Journal of Risk and Uncertainty, 18:321–325 (1999)

Reanalysis of the Chechile-Cooke Experiment: Correcting for Mismatched Gambles

RICHARD A. CHECHILE rchechil@emerald.tufts.edu
Tufts University, Medford, Massachusetts

R. DUNCAN LUCE
University of California, Irvine

Abstract

Chechile and Cooke (1997) experimentally tested a broad class of utility models subsumed under the Miyamoto (1988, 1992) generic utility theory. The Chechile and Cooke study required participants to match on each trial, a fully specified reference gamble to a partially specified comparison gamble by adjusting the probability of a win on the comparison gamble. The Chechile and Cooke experiment, however, contained a subset of trials which were intrinsically unmatchable. In such cases, the participants could only give an extreme probability (either 0 or 1). In this paper, those extreme trials were omitted and the results from the experiment reanalyzed. Despite the mismatch problem, the conclusions of the Chechile and Cooke experiment were again supported. For nine implementations of generic utility there is model failure due to the systematic variation of a parameter that should be a constant.

Key words: generic utility theory, subjective equivalence of gambles, context-dependent utility scaling
JEL Classification: C20, D81, C91

Chechile and Cooke (1997) tested the Miyamoto (1988, 1992) generic utility theory (GUT), for binary lotteries of a gain and a loss. The GUT class of models is very broad and includes subjective expected utility (von Neumann and Morgenstern, 1947), cardinal utility (Handa, 1977), prospect theory (Kahneman and Tversky, 1979), and rank- and sign-dependent utility (Luce and Fishburn, 1991, 1995).

Following Chechile and Cooke (1997), let the utility of a two-outcome mixed gamble be expressed as

$$U(g) = \alpha_0 \gamma_0 f_1(p) g(V_1) - \beta_0 \delta_0 f_2(1-p) h(|V_2|), \tag{1}$$

where α_0, β_0, γ_0, and δ_0 are positive (scale factor) constants, where f_1 and f_2 are positive functions of the probability of the respective gain and loss, and where g and h are positive functions, respectively, of the gain V_1 and the absolute value of the loss V_2. Defining x and y transforms as

$$x = f_1(p) g(V_1), \tag{2}$$

$$y = f_2(1-p) h(|V_2|), \tag{3}$$

results in a linear form for GUT, i.e., $y = \lambda x - \psi$, where -1 times the y intercept is a measure of the utility of the gamble, i.e., $\psi = U(G)/\beta_0 \delta_0$, and the slope is

$$\lambda = \alpha_0 \gamma_0 / \beta_0 \delta_0. \tag{4}$$

Chechile and Cooke reported that λ varied systematically, and such variation is inconsistent with the requirement from (4) that λ remain a constant.

The Chechile-Cooke experimental procedure required participants to match two gambles for each of 252 trials. The first gamble of a pair was called the reference gamble, and it was completely specified (i.e., $V_1 = \$50$, $V_2 = -\$50$, and the probability for the gain, p_r, either was .05, .1, .3, .5, .7, .9, or .95). The second gamble was called the comparison gamble, and the participant was provided with a V_1 value from the set $\{\$0, \$20, \$40, \$60, \$80, \$100\}$ and a V_2 value from the set $\{\$0, -\$20, -\$40, -\$60, -\$80, -\$100\}$. The participant was required to specify a probability for the win on the comparison gamble that provided a psychological match between the comparison gamble and the reference gamble. All 36 combinations of V_1 and V_2 were provided for each of the seven reference gambles. Consequently, for their most favorable reference gamble, $G_7(V_{1r}, p_r, V_{2r}) = (\$50, .95, -\$50)$, there was a trial where the participant was required to select a probability value for the comparison value which had a "gain" of \$0 and a loss of $-\$100$. It is unlikely even with a probability of 1.0 that the comparison gamble was subjectively equivalent to the highly favorable reference gamble. The recognition of this mismatch problem was the rationale for the current reanalysis of the data from the Chechile and Cooke study.

1.0. Statistical correction and analysis

Clearly, when a participant uses an extreme value for the comparison gamble probability, the individual is demonstrating difficulty matching the comparison gamble to the reference gamble. These trials ought to be removed from consideration. The reanalysis is predicated on omitting the data from such trials and thereby statistically correcting for the mismatch problem. We define an extreme probability as a $< .001$ or $> .999$ because individuals rarely used the values of either 0 or 1. More typically, they would give a very low or high probability value and the aforementioned criteria would detect those cases. This correction assumes that the participants, when confronted with an impossible task, in fact selected the extreme probability that resulted in the closest certainty equivalent rather than exhibiting a random response. Also the reanalysis omitted all the trials where both V_1 and V_2 were \$0. A total of 258 trials was culled by these considerations, which represented 8.5% of the total trials. Because the $V_1 = V_2 = \$0$ trial is removed for each of the seven reference gambles, there are 245 gamble pairs per participant. The mean probability (averaging across the participants who demonstrate a non-extreme probability for that trial) was computed. There were 12 participants in the Chechile and Cooke (1997) data base. All the subsequent analyses were based on these 245 means.

