Braga, Jacinto

Working Paper

Is preference reversal just stochastic variation?

CeDEx Discussion Paper Series, No. 2003-05

Provided in Cooperation with:
The University of Nottingham, Centre for Decision Research and Experimental Economics (CeDEx)

Suggested Citation: Braga, Jacinto (2003) : Is preference reversal just stochastic variation?, CeDEx Discussion Paper Series, No. 2003-05, The University of Nottingham, Centre for Decision Research and Experimental Economics (CeDEx), Nottingham

This Version is available at:
https://hdl.handle.net/10419/67964

Is Preference Reversal Just Stochastic Variation?*

Jacinto Braga
CeDEx, University of Nottingham
lexjcb@nottingham.ac.uk

Abstract. This paper investigates whether the preference reversal phenomenon can be accommodated by a stochastic model of expected utility. The model is based on Loomes and Sugden's (European Economic Review, 1995) theory of random preference. Its central assumption is that each individual has a set of preference orderings and a probability distribution over that set. Each decision is made according to a preference ordering drawn at random from that set. There are probability distributions over sets of preference orderings, call that *preference distributions*, that predict the observed asymmetric reversal patterns. These preference distributions are however hard to justify. Moreover they cannot explain the symmetric patterns of reversal that have been observed after repetition and feedback in some experiments (which different, more easily justifiable preference distributions can explain). The model casts doubts on a widely used measure of reversals, and on some conclusions based on that measure, such as the famous observation by Grether and Plott (American Economic Review, 1979) that incentives made preference reversal stronger.

There is abundant, mainly experimental observation of behaviour deviating from the predictions of the most widely used theory of rational risky choice in economics, expected utility theory. Attempts to accommodate these deviations within a rational choice framework have followed two approaches. The most common was the development of new theories along the traditional deterministic line, such as regret theory (Loomes and Sugden 1983) or rank dependent expected utility (Quiggin 1982). Another was the addition of stochastic elements to existing deterministic theories (Loomes and Sugden 1995, Hey and Orme 1994, and Harless and Camerer 1994).

We are concerned here with one of the most notorious deviations of observed behaviour from that predicted by theories: preference reversal. In the typical preference reversal experiment subjects are asked to choose between one safe and one risky gamble and to place monetary values on them. Many subjects choose the safe gamble but value the risky one more highly. This has been called *standard reversal*.

* I would like to thank Prof. Chris Starmer and Dr Henrik Orzen for their helpful comments. Any mistakes are solely of my own responsibility.

Choosing the risky bet while valuing the safer one more highly is far less frequent, and has been called *non-standard reversal*.

Several explanations for preference reversal based on deterministic preference theories have been proposed by Loomes and Sugden (1982), Holt (1986), and Karni and Safra (1987) among others. These explanations stimulated further experimental research and were eventually contradicted (see Tversky and Kahneman 1990, and Cubitt et al forthcoming, or Braga 2003 for a review of that literature). Sugden (forthcoming) proposes an explanation for preference reversal based on his newly developed deterministic theory of reference-dependent expected utility. Empirical research is yet to address this new explanation.

The predominant view among experimentalists seems to be that it is the asymmetric pattern of inconsistencies between choice and valuation ranking that challenges preference theory, not each inconsistency itself. Therefore the presence of stochastic elements in decision making has been accepted, at least implicitly. However the potential, or lack of it, of these stochastic elements in explaining preference reversal is largely unexplored. Lichtenstein and Slovic (1971), the discoverers of preference reversal, proposed and rejected a model of errors in choices and valuations. Since then experimental results have been largely accepted as systematic deviations from preference theory without evaluation against a stochastic model.

To ascertain that preference reversals were not mere random deviations from deterministic theoretical predictions most authors, in the absence of a formal statistical test, have relied on two features of the experimental results: in many studies standard reversals far outnumber those of the non-standard type, thus, on face value, the pattern appears not to be random; when a standard reversal occurs the difference between the values placed on the risky and on the safe bets is often larger than one might expect from randomness in decisions alone. Evaluation of these two features has thus provided a surrogate, informal test of the significance of preference reversal. Convincing as this informal test may have seemed, it amounts to testing a stochastic model without specifying it.

This paper is an attempt to improve this state of affairs. We will revisit Lichtenstein and Slovic's (1971) error model, develop a random expected utility model of choices and valuations, and test whether the data obtained from preference reversal experiments could be generated by those models. The random preference model to be developed here is rooted at Loomes and Sugden's (1995) theory of random preference. We will see that among the stochastic models of choice proposed recently in the literature Loomes and Sugden's theory is the one that can be extended more directly to valuations. We will also see that two special cases of the Lichtenstein and Slovic's and the random preference models of choices and valuations are observationally equivalent.

The remainder of this paper is as follows. Section 1 presents formally the preference reversal phenomenon, specifies the type of data that the models need to predict, and sets out the testing procedure. Section 2 briefly reviews the literature on stochastic decision models. Section 3 revisits the Lichtenstein and Slovic's error model. Section 4 develops the random preference model of choices and valuations. Section 5 fits the two models to preference reversal data. And section 6 concludes.

1. Preference Reversal

In a typical preference reversal experiment each subject makes a triple of decisions for each pair of lotteries: a choice between the two lotteries and two, usually monetary, valuations, one for each lottery. Thus we obtain N triples of decisions, where N is the product of the number of subjects multiplied by the number of pairs of lotteries. If we ignore instances of equal values given to both bets, which usually account for less than 5% of all triples, each of these triples of decisions can be placed into one of four categories, as in the table below. *PP* is the number of triples where the *P* bet is chosen over and valued above the *$* bet; *$$* is the reverse; *SR*, the number of standard reversals; and *NSR* the number of non-standard reversals.

Table 1: Categories of triples

Choices	Highest Value		total
	P	*$*	
P	*PP*	*SR*	*P + SR*
$	*NSR*	*$$*	*NSR + $$*
total	*P + NSR*	*SR + $$*	*N*

Only categories *PP* and *$$* are compatible with deterministic expected utility theory. Some other deterministic preference theories allow categories *SR*, *NSR* or both, but, as mentioned above, of these theories only Loomes and Sugden's (forthcoming) reference-dependent expected utility provides an explanation for preference reversal that has not been refuted by experimental research.

Models of stochastic choices and valuations may allow all four categories. A model that predicts the expected frequencies of the four categories is testable. The expected frequencies predicted by a model can be compared with the observed frequencies by means of a chi-squared test of goodness-of-fit to test whether the observed frequencies could have been generated by the model. That is, using the subscripts O and E to denote observed and expected frequencies respectively, and if all expected frequencies are at least five, the variable

$$\chi^2 = \frac{(PP_O - PP_E)^2}{PP_E} + \frac{(SR_O - SR_E)^2}{SR_E} + \frac{(NSR_O - NSR_E)^2}{NSR_E} + \frac{(\$\$_O - \$\$_E)^2}{\$\$_E} \qquad (1)$$

follows approximately a chi-square distribution with degrees of freedom equal to three (number of independent categories) minus the number of model parameters estimated from the observed frequencies.

The statistical test is feasible if there are at most two parameters to be estimated from the data. This poses a problem to both models under consideration here, as they have both three parameters. In section 3 we apply a testing procedure to the Lichtenstein and Slovic's (1971) error model that, in some circumstances, is able to reject the model. In section 5 we do the same with the random preference model of choices and valuations.

2. Stochastic Models of Decisions

Three stochastic theories of choice under uncertainty have emerged in recent literature: the random error theories of Hey and Orme (1994) (HO) and Harless and Camerer (1994) (HC), and the random preference theory (RP) of Loomes and Sugden (1995). All three theories may be seen as a combination of a deterministic preference theory and a stochastic element. Loomes and Sugden (1995, 1998) named the deterministic component of a stochastic theory the *core theory*, and the stochastic element, the *stochastic specification*. We will adopt these terms. The HO and RP models with expected utility as their core theories were first presented by Becker et al (1963). Hey and Orme (1994) and Loomes and Sugden (1995) generalised those models to other core theories.

These three theories are theories of choice only, not of valuations. We will see in section 4 that in most preference reversal experiments valuations were, according to preference theory, equivalent to choices of one gamble from a set of many gambles. Hey and Orme´s (1994) and Harless and Carmerer's (1994) theories apply to pairwise choices only. Therefore they cannot be extended in an obvious way to valuations. Loomes and Sugden's (1995) theory, as long as the core theory is transitive, applies to choices from sets of any number of gambles. Therefore its extension to valuations is straightforward.

The need to review the Loomes and Sugden's (1995) random preference theory of choice is obvious. We will also review the two error theories to put the Lichtenstein and Slovic's and the random preference models of choices and valuations in context.

2.1. Three Stochastic Specifications

In the HC and HO error models each individual is thought of as having unique deterministic preferences, as in any deterministic preference theory, but his decisions may deviate from his unique preferences because of a random error. This random error may be caused by failure to understand the tasks, inattention, or mistakes in processing complex information. In Loomes and Sugden's (1995) random preference theory, an individual, rather than unique preferences, has a collection of preferences,

and makes each decision according to some preferences drawn at random from his collection. In this specification individuals do not make errors, but their preferences are variable or imprecise.

In the HC model, an error is the choice of the truly least preferred option. This happens with a probability that is constant across individuals and choice tasks. Thus this model may be called *the constant probability error model*.

In the HO error model the probability that the least preferred gamble be chosen varies across individuals and choice tasks. The true difference in subjective value between two gambles X and Y is given by $V(X, Y)$, with X preferred to Y if and only if $V(X, Y) > 0$.[1] X is chosen over Y if and only if $V(X, Y) + \varepsilon > 0$, where ε is normally distributed with mean zero. The term ε may be interpreted as an error in the evaluation of the relative value of X and Y. Let the operator $\succ_c$ mean *is chosen over*. X is chosen over Y with probability $\Pr(X \succ_c Y) = \Pr(V(X, Y) + \varepsilon > 0) = \Pr(\varepsilon > -V(X, Y)) = \Pr(\varepsilon < V(X, Y))$ (because ε is symmetrically distributed around zero). Thus the larger $V(X, Y)$, that is, the more strongly X is preferred to Y, the more likely that X be chosen over Y.

Loomes and Sugden's (1995) theory of random preference is as follows. Consider a set A of conceivable prospects, and a complete, reflexive, and transitive binary preference relation, $\geq_p$, on A, that is, $\geq_p$ is a preference ordering on A. Let R be the set all possible preference orderings on A. The random preference model assumes that for each individual there is an additive probability measure f on R, so that any subset R' of R has the probability $f(R')$. Whenever an individual has to choose an element from a non-empty subset S of A, the elements of S are ranked according to a preference ordering r_i drawn at random from R. The individual will then choose the highest ranked element of S, or one of the highest ranked elements with equal probability if there is more than one element tied at the top of the preference ordering. As r_i is a preference order there will be at least one highest ranked element in S. A given individual could make all its conceivably possible choices according to only a small subset R^* of R. This means that $f(R^*) = 1$, and R^* defines the core theory

[1] In most of the commonly known core theories $V(X, Y)$ can be expressed as $u(X) - u(Y)$, that is, the difference between two utilities, but in others, as regret theory, it cannot.

governing that individual's preferences. Given two gambles X and Y, let $R_{X>Y}$ and $R_{X=Y}$ be the subsets of R that rank X above Y and X equally with Y respectively. Then, in a pairwise choice, $\Pr(X \succ_c Y) = f(R_{X>Y}) + 0.5 \times f(R_{X=Y})$.

2.2. Experimental Evidence

Before exploring how these models could apply to valuations, we will review how they have fared in explaining choice data.

Hey and Orme's (1994) fitted preference functions, coupled with the normal random error, separately to each individual's choices. Their aim is to compare the predictive power of expected utility and its generalisations, not evaluate the HO stochastic specification itself. Still their conclusion that behaviour can be reasonably approximated by expected utility plus noise implicitly approves the HO specification.

Harless and Camerer (1994) also compare the predictive power of expected utility and its generalisations, but associated with the constant probability error. However their method is different, and allows them to reject the model, regardless of the core theory.

Carbone and Hey (1998) used the same method as in Hey and Orme (1994) to compare the two random error specifications coupled with several core theories. They concluded that the stochastic specification that performs best depends on the core theory and on the individual.

Loomes and Sugden (1998) derive from the three stochastic specifications predictions that are independent of the core theory. None of the predictions was supported by the data, but the predictions of a model combining random preference and a constant probability error seemed compatible with the data. Additionally the authors derived implications of the three stochastic specifications when coupled with expected utility as the core theory. All models were rejected on this score, but the deviation from the predictions of the models appeared to be subsiding during the course of the experiment.

