Sources of Variability in Choice Tasks and Criteria for Evaluation of Error Models—TE part 2

Michael H. Birnbaum

California State University, Fullerton

Summer School Workshop, Konstanz

September, 2019

Outline

- Sources of Variability in Choice Tasks
- Criteria for Evaluation of Theories of Variations (Error Models are theories)
- Response Independence: A property that distinguishes classes of models
- Example of TE with test refuting CPT (& EU) that cannot be explained by any error model known to me.

Sources of Variation

- Read information and remember it.
- Evaluate information
- Aggregate the information
- Compare Aggregations or Aggregate comparisons -- & remember results
- Map to a response system
- Change one’s mind about importance
- Motor Response

Criteria for Evaluation

- Philosophical criteria for theory: deductive, meaningful, predictive, causal, & general
- Empirically Testable: evidence to test and possibly reject the error model
- Additional considerations: separate components of error, neutrality with respect to issue to be tested.

Response Independence

- Do binary choice proportions contain all the relevant information, so we can reduce our analysis?
- For example: S vs. R, and S’ vs. R’:
- Does P(RS’) = P(R)P(S’)?
- Certain Random Preference or Random Utility models imply indep, TE models typically violate response independence.

Mixtures

- Different people may have different models, different functions, different parameters (group analysis)
- Same person may differ from time to time in models, functions, and parameters. (individual analysis)
- Mixtures ~ variations in “true” preferences.

Changes within a Person

- People change “true” preferences: new information, new context, new attitude, new considerations. These changes should be minimal within a controlled experimental session with a practiced participant.
- Idea: Preference reversals within a short session (“block” of trials) due to error.

Sources of Random Error

- Misread, misremember information.
- Variation in the evaluation of a consequence. E.g., $u(x) = f(x) + e$
- Neural aggregation. E.g., random walk model of comparision or contrast model.
- Variation in response mapping
- Psychomotor response error

Rival Error Models

- Model + error (Fechner, Thurstone, Luce, EU + error, etc.)
- Model with variable parameters
- Mixture of Models (including functions & parameters) with random selected preference
- TE models– variation sorted to "true" changes and "error".

TE models

- TE models like ANOVA. Requires at least 2 reps per person per session.
- *i*TET—requires multiple sessions (blocks) of trials.
- gTET—requires many people.
- Variations: 1 error (“tremble”), n errors for n item), or 2n errors (error depends on “true” state.

Some advantages of TE

- Models are testable. Two statistical tests: test the error model, and test the substantive model as a special case of the general model.
- Decompose variation into “true” and “error” parts. Estimate probabilities of “true” patterns.
- Like ANOVA: we need replicates to estimate error.

Key Assumptions

- Error is estimated from preference reversals WITHIN a session by the same practiced person.
- "True" variation is estimated from changes between sessions not attributable to error within sessions.

Limitations/Advantages

- Does not assume transitivity. Transitive model is a special case. Thurstone Case V is a special, special case of transitive TE, for example.
- TE models imply violations of response independence: people predicted to be more consistent than predicted by independence. (empirical results)
- Does not predict or explain sequential effects or violations of stationarity. Add random walk on parameters?

