# ERRATA FOR
# GENERALIZED EXPECTED UTILITY, HETEROSCEDASTIC ERROR, AND PATH DEPENDENCE IN RISKY CHOICE.

ORIGINAL PAPER PUBLISED JANUARY, 2000

ERRATA DATE OCTOBER, 2007

By

David Buschena[1]
Department of Agricultural Economics and Economics
306 Linfield Hall
Montana State University
Bozeman, MT 59717
(406)994-5623

and

David Zilberman
Department of Agricultural and Resource Economics and Policy
207 Giannini Hall
University of California
Berkeley, CA 94720
(510)642-6570

*Authors' Note.*

We were contacted in March, 2007 by Nicky Grant, a Cambridge University undergraduate student working with John Hey, about a discrepancy between the results in Hey and Orme's 1994 *Econometrica* paper and ours in the January, 2000 *Journal of Risk and Uncertainty (JRU).* As both papers used the same data (Hey graciously made his data available to us), we took this question very seriously and carried out extensive checks and re-estimations of our work.

Although the estimation code we used previously was no longer available due to multiple computer replacements, we could not reproduce the results in our original paper. There appears to have been an error in the recording of our original estimates under homoscedastic error that subsequently were used in evaluations throughout our 2000 paper. Although we have not pin-pointed the reason for our error, we offer this errata as a corrected version.

The results in this corrected paper closely correspond with those in Hey and Orme (1994), with any small differences attributable to the different estimation routines used. We use the same estimation package (LIMDEP 7.0) and algorithm as for the original paper. As in our earlier paper, these re-estimates show the importance of the error structure in EU/GEU model evaluation. The re-estimates do however show considerably more support for the EU model under homoscedastic error than our paper's estimates do.

The editor of the *JRU* shared our goal of making these new estimates available, and suggested a dedicated and maintained webpage for this errata.

Our original paper addressed three questions in the conclusions. The answers to the first two are largely intact in this errata.

***Does the introduction of heteroscedastic error into both EU and GEU formulation substantively increase the relative fit of EU?*** *Introducing various heteroscedastic error structures substantially increased the relative fit of EU, consistently reducing the number of subjects whose choices violated EU under heteroscedasticity.*

***What is the performance of EU under heteroscedastic error relative to that for GEU models with homoscedastic error?*** *Consistent with the results in Hey (1995), models of EU with heteroscedasticity performed comparably and often better than GEU models with homoscedastic error specifications.*

The answer to our third question has changed from the original.

***What is the likely outcome of standard sequential model selection procedures when both heteroscedastic error and GEU formulations are separately tested for, and where inclusion of these components in a choice model depends on these sequential tests?*** *The majority of subjects did not exhibit path dependence in model selection. We did however found many respondents who exhibited path dependence for GEU and heteroscedastic error specifications under such a sequential testing and selection procedure. We did not find one model that is dominant for every person. Both GEU and heteroscedasticity are important components to consider, and the selection procedure used to include or exclude them must be evenhanded, perhaps using a method such as the generalized Cox approach in Pesaran and Deaton (1978).*

We thank and commend Nicky Grant and John Hey raising the question of the result discrepancy, and for their congenial approach to the issue. We also thank Kip Viscusi for his help and commitment to the *JRU*.

*Abstract*

We evaluate the fit of several generalized expected utility models under homoscedasticity and three different heteroscedastic error structures for the data set first reported in Hey and Orme (1994). Standard chi-squared tests are used for nested tests, and both the Akaike (1973) information criterion and its consistent version (Hurvich and Tsai, 1989) are used for non-nested ranking of these models. A testing framework is developed that explicitly accounts for the path-dependent nature of the model selection problem. Not only does the selection of preference models depend on the error structure assumed, but the reverse is also true: the selection of the error structure depends on the preference structure assumed.

**Keywords:** Generalized-expected utility, heteroscedastic error, path-dependence.

**JEL Classification:** D81, D83

# Generalized Expected Utility, Heteroscedastic Error, And Path Dependence in Risky Choice

Recent work has introduced heteroscedastic error structures into models for risky choice (Hey and Orme, 1994; Hey, 1995; Loomes and Sudgen, 1995; Ballinger and Wilcox, 1997; Carbone, 1997). In particular, Hey's (1995) results using individual choice data suggest that a model of Expected Utility (EU) with heteroscedastic errors reflecting decision difficulty, importance, or both may well be superior to Generalized Expected Utility (GEU) formulations under homoscedasticity. Hey's conclusion is strong:

> "*It may be the case that these further explorations may alter the conclusion to which I am increasingly being drawn: that one can explain experimental analyses of decision making under risk better (and simpler) as EU plus noise — rather than through some higher level functional — as long as one specifies the noise appropriately.*"

Although additional study is clearly required — and is the justification for this paper — in our reading heteroscedastic error has not become prevalent in modeling risk despite Hey's findings. We see three reasons for this:

1. Hey did not test the performance of GEU models under a heteroscedastic error structure versus EU with the same heteroscedastic structure. It may be that although GEU with homoscedasticity offers little predictive superiority to EU with heteroscedasticity, GEU is still (statistically) significantly better than EU when both models have the same heteroscedastic error structure.

2. Hey used a stochastic variable (response time) to define error variance in his highest ranked heteroscedastic error formulations. The likelihood function arising from using this stochastic variable to define the heteroscedastic error is complicated, as it is unlikely to yield a homoscedastic error structure.

3. Hey did not relate his heteroscedastic error formulations in of experimental evidence that has been so important to the development of GEU models. More direction

> into how this experimental evidence may help to define the "appropriate" heteroscedastic error specification would be of great value.

We evaluate the relative predictive power of heteroscedastic error for risky choice as advanced by Hey (1995) in light of these three concerns. Regarding the first concern, we evaluate statistically the degree to which the error specification affects the evaluation of GEU preference functions. We also address a related issue where the preference model specification affects the evaluation of the error specification.

Regarding the second and third concerns, we define three heteroscedastic error forms defined using exogenous variables. These exogenous variables were statistically significant as instruments for explaining decision time for Hey and Orme's data (results available upon request) and for response times from a separate experiment (Buschena and Zilberman, 1999). These exogenous variables relate to behavioral similarity models (Rubinstein, 1988; Leland, 1994; Buschena and Zilberman, 1995, 1999) that were designed to model the same behavioral regularities violating EU that have driven the development of the GEU models.

We find clear support for Hey's conclusion. While our estimates show EU to be clearly inferior in prediction to GEU models when only *homoscedastic* error is considered, for the majority of respondents GEU offers no statistically significant improvement in predictive power over EU when both models have identical *heteroscedastic* error structures.

In the course of our analysis, the relative predictive performance of two restricted models became of interest: EU under heteroscedastic error versus GEU models with homoscedastic error. Our empirical results show that these two restricted models largely act as substitutes for predicting choice, and the final sections of this paper focus on this substitution. Related to this substitutability in prediction is an evaluation of the likely result of standard model selection procedures when both heteroscedastic error and GEU formulations are separately and sequentially tested for and each separate component is retained in the model if it proves to offer a significant improvement in model fit.

We evaluate the performance of EU and six GEU models using one of four error structures (three heteroscedastic specifications and homoscedasticity). The six GEU models are well known in the risk literature and are all estimable through continuously differentiable parametric specifications. The heteroscedastic specifications reflect previously suggested measures for the differences between the risky pairs (Rubinstein, 1988; Leland, 1994; Buschena and Zilberman, 1995), and were selected for parsimony. Our tests rely on a data set from Hey and Orme (1994), to date one of the most extensive surveys with real payoffs. Their data set is unique, having enough observations from each respondent (80 each from two experiments) to allow estimation of choice models separately for each respondent. Indeed, we could not have individually evaluated the error structure and model selection question for risky choice without this data set. This intrapersonal estimation allows for differences in respondents' risk attitudes and importantly in their error structures (shown to be so important in Ballinger and Wilcox, 1997). Additional data sets such as Hey and Orme's are needed.

We use three testing approaches. Standard chi-squared tests are carried out for nested models separately to evaluate both the error structure and the preference model. Non-nested models are evaluated using the Akaike information criterion (Akaike, 1973) and its consistent correction (Hurvich and Tsai, 1989). Finally, we evaluate more completely the nested tests through a framework that explicitly accounts for the path-dependent nature of the model selection problem. Our objective in this final testing framework is to determine whether or not a single modeling approach will dominate regardless of the path.

