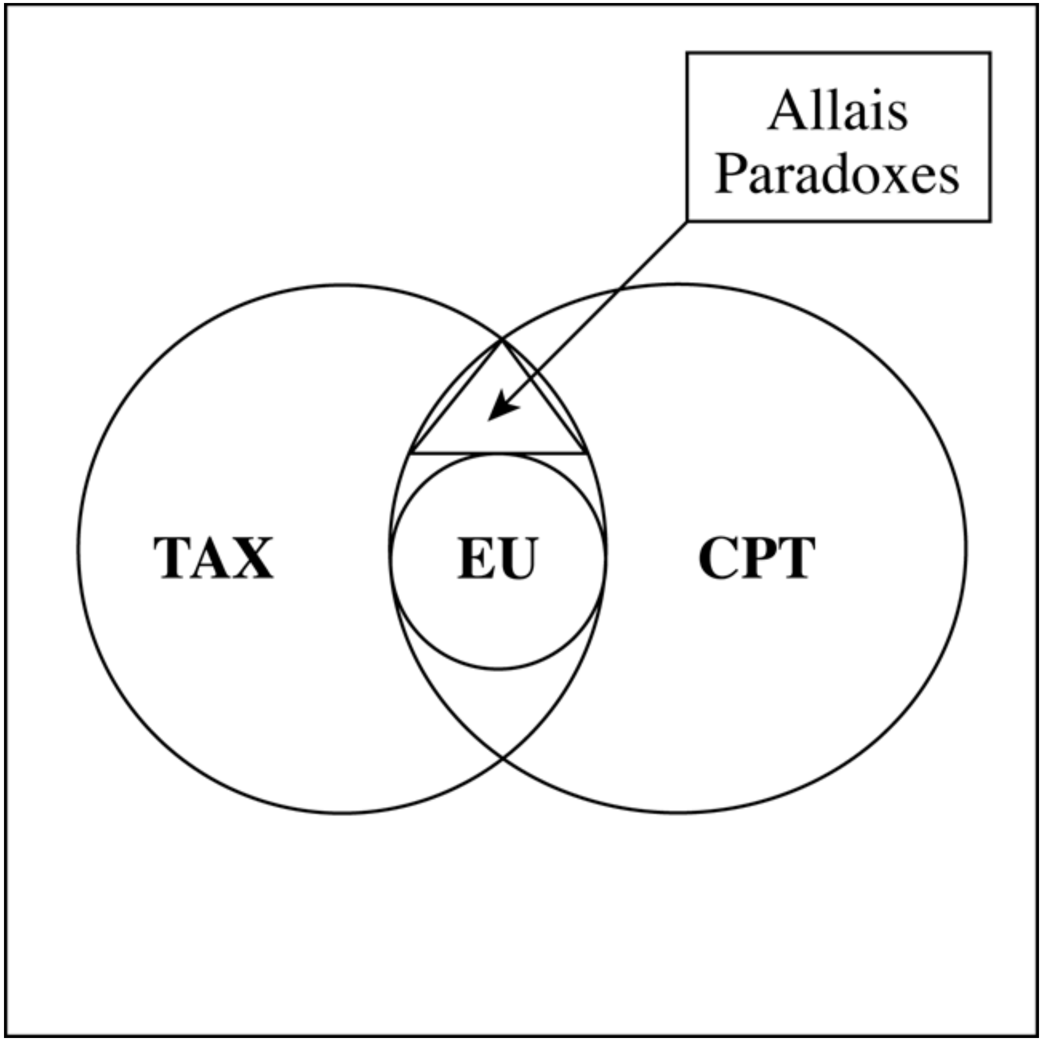


For $0 < x < \$150$
$u(x) = x$
Gives a decent approximation.
Risk aversion produced by δ.
$\delta = 1$.

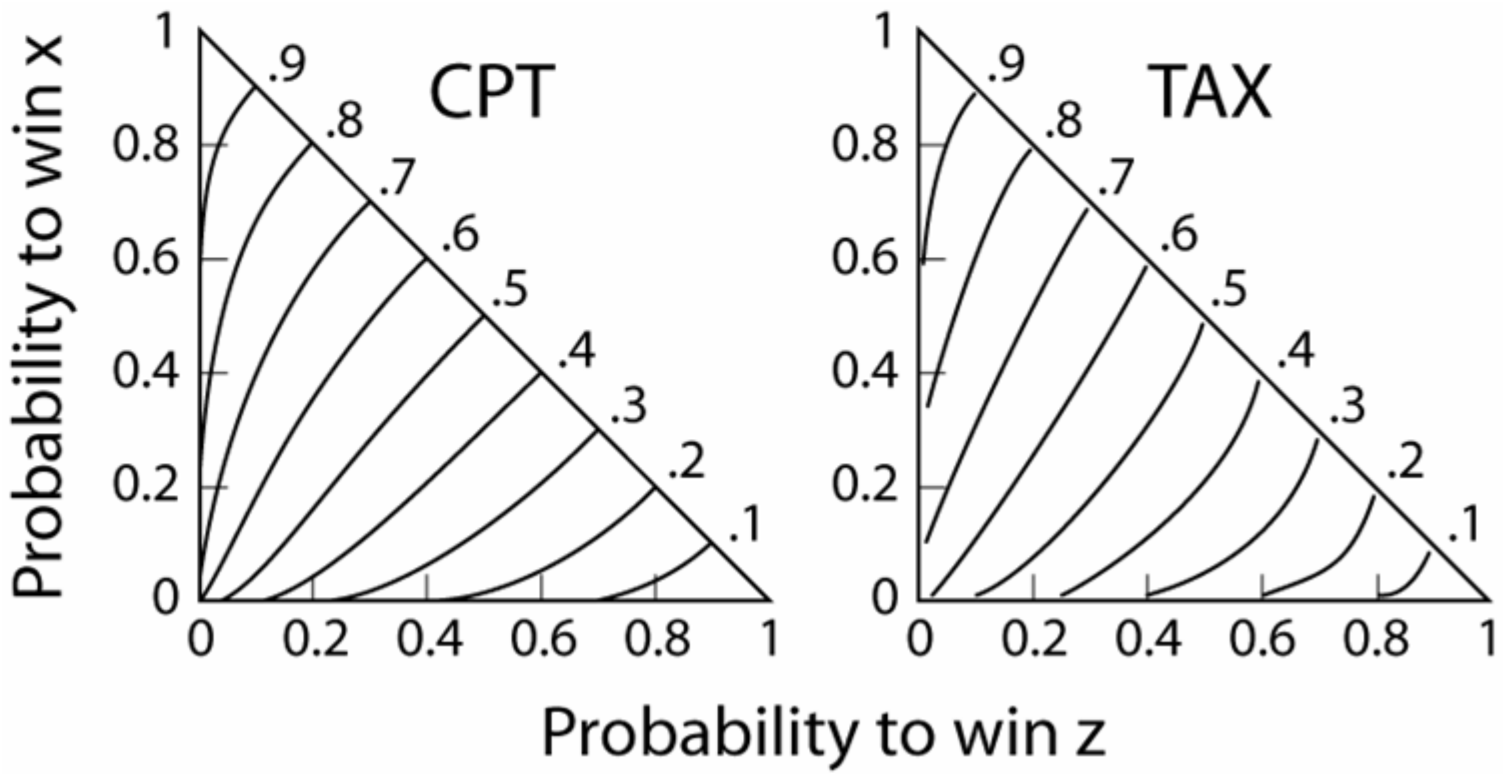
TAX and CPT nearly identical for binary (two-branch) gambles