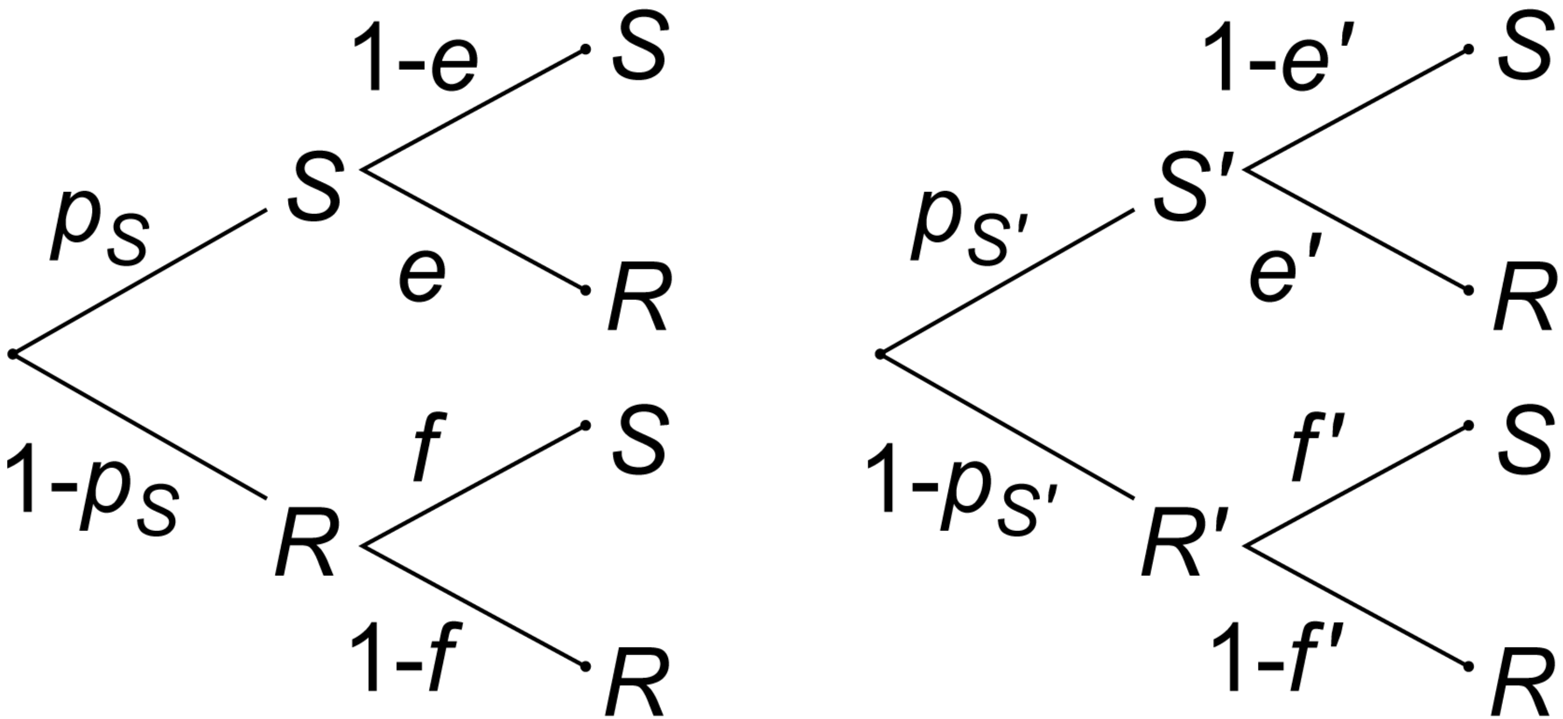
Critical Tests

- Critical tests are theorems of one model that can be violated by another: transitivity, first order stochastic dominance, restricted branch independence, coalescing, lower and upper cumulative independence, etc.
- Use parameters of rival model to design the test to generate predicted violations.

CPT implies FOSD

- First Order Stochastic Dominance: If Gamble F has at least as high (and sometimes higher) probability to win a prize greater or equal to t, for all t, than Gamble F dominates G by FOSD.
- CPT implies FOSD for ANY choice of monotonic utility function and ANY monotonic decumulative weighting function.

TE-4 model of 2 choice problems



Two Choices, Two Reps

- Two Choices, *S* vs. *R* and *S'* vs. *R'*
- Each choice problem has errors
- Data: 4 X 4 = 16 cells with 15 *df*
- Model: 4 "true" probs: *SS*' , *SR*' , *RS*' , *RR*'
- Model can have 1, 2, or 4 errors.
- For example, TE-4 model: 15 – 7 = 8 *df*

We test both TE models and special cases

- Fit and Test general TE model
- Test CPT & EU (substantive theory) as special cases of general TE model.

CPT and TAX are Special Cases of TE models

- The CPT model is a special case of TE Models.
- CPT Assumption: “true” patterns *SR*’ and *RS*’ can not occur.
- This model (CPT-TE-4) uses 5 parameters: 4 errors and p_S = “true” probability prefer the “Safe” gamble; $p_R = 1 - p_S$

Choice between *F* and *G*

- *F:* 90 tickets to win $96
  5 tickets to win $14
  5 tickets to win $12

- *G:* 85 tickets to win $96
  5 tickets to win $90
  10 tickets to win $12

Choice between *F’* and *G’*

- *F’*: 85 tickets to win $96
  5 tickets to win $96
  5 tickets to win $14
  5 tickets to win $12
- *G’*: 85 tickets to win $96
  5 tickets to win $90
  5 tickets to win $12
  5 tickets to win $12

Theories Compared

	FOSD holds	FOSD fails
FOSD holds (split form)	CPT & EU OPT + dom. detector	TAX, OPT + dom. Detector; no combination
FOSD fails (split form)		*OPT* + *combination, no dom. det.*

Testing CPT/EU versus TAX

- Both CPT and EU imply FOSD and coalescing.
- People should prefer F to G and prefer F’ to G’.
- TAX model with previous parameters predicts violation for F versus G and NO violation for F’ versus G’. With other parameters, TAX reduces to EU.

FOSD & Coalescing, n = 223

	FF'	*FG'*	*GF'*	*GG'*
FF'	**34**	4	22	0
FG'	2	**0**	1	1
GF'	27	3	**101**	11
GG'	2	2	11	**1**

FOSD & Coalescing

Model	TE-4	TE-2	TE-1	Indep.
FULL	7.32	9.39	18.52	45.80
TAX	15.38	15.38	19.58	
CPT/EU	123.47	123.47	682.46	

TE-2 Estimates

Model	P11	P12	P21	P22	e	e’
TE-2	.22	.00	.75	.02	.16	.09
TE_2 + TAX	.23	(0)	.77	(0)	.17	.12

Summary

- TE-4 model can be retained
- TE-2 model good approximation
- TE-1 (marginal, better than indep.)
- CPT/EU model rejected, even with 4 errors
- Two Violations: FOSD and Coalescing
- TAX model can be retained.

Random Parameters Cannot Save CPT or EU Models

- Under ANY parameters, & ANY monotonic utility and probability weighting functions, CPT and EU imply FOSD and Coalescing.
- Violations cannot be explained by variation in parameters.
- This *critical test* (FOSD) bypasses that approach to error (error in utility or weights).

Related Papers

- Birnbaum, M. H., Navarro-Martinez, D., Ungemach, C., Stewart, N. & Quispe-Torreblanca, E. G. (2016) in *JDM*.

http://journal.sjdm.org/15/15615a/jdm15615a.pdf

- Birnbaum, Schmidt, & Schneider (2017) in *JRU*

Website URL

- http://psych.fullerton.edu/mbirnbaum/birnbaum.htm
- Email: mbirnbaum@fullerton.edu

Also Available at my Website

- *Psych Review* Comment iid (2011)
- Birnbaum (2012). *JDM* Testing iid.
- Birnbaum & Bahra (2012a, 2012b): Testing transitivity in Linked Designs (2012) & Testing SD and RBI (2012) both papers in *JDM*
- Birnbaum (2013) in *JDM*
- Birnbaum & Diecidue (2015) in *Decision*
- Birnbaum, Schmidt, & Schneider (*JRU* in press) Allais paradoxes