Our results demonstrate the problems in testing deterministic choice models for risk using a homoscedastic error assumption. The number of respondents whose choices lead to rejection of EU under an error structure are substantially reduced under any one of the three heteroscedastic specifications and across all GEU specifications. When used exclusively, for many respondents both GEU and heteroscedastic specifications offer improvements in model fit relative to EU with homoscedasticity. When placed within a competitive atmosphere of allowing only either a GEU model (with homoscedasticity) or a heteroscedastic specification (with EU), the heteroscedastic specifications offer comparable and often superior fit to GEU models. For many respondents, these models appear to be substitutes; either GEU with homoscedasticity or EU with a heteroscedastic error structure suffice to explain choice patterns.

The path dependency problem for model selection is evident in these data. If GEU is assumed, then homoscedasticity is often not rejected; if heteroscedasticity is assumed, then EU is often not rejected. Therefore, not only does the selection of preference models depend on the error structure, but *vice versa* that the selection of the error structure depends on the preference structure assumed. For respondents who do not exhibit this path dependence (they show a greater degree of model selection consistency), homoscedastic error with either EU or a GEU model was most frequently supported.

Section 2 briefly summarizes the experimental data from Hey and Orme (1994) and describes the heteroscedastic specifications we use. Section 3 presents performance comparisons for several GEU models and for EU with heteroscedasticity. We conclude in Section 4.

## 1. Data and evaluation methods

In Hey and Orme's Data Set I, eighty respondents were asked to select their preferred alternative from 100 pairs of lotteries (indifference between the alternatives was allowed). The same 100 pairs were presented in a different order to each respondent again at least one day later (Data Set II). Some pairs offered a risk-return trade off, with one gamble carrying a larger variance and a larger expected value, while other pairs used gambles with equal means and unequal variances (one alternative was a mean-preserving spread of the other). In both experiments, each respondent had an opportunity to win money, when they played their preferred gamble from a randomly selected pair for real cash (Hey and Orme report an average per hour payment of £17.50).[1]

Hey and Orme utilized these data to estimate a number of preference models within a multinomial probit specification. This estimation was carried out separately for each respondent under considerable generality in preferences and in errors, an approach possible with Hey and Orme's design where each respondent faced a large number of risky choice pairs. Several GEU models were specified through the value function, $V_j\,(p, q, x, \alpha)$, defined over the risky alternatives' probability vectors, $p = (p_1, p_2, p_3, p_4)$ and $q = (q_1, q_2, q_3, q_4)$, for outcomes $x = (x_1, x_2, x_3, x_4) = (0, £10, £20, £30)$, for a parameter vector α, and for GEU model *j*. For example, the value function for EU maximization is the estimated difference in the gambles' expected utilities ($u(0) = 0$ without loss of generality), using *p, q,* and *x*, and where $\alpha_i x_i = u(x_i)$:

$$V_{\text{EU}}\left(p, q, x, \alpha\right) = \alpha_2 \cdot x_2 \cdot \left(p_2 - q_2\right) + \alpha_3 \cdot x_3 \cdot \left(p_3 - q_3\right) + \alpha_4 \cdot x_4 \cdot \left(p_4 - q_4\right). \tag{1}$$

At most, three outcomes were possible for each lottery pair; the lotteries were defined by either ($x_1$, $x_2$, $x_3$), ($x_1$, $x_2$, $x_4$), ($x_1$, $x_3$, $x_4$), or ($x_2$, $x_3$, $x_4$). Our analysis (Section 3) extends that in Hey and Orme and in Hey for a subset of the GEU models that showed promise in predicting choice. Readers are directed to Appendix A for parametric specifications of the GEU models we tested.

An individual's likelihood of selecting gamble *p* over *q* in pair *i*, given the preference model *j* as used in equation (1), disallowing indifference, and under choice error $\sigma^2_i$ is:

$$F\left(V_j\left(p, q, x, \alpha\right)+\varepsilon_i\right), \qquad \text{for} \quad \varepsilon_i \sim N\left(0, \sigma_i^2\right) \tag{2}$$

Heteroscedastic error specifications are defined below. Hey (1995) reports the results of three specifications for this heteroscedasticity; variants of two of these specifications ($\sigma_1$ and $\sigma_2$) are reconsidered here.[2]

We explore for each individual three forms for heteroscedastic error under EU plus numerous GEU models. Although there may be other useful specifications, these three forms allow us to assess, in a parsimonious manner, the effects of adding heteroscedastic error to model choice using specifications appropriate for the pairs in the experimental design. All three specifications are measures of the difficulty or the importance of the decision, or both and have proven to be effective instruments for understanding response time for Hey and Orme's data (results available upon request) and for a separate data set (Buschena and Zilberman, 1999).

The first error model is homoscedasticity:

$$\sigma_0 = 1. \tag{3}$$

The second error model is heteroscedastic, predicting that the standard deviation decreases ($\beta_1 < 0$) with model *j*'s value difference between the lotteries:

$$\sigma_1 = \exp\left(\beta_1 * \left|V_j\left(p,\ q,\ x,\ \alpha\right)\right|\right). \quad (4)$$

where $V_j$ is the level of the value function from preference model $j$. In Hey, $j$ = EU. In our analysis in Section 4, $j$ indicates the appropriate value function for either EU or one of the six GEU models.

Another heteroscedastic formulation is:

$$\sigma_2 = \exp\left(\beta_2 * N\right), \quad (5)$$

where $N$ is the average number of outcomes for which the two lotteries have positive probabilities, a measure proposed by Hey (1995) that takes values from 1.5 to 3 in the experimental data.[3]

Another heteroscedastic measure is a model-free measure for the difference between the lotteries. A general measure of this difference is the area (see standard calculus books such as Salas and Hille (1982) and see Buschena (1993) for an empirical application) between the lotteries' cumulative distribution functions (CDFs) over the range of outcomes:

$$\sigma_3 = \exp\left(\beta_3 * \int_x \left|\mathrm{CDF}_p\left(x\right) - \mathrm{CDF}_q\left(x\right)\right| dx\right) \quad (6)$$

This absolute CDF difference measure is more appropriate than a discrete measure such as the Euclidian distance for Hey and Orme's data set because it allows for differences between the outcome vectors among lotteries. Recall that not all outcomes are common to each lottery within the pairs.[4]

## 2. Analysis of choice models

The log-likelihood function (LLF) of observation $i$ for preference model $j$ and error structure $k$ is given by the following equation under a normality assumption, where $V_j = V_j(p, q, x, \alpha)$ is model $j$'s valuation function, $t_{jk}$ is a threshold and φ(·) is the standard normal CDF:

$$\mathrm{LLF}_{ijk}\left(p, q, x. \alpha, t_{jk}, \sigma_k\right) =$$

$$z_{0i} * \log\left[1-\phi\left(\frac{t_{jk}-V_j}{\sigma_k}\right)\right]$$

$$+ z_{1i} * \log\left[\phi\left(\frac{t_{jk}-V_j}{\sigma_k}\right)-\phi\left(\frac{-t_{jk}-V_j}{\sigma_k}\right)\right] \tag{7}$$

$$+ z_{2i} * \log\left[\phi\left(\frac{-t_{jk}-V_j}{\sigma_k}\right)\right].$$

In equation (7) and for observation $i$, $z_{0i}$ indicates selection of the left-hand side alternative, $z_{2i}$ indicates selection of the right-hand side alternative, and $z_{1i}$ indicates indifference. Note that in this formulation, there is an identifiable difference between the preference model ($V_j$), heteroscedasticity ($\sigma_k$), and the threshold ($t_{jk}$).

Seven preference models were evaluated: EU; Quiggin's (1982) Rank Dependent EU model (RDQ); Segal's (1987) Power Rank Dependent model (RDP); Viscusi's (1989) Prospective Reference Theory (PRT); Weighted Utility (WU) from Chew (1985) and Dekel (1986); Quadratic Utility (QU) from Chew, Epstein, and Segal (1991); and Regret Theory with Independence (RI) from Loomes and Sugden (1982).[5] A parametric specification is given for each of these models in Appendix A. Four error structures were tested for each preference model, where the error structure was defined through $\sigma_0$, $\sigma_1$ (using the value $V_j$ from the appropriate model), $\sigma_2$, or $\sigma_3$. For each respondent within one of the two data sets, we estimated 28 specifications (seven risk models each using four different error structures) and compared them intrapersonally. Since each of the 80 respondents had two sets of data, there were 4,480 LLFs estimated in total.