- CE (x, p; y) is an inverse-S function of p according to BOTH TAX and CPT, given their typical parameters.
- Therefore, there is little point trying to distinguish these models with binary gambles.

Non-nested Models



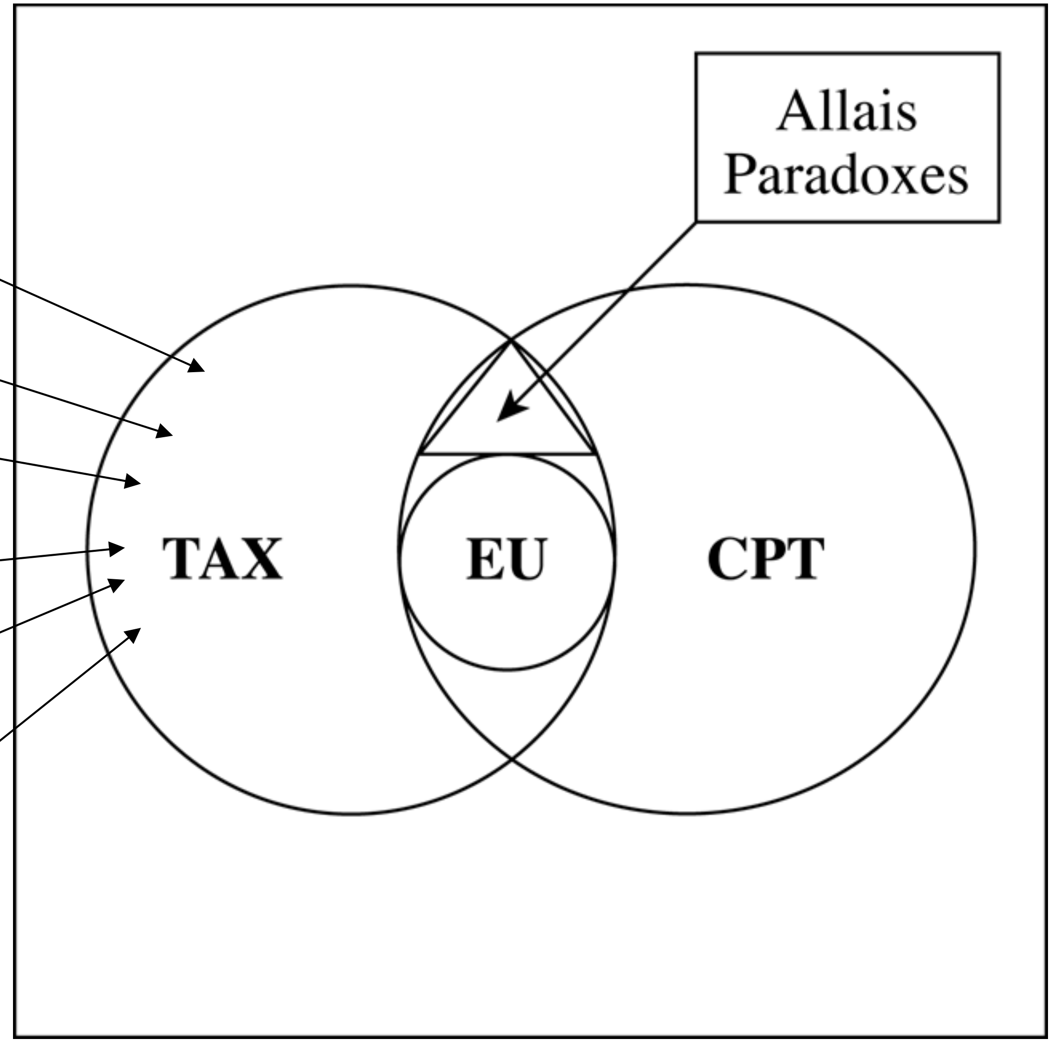
CPT and TAX nearly identical inside the prob. simplex