Loomes, Moffat, and Sugden (2002) conducted an econometric analysis of Loomes and Sugden's (1998) data. They estimated stochastic versions of expected and

rank-dependent utility. The stochastic specifications were random preference plus constant probability error, and the HO specification with and without a constant probability error. They conclude that rank dependent performs better than expected utility; the HO plus constant probability error performs better than the HO model alone; and the RP plus constant probability error model performs best of all. They also find evidence, in line with Loomes and Sugden (1998), that the various rank-dependent utility models were converging over time towards expected utility.

The balance of evidence appears to be that the best results will be obtained with a combination of two stochastic combinations (Loomes and Sugden 1998, Loomes et al 2002). The constant probability error specification in isolation appears to be particularly inadequate. Among the studies reviewed, this model finds some support only in Carbone and Hey's (1998) findings. Note that this study only compares the two error specifications. It includes no test to reject or accept either of them, contrary to Harless and Camerer (1994) or Loomes and Sugden (1998).

2.3. Valuations

As we mentioned above, and as we will see below, under the elicitation methods used in most preference reversal experiments, a valuation task was equivalent to a choice of a gamble from a set a many gambles, specifically, more than two. This means that theories of pairwise choice, as the HO or the HC error models, are not directly applicable to valuations.

One could think that a model of pairwise choice could determine a choice of one from more that two gambles by means of a chain of pairwise choices. Say, to choose a gamble from $\{X, Y, Z\}$, one could first choose a gamble from $\{X, Y\}$. Suppose X was chosen. Then one would chose from $\{X, Z\}$. The HO and HC stochastic specifications pose a problem to this chain method: given any three or more gambles, both models entail intransitive choice cycles with some positive probability. For the chain method to be guaranteed to arrive at a final choice one would have to make further assumptions. For instance one could rule out choice cycles. That is, one could assume that after a gamble had been rejected in a pairwise choice it would not be

object of any further choices. The random preference model also allows intransitive choice cycles. This is not a problem though, because a single choice is enough to select one from many options.

Interpreting a valuation as a choice is not merely a convenient way of applying the random preference theory to valuations. For reasons that will be clearer below we believe that this interpretation of valuations is precisely what preference theory prescribes.

If the type of errors assumed in the HC and HO stochastic specifications exist and influence people's choices and valuations it must be possible to model such errors in valuations, or, more generally, in choices of one from more than two gambles. Lichtenstein and Slovic's (1971) modelled error in valuation ranking (but not in non-binary choice), and we now turn to their model.

3. Lichtenstein and Slovic's Error Model

Lichtenstein and Slovic (1971) assume that the probability that a subject truly prefer the *P* bet over the *$* bet in a choice task is the same as the probability that he truly value *P* above *$* in the valuation tasks. Value here refers to the true subjective value. These may, because of error, differ from the monetary value actually placed on a gamble. Call this single probability *p*. Because of random errors subjects choose their least preferred bet with probability *r*, and value their least preferred bet above their preferred bet with probability *s*.

Note that a single probability *p* is an implication of deterministic expected utility, but not of other deterministic preference theories, such as rank-dependent utility, or regret theory.

3.1. Behavioural Rationale and Error Range

Lichtenstein and Slovic (1971) do not offer any rationale for their stochastic specification. Nor do they indicate the possible range of *r* and *s*. The authors did not need to specify these ranges, as they derived a testable hypothesis that is independent

of r and s. Before moving to the predictions of their model, we will discuss what rationale may be behind the model, and what limits can be imposed on r and s.

The error specification in choices is the same as that in Harless and Camerer (1994). The error in a valuation task is not specified. What the authors specify is an error in the valuation ranking, which depends on the comparison of two independent valuations.

Loomes et al (2002) suggest that a constant probability error in binary choices could arise from a subject failing to understand the task or suffering a lapse of concentration, in which case his choice would be unconnected with his preferences, that is, each option would be chosen with 50% probability. Then the probability that the least preferred bet be chosen, r, is at most 50%, in the extreme case where all choices are random.

Adopting the same rationale for valuations, a subject's stated value for a lottery could sometimes be unconnected with his true value. For instance, each possible value could be selected with equal probability. Consider the lottery $P = (y_P, p_P)$. It offers an amount of money y_P with probability p_P, and zero with probability $1 - p_P$. The paired lottery is $\$ = (y_\$, p_\$)$. Suppose that when a valuation is random the value placed on the bet is uniformly distributed between zero and the winning amount. This hypothesis implies different probabilities of error in valuation rankings, depending on whether P or $\$$ is preferred, contrary to Lichtenstein and Slovic's (1971) assumption of a single such probability.

Denote v_P and $v_\$$ the values stated by the subject. When a valuation is random the probability density functions are $f_P(v_P) = 1/y_P$ with $0 \le v_P \le y_P$, and $f_\$(v_\$) = 1/y_\$$ with $0 \le v_\$ \le y_\$$. If both valuations are random, under the natural assumption that the two valuations are independent, the joint probability density function is $f(v_P, v_\$) = 1/(y_P \times y_\$)$. Denote q the probability that a valuation is random, and v_P^* and $v_\$^*$ the true values. Then, if P is truly preferred over $\$$, the probability of error in the valuation ranking is

$$s_P = \Pr(v_\$ > v_P \mid v_P^* > v_\$^*) = q^2 \Pr(v_\$ > v_p \mid v_\$ \text{ and } v_\$ \text{ are random}) +$$

$$q(1-q)\Pr(v_\$ > v_P^* \mid v_\$ \text{ is random and } v_P = v_P^*) +$$

$$q(1-q)\Pr(v_P < v_\$^* \mid v_P \text{ is random and } v_\$ = v_\$^*).$$

$$\Pr(v_\$ > v_P \mid v_P^* > v_\$^*) = \int_0^{y_P}\int_{v_P}^{y_\$} \frac{1}{y_P \times y_\$} dv_\$ dv_P = 1 - \frac{y_P}{2y_\$}.$$

$$\Pr(v_\$ > v_P^* \mid v_\$ \text{ is random and } v_P = v_P^*) = \frac{y_\$ - v_P^*}{y_\$}.$$

$$\Pr(v_P < v_\$^* \mid v_P \text{ is random and } v_\$ = v_\$^*) = \frac{v_\$^*}{y_P}.$$

Thus

$$s_P = \Pr(v_\$ > v_P \mid v_P^* > v_\$^*) = q^2\left(1 - \frac{y_P}{2y_\$}\right) + q(1-q)\left(\frac{y_\$ - v_P^*}{y_\$} + \frac{v_\$^*}{y_P}\right).$$

It can be shown in the same manner that if $ is truly preferred over *P*, the probability of error in valuation ranking is

$$s_\$ = \Pr(v_P > v_\$ \mid v_\$^* > v_P^*) = q^2\frac{y_P}{2y_\$} + q(1-q)\left(\frac{v_P^*}{y_\$} + \max\left(\frac{y_P - v_\$^*}{y_P}, 0\right)\right).$$

The probability of error in valuation ranking when *P* is truly preferred to $ may be extremely high. For instance for the pair of bets *P* = (£8, 97%) and $ = (£32, 31%), which has been used in several experiments, s_P can be as high as 0.875, when $q = 1$. When $ is preferred over *P* that probability is at most 0.125, again when $q = 1$. It can be checked that with this pair of bets, for all values of q and any minimally plausible pairs of true values (say, $v_\$^*, v_P^* >$ £3,) s_P is substantially higher than $s_\$$.

This stochastic specification of the valuations suggests a model of errors that is different from Lichtenstein and Slovic's (1971). In this model, s would be replaced by s_P and $s_\$$, derived above, and $r = q/2$. We will not pursue this model for two reasons.

This model needs at least six parameters: p, the probability that a subject truly prefer the *P* bet, q, and two pairs of certainty equivalents, one with $v_\$^* > v_P^*$, and another with $v_P^* > v_\$^*$. Testing against our categorical data requires that no more than two parameters be estimated. This problem could possibly be overcome by assuming

values for the certainty equivalents, especially as their variation within plausible ranges causes little change in s_P and $s_\$$.

A more fundamental reason not to pursue this model is that our stochastic specification is an unrealistic account of all the errors that may occur in valuations. The misunderstandings and lapses of concentration suggested by Loomes et al (2002) as a cause of the constant probability error may occur in valuations. It is fair to assume that these misunderstandings and lapses lead to choosing each of the options in a binary choice with equal probability, but it is extreme to suppose that they lead to valuations uniformly distributed between zero and the winning amount. To put it in another way, lapses and misunderstandings leading to those uniformly distributed valuations should be infrequent. Most errors are more likely to be smaller, more symmetrically distributed deviations around the true value than implied by valuations uniformly distributed between zero and the winning amount.

These considerations suggest that the s_P derived above constitutes an upper bound for a probability of error in valuation rankings implied by a more realistic stochastic specification of errors in valuations. When $v_\$$ is uniformly distributed between 0 and $y_\$$, \$ is very likely to be valued above v_P, because $v_P \leq y_P$, and generally y_P is much smaller than $y_\$$. This leads to a high s_P. When P is preferred to \$, $v^*_\$ < v^*_P \leq y_P$. Thus, smaller, more symmetrically distributed deviations around $v^*_\$$ should lead to a smaller s_P than a uniform distribution of $v_\$$ between 0 and $y_\$$. Imposing in Lichtenstein and Slovic's (1971) model, $s \leq s_P$ is possibly a very loose constraint, but without a stochastic specification of error in each valuation it is not possible to characterise s precisely.

While a valuation distributed uniformly between zero and the winning amount when an error occurs is unrealistic, it is not clear what a realistic specification of errors in valuations should be. It is also not clear what rationale may lie behind the Lichtenstein and Slovic's (1971) assumption of a single probability of error in valuation ranking. We will derive the predictions of Lichtenstein and Slovic's (1971) model, and fit it to data, as it was used before in the literature, and bears some similarities with the random preference model of choices and valuations. We will impose $r \leq 0.5$, as discussed above. We could compute an upper bound for s, based on

s_P, for each pair of bets. However we will ignore this bound, as it is never binding in our estimates (s is zero in all but one estimate).

3.2. Predictions of the Lichtenstein and Slovic's Error Model

According to the model a reversal of one type or the other will occur whenever a subject makes a mistake either in the valuation ranking or in the choice, but not in both. The model predicts the probability that a triple of decisions will fall on each of the four categories. We shall call these four probabilities the category probabilities. The category probability distribution is given in table 2.

Table 2: Category probability distribution under the error model

Choices	Highest Value		total
	P	*$*	
P	$p(1-r)(1-s) + (1-p)rs$	$p(1-r)s + (1-p)r(1-s)$	$p(1-r) + (1-p)r$
$	$(1-p)(1-r)s + pr(1-s)$	$(1-p)(1-r)(1-s) + prs$	$(1-p)(1-r) + pr$
total	$p(1-s) + (1-p)s$	$(1-p)(1-s) + ps$	1

Key: p is the probability that P is truly preferred. r and s are the probabilities of error in choice and valuation ranking respectively.

For instance a standard reversal occurs if a subject prefers the P bet, chooses correctly, and ranks the two bets by value incorrectly (this will happen with probability $p(1-r)s$), or if he prefers the $ bet, chooses incorrectly, and ranks the paired bets by value correctly (which will happen with probability $(1-p)r(1-s)$). The other category probabilities may be found in a similar manner.

In any random sample the number of triples in each of the four categories is random. Their expected value predicted by the model is the product of the number of triples by the category probability. These predictions can be compared in a statistical test with the corresponding, observed sample frequencies. For that one needs to find values for the parameters p, r, and s. Lichtenstein and Slovic (1971, p. 51), tried to find parameters that predicted exactly the proportions observed in their samples. As only complex values for p satisfied that condition, they settled for what they deemed to be

an approximate real solution,[2] $p = 0.5$. If $p = 0.5$ the expected frequencies of standard reversals are equal to those of non-standard reversals for all r and s. This hypothesis was then rejected by the McNemar's test for correlated proportions with probability below 1% in all their samples.

This procedure is by no means satisfactory. The authors merely tested the hypothesis that $p = 0.5$. Instead we will test the maximum likelihood estimates of all three parameters. The likelihood function was maximised subject to $0 \leq p \leq 1$, $0 \leq r \leq 0.5$, and $s \geq 0$. Table 3 shows the maximum likelihood estimates of all parameters for Lichtenstein and Slovic's (1971) three treatments. The maximum likelihood estimates of p turn out to be quite different from 0.5.

Table 3: Maximum likelihood estimates of the error model, Lichtenstein and Slovic's (1971) data

Treatment	Incent compat[a]	Sample size	p	r	s	χ^2	Pr (1 degree of freedom)[b]
I – Selling prices	No	1038	0.12	0.46	0.00	19.8	0.00
II – Buying prices	No	444	0.39	0.40	0.00	5.4	0.02
III – Selling prices	Yes	84	0.30	0.37	0.00	6.7	0.01

[a] An experiment is incentive compatible if subjects' decisions have economic consequences for them.

[b] The assumption of one degree of freedom is too favourable to the model, but still rejects it. See text.