The Davidson-Flecter-Powell algorithm was used to maximize the LLF for each individual's choices using LIMDEP version 7.0's maximization routine (see Greene, 1995). The analysis thus allows for interpersonal differences in preferences and in error. The starting values were the estimated parameters of these models from fits over the entire eighty subjects.

*2.1 Testing EU vs. GEU with and without a heteroscedastic error structure*

Our first test assesses the fit of EU relative to GEU models *given* a heteroscedastic error structure for both EU and the GEU model. Note that this is a stronger test of Hey's claim that EU plus noise is sufficient in modeling choice than in his work. Specifically, we test the following using separately each one of three different heteroscedastic error formulations ($\sigma_1$, $\sigma_2$, or $\sigma_3$):

$H_{1b}$: Given a heteroscedastic error structure, EU is equivalent in fit to a GEU model, vs.

$H_{1a}$: Given a heteroscedastic error structure, EU gives fit inferior to a GEU model.

Table 1 lists for each data set the number and percentage of respondents for which the null hypothesis was rejected, supporting a GEU preference model over EU. All tests were at the 5 percent level using $\chi^2$ tests (the degrees of freedom for these tests is given in the df column of Table 1). These tests were run under the three heteroscedastic error specifications and also under homoscedasticity for comparison.

The results under homoscedastic error ($\sigma_0$) are consistent with those results in Hey and Orme (1994) and with those from over four decades of analysis of behavior under risk (e.g., Allais, 1953; Kahneman and Tversky, 1979). At most 50% (51%) of respondents' for RDQ under Data Set I (RI under Data Set II) choices show rejection of EU. Note, there is considerable difference in the degree of EU rejection across GEU models.

The occurrence of rejection of EU under the three heteroscedastic specifications is generally less than the rejection rates under homoscedastic error, with the lone exception being the RDP model under $\sigma_1$ for Data Set II.

*2.2 Comparing homoscedastic vs. heteroscedastic error, various preference models*.

For completeness, we also test the following hypotheses to assess the importance of including heteroscedasticity:

$H_{2b}$: Given a preference model, homoscedastic error is equivalent in fit to heteroscedastic error,

vs.

$H_{2a}$: Given a preference model, homoscedastic error is inferior in fit to heteroscedastic error.

Table 2 reports the results of intra personal $x_{(1)}^2$ tests (5% significance level) of these hypotheses for EU and the six GEU models. These results differ somewhat from those in Table 1, reflecting the importance of the testing order for GEU models and heteroscedastic error specifications. The rates of rejection for homoscedastic error given either EU or a GEU model do not systematically differ across the EU/GEU models as one moves down the rows of the table. The incidence of rejecting homoscedastic error under EU lies within the rejection range for the GEU models.

The results in Tables 1 and 2 demonstrate two important findings. First, under homoscedastic error GEU models are superior to EU for many respondents, but that the number of subjects for which GEU models offer substantial improvement in fit over EU is consistently reduced under heteroscedastic error. These results extend Hey's (1995). The second result, is that the benefits from introducing a GEU specification as an alternative to EU differ little under either a homoscedastic or a heteroscedastic model. These results raise the issue of how important path dependence in model selection might be for EU vs. GEU models under various error structures.

The modeling problem of determining the error structure in this case differs from the one usually considered where (if present) heteroscedasticity is first corrected for and then model selection occurs (at least for the case of formulations such as $\sigma_2$ and $\sigma_3$ that are independent of the preference model). However, because GEU models have been designed to address empirical behavior that also has been related to decision models reflecting choice difficulty, importance, or both (Rubinstein, 1988; Leland, 1994; Buschena and Zilberman, 1995), there is potential for path dependence in empirical model specification, tests, and selection. We address this problem extensively below (Section 3.5).

### *2.3 Comparing GEU with homoscedasticity with EU and heteroscedasticity.*

Taking an agnostic view with a goal of only determining which model has the best fit, the two modeling approaches of GEU preferences and heteroscedastic error need not be inherent competitors. The theoretical underpinnings of these two approaches differ, and it should prove useful to have some objective analysis of their performance exclusive of one another. Considering instead an absolutist view, if heteroscedastic error based on decision difficulty is a predictive substitute for GEU preference formulations, a heteroscedastic error structure may allow us to retain a relatively simple preference structure and still address behavioral regularities taken to violate EU preference axioms. We test the following hypotheses for each GEU model and all three of our heteroscedastic specifications, $\sigma_1$, $\sigma_2$, $\sigma_3$:

*$H_{3p}$:* Predictive Power of GEU with homoscedasticity is indistinguishable from that from EU with heteroscedastic error, where the error formulation reflects decision difficulty or importance.

*$H_{3a}$:* Predictive Power of GEU with homoscedasticity is distinguishable from that from EU with heteroscedastic error, where the error formulation reflects decision difficulty or importance.

Following Hey and Orme (1994) and Hey (1995), we first compare the fit of each of the GEU models under homoscedastic error with one another and with the EU model with each one of the three heteroscedastic formulations. These comparisons use Akaike's measure of the model LLFs from equation (7), with an adjustment (penalty) for the total number of parameters ($k_j$) in the model. This Akaike information criteria (AIC) for model $j$ and for each respondent $i$ is given by

$$\text{AIC}_{ji} = 2 * \left[ \text{-LLF}_{ji} \left( p, q, x, \alpha, t_{ji} \right) + k_{ji} \right] / n, \tag{8}$$

where $n$ is the number of observations (100 for each individual within a single data set). The model with the lowest AIC for an individual's choices receives the first rank, the model with the second lowest receives the second rank, and so on.

The AIC measure, while asymptotically efficient, is potentially biased toward retaining models with higher dimension (Hannan and Quinn, 1979; Hurvich and Tsai, 1989; Schwartz, 1978). This bias is of particular concern for small samples, or when the number of fitted parameters is a moderate to large fraction of the sample size (Hurvich and Tsai). Because we are using non-nested and nonlinear forms, it is difficult to be certain that the AIC measure's bias is benign. Consequently, we also report the ranking results of a bias-corrected version of the AIC, the AIC C measure originally proposed in Suguria (1978) and extended to nonlinear models by Hurvich and Tsai:[6]

$$\text{AIC C} = \text{AIC} + \frac{2(\mathrm{k}+1)(k+2)}{n-k-2} \tag{9}$$

Tables 3A and 3B list the first five AIC rankings and the summary ranks for both the AIC and the AIC C for each data set.[7] We add a column to indicate when the model had a singular variance/covariance matrix or was otherwise not successfully estimated. The average rank summary columns include two averages for each criteria; the first column ("No S") excludes from the average the singular estimation cases, while the second ("S=9") includes these cases and gives them a rank of 9. The appropriate treatment of these singular cases is undefined in the AIC approach, and we are reluctant to ignore them altogether. We also recognize that giving these singular cases a rank of 9 may be a harsh penalty. We include both treatments so the reader may form his or her own conclusions. The average rankings for favorably ranked models are given in bold font.

The bias-corrected average AIC C ranks for the highest ranking models differ somewhat from their ranks using the AIC criteria. As expected given the AIC measure's potential bias toward models with more parameters, models such as weighted utility, quadratic utility, regret with independence have higher (less supported) average ranks under the AIC C criteria.

No single model has the top rank for every respondent, but the performance of some models is better than others. Considering the models' number of "firsts," two models of EU with heteroscedasticity ($\sigma_2$ and $\sigma_3$) and RDQ with homoscedasticity are supported. Based on the models' average ranks, RDQ is supported under the AIC criteria while EU with heteroscedastic formulations $\sigma_2$ and $\sigma_3$ is supported under the AIC C criteria for both data sets.