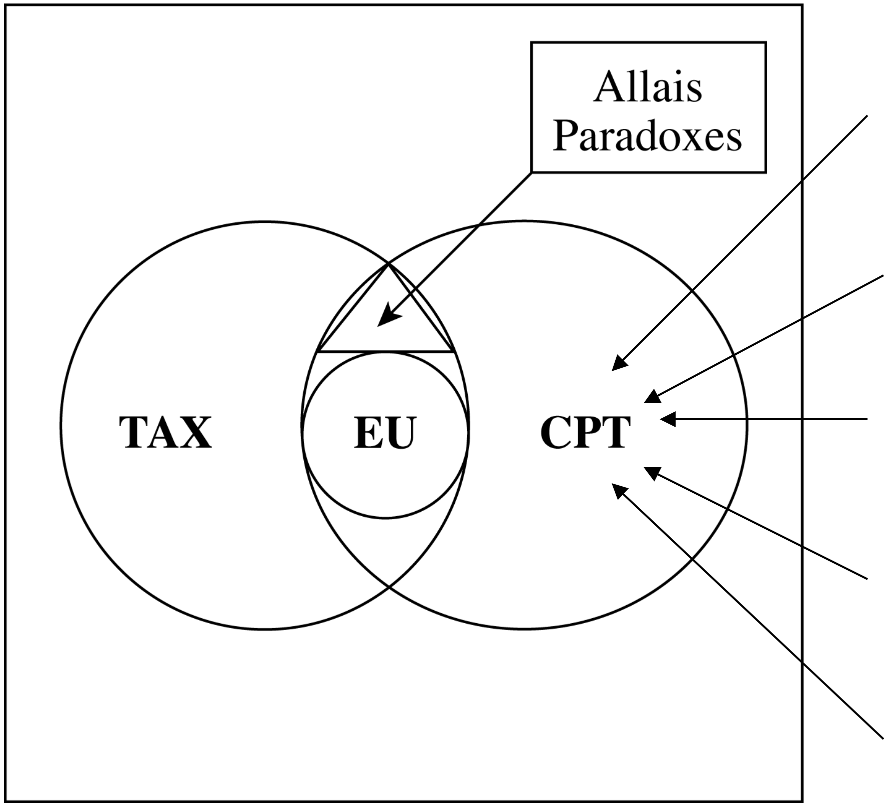
Testing CPT

TAX:Violations of:

- Coalescing
- Stochastic Dominance
- Lower Cum. Independence
- Upper Cumulative Independence
- Upper Tail Independence
- Gain-Loss Separability



Testing TAX Model



CPT: Violations of:

- 4-Distribution Independence (RS')
- 3-Lower Distribution Independence
- 3-2 Lower Distribution Independence
- 3-Upper Distribution Independence (RS')
- Res. Branch Indep (RS')

Stochastic Dominance

- A test between CPT and TAX:

$G = (x, p; y, q; z)$ vs. $F = (x, p - s; y', s; z)$

Note that this recipe uses 4 distinct consequences: $x > y' > y > z > 0$; outside the probability simplex defined on three consequences.

CPT $\Rightarrow$ choose *G*, TAX $\Rightarrow$ choose *F*

Test if violations due to "error."