With three independent data categories and three estimated parameters we have no degrees of freedom. Even if we had an extra category the distribution of the χ^2 statistic would be unknown because the derivative of the likelihood function with respect to s at the optimal solution is not zero (the constraint $s \geq 0$ is binding). Nevertheless we are still able to reject the model with all three samples. If we assume

[2] Equating the proportions in table 2 to observed proportions in their three samples yielded the following values for $p(1\text{-}p)$: 0.295, 0.315, and 0.27. The authors thought that these values were not too far from 0.25, which is the maximum of $p(1\text{-}p)$ for any real p (p=0.5). It may look so, but the true sample solutions are $0.5\pm0.21i$, $0.5\pm0.26i$, and $0.5\pm0.14i$. To better appreciate the approximation error, if the sample values for $p(1\text{-}p)$ had been 0.205, 0.185, and 0.23, which are as far from 0.25 as the actual sample values, but yield real solutions for p, changing them to 0.25 would give rise to approximation errors in the value of p of 0.21, 0.26, and 0.14. These do not look small at all.

some value for one of the parameters, and estimate only the remaining two the χ^2 statistic will gain one degree of freedom. Suppose we assume a value for s. Assuming a value for s instead of for one of the other two parameters solves the problem of the non-null derivative. Regardless of the value we assume, the χ^2 statistic will not be smaller than values shown on the table, which were obtained with the estimation of all three parameters. Therefore if these values of χ^2 lead to the rejection of the model assuming one degree of freedom, as it is the case, the model may be confidently rejected. If they did not the test would be inconclusive.

Many statistics textbooks caution that failing to reject the null hypothesis should not be interpreted as its acceptance. This word of caution has a special importance in the testing procedure we have just applied. This test uses a chi-squared statistic based on three estimated parameter as if only two of them had been estimated. Therefore this test could never be used to accept a model. We will use this testing procedure with minor variations in section 5.

Instead of maximising the likelihood function we could minimise the χ^2 statistic on p, r, and s. This would give the model its best chance of not being rejected. As it happens the minimum χ^2 estimates are only negligibly different from the maximum likelihood estimates.

4. A Random Preference Model of Choices and Valuations

We will begin by examining which elicitation methods and core theories allow the Loomes and Sugden's (1995) random preference theory to apply to valuations. Next we derive the category probabilities predicted by the model. In the most general model, the category probabilities depend on three parameters. We will then restrict our core theory to expected utility, and derive the constraints it imposes on the model parameters. These constraints turn out to be rather weak if one accepts all manner of sets of expected utility preference orderings (subsets of R) and probability measures (f) over those sets. In a final subsection we restrict our core theory further to preferences described by power functions, and the probability measure f to truncated normal distributions (of the power parameter).

4.1. Valuations

If valuations are elicited in a second-price auction or with the BDM procedure (Becker, DeGroot, and Marshak 1964) application of the model is straightforward, as such a valuation can be seen as a choice among many gambles. The same is true of a second-to-last price auction, used in Braga and Starmer (2001), or, generally, of an n^{th}-price auction. In an n^{th}-price auction, in a group of n or more subjects, $n-1$ sell their lotteries for a price equal to the n^{th} lowest bid. From the point of view of the subject, the BDM procedure is equivalent to a second-price, selling auction where he bids against a device that makes random bids. Therefore the following argument is also valid for the BDM.

Suppose a subject is bidding to sell a lottery X in an n^{th}-price auction. For ease of exposition assume that the set of admissible bids is $B_X = \{b_1, b_2, \ldots, b_N\}$, with $b_i < b_{i+1}$ for $i = 1, \ldots, N-1$. These restrictions on the admissible bids, discreteness and upper and lower limits, are common in experiments that use auctions to elicit valuations. For instance, bids must usually be a whole number of pence (or the lowest subdivision of the relevant currency) and no higher than the highest outcome of the gamble. Elicitation with the BDM procedure usually imposes no restrictions on subjects' bids. The random counter bids are however drawn from a set as B_X. Therefore the subject will be indifferent between his unrestricted bid and the restricted bid immediately bellow or above, depending on the rule to resolve ties. Let w be the market price, which is the n^{th} lowest bid, therefore $w \in B_X$. Let $G(w)$ be the cumulative probability distribution the bidder attributes to w, that is $G(w_0) = \Pr(w \leq w_0)$, and $G(b_N) = 1$. For simplicity assume that all subjects make different bids. Suppose the individual bids b_k to sell X. If the market price is lower than or equal to b_k the individual keeps and plays X. The individual attributes to this event the probability $\Pr(w \leq b_k) = G(b_k)$. If the market price is higher than b_k, for instance $w = b_l > b_k$, the individual sells X for b_l. The individual attributes to this the probability $\Pr(w = b_l) = G(b_l) - G(b_{l-1})$.Then when a subject bids b_k to sell a lottery X he is in effect choosing to play the following compound lottery:

$$C(b_k, X) = [X, G(b_k); b_{k+1}, G(b_{k+1}) - G(b_k); \ldots ; b_N, 1 - G(b_{N-1})]. \quad (2)$$

Therefore $C(.,X)$ transforms the N admissible bids of B_X into a set of N different gambles. Denote S_X this set of gambles. A valuation task is then in effect a choice of one element from S_X. If the core theory stipulates independence the optimal choice of a valuation will be the same regardless of $G(w)$: the highest among the admissible bids that are no higher than the certainty equivalent of X.[3] (See Karni and Safra 1987, or Braga 2003)

Cox and Grether (1996) elicited valuations in an English clock auction. Here subjects can arrive at their valuations through a series of choices between the lottery and a decreasing price, therefore this elicitation method raises no questions under the random preference model.

Some experiments elicited valuations with a so-called ordinal payoff scheme (Tversky et al 1990, Cubitt et al forthcoming). With this scheme subjects face some probability (usually 50% divided by the number of pairs of lotteries) of playing, of any two paired bets, the bet they valued most highly. Thus subjects would like to value their preferred bet in a pair more highly than the paired bet. If preferences obey independence and transitivity, this can easily be achieved by valuing each bet at its certainty equivalent. Otherwise coordinating two valuation tasks may be very difficult (see Braga 2003), and it is not clear what a subject's best strategy should be. But it seems clear that the random preference model applies to valuations elicited under the ordinal payoff scheme as much as its deterministic core theory does.

Loomes and Sugden (1995) also considered non-transitive preference relations when all choices are binary. In choices among three or more elements, a non-transitive preference relation may leave one not knowing which element to choose. As valuations are choices of one among many gambles, a random preference model of choices and valuations appears to exclude intransitive preference relations from its possible core theories. Indeed it appears that any theory of binary preference relations aiming to explain preference reversal, a phenomenon for which violation of

[3] If the certainty equivalent happens to coincide with an admissible bid the individual will be indifferent between that admissible bid and the one immediately below it. We will ignore the possibility of the subject choosing the latter.

transitivity has been seen as a possible explanation, needs to impose transitivity of preferences so that elicitation of valuations are possible in auctions or with the BDM procedure. However when the intransitivity arises only from statewise comparisons of the outcomes of the gambles, as is the case of regret theory with dependent gambles, the preference relation will be transitive on S_X, even if it leads to intransitive cycles among other elements of A (see Appendix A for proof).

The random preference model of choice and valuation, therefore, imposes very few restrictions on the preference relation of its core theory. The preference relation need not be a preference ordering on A; it needs only to be complete and reflexive in A, and transitive in S_X, for any X of A.

4.2. Category Probabilities

A random preference model will predict a probability g that P be chosen over $\$$, and a probability h that P be valued above $\$$. We shall call these the *ranking probabilities*. The first thing to note is that the ranking probabilities may be different. We will show this with an example that assumes expected utility as the core theory. If expected utility allows different ranking probabilities, so will their generalisations, which account for most theories of preference under uncertainty.

When expected utility is the core theory, all we need to know to determine an individual's choice and valuations concerning a pair of bets is the pair of certainty equivalents. The individual will chose the bet with the highest certainty equivalent, and will value each bet at the highest of the admissible bids that are no higher than the certainty equivalent. This is the optimal bid only if the individual never has to sell the bet for a price equal to his own bid. Above we ruled out this possibility by assuming that all subjects made different bids. This is an unrealistic assumption. It is however unlikely that anyone's certainty equivalents might have a higher resolution than that allowed by admissible bids. Thus we may simply assume that certainty equivalents are one of the admissible bids. Therefore the optimal bid will be the certainty equivalent. Thus, whenever we assume expected utility as the core theory, we may represent each preference ordering as a pair of certainty equivalents, which is

the same as the pair of optimal bids. This will contain all the information we need to determine choices and valuations. We will denote b_{Pi} and $b_{\$i}$ the optimal bids for the *P* and *$* bets under preference ordering r_i.

Suppose expected utility is the core theory, and that an individual makes all his decisions according to one of two preference orderings, each with 50% probability. Let $b_{P1} = 1$, $b_{\$1} = 2$, $b_{P2} = 3$, and $b_{\$2} = 4$. *$* is preferred to *P* under both preference orderings. Thus *P* will never be chosen over *$* ($g = 0$). The valuations of the two bets will be determined by two independent draws of preference orderings. Thus the individual will value *P* above *$* with 25% probability ($h = 0.25$), when *P* is valued under r_2, and *$*, under r_1. Further down we will derive the limits placed by expected utility on the difference between the ranking probabilities.

For the time being assume that the ranking probabilities are the same for all subjects. These probabilities determine the category probability distribution, which is given in table 4. As in the error model the categories probabilities multiplied by the number of decision triples yield the categories expected frequencies. These can then be compared with the observed frequencies to test the model.

Table 4: Category probabilities under the random preference model, equal individuals

Choices	Highest Value		total
	P	*$*	
P	gh	$g(1-h)$	g
$	$(1-g)h$	$(1-g)(1-h)$	$1-g$
total	h	$1-h$	1

Key: g and h are the probabilities that *P* will be ranked above *$* in choices and valuations respectively.

Note that in the Lichtenstein and Slovic's (1971) error model if p, the probability that a subject truly prefer the *P* bet, is nil, any choice of *P* or any valuation of *P* above *$* will be an error, and any error will result in the choice of *P* or in the valuation of *P* above *$*. Thus substituting in table 2 zero for p, g for r, and h for s yields the contents of table 4. For a similar reason, the contents of table 4 will also be obtained by

substituting, in table 2, 1 for p, $1 - g$ for r, and $1 - h$ for s. Although the theoretical foundations of the two models are different, the random preference model of choices and valuations with equal individuals is, from an observational point of view, a special case of the Lichtenstein and Slovic's (1971) error model. The theoretical foundation makes a difference, though. As seen above, it is not possible to derive any relationship between the two error probabilities, whereas it is possible to derive constraints on the difference between the ranking probabilities, at least under the assumption of expected utility.

The claim that preference reversal is a systematic deviation from preference theory rested to a large extent on the asymmetry of rates of reversal conditional on choice. That is, standard reversals as a proportion of P choices, $SR/(PP + SR)$, usually between 0.5 and 0.8, and non-standard reversals as a proportion of $\$$ choices, $NSR/(NSR + \$\$)$, usually below 0.2. According to our model (of equal individuals) this rate of standard reversal is $g(1 - h)/g = 1 - h$, and that of non-standard reversal is $(1 - g)h/(1 - g) = h$. That is, the usual asymmetric pattern of reversal will appear if h is low, regardless of g. An unbiased stochastic element in decisions may give rise to decision patterns that seem at first sight highly non-random. We may not conclude that this model will explain preference reversal. For instance, according to the model the sum of those two rates of reversal is one. In most empirical studies that sum lies between 0.7 and 0.9. What we may conclude is that the rates of reversal conditional on choice are not a meaningful measure of systematic, that is, non-random, deviation from all stochastic preference theories, as it has implicitly been assumed.

We shall now explore the implications of allowing ranking probabilities vary across subjects and pairs of lotteries. The category probabilities are now the expected values of the expressions in table 4. Assume first that the ranking probability in choices is independent of the ranking probability in valuations. Then simply replacing g and h in table 4 by the respective means yields the new predicted category probabilities. This assumption is, however, no more sensible than that of equal ranking probabilities for everyone. Independence between the ranking probabilities means that a subject with a high ranking probability in choices and a low ranking probability in valuations is as likely as a subject with high ranking probabilities both

in choices and valuations. We would naturally expect the latter to be more likely than the former. Assume then that the ranking probabilities are positively correlated. The category probabilities can now be expressed in terms of the mean ranking probabilities and their covariance. Denoting μ_g the mean of g, μ_h the mean of h, and Cov(g,h) the covariance, the category probability distribution is as shown in table 5.