*2.4 Comparing EU and GEU using both homoscedastic and heteroscedastic error structures*

In order to further evaluate heteroscedastic error structures and GEU models, we rank 14 models for each data set: EU and 6 GEU models under homoscedastic error ($\sigma_0$), and EU and the 6 GEU models under the heteroscedastic error structure $\sigma_3$.[8] This ranking allows a more thorough evaluation of heteroscedastic error than in Hey since both the GEU formulation and heteroscedastic error are allowed in the same model. These rankings allow us to further evaluate improvements in fit relative to EU and homoscedastic error. Ranking results are summarized in Table 4. Models with the greatest number of "first" rankings, and those with the lowest average AIC and AIC C rankings, have these entries in boldface type. We report for clarity and parsimony only the incidence for each model receiving one of the top three AIC ranks, the occurrence of singular variance/covariance estimation (S), and each model's average ranking under both the AIC and the AIC C criteria. EU under homoscedastic error ($\sigma_0$) compares favorably to the other models, under the AIC C criteria. This favorable ranking reflects those in prior tables, and speaks to the empirical power of EU for many respondents. We are also interested in the relative rankings for the higher-order models, reflecting subjects whose choice patterns deviate significantly from EU under homoscedastic error.

The choice model of EU with $\sigma_3$ is in the top two models in both data sets under the criteria of the number of "firsts" and has a favorable average ranking under either criteria and in both data sets. The RDQ model under homoscedastic error has favorable average rankings in both Data Sets, and a high number of "firsts" in Data Set II.  As expected, the average rankings under the AIC C criteria in general favor models with fewer parameters.  Taken in total, the results from this evaluation of average rankings show support for Hey's conclusion of the importance of including heteroscedastic error models, EU with homoscedastic error has considerable support, EU under heteroscedastic error has some support, and RDQ under both homoscedastic and heteroscedastic error formulations also has statistical support.

### *2.5  Implications for "Standard" model selection methods*

The results in Tables 1-4 suggest trade-offs in modeling choice through error and preference formulations for subjects with choice significantly different from EU under homoscedastic error. The results also suggest that no single approach is correct for all subjects; modeling some respondent's choices may require heteroscedastic error, while others may require a GEU preference structure.

We couch and further explore the implications of these findings within the context of a sequential model selection procedure using standard likelihood-ratio tests.  In this interpretation, a researcher first tests for homoscedastic error *given* a GEU structure.  If homoscedasticity is not rejected, then a second stage assesses the occurrence of subjects who require the GEU preference model under homoscedastic error.

This sequential modeling procedure could be carried out in reverse order.  First, the GEU specification could be tested against EU *given* a heteroscedastic formulation.  If EU is not rejected in the first step, then the heteroscedastic model could be tested against homoscedasticity (now both under EU).

Consider Figure 1. Researchers who test first for homoscedasticity given GEU and then for EU are in the upper branch of the figure. One result supports the full model with both a GEU specification and heteroscedasticity (Node 1). Another is that EU with heteroscedasticity is called (Node 2). Additionally, GEU with homoscedasticity (Node 3) or EU with homoscedasticity (Node 4) could be supported.

The lower branch describes the potential results when a researcher first tests between one GEU model and EU given heteroscedasticity and then tests for homoscedasticity (given EU). In this branch the possibilities are: GEU and heteroscedasticity (Node 5), GEU with homoscedasticity (Node 6), EU with heteroscedasticity (Node 7), or EU with homoscedasticity (Node 8).

The results of these sequential testing procedures may be path dependent. Path independence with respect to the preference model and the error structure occurs when a subject falls into one of the node pairs (1, 5), (2, 7), (3, 6), or (4, 8). Such path independence would give us a clear modeling recommendation for the preference structure and error specification. For example, if most respondents fell into node pair (2, 7), EU with the heteroscedastic formulation would be consistently called for. Alternatively, if most respondents fell into node pair (3,6), a GEU model with homoscedasticity would be called for.

We use tests over the LLFs from our model estimates to find the number of respondents who fall into each node illustrated in Figure 1. Since in each branch the competing models for each test are nested, we use standard likelihood-ratio tests using a 5% level of significance. Tables 5A through 5F list the node resulting from this sequential estimation for both data sets, for the six GEU models, and for the three heteroscedastic formulations. Note that the full model, including both a GEU preference model and a heteroscedastic specification (Nodes 1 and 5), occurs very rarely.

These tables also list the number of respondents who were or were not path consistent in total and for the patterns of primary interest. Specifically, the tables list those who were path consistent in the node pairs of (3, 6) or (2, 7), where such observations allow separability of GEU models and heteroscedastic specifications regardless of the path. The tables also list subjects who were not path consistent where a GEU model would be selected in the upper branch but EU would be selected in the lower branch.

The positive news from these tables is that many respondents exhibit path consistency, particularly under heteroscedastic error as formulated through $\sigma_1$ for Data Set I and all error structures for Data Set II.

There are a large number of respondents who exhibit path inconsistency for Data Set I under both $\sigma_2$ and $\sigma_3$, across all GEU preference models. Recall from prior tables that these two heteroscedastic error structures had more statistical support that $\sigma_1$ did. The incidence of path inconsistency varied somewhat across the GEU models, with the GEU models which have a larger number of parameters generally exhibiting more of this inconsistency.

## 3. Conclusions

The data set developed by Hey and Orme allows for a very general treatment of preferences and error structure across individuals and gives insights into the performance of various models for risky choice. We were fortunate to have access to Hey and Orme's data. Additional intensive data sets are needed for further study of the relationships between heteroscedastic error and preference structures and to distinguish them in predictive performance. This data allowed us to answer three questions.

*Does the introduction of heteroscedastic error into both EU and GEU formulation substantively increase the relative fit of EU?* Introducing various heteroscedastic error structures substantially increased the relative fit of EU, consistently reducing the number of subjects whose choices violated EU under heteroscedasticity. It is noteworthy that EU under homoscedastic error was not rejected for the majority of subjects

*What is the performance of EU under heteroscedastic error relative to that for GEU models with homoscedastic error?* Consistent with the results in Hey (1995), models of EU with heteroscedasticity performed comparably and often better than GEU models with homoscedastic error specifications.

*What is the likely outcome of standard sequential model selection procedures when both heteroscedastic error and GEU formulations are separately tested for, and where inclusion of these components in a choice model depends on these sequential tests?* The majority of subjects did not exhibit path dependence in model selection. We did however found many respondents who exhibited path dependence for GEU and heteroscedastic error specifications under such a sequential testing and selection procedure. We did not find one model that is dominant for every person. Both GEU and heteroscedasticity are important components to consider, and the selection procedure used to include or exclude them must be evenhanded, perhaps using a method such as the generalized Cox approach in Pesaran and Deaton (1978).

We have reached two primary conclusions. First, selecting a deterministic preference structure such as EU or GEU considering only a homoscedastic error is problematic for many respondents. Second, adding heteroscedastic error to the stochastic structure considerably reduces the advantage of GEU over EU in model fit for many subjects; alternatively, adding a GEU specification to the deterministic structure considerably reduces the advantage of heteroscedasticity over homoscedasticity in model fit for many subjects.

Our second conclusion is important for model selection. From Tables 5A through 5F, if homoscedasticity is imposed, GEU models will dominate EU for some subjects in both data sets. However, if a heteroscedastic error structure is allowed, choice patterns from fewer respondents reject EU in favor of GEU. To be sure, the EU model with homoscedastic error does quite well, and the relative number of respondents rejecting or not rejecting EU in favor of a GEU model differs between the data sets, among the GEU models, and among the heteroscedastic specifications. The ability of these simple one-parameter heteroscedastic formulations to significantly alter the incidence of rejection for EU is striking and deserves further consideration. This is not simply an issue of drawing fine distinctions between alternative models but rather that there are strong theoretical differences between models altering preferences structures to account for behavioral regularities versus models incorporating heteroscedastic error.

There are three implications of our findings for future study of the relationship between heteroscedastic error formulations and preference models. Additional research efforts and modeling innovations are needed to capture the underlying factors influencing choice that are empirically manifested through heteroscedastic error. In particular, additional effort should be given to relating heteroscedasticity to algorithm selection, where algorithm selection is a function of the characteristics of the alternatives. The complete modeling of a choice theory incorporating such algorithm selection is not at all straightforward; this difficulty is attested to in Conlisk (1996).

Additional development of GEU specifications is needed to further assess the predictive value of error specifications and preference structures. Specifically, components of GEU models that remain important under heteroscedastic error specifications must be identified, modeled, and tested. It will be in this process that the joint contributions of heteroscedastic error and GEU specifications can best be made.

ENDNOTES

[1] For the 2007 errata version, Nicky Grant and John Hey have our thanks and are to be commended for pointing out our errors. We are very grateful to John Hey and Chris Orme for providing their experimental data. John Hey, David Harless, and Nat Wilcox provided insightful comments regarding experimental risk modeling. Justin Tobias and Marcelo Navarro proved very helpful in evaluating the various information criteria. All errors remain our responsibility.