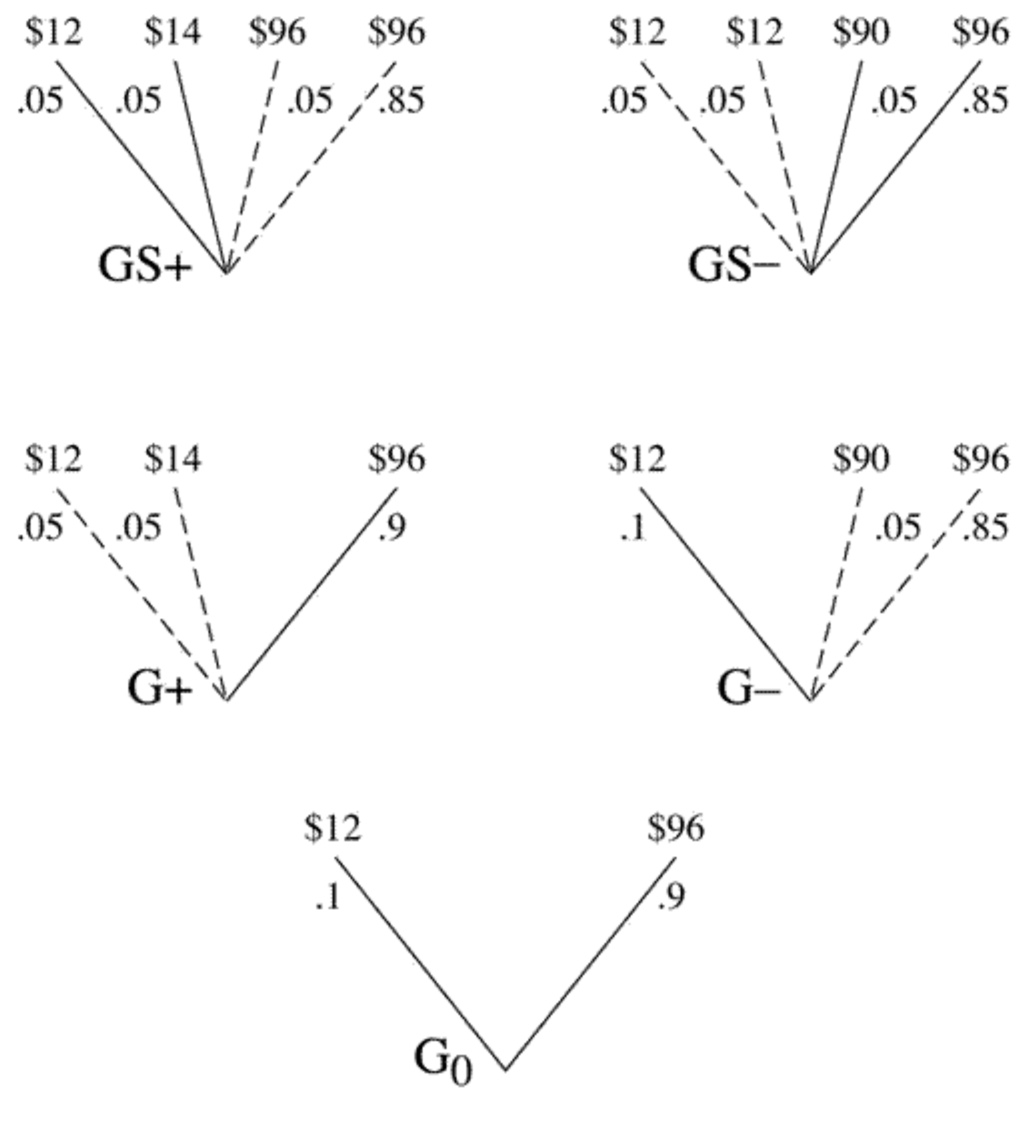
Error Model Assumptions

- Each choice pattern in an experiment has a true probability, *p*, and each choice has an error rate, *e*.
- The error rate is estimated from inconsistency of response to the same choice by same person over repetitions. The "true" *p* is then estimated from consistent (repeated) responses to same question.

Violations of Stochastic Dominance

A: 5 tickets to win $12 5 tickets to win $14 90 tickets to win $96	**B: 10 tickets to win $12 5 tickets to win $90 85 tickets to win $96**

122 Undergrads: 59% TWO violations (BB)
28% Pref Reversals (AB or BA)
Estimates: $e = 0.19$; $p = 0.85$
170 Experts: 35% repeated violations
31% Reversals
Estimates: $e = 0.20$; $p = 0.50$


$12
$14
$96
$96
.05
.05
.05
.85
GS+
$12
$12
$90
$96
.05
.05
.05
.85
GS−
$12
$14
$96
.05
.05
.9
G+
$12
$90
$96
.1
.05
.85
G−
$12
$96
.1
.9
G0

42 Studies of Stochastic Dominance, *n* = 12,152

- Large effects of splitting vs. coalescing of branches
- Small effects of education, gender, study of decision science
- Very small effects of 15 probability formats and request to justify choices.
- Miniscule effects of event framing (framed vs unframed)

Allais Paradox Dissection

	Restricted Branch Independence	
Coalescing	Satisfied	Violated
Satisfied	EU, PT*,CPT*	CPT
Violated	PT	TAX

Summary: Prospect Theories not Descriptive

- Violations of Coalescing
- Violations of Stochastic Dominance
- Violations of Gain-Loss Separability
- Dissection of Allais Paradoxes: viols of coalescing and restricted branch independence; RBI violations opposite of Allais paradox; opposite CPT.

Results: CPT makes wrong predictions for all 12 tests

- Can CPT be saved by using different formats for presentation?
- Violations of coalescing, stochastic dominance, lower and upper cumulative independence replicated with 14 different formats and thousands of participants.
- See Birnbaum, Psych Review 2008, & papers 2008-2017 in JDM.

Lexicographic Semiorders

- Intransitive Preference.
- Priority heuristic of Brandstaetter, Gigerenzer & Hertwig is a variant of LS, plus some additional features.
- In this class of models, people do not integrate information or have interactions such as the probability X prize interaction in family of integrative, transitive models (CPT, TAX, GDU, EU and others)

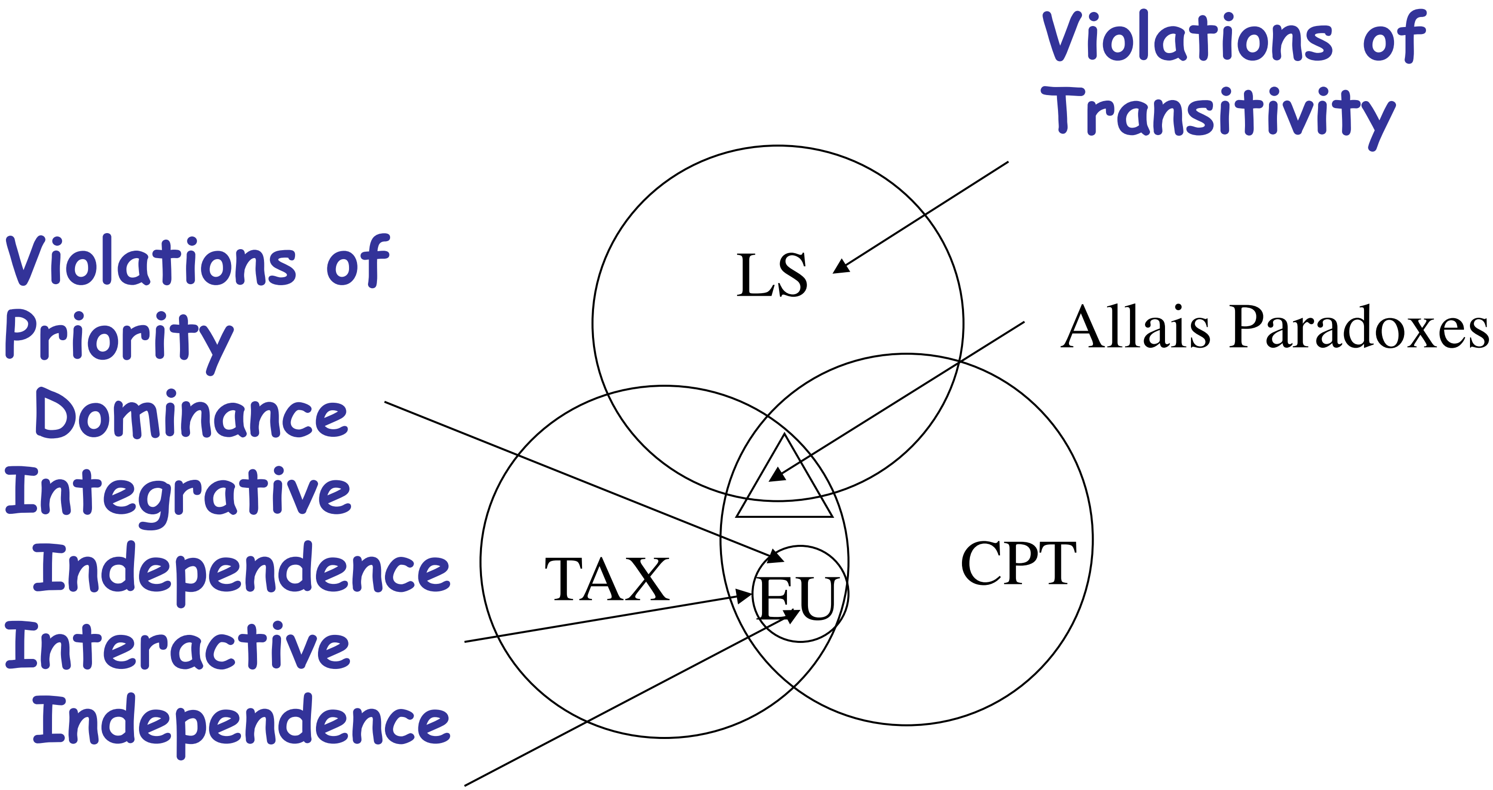
LPH LS: $G = (x, p; y)$ $F = (x', q; y')$