Table 5: Category probabilities under the random preference model, correlated ranking probabilities

Choices	Highest value		total
	P	*$*	
P	$\mu_g\mu_h + \text{Cov}(g, h)$	$\mu_g(1 - \mu_h) - \text{Cov}(g, h)$	μ_g
$	$(1 - \mu_g)\mu_h - \text{Cov}(g, h)$	$(1 - \mu_g)(1 - \mu_h) + \text{Cov}(g, h)$	$1 - \mu_g$
total	μ_h	$1 - \mu_h$	1

Key: μ is the mean; Cov, the covariance.

This is no longer, even from an observational point of view, a special case of the original Lichtenstein and Slovic's (1971) error model. It would be a special case of an augmented error model where correlated error rates vary across subjects.

The model is able to predict exactly the category frequencies of many datasets. The patterns of category frequencies allowed by the model depend on the admissible range of the parameters. We have not derived them yet, but we may anticipate some conclusions. Denote *pp*, *sr*, *nsr*, and *dd* the relative frequencies of the four categories observed in a sample (that is, *PP*, *SR*, *NSR*, and *$$* divided by the total number of triples). Equating the category probabilities predicted by the model to the sample relative frequencies (only three of the four equations are independent), and noting that $dd = 1 - pp - sr - nsr$ yields:

$$\begin{cases} \mu_g\mu_h + \text{Cov}(g,h) = pp \\ \mu_g(1-\mu_h) - \text{Cov}(g,h) = sr \\ \mu_h(1-\mu_g) - \text{Cov}(g,h) = nsr \\ dd = 1 - pp - sr - nsr \end{cases} \Leftrightarrow \begin{cases} \mu_g = pp + sr \\ \mu_h = pp + nsr \\ \text{Cov}(g,h) = pp \times dd - sr \times nsr \\ dd = 1 - pp - sr - nsr \end{cases}$$

We will call these values for μ_g, μ_h, and Cov(g, h) the *exact solution*, as with them the model matches the observed frequencies exactly. For some datasets the exact solution will not be feasible. If *sr* and *nsr* are sufficiently high the exact solution will violate our assumption that the covariance is non-negative. Additionally, in the exact solution, $\mu_g - \mu_h = sr - nsr$. We will see that expected utility theory places limits on $\mu_g - \mu_h$. Therefore if the reversal pattern is highly asymmetric the exact solution will not be feasible. Note that we are measuring the asymmetry with *sr* and *nsr*. These are *not* the rates of reversal conditional on choice, on which evaluation of the significance of preference reversal has to a large extent relied. The rates of reversal conditional on choice are $sr/(pp + sr)$ and $nsr/(nsr + dd)$. The rates of reversal conditional on choice may be very different from each other when *sr* and *nsr* are similar, and vice-versa.

In terms of the parameters of the model the rate of standard reversal conditional on choice is $1 - \mu_h - \text{Cov}(g, h)/\mu_g$, and the rate of non-standard reversal conditional on choice is $\mu_h - \text{Cov}(g, h)/(1 - \mu_g)$. If the covariance is high (we will see below that $\text{Cov}(g, h) \leq \mu_h(1 - \mu_g)$, $\mu_g(1 - \mu_h)$) the usual asymmetric pattern between the rates of standard and non-standard reversal conditional on choice will appear if μ_h is low and μ_g is high (that is, if *sr* is high and *nsr* is low). If the covariance is small, the usual asymmetric pattern will appear if μ_h is low, regardless of μ_g. The influence of the covariance can be appreciated with the following example. Suppose $\mu_g = \mu_h$. If the covariance is zero, all individual ranking probabilities coincide with the means, or are independent, and the model reduces to the model of equal individuals. With $\mu_g = \mu_h$ the covariance is at its highest when for some individuals $g = h = 0$, and for others, $g = h = 1$. Then no reversals ever occur.

The parameters of the model are constrained by basic statistical principles and by the core theory. Thus any estimates for μ_g, μ_h, and Cov(g, h) must satisfy two conditions. Firstly, those estimates must actually be the means and covariance of some joint distribution of g and h, with $0 \leq g, h \leq 1$. We will call such joint distributions *suitable distributions of (g, h)* (for some given set of μ_g, μ_h and Cov(g, h)). Secondly, at least one suitable distribution of (g, h) must be such that every pair (g, h) with positive probability density must actually be the ranking probabilities implied by some probability measure over some set of preference relations allowed by the

core theory. We will call such probability measures *suitable preference distributions* (for some given pair (g, h)).

One may want to impose a third condition: that for every pair (g, h) with positive probability density at least one suitable preference distribution be "reasonable". Consider $\mu_g = 2/3$, $\mu_h = 1/3$, and Cov(g, h) = 0. A suitable distribution for these parameter values is Pr(g = 2/3, h = 1/3). Suppose the core theory is expected utility. Thus a pair of optimal bids specifies a preference orderings for our purposes. The preference distribution shown on table 6 is suitable for g = 2/3 and h = 1/3, but one might question its reasonability. Not many people would value a P bet, which is always very safe, at £1, £3, or £5 with equal probability.

Table 6: An unreasonable preference distribution?

Preference ordering	b_P	b_S	Probability
r_1	£3	£2	1/3
r_2	£5	£4	1/3
r_3	£1	£6	1/3

The assumption of positive covariance is in effect a restriction on the parameters on grounds of reasonability. We will now address each of these conditions. Naturally the first two conditions are derived from basic statistical principles and the axioms of the core theory, whereas the third is a matter of individual judgement.

4.3. Suitable Distributions of (g, h)

A suitable distribution of (g, h) will exist if and only if the following restrictions are imposed on μ_g, μ_h, and Cov(g,h): $0 \le \mu_g, \mu_h \le 1$, $\text{Cov}(g,h) \le \mu_g(1 - \mu_h)$, $\text{Cov}(g,h) \le (1 - \mu_g)\mu_h$, $\text{Cov}(g,h) \ge -\mu_g\mu_h$, and $\text{Cov}(g,h) \ge -(1 - \mu_g)(1 - \mu_h)$. The first restriction is obvious. The other four are necessary to assure that the four category probabilities are non-negative. The last two restrictions are made redundant by the more stringent $\text{Cov}(g,h) \ge 0$. Thus we will we need only to prove that $0 \le \mu_g, \mu_h \le 1$, $\text{Cov}(g,h) \le \mu_g(1 - \mu_h)$, $\text{Cov}(g,h) \le (1 - \mu_g)\mu_h$, and $\text{Cov}(g,h) \ge 0$ are sufficient conditions for the existence of a

suitable distribution. Consider first the case where μ_g, μ_h, or both are either zero or one, say μ_g is either zero or one. Then as Cov(g, h) ≤ $\mu_g(1 - \mu_h)$ and Cov(g,h) ≤ $(1 - \mu_g)\mu_h$, Cov(g,h) = 0, which means that Pr(g=μ_g, h=μ_h) = 1 is a suitable distribution. This would be a case of invariant ranking probabilities across subjects that we saw first. It would be the only suitable distribution if $\mu_h = 0$ or $\mu_h = 1$, otherwise there would be infinite suitable distributions. For the case where $0 < \mu_g, \mu_h < 1$, suppose that $\mu_g \le \mu_h$ (the proof is similar for $\mu_g \ge \mu_h$). Then $\mu_g(1 - \mu_h) \le (1 - \mu_g)\mu_h$, and Cov($g$,$h$) ≤ $\mu_g(1 - \mu_h)$ is the active restriction. Define a = Cov(g,h)/[$\mu_g(1 - \mu_h)$]. Then for all μ_g, μ_h, and Cov(g,h) such that $0 < \mu_g \le \mu_h < 1$, Cov(g,h) > 0, and Cov(g,h) ≤ $\mu_g(1 - \mu_h)$, the distribution of (g, h) shown in the table below is suitable. It may be easily checked that all conditions are met: note that $0 < a \le 1$, therefore $0 \le g, h \le 1$; the means of g and h are μ_g, and μ_h; the covariance is Cov(g,h); and all probabilities are non-negative and add up to unity (under the assumption that $\mu_g \le \mu_h$)

Table 7: A suitable distribution of (g, h)

g	h	
	$\mu_h - \mu_h\sqrt{a}$	$\mu_h + (1 - \mu_h)\sqrt{a}$
$\mu_g - \mu_g\sqrt{a}$	$1 - \mu_h$	$\mu_h - \mu_g$
$\mu_g + (1 - \mu_g)\sqrt{a}$	0	μ_g

Note: a = Cov(g,h)/[$\mu_g(1 - \mu_h)$], $0 < \mu_g \le \mu_h < 1$.

If the covariance is at the maximum value allowed by the restrictions, Cov(g,h) = $\mu_g(1 - \mu_h)$, in which case $a = 1$, and if $\mu_g = \mu_h = 0.5$, then g and h take only the values one and zero, and the distribution above is the only suitable one; otherwise there are infinite suitable distributions.

4.4. Suitable Preference Distributions

Each individual has a preference distribution. Then the entire population has a set of preference distributions. Not every suitable distribution of (g, h) will have a set of suitable preference distributions. For instance, if in a random sample all triples are

standard reversals, the maximum likelihood estimates are $\mu_g = 1$, $\mu_h = 0$, and Cov(g,h) = 0. Then Pr(g=1, h=0) = 1 is a suitable distribution, the only one, for those estimates, but no preference distribution suits (g=1, h=0). If g=1 the P bet is ranked above the $\$$ bet in all preference orderings, therefore there must be some positive probability that P be valued above $\$$.

We will now restrict our core theory to expected utility, and derive the limits this theory imposes on the difference between g and h.

Consider a pair of P and $\$$ bets. As we noted above, under the assumption of expected utility all we need to know of a preference ordering are the certainty equivalents of the paired bets, which we assume to coincide with some admissible bids. Thus we will use the terms *certainty equivalent* and *optimal bid* interchangeably. We denote (b_{Pi}, $b_{\$i}$) the pair of optimal bids in preference ordering r_i, and denote (b_P, $b_\$$) two optimal bids determined by two independently drawn preference orderings.

In a typical preference reversal experiment the set of all possible pairs (b_{Pi}, $b_{\$i}$) is finite. For instance in the experiment reported in Braga and Starmer (2001) some subjects dealt with the following bets: P = (£8, 97%), $\$$ = (£32, 31%). Subjects' bids when expressed in pence had to be integers, no less than zero, and no higher than the positive outcome of the bet. Therefore there were 801 admissible bids for P, and 3201 for $\$$. This makes for 801 × 3201 = 2 564 001 pairs of admissible bids. Let the admissible bids for the P bet be $b_P \in B_P = \{0, 1,\ldots, m_P\}$, and those for the $\$$ bet be $b_\$ \in B_\$ = \{0, 1,\ldots,m_\$\}$, with $m_P \leq m_\$$ as in the typical experiment.

The unit b_P and $b_\$$ are expressed in is the lowest admissible increment. Typically the lowest increment is the lowest subdivision of the relevant currency, say, one penny or one cent; but if bids must be multiples of ten pence, for instance, the elements of $B_\$$ and B_p denote multiple of ten pence.[4] Of course if the lowest possible increment of admissible bids is large the assumption that certainty equivalents always coincide with some admissible bids will be unrealistic.

Denote a pair of admissible bids (i, j), with (i, j) $\in B_P \times B_\$$. Two preference orderings may specify the same certainty equivalents for the P and $\$$ bets, but differ

[4] In Harrison (1994) the minimum increment was higher than one dollar cent, as much as 2 US dollars in one treatment.

with respect to other gambles. Thus the set of relevant preference orderings R^*, that is, $f(R^*) = 1$, may have more elements than $B_P \times B_\$$, but each preference ordering of R^* specifies optimal bids for P and $\$$ that belongs to $B_P \times B_\$$. The probability distribution f over R^* then implies a probability distribution p over $B_P \times B_\$$. We shall denote p_{ij} the probability that (i, j) be drawn from $B_P \times B_\$$. Thus $p_{ij} = f(R_{ij})$, where R_{ij} is the set of all preference orderings that specify (i, j) as the optimal bids.

We will assume, for the sake of simplicity, that the subject is never indifferent between P and $\$$ in any preference ordering. That is, $p_{ii} = 0$ for all i. The probability distribution p determines g and h. We defined h as the probability that P be valued above $\$$. To establish the bounds on h given any g, we will minimise $h = \Pr(b_P > b_\$)$, but maximise $\Pr(b_P \geq b_\$)$, which is the same as to minimise $\Pr(b_P < b_\$)$. We will show that the difference predicted by the model between the maxima of $\Pr(b_P > b_\$)$ and $\Pr(b_P \geq b_\$)$ is negligible, less than 0.002, under reasonable assumptions. The reason why we follow this procedure is that the symmetry between the minimisation of $\Pr(b_P > b_\$)$ and maximisation of $\Pr(b_P \geq b_\$)$ (minimisation of $\Pr(b_p < b_\$)$) simplifies the analysis.