1. This per hour average payment was the average payment over the two experiments evaluated here plus two unrelated dynamic choice experiments.

2. Hey also considered an EU model where heteroscedasticity depends on the time the respondent took to select the alternative. While this model was superior to other choice models in fit, the nature of the log-likelihood function that results is unclear since response time itself is a variable measured with error (as shown in Wilcox (1993) and in Buschena and Zilberman (1997)). In particular, if the error variance of the probit model and the time relationship have positive covariance, the resulting log-likelihood function would be artificially enhanced by a specification such as $\sigma_4 = \exp(\beta_3 * \text{time})$.

3. Clearly $N$ is subject to manipulation outside the sample if extremely close outcomes are considered. For an extreme example, a lottery pair consisting of gamble A giving \$100 with a 50% chance (otherwise \$0) and gamble B giving \$100.01 with probability .25, \$99.99 with probability .25 (otherwise \$0) would give $N$=2.5, but most respondents would likely view

\$99.99, \$100, and \$100.01 as the same outcome. No such lottery pairs exists in the data set we use, and are quite uncommon in risk experiments to avoid alienating respondents.

4. As for *N* defining the heteroscedastic error in (5), the appropriateness of the absolute difference in (6) depends on the pairs in the experimental design. Again, out-of-sample manipulation of the risky pairs can lead to non-sensical rankings for pairs' error variance. For example, consider three probability vectors over the outcome vector $x$ = (0, \$100, \$200): $p$ = (0, 1.0, 0), $q$ = (1.0, 0, 0) and $r$ = (.5, 0, .5). The absolute CDF differences for pair *pq* and pair *pr* are identical (as are their Euclidian Distances), but the error variance for choice is expected to be larger for pair *pr* than for pair *pq*. Pairs such as *pq* are not included in the data set.

5. The respective number of parameters in these models under homoscedastic error were 5, 6, 6, 6, 7, 10, and 8. The inclusion of a heteroscedastic formulation adds an additional parameter to the estimation, as does the inclusion of the threshold.

6. The ranking differences between the AIC and the AIC C criteria were comparable for this data set to the ranking differences between the AIC and Schwartz's Bayesian-based consistent information criteria.

7. The differences in these "top five" ranks between those reported and those from the AIC C criteria are very small. The primary effect is to reduce the occurrence of these favorable rankings for the higher dimension models.

8. This single heteroscedastic error structure was considered for parsimony and since it had argueably the best performance of the heteroscedastic error forms in Table 3A and 3B.

## REFERENCES


Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principle," in *Second International Symposium on Information Theory*, ed. by B. N. Petrov and F. Csaki, pp. 267-281. Budapest: Akademiai Kiado.

Allais, M. (1953). "Le Comportement de L'homme Rationnel Devant le Risque: Critique de Postulats et Axiomes de L'École Américaine," *Econometrica,* 21, 503–546.

Amemiya, T. (1980). "Selection of Regressors," *International Economic Review*, 21, 331-353.

Azipurua, J.M., T. Ichiishi, J. Nieto, and J.R. Uriarte. (1993). "Similarity and Preferences in the Space of Simple Lotteries," *Journal of Risk and Uncertainty,* 6, 289–297.

Ballinger, T. P., and N. T. Wilcox. (1997). "Decisions Error, and Heterogeneity," *Economic Journal*, 107, 1090–1105.

Buschena, D.E. (1993). "The Effects of Alternative Similarity on Choice Under Risk: Toward a Plausible Explanation of Independence Violations of the Expected Utility Model," Ph. D. Dissertation, University of California at Berkeley.

Buschena, D.E. and D. Zilberman. (1995). "Performance of the Similarity Hypothesis Relative to Existing Models of Risky Choice," *Journal of Risk and Uncertainty*, 11, 233–262.

____________. (1997). "Predictive Value of Incentives, Decision Difficulty, and Expected Utility Theory for Risky Choices," Working paper, Department of Agricultural Economics and Economics, Montana State University-Bozeman.

____________. (1999). "Testing the Effects of Similarity and Real Payoffs on Choice." Forthcoming in the Foundations of Risk and Uncertainty Volume VIII, Edited by Mark Machina and Bertrand Munier.

Carbone, E. (1997). Investigation of Stochastic Preference Theory Using Experimental Data," *Economics Letters*, 57, 305–311.

Chew, S. H. (1985). "From Strong Substitution to Very Weak Substitution: Mixture-Monotone Utility Theory and Semi-Weighted Utility Theory," Working Paper, Department of Political Economy, Johns Hopkins University.

Chew, S.H., L. Epstein, and U. Segal. (1991). "Mixture Symmetry and Quadratic Utility," *Econometrica*, 59, 139–163.

Conlisk, J. (1996). "Why Bounded Rationality?" *Journal of Economic Literature*, 34, 669–700.

Dekel, E. (1986). "An Axiomatic Characterization of Preferences Under Uncertainty Weakening the Independence Axiom," *Journal of Economic Theory,* 40, 304–318.

Greene, W. H. (1995). Users Manual, LIMDEP Version 7.0. Econometric Software, Inc. Bellport, N.Y.

Hannan, E. J., and B.G. Quinn. (1979). "The Determination of the Order of an Autoregression." *Journal of the Royal Statistical Society*, 41, 190-195.

Harless, D. and C. F. Camerer. (1994). "The Predictive Utility of Generalized Expected Utility Theory," *Econometrica*, 62, 1251-1289.

Hey, J. D. (1995). "Experimental Investigations of Errors in Decision Making Under Risk," *European Economic Review*, 39, 633–640.

Hey, J.D. and C. Orme. (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica,* 62, 1291–1326.

Hurvich, C. M. and G.-L. Tsai. (1989). "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76, 297-307.

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

Leland, J.W. (1994). "Generalized Similarity Judgments: An Alternative Explanation for Choice Anomalies," *Journal of Risk and Uncertainty*, 9, 151–172.

Loomes, G. and R. Sugden. (1982). "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," *Economic Journal*, 92, 805–824.

________________. (1995). "Incorporating a Stochastic Element Into Decision Theories," *European Economic Review*, 39, 641-648.

Pesaran, M. H., and A.S. Deaton. (1978). "Testing Non-Nested Nonlinear Regression Models," *Econometrica*, 46, 677-694.

Quiggin, J. (1982). "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

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.

Salas, S.L. and E. Hille. (1982). *Calculus: One and Several Variables with Analytic Geometry*, Fourth Edition. New York: Wiley.

Schwartz, G. (1978). "Estimating The Dimension of a Model," *The Annals of Statistics*, 6, 461-464.

Segal, U. (1987). "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202.

Suguira, N. (1978). "Further Analysis of the Data by Akaike's Information Criterion and the Finite Corrections," *Communications in Statistics*, A7, 13-26.

Viscusi, W.K. (1989). "Prospective Reference Theory: Toward an Explanation of the Paradoxes," *Journal of Risk and Uncertainty*, 2, 235–264.

Wilcox, N.T. (1993). "Lottery Choice: Incentives, Complexity and Decision Time," *Economic Journal,* 103, 1397–1417.

ERRATA Table 1. EU vs. GEU, HOMOSCEDASTIC AND HETEROSCEDASTIC ERROR

N=80

| | | Subjects with EU rejected in favor of GEU | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Data Set I | | | | Data Set II | | | |
| | | error specification | | | | error specification | | | |
| | df | $\sigma_0$ | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ | $\sigma_0$ | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ |
| Rank Dependent, Quiggin's | 1 | 39 (50) | 21 (28) | 33 (46) | 30 (43) | 37 (49) | 21 (33) | 26 (41) | 31 (47) |
| Rank Dependent, Power | 1 | 21 (27) | 13 (18) | 18 (25) | 9 (13) | 24 (18) | 15 (23) | 12 (19) | 2 (3) |
| Prospective Reference Theory | 1 | 21 (27) | 12 (17) | 18 (25) | 11 (16) | 21 (28) | 5 (8) | 7 (19) | 9 (14) |
| Weighted Utility | 2 | 29 (37) | 15 (20) | 22 (31) | 18 (26) | 22 (29) | 5 (8) | 12 (19) | 17 (26) |
| Quadratic Utility | 6 | 32 (41) | 10 (14) | 8 (9) | 9 (13) | 28 (37) | 6 (9) | 7 (11) | 16 (24) |
| Regret with Independence | 3 | 39 (49) | 22 (30) | 28 (34) | 24 (34) | 39 (51) | 12 (19) | 17 (27) | 23 (35) |
| EU not successfully estimated | | 1 | 6 | 8 | 10 | 4 | 16 | 16 | 14 |

Percentages in parenthesis.