- If $(y - y' > \Delta)$ choose G
-     Else if $(y' - y > \Delta)$ choose F
- Else if $(p - q > \delta)$ choose G
-     Else if $(q - p > \delta)$ choose F
- Else if $(x - x' > 0)$ choose G
-     Else if $(x' - x > 0)$ choose F
- Else choose randomly

Family of LS

- In two-branch gambles, G = (x, p; y), there are three dimensions: L = lowest outcome (y), P = probability (p), and H = highest outcome (x).
- There are 6 orders in which one might consider the dimensions: LPH, LHP, PLH, PHL, HPL, HLP.
- In addition, there are two threshold parameters (for the first two dimensions).

Testing Lexicographic Semiorder Models



New Critical Tests distinguishing family of LS from {TAX, CPT, EU}

- *Dimension Interaction:* Decision should be independent of any dimension that has the same value in both alternatives.
- *Dimension Integration:* indecisive differences cannot add up to be decisive.
- *Priority Dominance:* if a difference is decisive, no effect of other dimensions.

Taxonomy of choice models

	Transitive	Intransitive
Interactive & Integrative	EU, CPT, TAX	Regret, Majority Rule
Non-interactive & Integrative	Additive, CWA	Additive Diffs, SDM
Not interactive or integrative	1-dim.	LS, PH*

Dimension Interaction

Risky	Safe	TAX	LPH	HPL
($95,.1;$5)	($55,.1;$20)	S	S	R
($95,.99;$5)	($55,.99;$20)	R	S	R

Family of LS

- 6 Orders: LPH, LHP, PLH, PHL, HPL, HLP.
- There are 3 ranges for each of two parameters, making 9 combinations of parameter ranges.
- There are 6 X 9 = 54 LS models.
- But all models predict SS, RR, or ??.

Results: Interaction n = 153

Risky	Safe	% Safe	Est. p
($95,.1;$5)	($55,.1;$20)	71%	.76
($95,.99;$5)	($55,.99;$20)	17%	.04

Analysis of Interaction

- Estimated probabilities:
- P(SS) = 0 (prior PH)
- P(SR) = 0.75 (prior TAX)
- P(RS) = 0
- P(RR) = 0.25
- Priority Heuristic: Predicts SS

Probability Mixture Model

- Suppose each person uses a LS on any trial, but randomly switches from one order to another and one set of parameters to another.
- But any mixture of LS is a mix of SS, RR, and ??. So no LS mixture model explains SR or RS.

Results: Dimension Integration

- Data strongly violate independence property of LS family
- Data are consistent instead with dimension integration. Two small, indecisive effects can combine to reverse preferences.
- Observed with all pairs of 2 dims.

Studies of Transitivity

- LS models violate transitivity: A > B and B > C implies A > C.
- Birnbaum & Gutierrez (2007) tested transitivity using Tversky's gambles, using typical methods for display of choices.
- Text displays and pie charts with and without numerical probabilities. Similar results with all 3 procedures.

Replication of Tversky (‘69) with Roman Gutierrez

- 3 Studies used Tversky’s 5 gambles, formatted with tickets and pie charts.
- Exp 1, n = 251, tested via computers.

Three of Tversky's (1969) Gambles

- *A* = ($5.00, 0.29; $0)
- *C* = ($4.50, 0.38; $0)
- *E* = ($4.00, 0.46; $0)

Priority Heurisitc Predicts:

A preferred to *C*; *C* preferred to *E*,

But *E* preferred to *A*. Intransitive.