The probability g that the P bet be chosen over the $\$$ bet is the sum of the probabilities of all the pairs (i, j) where $j < i$ (if we had not ruled out indifference we would have to add half the sum of the probabilities of all pairs (i, i)):

$$g = \sum_{i=1}^{m_P} \sum_{j=0}^{i-1} p_{ij} \ .$$

To compute the probability h is useful to begin with the marginal probability distributions. The probability that the optimal bid for the P bet be i is

$$\Pr(b_P = i) = \sum_{j=0}^{m_\$} p_{ij} \ ,$$

and the probability that the $\$$ bet be valued at l is

$$\Pr(b_\$ = l) = \sum_{k=0}^{m_P} p_{kl} \ .$$

The probability that $ be valued below *i* is then

$$\Pr(b_{\$} < i) = \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p_{kl} \, .$$

The optimal bids for *P* and $ are determined by two pairs of admissible bids independently drawn. Therefore the probability that a subject bid *i* for the *P* bet and less than *i* for the $ bet is

$$\Pr(b_P = i \text{ and } b_{\$} < i) = \Pr(b_P = i) \times \Pr(b_{\$} < i) = \sum_{j=0}^{m_{\$}} p_{ij} \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p_{kl} \, .$$

Summing across all admissible bids for *P* yields the probability that *P* be valued above $:

$$h = \sum_{i=1}^{m_P} \sum_{j=0}^{m_{\$}} p_{ij} \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p_{kl} \, . \tag{3}$$

The probability distribution over $B_P \times B_{\$}$ that minimises *h* given *g* is a solution to the following problem:

$$\min \; h = \sum_{i=1}^{m_P} \sum_{j=0}^{m_{\$}} p_{ij} \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p_{kl}$$

subject to

$$\sum_{i=1}^{m_P} \sum_{j=0}^{i-1} p_{ij} = g \, ,$$

$$\sum_{i=0}^{m_P} \sum_{j=0}^{m_{\$}} p_{ij} = 1,$$

$p_{ij} \geq 0, \quad i = 0,\ldots, m_P$ and $j = 0,\ldots, m_{\$}$.

$p_{ii} = 0, \quad i = 0,\ldots, m_P$

The first constraint states that the probability that *P* be chosen over $ be *g*; the second, that the probabilities of all pairs of admissible bids add to up to 1, and the

third is obvious. The constraint that each probability must not exceed 1 is guaranteed by the second and third constraints. The last constraint could simply be substituted in the objective function and the second and third constraints (it is already in the first), but this would make the expressions even more cumbersome.

The global minimum is obtained with any probability distribution that obeys conditions (4) to (6) below. We leave the proof to Appendix B, and will give here some interpretation of those conditions. Note that although there are many probability distributions that obey those conditions, the arguments of function *h* are the same in all of them. The variability lies in the *$* bids of the pairs defined by (4). As they are always no less than the highest admissible *P* bid, their precise values do not influence *h*. Therefore we will refer to *the probability distribution* (and not *distributions*) defined by conditions (4) to (6), to be called the *p** probability distribution, even though many different probability distributions obey conditions (4) to (6).

$$\sum_{j=m_P}^{m_\$} p^*_{0n} = 1 - g\,, \quad \text{for } n = m_P, \ldots, m_\$, \tag{4}$$

$$p^*_{m,m-1} = \frac{g}{m_P} \quad \text{for } m = 1, \ldots, m_P, \tag{5}$$

$$p^*_{mn} = 0 \quad \text{for all other } (m, n). \tag{6}$$

Table 8 shows, somewhat graphically, an example of a *p** probability distribution. Suppose the pair of bets is (£5, 81%) and (£18, 19%), and that the admissible bids and the certainty equivalents are integer numbers of pounds, no less than zero, and no more than the positive outcome of the bet. Then $m_P = 5$, and $m_\$ = 18$. These are unrealistic assumptions, particularly that certainty equivalents are whole number of pounds, but they are useful for illustration purposes. In an actual experiment it would be more likely that $m_P = 500$ and $m_\$ = 1800$, as in Braga and Starmer's (2001) experiment, where this pair of bets was actually used. The table shows each pair (b_{Pi}, $b_{\$i}$) with positive probability in a column. The position of the characters *P* and *$* relative to the values on the left-most column indicate the value of b_{Pi} and $b_{\$i}$ respectively. For simplicity only one of the pairs defined by (4) has positive probability, namely the pair (£0, £5), in column 6.

Table 8: A p^* probability distribution, $m_P = 5$

	Pairs (b_{Pi}, $b_{\$i}$) and their probabilities					
Value	1	2	3	4	5	6
	$g/5$	$g/5$	$g/5$	$g/5$	$g/5$	$1-g$
£0	$					P
£1	P	$				
£2		P	$			
£3			P	$		
£4				P	$	
£5					P	$

Each column represents a pair (b_{Pi}, $b_{\$i}$). The positions of P and $ indicate the values of b_{Pi} and $b_{\$i}$.

In preference orderings that rank $ above P, b_{Pi} is zero and $b_{\$i}$ is no less than the maximum possible optimal bid for P. Thus when the valuation of either P or $ is determined by one such preference ordering, and regardless of the preference ordering determining the other valuation, P will never be valued above $. Or, putting it the other way, P will be valued above $ only if both valuations are obtained from preference orderings that rank P above $. In these preference orderings the difference between the optimal bids is the smallest possible. Having the probability of pairs that rank P above $, g, allocated evenly to pairs running all the way from (1, 0) to (5, 4) leads to a smaller h than having it concentrated on, say, a single pair. If b_P is obtained from the first pair (column 1), P is valued above $ only if $b_\$$ is also obtained from that pair. This happens with probability (g/5) × (g/5). If b_P is obtained from the second pair, P is valued above $ only if $b_\$$ is obtained from the first or second pairs. This happens with probability (g/5) × (2g/5). Thus, in this example, given g, the value of h is

$$h = \frac{g}{5}\left(\frac{g}{5} + 2\frac{g}{5} + 3\frac{g}{5} + 4\frac{g}{5} + 5\frac{g}{5}\right) = \frac{3}{5}g^2.$$

Thus h may be much smaller than g. Say, for g = 0.5, and m_P = 5, h can be as small as 0.15. In this example there is a non-negligible probability that the two bets be valued equally. Again for g = 0.5, Pr($b_P = b_\$$) = 0.14. With the more realistic m_P = 500, that probability would be 0.001499. To compute h under the p^* distribution for a general m_P, one may substitute p^* in the objective function (3). Only pairs defined by

(5) matter, pairs where $m \geq 1$ or $n \leq m_P - 1$. Then in expression (3) for each k there is only one $p_{kl} > 0$, namely, $p_{k,k-1}$, and only if $1 \leq k \leq i$. Also for each i there is only one $p_{ij} > 0$, namely $p_{i,i-1}$. Thus expression (3) simplifies to

$$h = \sum_{i=1}^{m_P} p_{i,i-1} \sum_{k=1}^{i} p_{k,k-1} \,. \tag{7}$$

Then according to (5)

$$\min h = h^* = \sum_{i=1}^{m_P} \frac{g}{m_P} \sum_{k=1}^{i} \frac{g}{m_P} = \frac{g^2}{m_P^2} \sum_{i=1}^{m_P} i = \frac{g^2}{2} + \frac{g^2}{2m_P} \,. \tag{8}$$

The solution to the maximisation of $\Pr(b_P \geq b_\$)$ given g may be obtained from the solution to the minimisation of $\Pr(b_P > b_\$)$ given g. Appendix B shows the details. The probability distribution that maximises $\Pr(b_P \geq b_\$)$ given g, to be called $p^\#$, is shown in (9) to (11).

$$p^\#_{m_P,0} = g, \tag{9}$$

$$p^\#_{m-1,m} = \frac{1-g}{m_P} \quad \text{for } m = 1, \ldots, m_P, \tag{10}$$

$$p^\#_{mn} = 0 \quad \text{for all other } (m, n). \tag{11}$$

Note that this distribution may be obtained by swapping the indices in p^* (taking into account in (4) and (9) that b_P cannot exceed m_P) and substituting $1 - g$ for g. In the same manner, a matrix representation of $p^\#$ may be obtained from table 8 by swapping the P's and \$'s, and substituting $1 - g$ for g. The value of $\Pr(b_P \geq b_\$)$ under this distribution may be obtained by substituting in (8) $1 - g$ for g (again see Appendix B):

$$\max \Pr(b_P \geq b_\$) = h^\# = g + \frac{1-g^2}{2} - \frac{1-g}{2m_P} \,. \tag{12}$$

Under $p^\#$, as under p^*, $\Pr(b_P = b_\$) = g/m_P - (g/m_P)^2$. This decreases with m_P, but even for a P bet offering a low amount to win, say, £5, if any number of pence up to £5 is an admissible bid, this probability is at most 0.002 (when $g = 1$). This means that,

for m_P = 500, the maximum of Pr($b_P \geq b_\$$) exceeds the maximum of Pr($b_P > b_\$$) by at most 0.002 (by less if Pr($b_P > b_\$$) is maximised by a probability distribution other than $p^\#$), and obviously Pr($b_P > b_\$$) cannot exceed Pr($b_P \geq b_\$$). Therefore the difference between the two maxima is negligible.

As Pr($b_P > b_\$$) ≤ Pr($b_p \geq b_\$$), h = Pr($b_P > b_\$$) ≤ $h^\#$. Therefore we may write, from (8) and (12),

$$\frac{g^2}{2} + \frac{g^2}{2m_P} \leq h \leq g + \frac{1-g^2}{2} - \frac{1-g}{2m_P}. \tag{13}$$

We have so far assumed that admissible bids are whole numbers of pence. We may now relax this assumption. If the only restrictions on admissible bids are a lower and an upper limit m_P is infinite. As m_P approaches infinity the lower and upper bounds of h approach

$$\frac{g^2}{2} < h < g + \frac{1-g^2}{2}. \tag{14}$$

Note that even for a finite m_p, h will be within these bounds, and for the values of m_P commonly found in experiments h will be very close to the bounds.[5] Therefore, from now on, unless stated otherwise, we will consider only the bounds of h for an infinite m_P. Figure 1 shows all the combinations of g and h that are suited by some preference distribution. For instance, if g = 0, h can vary between 0 and 0.5; and if g = 0.5, h can vary between 0.125 and 0.875. These differences between the ranking probabilities allowed by random expected utility are probably larger than most people would have expected.

[5] Harrison (1994) is an exception. In one treatment admissible bids were multiples of 2 US dollar. The BDM procedure was used with a maximum counter price of $10. Thus m_P was 5 in that treatment.

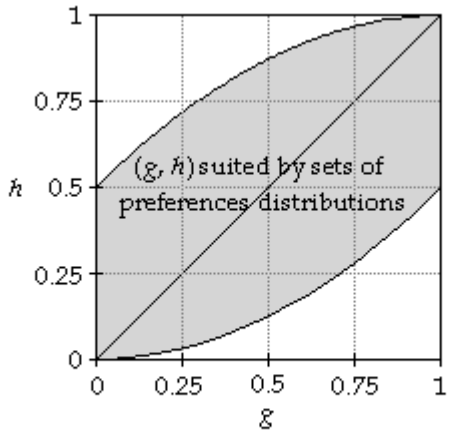

Figure 1: (*g*, *h*) suited by expected utility distributions

These bounds apply to each person, that is, to each pair (g, h) with positive probability density. This implies corresponding bounds on the population means, μ_g and μ_h. If for all pairs (g, h) with positive probability density, $h > g^2/2$ then

$$\mu_h = \mathrm{E}(h) > \mathrm{E}\left(\frac{g^2}{2}\right) = \frac{\mathrm{E}\left[\left(\mu_g + (g - \mu_g)\right)^2\right]}{2} =$$

$$= \frac{\mu_g^2 + 2\mu_g \mathrm{E}(g - \mu_g) + \mathrm{E}\left((g - \mu_g)^2\right)}{2} = \frac{\mu_g^2 + \sigma_g^2}{2}, \qquad (15)$$

where σ_g^2 is the variance of g. It can be shown in a similar manner that if for all pairs (g, h) with positive probability density, $h < (1 - g^2)/2 + g$, then

$$\mu_h < \mu_g + \frac{1 - \mu_g^2 - \sigma_g^2}{2}. \qquad (16)$$

That μ_h lies in the interval defined by (15) and (16) is a necessary and sufficient condition for a set of preference distributions to exist that suits μ_h, μ_g, and σ_g^2. Our model does not explicitly involve σ_g^2, but places restrictions on it. We can derive lower and upper bounds for σ_g^2. By substituting the lower bounds in (15) and (16) we will obtain a wider interval, and therefore necessary but not sufficient conditions for the existence of a set of suitable sets of preference orderings; by substituting the upper bounds will obtain a narrower interval, and therefore sufficient but not necessary conditions.