ERRATA TABLE 2
HOMOSCEDASTIC vs. HETEROSCEDASTIC ERROR, VARIOUS PREFERENCE MODELS

N=80

| | | Subjects with Homoscedastic Error Rejected | | | | | |
|---|---|---|---|---|---|---|---|
| | | Data Set I Error Specification | | | Data Set II Error Specification | | |
| Risk Model | df | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ |
| Expected Utility | 1 | 15 (19) | 8 (10) | 9 (11) | 15 (20) | 12 (16) | 15 (20) |
| Rank Dependent, Quiggin's | 1 | 12 (15) | 13 (16) | 13 (16) | 9 (12) | 9 (12) | 15 (20) |
| Rank Dependent, Power | 1 | 10 (13) | 9 (11) | 10 (13) | 15 (20) | 11 (15) | 10 (13) |
| Prospective Reference Theory | 1 | 14 (18) | 12 (15) | 10 (13) | 17 (23) | 15 (20) | 15 (20) |
| Weighted Utility | 1 | 16 (20) | 13 (16) | 14 (18) | 13 (17) | 11 (15) | 18 (24) |
| Quadratic Utility | 1 | 11 (14) | 5 (6) | 15 (19) | 11 (15) | 7 (9) | 16 (21) |
| Regret with Independence | 1 | 16 (20) | 11 (14) | 11 (14) | 9 (12) | 7 (9) | 15 (20) |

Percentages in parentheses

ERRATA TABLE 3A
AIKAKE INFORMATION CRITERIA RANKINGS, DATA SET I

N = 80

| Model | AIC Rank | | | | | Singular Covariance | Average Rank AIC | | Average Rank, AIC C | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | | No S | S = 9 | No S | S = 9 |
| EU, $\sigma_1$ | 6 | 11 | 15 | 10 | 8 | 11 | 4.19 | 4.85 | 2.09 | 3.04 |
| EU, $\sigma_2$ | 12 | 9 | 10 | 9 | 8 | 8 | 4.21 | 4.69 | **2.00** | **2.70** |
| EU, $\sigma_3$ | 13 | 11 | 7 | 9 | 9 | 10 | **3.97** | 4.60 | **1.80** | **2.70** |
| RDQ | 17 | 15 | 7 | 12 | 12 | 1 | **3.65** | **3.58** | 4.28 | 4.22 |
| RDP | 7 | 7 | 6 | 9 | 14 | 1 | 5.08 | 5.13 | 4.73 | 4.79 |
| PRT | 7 | 8 | 9 | 6 | 9 | 1 | 5.31 | 5.35 | 4.99 | 5.04 |
| WU | 2 | 9 | 10 | 13 | 5 | 1 | 5.39 | 5.24 | 6.67 | 5.24 |
| QU | 8 | 3 | 7 | 4 | 5 | 1 | 6.34 | 6.38 | 8.66 | 8.66 |
| RI | 8 | 6 | 8 | 7 | 8 | 1 | 5.44 | 7.54 | 7.65 | 7.66 |

If a higher order (e.g., EU with heteroscedastic error, GEU with homoscedastic error) was not significantly superior to EU with homoscedastic error, the ranking criteria used EU with homoscedastic error.

ERRATA TABLE 3B

AIKAKE INFORMATION CRITERIA RANKINGS, DATA SET II

N = 80

| Model | Rank | | | | | Singular Covariance | Average Rank | | Average Rank, AIC C | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | | No S | S = 9 | No S | S = 9 |
| EU, $\sigma_1$ | 8 | 12 | 12 | 8 | 8 | 16 | 3.91 | 4.93 | 2.03 | 3.28 |
| EU, $\sigma_2$ | 8 | 13 | 10 | 7 | 8 | 16 | 4.88 | 4.99 | 3.06 | 3.21 |
| EU, $\sigma_3$ | 16 | 10 | 7 | 5 | 11 | 14 | **3.82** | 4.73 | **1.71** | **2.99** |
| RDQ | 17 | 9 | 13 | 11 | 11 | 4 | **3.64** | **3.91** | 5.80 | 5.96 |
| RDP | 5 | 5 | 9 | 13 | 9 | 4 | 5.11 | 5.31 | 6.71 | 6.83 |
| PRT | 6 | 6 | 5 | 8 | 11 | 4 | 5.26 | 5.45 | 6.78 | 6.90 |
| WU | 6 | 2 | 9 | 8 | 9 | 4 | 5.17 | 5.36 | 6.18 | 6.33 |
| QU | 6 | 2 | 5 | 5 | 3 | 4 | 6.55 | 6.55 | 5.22 | 5.22 |
| RI | 6 | 10 | 6 | 11 | 6 | 4 | 5.09 | 5.09 | 5.39 | 5.39 |

If a higher order (e.g., EU with heteroscedastic error, GEU with homoscedastic error) was not significantly superior to EU with homoscedastic error, the ranking criteria used EU with homoscedastic error.

ERRATA TABLE 4
AIKAKE INFORMATION CRITERION RANKINGS USING HOMOSCEDASTICITY AND $\sigma_2$

| | Data Set I | | | | | | Data Set II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rank | | | | Average Rank[a] | | Rank | | | | Average Rank[a] | |
| Model | 1 | 2 | 3 | S | AIC | AIC C | 1 | 2 | 3 | S | AIC | AIC C |
| EU, $\sigma_0$ | 3 | 12 | 3 | 1 | 6.25 | **1.09** | 10 | 7 | 11 | 4 | 5.49 | **1.24** |
| RDQ, $\sigma_0$ | 2 | 16 | 7 | 3 | **4.15** | **2.84** | **19** | 8 | 9 | 6 | **4.43** | 3.59 |
| RDP, $\sigma_0$ | 8 | 3 | 2 | 3 | 6.09 | 3.46 | 7 | 8 | 5 | 9 | 6.40 | 4.48 |
| PRT, $\sigma_0$ | 5 | 6 | 4 | 3 | 6.44 | 3.75 | 9 | 8 | 9 | 8 | 6.61 | 4.01 |
| WU, $\sigma_0$ | 2 | 3 | 6 | 6 | 7.23 | 7.60 | 9 | 2 | 4 | 10 | 6.69 | 8.04 |
| QU, $\sigma_0$ | 5 | 5 | 2 | 9 | 8.73 | 12.0 | 6 | 4 | 2 | 10 | 8.35 | 11.5 |
| RI, $\sigma_0$ | 3 | 7 | 4 | 6 | 7.44 | 9.76 | 8 | 6 | 9 | 10 | 6.31 | 9.54 |
| EU, $\sigma_2$ | **17** | 4 | 7 | 10 | 6.86 | 3.91 | **12** | 5 | 3 | 14 | 6.11 | **3.29** |
| RDQ, $\sigma_2$ | **20** | 16 | 12 | 10 | **5.34** | 6.44 | 9 | 15 | 10 | 15 | 4.46 | 7.50 |
| RDP, $\sigma_2$ | 5 | 4 | 6 | 10 | 8.99 | 7.56 | 4 | 1 | 6 | 18 | 8.19 | 8.86 |
| PRT, $\sigma_2$ | 4 | 6 | 7 | 10 | 9.15 | 7.63 | 5 | 3 | 5 | 19 | 7.78 | 7.60 |
| WU, $\sigma_2$ | 4 | 5 | 10 | 13 | 8.90 | 9.98 | 5 | 5 | 0 | 18 | 7.75 | 11.01 |
| QU, $\sigma_2$ | 4 | 8 | 4 | 12 | 10.6 | 12.8 | 8 | 2 | 1 | 19 | 9.16 | 12.5 |
| RI, $\sigma_2$ | 11 | 8 | 3 | 10 | 8.84 | 11.1 | 8 | 3 | 2 | 16 | 7.09 | 11.2 |

Estimates with singular variance/covariance matrices were included in the average rankings, with these models ranked lowest. In the event that more than one model was so unsuccessfully estimated for a subject, each model was assigned the same rank value.