TAX (prior): *E* > *C* > *A*

Response Combinations

Notation	(A, C)	(C, E)	(E, A)	
000	A	C	E	* PH
001	A	C	A	
010	A	E	E	
011	A	E	A	
100	C	C	E	
101	C	C	A	
110	C	E	E	TAX
111	C	E	A	*

Results-ACE

pattern	Rep 1	Rep 2	Both
000 (PH)	10	21	5
001	11	13	9
010	14	23	1
011	7	1	0
100	16	19	4
101	4	3	1
110 (TAX)	176	154	133
111	13	17	3
sum	251	251	156

Comments

- Results were surprisingly transitive.
- Differences: no pre-test, selection;
- Probability represented by # of tickets (100 per urn); similar results with pies.
- Regenwetter and colleagues: studies and new analyses (random utility definition of transitivity); they also conclude that evidence against transitivity is extremely weak.
- With Jeff Bahra: individual data also transitive

Summary

- Priority Heuristic model's predicted violations of transitivity are rare.
- Dimension Interaction violates any member of LS models including PH.
- Dimension Integration violates any LS model including PH.
- Evidence of Interaction and Integration compatible with models like EU, CPT, TAX.
- Birnbaum, J. Mathematical Psych. 2010.

Integrative Contrast Models

- Family of Integrative Contrast Models
- Special Cases: Regret Theory, Majority Rule (aka Most Probable Winner)
- Predicted Intransitivity: Forward and Reverse Cycles
- Birnbaum, M. H., & Diecidue, E. (2015). Testing a class of models that includes majority rule and regret theories: Transitivity, recycling, and restricted branch independence. *Decision, 2,* 145-190.

Integrative, Interactive Contrast Models

$$A \succ B \Leftrightarrow \sum_{i=1}^{n} \phi(E_i)\psi(a_i, b_i)$$

$$A = (a_1, E_1; a_2, E_2; \ldots; a_n, E_n)$$

$$B = (b_1, E_1; b_2, E_2; \ldots; b_n, E_n)$$

Assumptions

$$\psi(a_i, b_i) = -\psi(b_i, a_i)$$

$$\psi(a_i, b_i) = 0 \Leftrightarrow a_i = b_i$$

$$Difference\ Model:$$

$$\psi(a_i, b_i) = f[u(a_i) - u(b_i)]$$

Special Cases

- Majority Rule (aka Most Probable Winner)
- Regret Theory
- Other models arise with different functions, *f*.

Regret Aversion

$$\psi[a,c] \geq \psi[a,b] + \psi[b,c], \quad u(a) > u(b) > u(c)$$

Regret Model

$$f[u(a)-u(b)] = |u(a)-u(b)|^{\beta}, \quad u(a) > u(b)$$

$$f[u(a)-u(b)] = -|u(a)-u(b)|^{\beta}, \quad u(b) > u(a)$$

$$\beta > 1$$

Majority Rule Model

$$f[u(a)-u(b)] = \begin{cases} 1 & if \;\; u(a) > u(b) \\ 0 & if \;\; u(a) = u(b) \\ -1 & if \;\; u(a) < u(b) \end{cases}$$

Predicted Intransitivity

- These models violate transitivity of preference
- Regret and MR cycle in opposite directions
- However, both REVERSE cycle under permutation over events; i.e., "juxtaposition."

Concrete Example

- Urn: 33 Red, 33White, 33 Blue
- One marble drawn randomly
- Prize depends on color drawn.
- *A* = ($4, $5, $6) means win $400 if Red, win $500 if White, $600 if Blue. (Study used values x 100).

Majority Rule Prediction

- A = (\$4, \$5, \$6)
- B = (\$5, \$7, \$3)
- C = (\$9, \$1, \$5)
- AB: choose B
- BC: choose C
- CA: choose A
- Notation: 222

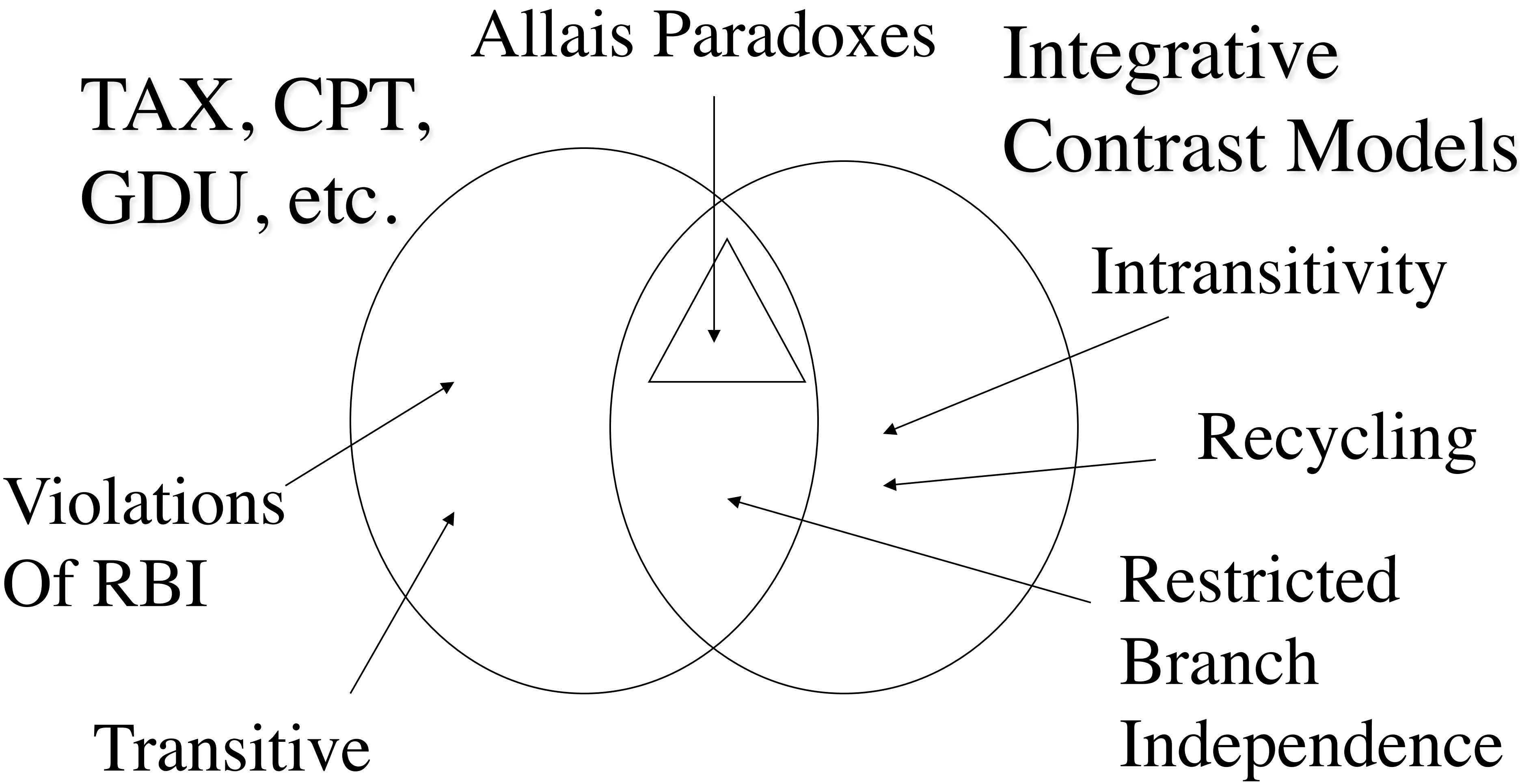
- A' = (\$6, \$4, \$5)
- B' = (\$5, \$7, \$3)
- C' = (\$1, \$5, \$9)
- A' B' : choose A'
- B' C' : choose B'
- C' A' : choose C'
- Notation: 111