As g is limited between zero and one, its variance is also limited: $\sigma_g^2 = E(g^2) - \mu_g^2$. As $0 \le g \le 1$, $E(g^2) \le E(g)$, and $\sigma_g^2 \le \mu_g - \mu_g^2$. Thus a sufficient condition for the existence of set of suitable preference distributions is

$$\frac{\mu_g}{2} < \mu_h < \frac{1+\mu_g}{2}. \tag{17}$$

The variances of g and h must be large enough to accommodate the covariance estimated by the model: $\mathrm{Cov}(g,h) \le \sigma_g\sigma_h$, or $\sigma_g \ge \mathrm{Cov}(g,h)/\sigma_h$. As with g it must be $\sigma_h^2 \le \mu_h - \mu_h^2$. Then

$$\sigma_g \ge \frac{\mathrm{Cov}(g,h)}{\sigma_h} \ge \frac{\mathrm{Cov}(g,h)}{\sqrt{\mu_h - \mu_h^2}}.$$

Therefore a necessary condition for the existence of a set of suitable preference distribution is

$$\frac{1}{2}\left(\mu_g^2 + \frac{\mathrm{Cov}^2(g,h)}{\mu_h - \mu_h^2}\right) < \mu_h < \mu_g + \frac{1}{2}\left(1 - \mu_g^2 - \frac{\mathrm{Cov}^2(g,h)}{\mu_h - \mu_h^2}\right). \tag{18}$$

Conditions (17) and (18) are expressed in terms of the parameters of the model, and can therefore be used to check the existence of a set of suitable preference distributions. Condition (18) may be expressed as restrictions on the covariance: The lower and upper bounds on h are respectively equivalent to

$$\mathrm{Cov}(g,h) < \sqrt{(\mu_h - \mu_h^2)(2\mu_h - \mu_g^2)} \quad \text{and} \tag{19a}$$

$$\mathrm{Cov}(g,h) < \sqrt{(\mu_h - \mu_h^2)(1 - \mu_g^2 + 2\mu_g - 2\mu_h)}. \tag{19b}$$

Figure 2 shows for two values of g all the restrictions on the means and the covariance that we have derived so far.

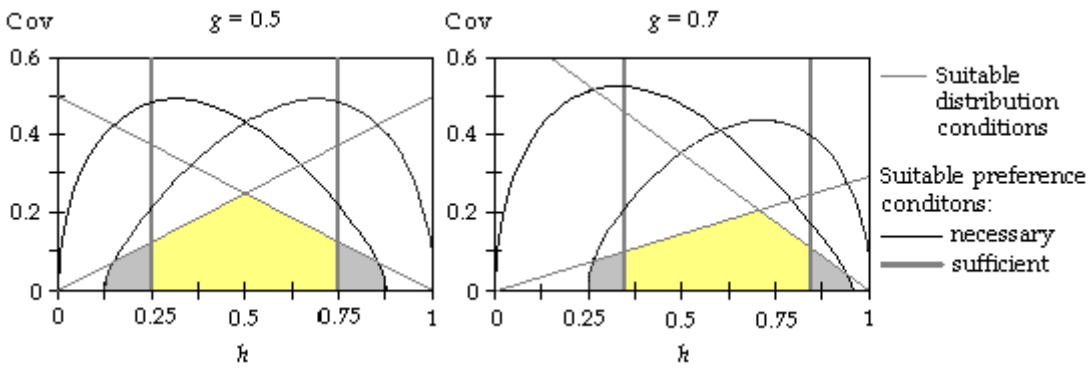

Figure 2: Restrictions on the means and the covariance

That the covariance lies underneath both slanting straight lines is a necessary and sufficient condition for the existence of a suitable joint distribution of (g, h); that it lies underneath the arches is a necessary condition for the existence of suitable preference distributions; and that h lies between the vertical lines is a sufficient condition for a set of preference distributions to exist that suits the two means, regardless of the covariance. If μ_g, μ_h, and Cov are such that (μ_h, Cov) lies inside the lighter shaded area then a set of preference distributions exists that suits these values; if (μ_h, Cov) lies outside the union of the three shaded areas then no such set suits the values.

These conditions are easy to check, and may in many cases be enough to accept or reject a set of estimates for μ_g, μ_h, and Cov(g, h) (on grounds of the existence or non-existence of suitable distributions of (g, h) and suitable preference distributions only). However of the conditions for the existence of suitable sets of preference distributions, the sufficient conditions, the vertical lines, are too demanding, and the necessary conditions, the arches, are too lax, thus leaving a grey area (dark grey in the figure above) where these conditions will not be enough to either accept or reject a set of estimates.

When this happens the following sufficiency conditions may be used. These are less demanding than conditions (17), but also more laborious to check. A set of suitable sets of preference distributions will exist for the parameters μ_g, μ_h, and Cov(g, h) if values for σ_g, σ_h, and p can be found, such that

$$\frac{\sigma_h^2}{(1-\mu_h)^2+\sigma_h^2} \le p \le \frac{\mu_h^2}{\mu_h^2+\sigma_h^2}, \tag{20a}$$

$$\frac{\sigma_g^2}{(1-\mu_g)^2+\sigma_g^2} \le p \le \frac{\mu_g^2}{\mu_g^2+\sigma_g^2}, \tag{20b}$$

$$\sigma_g\sigma_h = \mathrm{Cov}(g, h), \tag{20c}$$

$$\mu_h - \sigma_h\sqrt{\frac{p}{1-p}} > \frac{1}{2}\left(\mu_g - \sigma_g\sqrt{\frac{p}{1-p}}\right)^2, \tag{20d}$$

$$\mu_h + \sigma_h\sqrt{\frac{1-p}{p}} > \frac{1}{2}\left(\mu_g + \sigma_g\sqrt{\frac{1-p}{p}}\right)^2, \tag{20e}$$

$$\mu_h - \sigma_h\sqrt{\frac{1-p}{p}} < \mu_g - \sigma_g\sqrt{\frac{1-p}{p}} + \frac{1}{2} - \frac{1}{2}\left(\mu_g - \sigma_g\sqrt{\frac{1-p}{p}}\right)^2, \tag{20f}$$

$$\mu_h + \sigma_h\sqrt{\frac{1-p}{p}} < \mu_g + \sigma_g\sqrt{\frac{1-p}{p}} + \frac{1}{2} - \frac{1}{2}\left(\mu_g + \sigma_g\sqrt{\frac{1-p}{p}}\right)^2. \tag{20g}$$

Conditions (20a), (20b), and (20c) simply guaranty that the following is a suitable joint distribution of (g, h):

Table 9: Joint distribution of (g, h)

	h	
g	$\mu_h - \sigma_h\sqrt{\frac{p}{1-p}}$	$\mu_h + \sigma_h\sqrt{\frac{1-p}{p}}$
$\mu_g - \sigma_g\sqrt{\frac{p}{1-p}}$	$1-p$	0
$\mu_g + \sigma_g\sqrt{\frac{1-p}{p}}$	0	p

It may be checked that the means and standard deviations of g and h are actually μ_g, μ_h, σ_g, and σ_h, and that $\mathrm{Cov}(g, h) = \sigma_g\sigma_h$. Conditions (20a) and (20b) simply limit g and h between zero and one. Condition (20c) guarantees that the estimated $\mathrm{Cov}(g, h)$ is feasible. The remaining conditions guarantee that all pairs (g, h) with positive probability satisfy conditions (14), the necessary and sufficient conditions for the existence a suitable preference distribution. Conditions (20) then amount to try to

find an actual suitable distribution of (g, h) where all pairs (g, h) with positive probability satisfy the conditions for the existence of a suitable preference distribution. The rationale for searching within that type of distribution is as follows.

The first thing to notice is that the above distribution is simply a joint distribution of two perfectly correlated, binary variables that happens to be fully specified in terms of the means and standard deviations of the variables, and of the probabilities.

Specification in terms of the means and standard deviation facilitates the search of the distribution as the model imposes restrictions on these parameters. Binary variables make the distribution simple. The simplest of all distributions is of course a single value with probability one. That is not possible if the covariance is positive. As $\sigma_g\sigma_h \geq \text{Cov}(g, h)$, if the covariance is positive both variances must be positive. Then a distribution with only two values with positive probability is the simplest possible.

Perfect correlation increases the chances that all pairs (g, h) satisfy conditions (14), the condition for the existence of a set of suitable preference distributions. Perfect correlation, $\sigma_g\sigma_h = \text{Cov}(g, h)$, allows both variances to be as low as possible in the sense that, given σ_g, the lowest possible σ_h is such that $\sigma_g\sigma_h = \text{Cov}(g, h)$. We know from condition (15) and (16) that a low σ_g^2 facilitates the existence of suitable preference distributions. So does a low σ_h^2. Suppose that the estimated μ_h is just above the estimated $\mu_g^2/2$, and thus the lower bound of h might be at risk of being violated. The reasoning is similar if the upper bound is the problem. Compliance with the lower bound is facilitated by a high σ_h^2 for values of h above the mean, but by a low σ_h^2 for values of h below the mean. As g spreads away from the mean the lower bound does not change much: $\partial(g^2/2)/\partial g = g < 1$ for $g < 1$; this is also evident in figure 1. Thus compliance with the lower bound is easier for values of h above the mean in spite of a low σ_h^2 than for values of h below the mean if σ_h^2 is high.

The use of this approach is illustrated with the following example. Suppose fitting the model to actual data gives rise to the following estimates: $\mu_g = 0.6$, $\mu_h = 0.24$, and $\text{Cov}(g, h) = 0.04$. There is a suitable distribution of (g, h) for these estimates: $\text{Cov}(g, h) < \mu_g(1 - \mu_h)$, $\mu_h(1 - \mu_g)$. The necessary conditions (19) for the existence of suitable sets of preference orderings require that $\text{Cov}(g, h) < 0.15$ and $\text{Cov}(g, h) < 0.5$,

and are satisfied. The sufficient conditions (17) require than $0.3 < \mu_h < 0.8$, and are violated. These estimates would then lie in left-hand side dark grey area of figure 2 (if we had drawn one for $\mu_g = 0.6$). Thus these conditions would not tell us whether any set preference distributions suits those estimates, but conditions (20) do. They are satisfied with, for instance, $\sigma_g = 0.25$, $\sigma_h = 0.16$, and $p = 0.5$. These give rise to the distribution below.

Table 10: Joint distribution of (*g*, *h*)

g	*h*	
	0.08	0.4
0.35	0.5	0
0.85	0	0.5

Note: $0.35^2/2 = 0.061$, $0.85^2/2 = 0.36$

We have derived easy to check but too strict and not so strict but harder to check sufficiency conditions for the existence of suitable preference distributions. Given the bizarre shape a preference distribution needs to have to produce big differences between the two ranking probabilities (see table 8), one might not be willing to accept a model that fails the strict sufficiency conditions. This brings us back to the issue of reasonable preference distributions.

4.5. Reasonable Preference Distributions

We have no intention of defining, or even of suggesting, *the* conditions a preference distribution should meet to qualify as reasonable. Instead we will take a particular type of preference distribution that we expect will be accepted as reasonable, and see by how much the two ranking probabilities diverge.

We chose preferences described by expected utility, with the utility of a monetary outcome, y, given by $u_a(y) = y^a$, $y \geq 0$, $a > 0$. The power function is a familiar utility function when its argument is the level of wealth, but whether utility given by the power of the change in wealth, as is the case here, is a reasonable assumption may depend on the point of view.

A power of the change in wealth implies constant relative risk aversion when aversion is measured in terms of changes in wealth, but dramatically variable relative and absolute risk aversion for small changes from initial wealth when aversion is measured in terms of levels of wealth. One might thus object to a power of the change in wealth on normative grounds.

On the other hand behaviour in individual choice experiments is more easily explained by utility given by a power of the change in wealth than by a power of the level of wealth. Sugden (forthcoming) notes that in some recent experiments where subjects make many choices between lotteries, behaviour converges towards a pattern that is compatible with stochastic expected utility, with the utility given by a power of the change in wealth. In addition if one assumes that utility is given by a power of the level of wealth it will be hard to explain the choice of the *P* bet in most pairs used in preference reversal experiments. Consider the pair of bets *P* = (£5, 94%) and *$* = (£17, 39%). Suppose utility is given by $W^{0.01}$, where *W* is the level of wealth. Even such a risk-averse individual would choose the *$* bet if his wealth was £10 or more. In several experiments fair proportions of subjects have chosen the *P* bet from that pair. Even though virtually all subjects were young students, it is hard to believe that any of them possessed less than ten pounds.

Our aim is to find a reasonable model of random preferences and test it against data. Then it will be inappropriate to use compatibility with the data as a criterion for reasonableness. Note however that the data against which the model will be tested results from crossing choices with valuations, whereas we are basing our choice of the utility function on choice data only. The results of preference reversal experiments have been seen as a challenge to preference theory because of inconsistencies between choices and valuations, not because many subjects choose the *P* bet. Therefore using a utility function that is compatible with the choices observed in those experiments is appropriate.