ERRATA TABLE 5A

MODEL SELECTION COUNTS FOR DATA SET I, HETEROSCEDASTIC MODEL $\sigma_1$

N = 79

| | Node | | | | | | | | Path Consistent | | | Path Inconsistent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GEU Model | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total | (3,6) | (2,7) | Total | GEU/EU |
| RDQ | 6 | 5 | 33 | 35 | 6 | 14 | 11 | 48 | 57 | 13 | 5 | 22 | 21 |
| RDP | 3 | 5 | 19 | 52 | 3 | 11 | 11 | 54 | 65 | 10 | 5 | 14 | 10 |
| PRT | 3 | 4 | 19 | 53 | 3 | 8 | 13 | 55 | 61 | 8 | 4 | 18 | 11 |
| WU | 5 | 2 | 20 | 52 | 7 | 13 | 8 | 51 | 55 | 8 | 0 | 24 | 19 |
| QU | 4 | 4 | 26 | 45 | 8 | 12 | 10 | 49 | 46 | 6 | 0 | 33 | 30 |
| RI | 11 | 2 | 33 | 33 | 12 | 24 | 5 | 38 | 55 | 20 | 0 | 24 | 20 |

ERRATA TABLE 5B
MODEL SELECTION COUNTS FOR DATA SET I, HETEROSCEDASTIC MODEL $\sigma_2$

N = 79

| | Node | | | | | | | | Path Consistent | | | Path Inconsistent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GEU Model | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total | (3,6) | (2,7) | Total | GEU/EU |
| RDQ | 8 | 5 | 36 | 30 | 8 | 29 | 26 | 16 | 40 | 18 | 3 | 39 | 29 |
| RDP | 4 | 5 | 33 | 34 | 4 | 29 | 27 | 18 | 42 | 20 | 4 | 37 | 26 |
| PRT | 6 | 6 | 36 | 31 | 6 | 25 | 28 | 19 | 39 | 15 | 5 | 39 | 30 |
| WU | 7 | 4 | 31 | 37 | 7 | 16 | 31 | 25 | 35 | 8 | 3 | 44 | 31 |
| QU | 1 | 1 | 33 | 41 | 2 | 13 | 31 | 31 | 31 | 9 | 0 | 46 | 29 |
| RI | 7 | 4 | 39 | 34 | 8 | 24 | 28 | 19 | 34 | 15 | 1 | 45 | 33 |

ERRATA TABLE 5C

MODEL SELECTION COUNTS FOR DATA SET I, HETEROSCEDASTIC MODEL $\sigma_3$

N = 79

| GEU Model | Node | | | | | | | | Path Consistent | | | Path Inconsistent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total | (3,6) | (2,7) | Total | GEU/EU |
| RDQ | 5 | 8 | 35 | 31 | 5 | 32 | 24 | 18 | 42 | 22 | 4 | 37 | 23 |
| RDP | 5 | 5 | 36 | 33 | 5 | 24 | 23 | 18 | 42 | 22 | 3 | 37 | 24 |
| PRT | 2 | 10 | 36 | 31 | 2 | 34 | 25 | 18 | 45 | 24 | 7 | 33 | 22 |
| WU | 6 | 9 | 30 | 33 | 7 | 27 | 24 | 21 | 42 | 17 | 6 | 37 | 24 |
| QU | 6 | 12 | 30 | 31 | 10 | 27 | 24 | 18 | 38 | 11 | 5 | 41 | 27 |
| RI | 6 | 7 | 34 | 32 | 8 | 29 | 24 | 18 | 36 | 20 | 1 | 43 | 31 |

ERRATA TABLE 5D

MODEL SELECTION COUNTS FOR DATA SET II, HETEROSCEDASTIC MODEL $\sigma_1$

N = 76

| GEU Model | Node | | | | | | | | Path Consistent | | | Path Inconsistent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total | (3,6) | (2,7) | Total | GEU/EU |
| RDQ | 15 | 0 | 31 | 29 | 15 | 16 | 11 | 32 | 56 | 15 | 0 | 18 | 17 |
| RDP | 3 | 0 | 26 | 47 | 3 | 9 | 14 | 50 | 57 | 9 | 0 | 19 | 17 |
| PRT | 13 | 5 | 11 | 47 | 13 | 2 | 13 | 47 | 60 | 2 | 4 | 14 | 8 |
| WU | 11 | 12 | 8 | 44 | 19 | 6 | 13 | 37 | 50 | 4 | 2 | 25 | 14 |
| QU | 15 | 24 | 8 | 27 | 36 | 11 | 10 | 19 | 43 | 8 | 2 | 33 | 24 |
| RI | 16 | 11 | 22 | 27 | 24 | 16 | 9 | 27 | 54 | 15 | 1 | 22 | 20 |

ERRATA TABLE 5E

MODEL SELECTION COUNTS FOR DATA SET II, HETEROSCEDASTIC MODEL $\sigma_2$

N = 76

| | Node | | | | | | | | Path Consistent | | | Path Inconsistent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GEU Model | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total | (3,6) | (2,7) | Total | GEU/EU |
| RDQ | 10 | 0 | 34 | 32 | 10 | 20 | 9 | 37 | 62 | 20 | 0 | 14 | 14 |
| RDP | 1 | 0 | 23 | 52 | 1 | 9 | 11 | 55 | 58 | 8 | 0 | 18 | 16 |
| PRT | 11 | 2 | 20 | 43 | 11 | 7 | 12 | 46 | 55 | 6 | 0 | 20 | 15 |
| WU | 8 | 5 | 15 | 48 | 12 | 12 | 11 | 41 | 57 | 8 | 1 | 19 | 15 |
| QU | 15 | 19 | 11 | 31 | 30 | 16 | 5 | 25 | 50 | 10 | 2 | 26 | 22 |
| RI | 8 | 5 | 32 | 31 | 12 | 25 | 9 | 30 | 59 | 23 | 0 | 17 | 15 |

ERRATA TABLE 5F

MODEL SELECTION COUNTS FOR DATA SET II, HETEROSCEDASTIC MODEL $\sigma_3$

N = 76

| | Node | | | | | | | | Path Consistent | | | Path Inconsistent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GEU Model | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total | (3,6) | (2,7) | Total | GEU/EU |
| RDQ | 6 | 0 | 43 | 27 | 6 | 27 | 11 | 32 | 60 | 27 | 0 | 16 | 16 |
| RDP | 8 | 0 | 13 | 55 | 8 | 2 | 13 | 53 | 63 | 2 | 0 | 13 | 11 |
| PRT | 10 | 4 | 20 | 42 | 10 | 5 | 14 | 47 | 55 | 5 | 1 | 21 | 15 |
| WU | 7 | 7 | 20 | 42 | 11 | 14 | 15 | 36 | 48 | 9 | 1 | 28 | 20 |
| QU | 6 | 12 | 22 | 36 | 17 | 24 | 11 | 24 | 46 | 16 | 1 | 30 | 25 |
| RI | 6 | 6 | 38 | 26 | 10 | 30 | 11 | 25 | 56 | 28 | 0 | 20 | 18 |