Regret Prediction

- A = ($4, $5, $6)
- B = ($5, $7, $3)
- C = ($9, $1, $5)
- AB: choose A
- BC: choose B
- CA: choose C
- Notation: 111

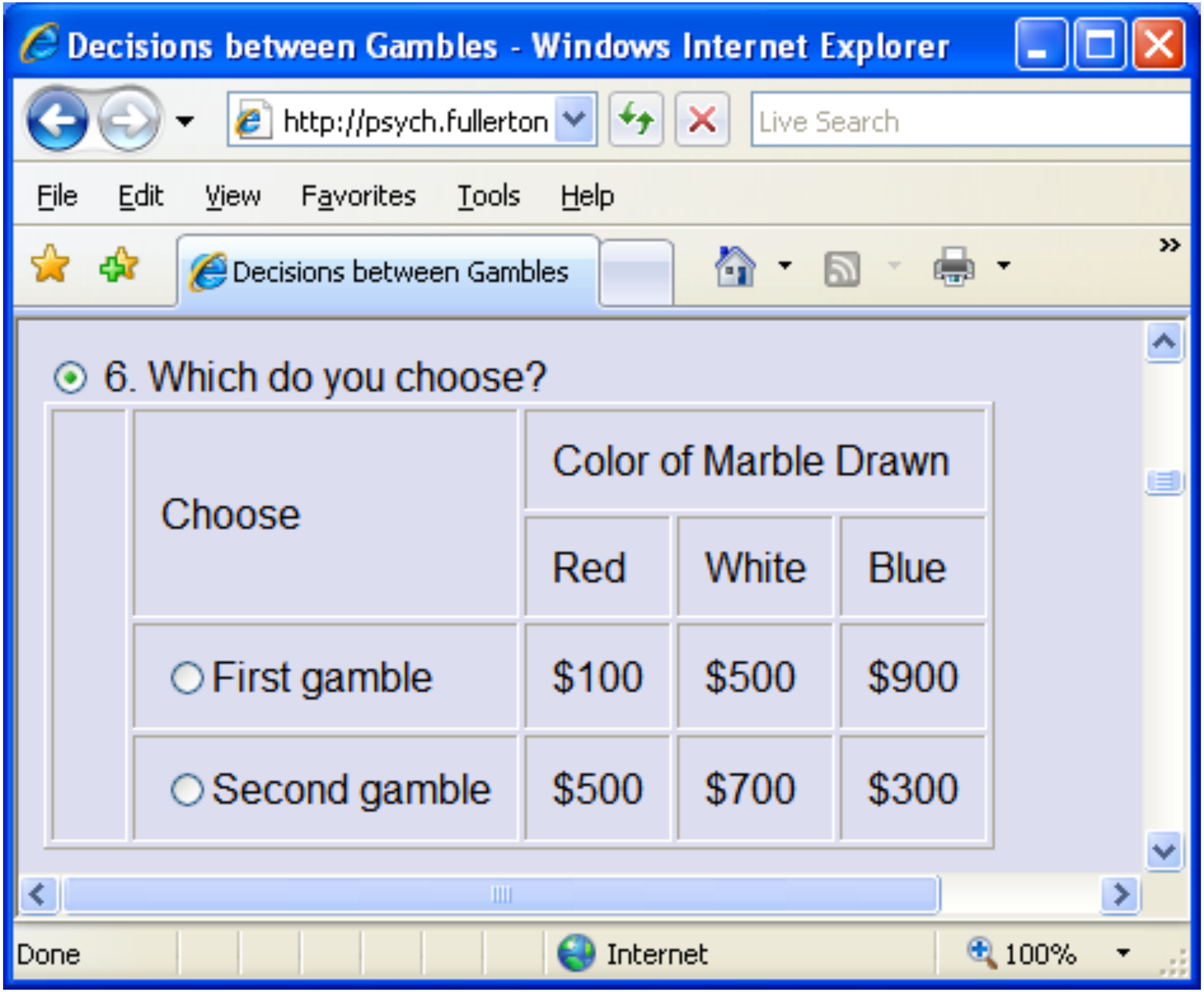
- A' = ($6, $4, $5)
- B' = ($5, $7, $3)
- C' = ($1, $5, $9)
- A' B' : choose B'
- B' C' : choose C'
- C' A' : choose A'
- Notation: 222