Chechile and Cooke examined nine GUT linearized models resulting from the factorial combination of three probability weighting methods for specifying the f_1 and f_2 functions and three utility scaling methods for the g and h functions. The f_1 and f_2 functions include, respectively, s and w fitting parameters. The g and h functions have c and d fitting parameters. Complete details about these nine models can be found in Chechile and Cooke (1997).

For any particular GUT linearized model, it is necessary to determine the model parameter values that are optimal for the group-averaged data (i.e., we need to specify the values for the s, w, c, and d parameters). A search of the parameter space for the s, w, c, and d parameters was conducted by examining 12 values for each of the parameters in an initial pass with a relatively large parameter increment between candidate values. A second pass and third pass also examined 12 values for each parameter, but these searches used increasingly narrowed ranges based upon the results from prior passes. For each combination of parameters, the x and y transforms are specified, and r^2 is computed for each reference gamble. The best fit corresponds to the s, w, c, and d parameters which result in the largest average r^2 value across the seven reference gambles. The optimal parameter values and the subsequent results for each model are contained in Table 1.

According to generic utility theory, λ should be a constant for a particular model. Yet for each model, there is a significant decline in λ as a function of the increasing favorableness of the reference gamble (i.e., there is a statistically significant Spearman rank-order correlation between the rank of the reference gamble probability, p_r, and the rank of λ). This pattern for λ is similar to that reported by Chechile and Cooke; although the ratio $\lambda(.05)/\lambda(.95)$ is now smaller for all models except the flexible/power model, it is still considerably greater than 1.0. Moreover, Chechile and Cooke reported that the corresponding Spearman correlation between the rank of the reference gamble and the rank of ψ was erratic. This correlation has generally increased in Table 1 but remains widely variable with the various models. Consequently, the major conclusions of the Chechile and Cooke experiment still hold when there are statistical controls imposed to mitigate the problem caused by the mismatched trials.

Finally, it should be pointed out that the statistical correction done in the present paper has two potential weaknesses. As noted earlier, it assumes that the participants did not simply respond unpredictably when confronted with a mismatched pair of gambles. If participants did respond randomly in such cases, then some of the non-extreme probability trials would also need to be removed. In addition, the reanalysis assumes that all trials with an extreme probability reflect a mismatch problem. Perhaps for some of these extreme probability trials, there might actually be a psychological match between the comparison and reference gambles, so those trials should not have been removed. The ideal solution to the mismatch problem is to conduct a new experiment which circumvents this difficulty

Table 1. Model parameters of nine GUT linearized models

	Model Probability weighting system/utility scaling system								
Parameter	Power/ power	Power/ −exp	Power/ mixed	Normed/ power	Normed/ −exp	Normed/ mixed	Flexible/ power	Flexible/ −exp	Flexible/ mixed
s	1.65	0.71	.90	1.80	0.95	1.03	2.16	1.37	1.46
w	2.62	2.35	2.45	2.62	2.27	2.62	2.54	1.65	1.91
c	2.17	.001	.001	2.08	.001	.001	2.16	.001	.001
d	1.83	.001	.011	1.60	.001	.011	1.73	.001	.011
r^2_{av}	.454	.389	.408	.501	.421	.448	.495	.413	.438
$\lambda(.05)/\lambda(.95)$	253	39	41	321	39	51	334	22	44
$r_s(\lambda)$	−1.00	−1.00	−.964	−1.00	−1.00	−.964	−1.00	−.964	−.964
$r_s(\psi)$	.536	.393	.893	.571	.464	.750	.857	.893	.964

Note. The r^2_{av} is the average (over the seven reference gambles) of the squared Pearson correlation between the y and x transformed variables. The rank-order correlations between reference gamble probability and λ and ψ, respectively, are $r_s(\lambda)$ and $r_s(\psi)$. See Chechile and Cooke (1997) for details about the models and the s, w, c, and d fitting parameters.

by adjusting the range of V_1 and V_2 values. Such an experiment has been conducted, and a report of this experiment is currently being written. Chechile and Butler (1998) provided an advanced report of this experiment. The new experiment also found that λ varied with the reference gamble context.

References

Chechile, R. A., & S. F. Butler. (1988, November). "Evaluating a Class of Utility Theories," Paper read at the Psychonomic Society meeting. Dallas, TX.

Chechile, R. A., & A. D. J. Cooke. (1997). "An Experimental Test of a General Class of Utility Models: Evidence for Context Dependency," *Journal of Risk and Uncertainty* 14, 75–93.

Handa, J. (1977). "Risk, Probabilities, and a New Theory of Cardinal Utility," *Journal of Political Economy* 85, 97–122.

Kahneman, D., & A. Tversky. (1979). "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263–291.

Luce, R. D., & P. C. Fishburn. (1991). "Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles," *Journal of Risk and Uncertainty* 4, 29–59.

Luce, R. D., P. C. Fishburn. (1995). "A Note on Deriving Rank-Dependent Utility Using Additive Joint Receipts," *Journal of Risk and Uncertainty* 11, 5–16.

Miyamoto, J. M. (1988). "Generic Utility Theory: Measurement Foundations and Applications in Multiattribute Utility Theory," *Journal of Mathematical Psychology* 32, 357–404.

Miyamoto, J. M. (1992). "Generic Analysis of Utility Models." In Ward Edwards (ed.), *Utility Theories: Measurements and Applications*. Boston: Kluwer Academic Publishers.

von Neumann, J., & O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.