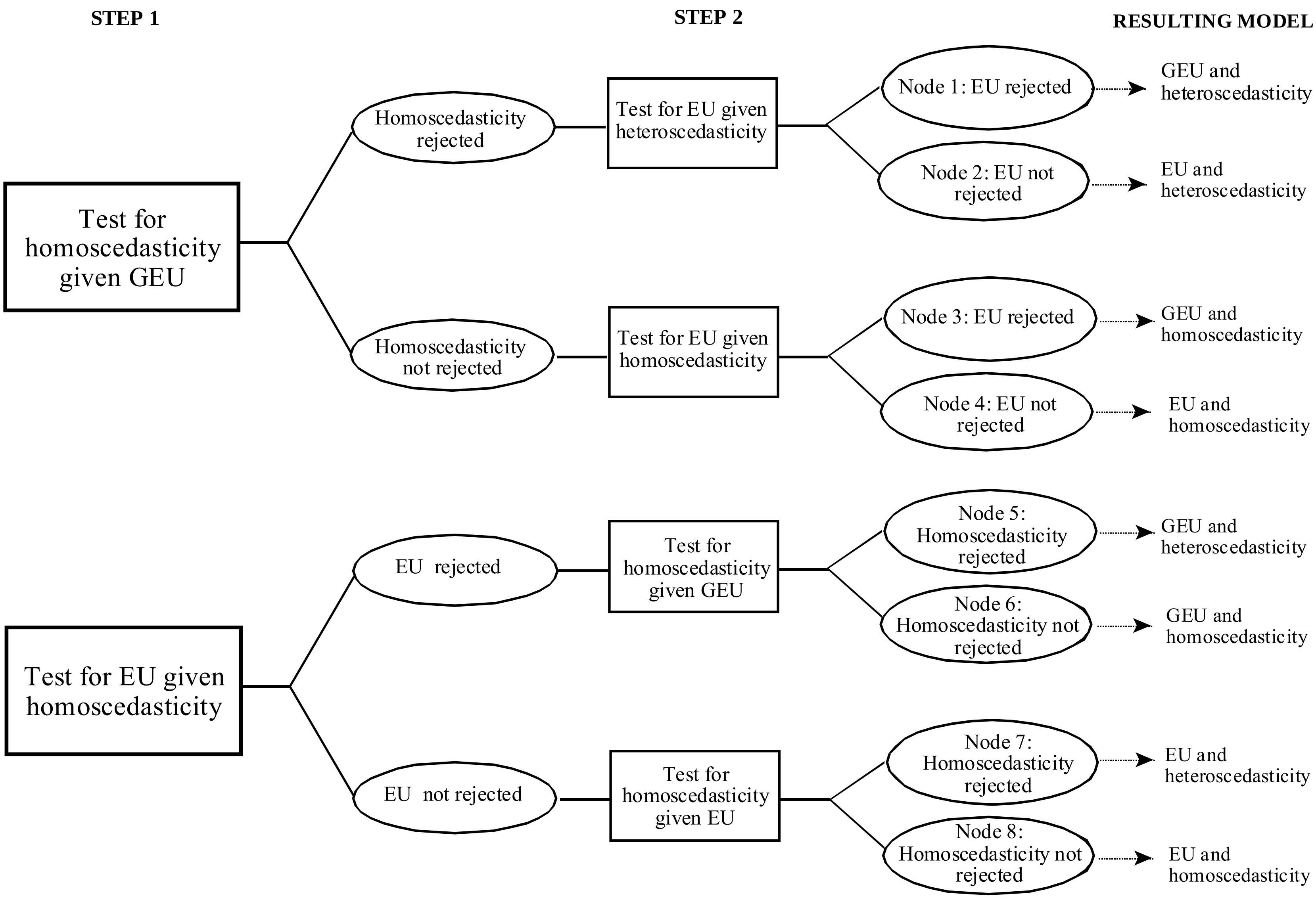
STEP 1
STEP 2
RESULTING MODEL
Test for homoscedasticity given GEU
Homoscedasticity rejected
Test for EU given heteroscedasticity
Node 1: EU rejected
GEU and heteroscedasticity
Node 2: EU not rejected
EU and heteroscedasticity
Homoscedasticity not rejected
Test for EU given homoscedasticity
Node 3: EU rejected
GEU and homoscedasticity
Node 4: EU not rejected
EU and homoscedasticity
Test for EU given homoscedasticity
EU rejected
Test for homoscedasticity given GEU
Node 5: Homoscedasticity rejected
GEU and heteroscedasticity
Node 6: Homoscedasticity not rejected
GEU and homoscedasticity
EU not rejected
Test for homoscedasticity given EU
Node 7: Homoscedasticity rejected
EU and heteroscedasticity
Node 8: Homoscedasticity not rejected
EU and homoscedasticity

**Figure 1**

**Appendix A:**

*Parametric representations of generalized expected utility preference functions*

For reference, we list each one of the six GEU preference functions we estimated for four-outcome lottery pairs, using the similar parametric forms as in Hey and Orme (1994). Each of these pairs nests EU as a special case. The parameters to be estimated for each model are also given, including the threshold parameter $t$ (defined as $t_{ij}$ in equation (7)).

For each pair of lotteries defined over the outcome vector $x$, the relative value of the lottery defined by $p$ over the lottery defined by $q$ is defined through model $j$ as $V_j(p, q, x, \alpha)$, using a vector of parameters α. The $V_j(\cdot)$ functions is defined for each GEU model $j$ below.

Note that $u(x_1)$ is normalized to equal 0 for each model.

1. *Rank-Dependent EU, Power formulation (RDP):* Quiggin (1982); Segal (1987); Chew, Karni, and Safra (1987); see also Yaari (1987)

$$V(p, q, x, \alpha) = \beta_2 x_2 \left[W_1(p) - W_1(q)\right] + \beta_3 x_3 \left[W_2(p) - W_2(q)\right] + \beta_4 x_4 \left[W_3(p) - W_3(q)\right].$$

where

$$W_1(r) = w(r_2 + r_3 + r_4) - w(r_3 + r_4),$$
$$W_2(r) = w(r_3 + r_4) - w(r_4), W_3(r) = w(r_4), w(r) = r^{\gamma}$$

Parameters to estimate: $\beta_2$, $\beta_3$, $\beta_4$, γ, and $t$. Reduces to EU if γ = 1.

2. *Rank-Dependent EU, Quiggin's "S-shaped" Formulation (RDQ):* Quiggin (1982)

As above for the power formulation, but with $w(r) = r^{\gamma} / [r^{\gamma} + (1-r)^{\gamma}]^{1/\gamma}$

3. *Prospective Reference Theory (PRT):* Viscusi (1989)

$$V(p, q, x, \alpha) = W(p) - W(q),$$

where

$$W(r) = \lambda\left[r_2\beta_2 x_2 + r_3\beta_3 x_3 + r_4\beta_4 x_4\right]$$
$$+ (1-\lambda)\left[c_2\beta_2 x_2 + c_3\beta_3 x_3 + c_4\beta_4 x_4\right]$$

where $c_i = 1/n(r)$ if $r_i > 0$, 0 otherwise. The function $n(r)$ is the number of non-zero elements in $r$, giving a Bayesian uniform prior over the outcomes.

Parameters to estimate: $\beta_2$, $\beta_3$, $\beta_4$, $\lambda$, and $t$. Reduces to EU if $\lambda$=1.

4. *Weighted Utility (WU):* Chew (1985), and Dekel (1986).

$$V(p, q, x, \alpha) = W(p) - W(q),$$

where

$$W(r) = \left[w_2 r_2 \beta_2 + w_3 r_3 \beta_3 + w_4 r_4 \beta_4 +\right] / \left[r_1 + r_2 w_2 + r_3 w_3 + r_4\right]$$

Parameters to estimate: $\beta_2$, $\beta_3$, $\beta_4$, $w_2$, $w_3$, and $t$. Reduces to EU if $w_2 = w_3 = 1$.

5. *Quadratic Utility (QU):* Chew, Epstein, and Segal (1991).

$$V(p, q, x, \alpha) = W(p) - W(q),$$

where

$$\sum_{i=1}^{4}\sum_{i=k}^{4}\beta(x_i, x_k) r_i r_k \text{ and } \beta(x_k, x_i) = \beta(x_i, x_k).$$

As in Hey and Orme, we normalize $\beta(x_1, x_1) = 0$. Define $\beta(x_i, x_k) = \beta_{ij}$.

Parameters to estimate: $\beta_{12}$, $\beta_{13}$, $\beta_{14}$, $\beta_{22}$, $\beta_{23}$, $\beta_{24}$, $\beta_{33}$, $\beta_{34}$, $\beta_{44}$, and $t$.

Reduces to EU if $\beta_{ij} = [\alpha_i x_i + \alpha_k x_k]/2$ for all $i$, $k$.

6. *Regret with Independence (RI):* Loomes and Sugden (1982)

$$V(p, q, x, \alpha) = (p_1 q_2 - p_2 q_1)\beta_{21} + (p_2 q_3 - p_3 q_2)\beta_{32}$$
$$+ (p_3 q_4 - p_{42} q_2)\beta_{43} + (p_1 q_3 - p_3 q_1)\beta_{31}$$
$$+ (p_2 q_4 - p_4 q_2)\beta_{42} + (p_1 q_4 - p_4 q_1)\beta_{41}$$

Parameters to estimate: $\beta_{21}$, $\beta_{32}$, $\beta_{43}$, $\beta_{31}$, $\beta_{42}$, $\beta_{41}$, and $t$. Reduces to EU if $\beta(x_i, x_k) = \alpha_i x_i - \alpha_k x_k$ for all $i$, $k$.