The power function defines only the core theory of the model. To complete the model one needs a set of power functions, and a probability distribution over that set. Each choice or valuation concerning monetary outcomes will be based on one power function drawn at random. We will assume here that the parameter a of the utility

function follows a truncated normal distribution. A normal distribution is perhaps the most natural of assumptions, and we will see the usefulness of the truncation below.

Let $a \in [a_1, a_2]$, and its truncated normal distribution be centred on the midpoint of $[a_1, a_2]$. Thus the mean of a is $\mu_a = (a_1 + a_2)/2$. The cumulative distribution of a, $F(a)$, will be obtained as follows. The same amount of probability, c, $0 < c < 0.5$, is cut off from both tails of the standard normal distribution. A scalar x is defined so that $\Phi[(a_1 - \mu_a)/x] = c$ (and $\Phi[(a_2 - \mu_a)/x] = 1 - c$) where $\Phi(.)$ is the standard normal cumulative distribution. This makes $x = (a_1 - a_2)/[2\Phi^{-1}(c)]$ or, more intuitively, $x = (a_2 - a_1)/[\Phi^{-1}(1 - c) - \Phi^{-1}(c)]$. Finally, the cumulative distribution of a is defined as

$$F(a) = \frac{\Phi\left(\frac{a - \mu_a}{x}\right) - c}{1 - 2c},$$

and its probability density function is

$$f(a) = \frac{\phi\left(\frac{a - \mu_a}{x}\right)}{1 - 2c},$$

where $\phi(.)$ is the standard normal probability density function.

Loosely speaking, this procedure takes a central portion of the standard normal distribution (between $\Phi^{-1}(1 - c)$ and $\Phi^{-1}(c)$), and compresses or stretches it so as to fit the interval $[a_1, a_2]$. x would then be the standard deviation of a if its distribution were not truncated. The probability that was cut off from the tails, $2c$, is then distributed *pro rata* by the interval $[a_1, a_2]$.

The truncation serves two purposes. It is needed to keep a positive. If a was zero, certainty equivalents would not exist, and if a was negative monotonicity would be violated. Secondly, although we favour a distribution close to the untruncated normal, truncation allows us to look at a continuum of probability distribution shapes: from an essentially untruncated normal, if c is close to zero, to an essentially uniform distribution, if c is close to 0.5.

Suppose an individual with such a preference distribution is to choose between a pair o *P* and *$* bets. Let each bet offer some positive amount of money with some

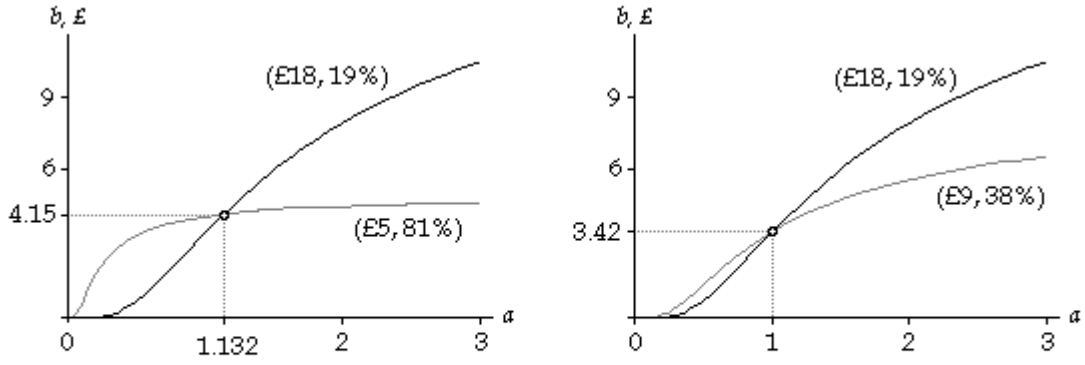

Figure 3: Optimal bids (b) as a function of a

probability and zero with the remaining probability. Denote $P = (y_P, p_P)$ and $\$ = (y_\$, p_\$)$, where y denotes the amount of money to win, and p, its probability. The expected utilities are $p_P y_P^a$ and $p_\$ y_\$^a$, and the certainty equivalents are $p_P^{1/a} y_P$ and $p_\$^{1/a} y_\$$.

Figure 3 shows the certainty equivalents of the lotteries of two pairs as a function of a. Preference reversals were observed with (£5, 81%) and (£18, 19%) in the experiment reported in Braga and Starmer (2001). The other pair is not typical of preference reversal experiments. Its lotteries can be thought of as a pair of P and $\$$ bets, as one offers less money and a higher winning probability than the other. However 38% is a winning probability typical of a $\$$ bet, not of a P bet. We are presenting this pair here because it will be useful in the discussion below.

The certainty equivalent of the safer bet varies less with a than that of the $\$$ bet. For each pair of P and $\$$ bets there is a value of a that leads to equal certainty equivalents for both bets: $p_P^{1/a} y_P = p_\$^{1/a} y_\$ \Leftrightarrow a = (ln\ p_P - ln\ p_\$)/(ln\ y_\$ - ln\ y_P)$, approximately 1.132 in the first pair, exactly 1 in the second. If a is less than that value, the certainty equivalent of P will exceed that of $\$$. If a is higher than that value, the opposite will happen.

In choice tasks both lotteries are evaluated with a single a drawn from $[a_1, a_2]$. Thus P will be chosen over $\$$ if a is less than $(ln\ p_P - ln\ p_\$)/(ln\ y_\$ - ln\ y_P)$. Thus the probability that P be chosen over $\$$ is

$$g = \Pr\left(a < \frac{ln\, p_P - ln\, p_\$}{ln\, y_\$ - ln\, y_P}\right) = F\left(\frac{ln\, p_P - ln\, p_\$}{ln\, y_\$ - ln\, y_P}\right).$$

The valuations are made with two independently drawn powers. Let P be valued with a, and \$, with a'. Then

$$h = \Pr(p_P^{1/a} y_P > p_\$^{1/a'} y_\$) = \Pr\left(a > \frac{a' \ln p_P}{\ln p_\$ + a' \ln(y_\$ / y_P)}\right) = \Pr[a > a(a')],$$

where $a(a') = \dfrac{a' \ln p_P}{\ln p_\$ + a' \ln(y_\$ / y_P)}$.

Defining $lb(a') = \min\{a_2, \max[a_1, a(a')]\}$, then

$$h = \int_{a_1}^{a_2} \int_{lb(a')}^{a_2} f(a) f(a') da da'.$$

As a and a' are independent, their joint probability density is simply the product of the individual probability densities. If a_1 and a' are low enough P will be valued above \$ for any $a \in [a_1, a_2]$. In these cases $a(a') < a_1$. Conversely, if a_2 and a' are high enough P will be valued below \$ for any $a \in [a_1, a_2]$. In these cases $a(a') > a_2$. The function $lb(.)$ guarantees the relevant lower bound of the integration with respect to a.

Figure 4 plots h against g, the curve hh, for the two pairs of bets of figure 3, and two values of c. Each point of each curve is obtained with a different interval $[a_1, a_2]$. In each curve we kept the ratio of a_2 to a_1 constant, 1.5 in all curves of figure 3, and computed pairs (g, h) for a few scores of intervals so as to obtain a full range of g between 0 and 1. For instance, for the lotteries (£5, 81%) and (£18, 19), figure 3 shows that $g = 0$ with $a_1 = 1.132$, and $g = 1$ with $a_2 = 1.132$. Thus, for g to vary between 0 and 1 with $a_2 = 1.5a_1$, the intervals varied from to [1.132, 1.132 × 1.5] to [1.132/1.5, 1.132].

The computation of g is straightforward. That of h is not. Therefore we computed discrete approximations to it numerically. As each curve is made of scores of pairs (g, h), and the computation of each value of h is long and tedious we wrote a computer programme to do the job. This allowed us to observe effortlessly curves for many pairs of lotteries, many values of c, and many a_2-to-a_1 ratios. We observed curves for, among others, the following pairs: (£5, 94%), (£17, 39%); (£4, 92%), (£10, 50%); (£6, 94%), (£13, 50%); (£8, 97%), (£32, 31%); and (£4, 81%), (£18, 19%). Many preference reversal experiments, including the best known ones, used these pairs or

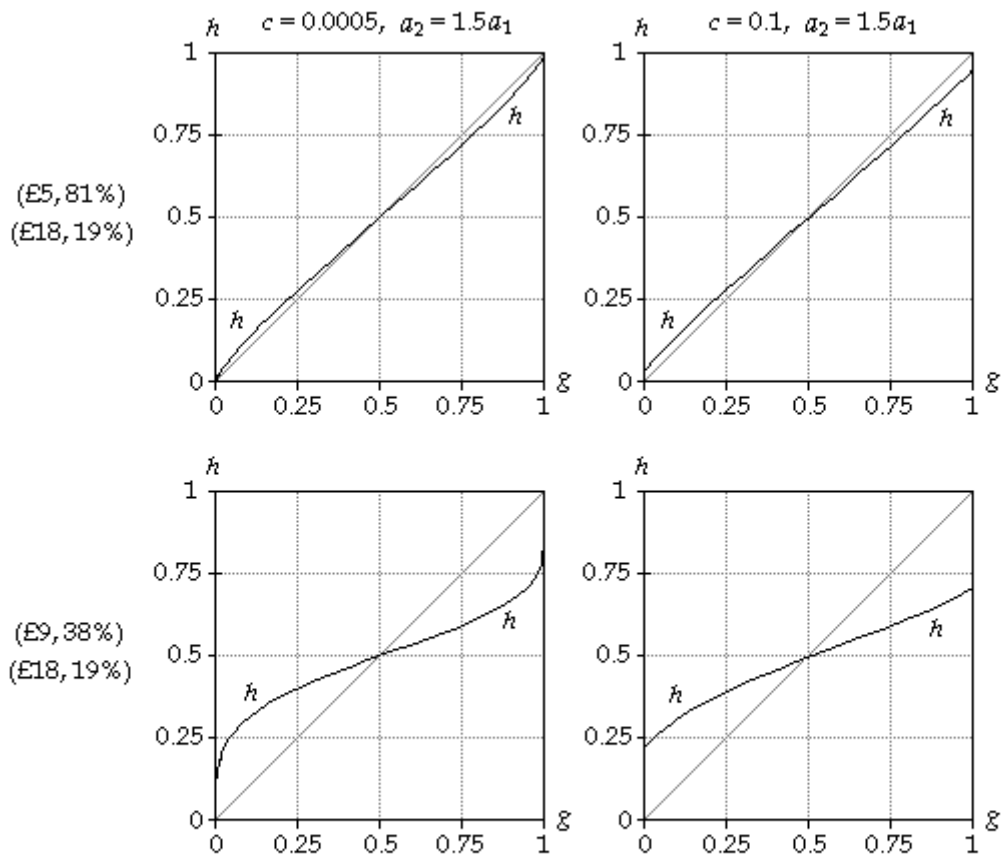

Figure 4: g and h under power-function expected utility

pairs very similar to these.[6] One of the pairs of figures 3 and 4 is a small variation of the last pair of the list, and was, together with the second-to-last pair of the list, used in Braga and Starmer (2001).

The main findings of this exercise may be summarised in five stylised facts.

1. The curve *hh* crosses the diagonal only once, always near its midpoint, and from above.

2. With the pairs listed above, and for any values of c and the a_2-to-a_1 ratio, the absolute difference between h and g is always less than 20 percentage points, and less

[6] Cubitt et al (forthcoming) used these pairs. Tversky et all (1990) used pairs with the same probabilities and proportional payoffs. Multiplying the payoffs in a pair by the same scalar does not alter g or h, as inspection of its analytical expressions reveals. The seminal experiment by Lichtenstein and Slovic (1971), and the landmark study by Grehter and Plott (1979) used pairs with the same probabilities, proportional winning amounts, and a small loss instead of the null outcome.