Non-Nested Models



Study

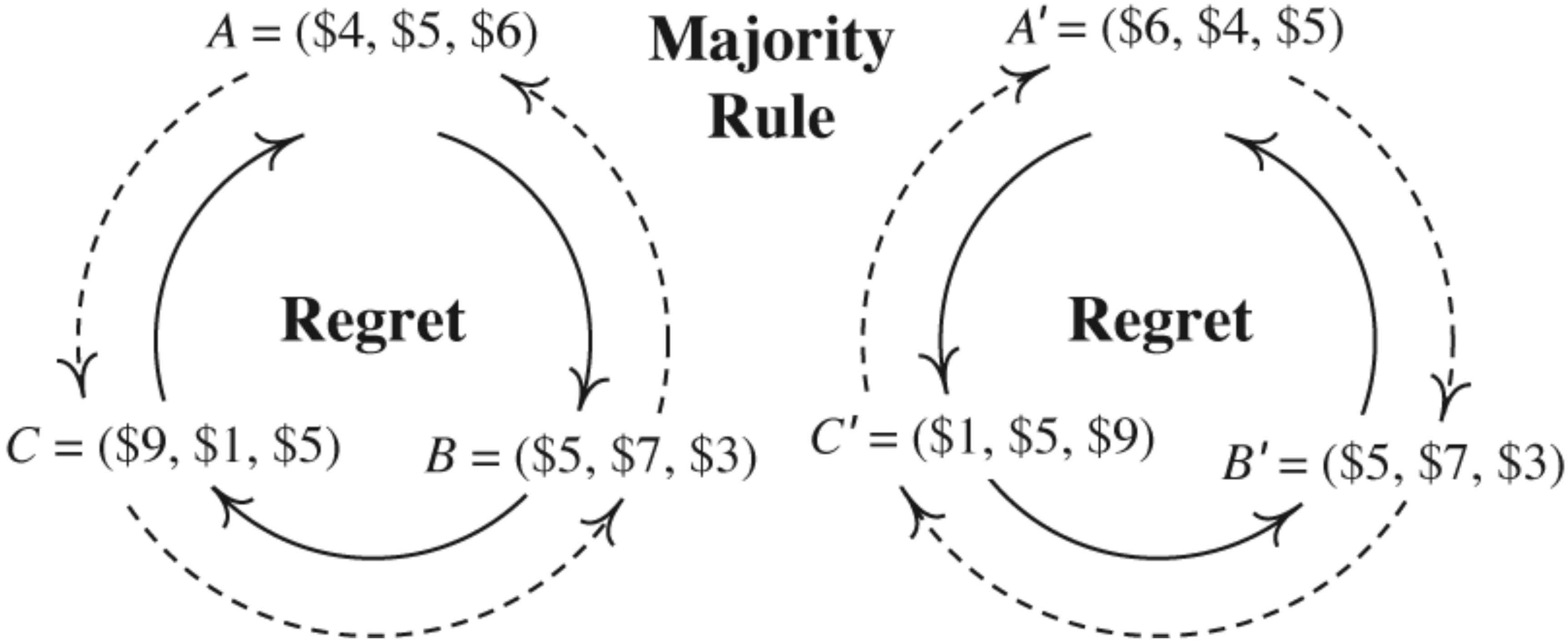
- 240 Undergraduates
- Tested via computers (browser)
- Clicked button to choose
- 30 choices (includes counterbalanced choices)
- 10 min. task, 30 choices repeated.



6. Which do you choose?

Choose	Color of Marble Drawn		
	Red	White	Blue
First gamble	$100	$500	$900
Second gamble	$500	$700	$300

Recycling Predictions of Regret and Majority Rule



Results

- Most people are transitive.
- Most common pattern is 112, pattern predicted by TAX with prior parameters.
- However, 2 people were perfectly consistent with MR on 24 choices (incl. Recycling pattern).
- No one fit Regret theory perfectly.

Results: Continued

- Among those few (est. ~10%) who cycle (intransitive), most have no regrets (i.e., they appear to satisfy MR).
- Systematic Violations of RBI.
- Suppose 5-10% of participants are intransitive. Do we think that they indeed use a different process? Can we increase the rate of intransitivity?

Conclusions

- Violations of transitivity predicted by regret, MR, LS appear to be infrequent.
- Violations of Integrative independence, priority dominance, interactive independence are frequent, contrary to family of LS, including the PH.
- "New paradoxes" rule out CPT and EU but are consistent with TAX.
- CPT, TAX, and EU are transitive and could have been refuted by systematic intransitivity, but data did not require rejection for most people.

Editing: Another way to become intransitive

- Original Prospect Theory— (KT 79). Maybe people edit some choices but not others.
- B, P, & L (1999). Perhaps people would "see" stochastic dominance in simple choices and not in more complex ones.
- This theory properly tested in JDM (2016).

Transitivity Analysis: R- scripts for Monte Carlo and Bootstrapping in TE models.

- Birnbaum, M. H., Navarro-Martinez, D., Ungemach, C., Stewart, N. & Quispe-Torreblanca, E. G. (2016). Risky decision making: Testing for violations of transitivity predicted by an editing mechanism. *Judgment and Decision Making, 11,* 75-91.
- http://journal.sjdm.org/vol11.1.html

Results

- Results did not provide much evidence (if any) for intransitivity predicted by editing.
- People violated stochastic dominance too often even in the "easy" choices we thought would be "transparent" or at least "translucent"
- Paper shows proper methods for analysis of transitivity/intransitivity.