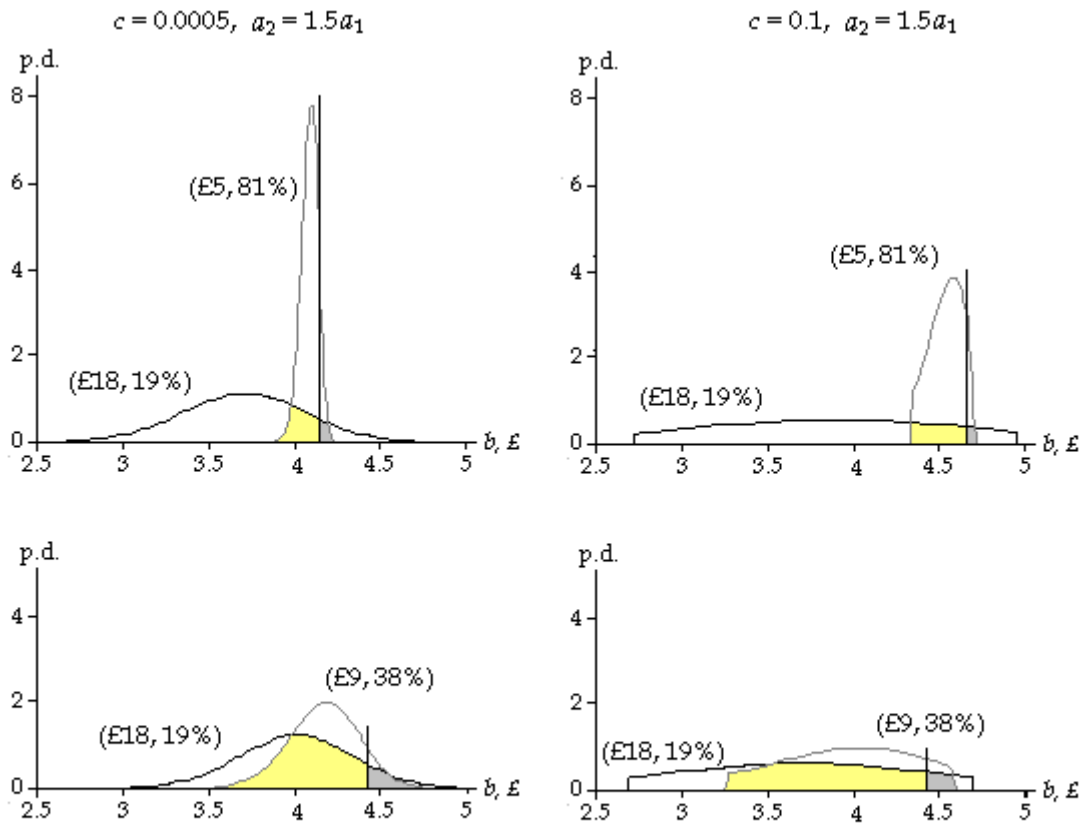

Figure 5: Optimal bids (*b*) probability density (p.d.) distribution with $g = 0.9$

than 10 percentage points for $g < 0.75$. This is much less than the differences under the p^* and $p^{\#}$ preference distributions seen above.

3. Keeping c and the a_2-to-a_1 ration constant, this difference increases as the *$* bet becomes safer given the *P* bet; as the *P* bet becomes riskier given the *$* bet; and as the two bets become riskier given the absolute difference between their winning probabilities.

4. The difference increases at the ends of the curve when c increases, that is, when the distribution of a becomes flatter.

5. Increasing the a_2-to-a_1 ratio slides the whole *hh* curve down, but mainly at the ends.

In what follows we will try to offer an explanation for these findings.

The stylised facts 2 and 3 hinge on the dispersion of the certainty equivalents of the *P* bet relative to those of the *$* bet. Figure 5 helps to understand why. The figure shows the probability distribution of the optimal bids corresponding to the point with abscissa $g = 0.9$ of the *hh* curves in figure 4. There is no particular reason to

choose that particular abscissa. The vertical lines mark the intersection points of the curves in figure 3. Any choice task compares two certainty equivalents that are on the same side of the vertical line, or, with probability zero, on the vertical line: if on the left-hand side, the *P* bet is chosen; if on the right-hand side the *$* bet is chosen.

h may differ from *g* only to the extent that the distribution of the *P* optimal bid overlaps that of the *$* bet. The reason for this is that the *$* certainty equivalents below the minimum *P* certainty equivalent rank necessarily the *$* bet below the *P* bet both in valuations and in choices, and the *$* certainty equivalents above the maximum *P* certainty equivalent rank *$* above *P* both in choices and valuations. Therefore, if the whole range of the *P* certainty equivalents overlaps only a small portion of the *$* certainty equivalent range, as is the case when the *P* bet is very safe relative to the *$* bet, then *g* and *h* must be very close. If the distribution of the *P* certainty equivalents collapses to the vertical line, if *P* is a sure amount, then *g* and *h* are equal. If the *P* certainty equivalent distribution has long but slim tails covering most of the *$* certainty equivalent distribution, the argument becomes a bit more complicated. In that case we would have to say that the *$* certainty equivalents to the left of the left-hand side intersection of the two distribution curves rank *$* below *P* in choices and have only a slim probability (as slim as the tail of the *P* certainty equivalent distribution) of ranking *$* above *P* in valuations. The converse is true to the right of the right-hand side intersection. This explains stylised facts 2 and 3, and is illustrated by comparing the top and bottom panels of figures 4 and 5.

h is lower than *g* if the light shaded areas of figure 5 are bigger than the dark shaded areas and vice-versa. The light shaded area is in the *P* choice area, but a *$* valuation drawn from the corresponding range has some probability of being higher than the valuation of the paired *P* bet. The converse is true of the dark shaded area. Note that the vertical lines divide the area under each of the curves into *g* on the left-hand side, and $1 - g$ on the right-hand side. Thus, when *g* is close to zero *h* is higher than *g*, and when *g* is close to 1 *h* is lower than *g*. When *g* is close to 0.5, the two areas have similar sizes, and *h* is close to 0.5 as well. This explains the first stylised fact.

The comparison of the left and right-hand side panels of figure 5 help understand stylised fact 4, if one imagines a *g* even closer to 1. For $0.1 < g < 0.9$ the

difference between g and h hardly changes with c. Increasing the parameter c makes the distribution of a, and consequently the distribution of the certainty equivalents flatter. This increases the size of the overlapping areas at the ends of the g spectrum, magnifying the difference that tends to exist there between g and h. If g is close to 0.5 the overlapping area decreases, but there the two halves of that area offset each other to a great extent, therefore the difference between g and h, small to begin with, does not change noticeably.

The distribution of the P certainty equivalents has a left-hand side tail that is hardly noticeable if the a_2-to-a_1 ratio is not too big, say no bigger than 2. Increasing this ratio makes the tail longer. This increases the light shaded areas relative to the dark shaded areas, resulting in a lower h for any value of g, which explains stylised fact 5. This effect is hardly noticeable with the pairs of lotteries listed above, unless coupled with a large c.

The general conclusion to draw from our simulations is that with this type of preference distributions and the pairs of lotteries that have been used in best known preference reversal studies the ranking probabilities should be very similar. Of the listed pairs, it is with the last one, (£4, 81%), (£18, 19%), that the hh curve is further from the diagonal (this hh curve cannot be distinguished at naked eye from that obtained with the lottery (£5, 81%)). Even with this pair, the difference between g and h never reaches 10 percentage points, unless c and the a_2-to-a_1 ratio are both large, say $c > 0.1$, and $a_2 > 3a_1$.These values generate implausible distributions of the certainty equivalents. For instance, with $c = 0.1$ and $a_2 = 3a_1$ the certainty equivalent of (£18, 19%) becomes roughly uniformly distributed between 0.4 and 4.8 for $g = 0.9$, and between 0.5 and 6.4 for $g = 0.5$. One would not expect the preferences of an individual to be so random.

The question naturally arises whether other types of core utility functions and other types of probability distributions over these functions will give rise to larger differences between g and h. We will not try to answer this question, but feel inclined to speculate that the difference will tend to remain small if the P optimal bids are very concentrated relative to the $\$$ optimal bids, and the $\$$ optimal bid is an increasing function of the P optimal bid, in the sense that it is so in figure 3. The importance of

the first condition was made explicit in our previous discussion. The second condition was also crucial, even if not so obvious. The second condition guarantees that in figure 5 *P* is chosen over *$* to the left of the vertical lines, and *$* is chosen over *P* to the right of that line. It is the two conditions together that with our type of utility function cause *g* and *h* to have similar values.

5. Testing the Stochastic Models with Actual Data

We will use the results of Braga and Starmer (2001) to test the error and the random preference models. In that experiment subjects were randomly allocated to two treatments: in one, valuations were elicited in a second-price auction, in the other, in a second-to-last price auction. In both markets each lottery was auctioned five times in a row. Table 11 shows the data obtained with the fifth valuations, and the single choice, made after all the auction sessions.

The relative frequencies are significantly different. If we assume that the category probabilities are the same in both markets, the maximum likelihood estimates of these probabilities, which are simply the relative frequencies of the aggregated data, are *pp* = 0.301, *sr* = 0.266, *nsr* = 0.127, and *dd* = 0.306. The chi-squared statistic, expression (1), may be computed for both markets. Their sum follows a chi-squared distribution with three degrees of freedom (number of independent categories, six, minus the number of independently estimated parameters, three). The sum of the chi-squared statistics is 17.14, and the probability value is 0.007.

Table 11: Category frequencies in Braga and Starmer (2001)

	2nd Price Auction		**2nd-to-last Price Auction**	
	Highest Price		Highest Price	
Choice	P	*$*	P	*$*
P	35	16	17	30
$	15	21	7	32

Therefore no model that assumes no differential effect of the type of auction can with the same set of parameters fit the two data sets. We will therefore fit the model separately to each of the datasets.

5.1. Testing the Lichtenstein and Slovic's Error Model

Table 12 shows the results of fitting the Lichtenstein and Slovic's (1971) error model to each of the markets. There is a set of plausible parameters with which the model fits the second-price auction perfectly. One cannot therefore reject the hypothesis that behaviour in that market can be represented by the error model.

In the second-to-last price auction the fit is not as good. In addition the estimate of r is implausibly high. With $r = 0.43$, the choices are close to random: we would expect $r = 0.43$ if 86% of all choices were random, and the remaining 14% always correct. The testing procedure carried out in section 3, when applied to the second-to-last price auction data, leads to a probability value of 0.106. One might not feel entirely confident about rejecting the model on the basis of this probability value. Therefore we carried out another test.

As 0.43 is an implausibly high value for r, we estimated the model subject to the constraint $r \leq 1/3$. An error probability of 1/3 in choices is still implausibly high (implying that 2/3 of all choices are random), but it is low enough for the procedure of section 3 to reject the model. The maximum likelihood estimates are $p = 0.274$, $r = 1/3$, and $s = 0.016$, which yield $\chi^2 = 6.48$. Imagine that the value $r = 1/3$ had been assumed and not estimated from de data. Then the χ^2 would have one degree of freedom, and the probability value would be 0.011. If we had assumed that r was any other value, lower than 1/3, $\chi^2 \geq 6.48$ (as $\chi^2 = 6.48$ was obtained with the constraint $r \leq 1/3$, not $r = 1/3$), and the model would also be rejected.

Table 12: Maximum likelihood estimates of the error model

Lichtenstein and Slovic's (1971) Error Model	**2nd Price Auction**	**2nd-to-last Price Auction**
	$\chi^2_1 = 0$	$\chi^2_2 = 2.61$
	$p = 0.65$, $r = 0.25$, $s = 0.21$	$p = 0.28$, $r = 0.43$, $s = 0.0$

		Highest Price		Highest Price	
	Choice	*P*	*$*	*P*	*$*
Observed	*P*	35	32	17	30
Frequencies	*$*	15	21	7	32
Expected	*P*	35.0	16.0	13.9	26.7
Frequencies	*$*	15.0	21.0	10.7	34.8

We conclude then that the second-price auction may be explained by the model, but the second-to-last price auction cannot.

5.2. Testing the Random Preference Model

To fit the random preference model to actual data we maximise the likelihood function subject to the necessary and sufficient conditions for the existence of a suitable distribution of (g, h), and the necessary conditions for the existence of a suitable set of preference distributions. For a quick reference we present these constraints together here, and will refer to them as the standard constraints:

$0 \le \mu_g, \mu_h \le 1,$

$\mathrm{Cov}(g, h) \ge 0,$

$\mathrm{Cov}(g, h) \le \mu_g(1-\mu_h), \ \mu_h(1-\mu_g),$

$\mathrm{Cov}(g,h) < \sqrt{(\mu_h - \mu_h^2)(2\mu_h - \mu_g^2)} \ , \ \sqrt{(\mu_h - \mu_h^2)(1 - \mu_g^2 + 2\mu_g - 2\mu_h)}\,.$

The demanding sufficient conditions for the existence of set of suitable preference distributions, $\mu_g/2 < \mu_h < (1 + \mu_g)/2$ (24), will not be imposed, but checked after the maximisation.

The exact solution (see section 4.2) is feasible in both datasets. That is, there is for each market a set of parameter with which the model predicts exactly the observed frequencies, making $\chi^2 = 0$. Both sets of parameters satisfy the demanding sufficient conditions for the existence of a set of suitable preference distributions. For the second-price auction the parameters are $\mu_g = 0.59$, $\mu_h = 0.58$, and $\mathrm{Cov}(g, h) = 0.065$ (sufficient conditions: $0.295 < \mu_h < 0.795$); for the second-to-last price auction the

parameters are $\mu_g = 0.547$, $\mu_h = 0.279$, and Cov(g, h) = 0.045 (sufficient conditions: $0.274 < \mu_h < 0.774$).

Of the two sets of parameters, that of the second-price auctions looks more sensible. The small difference between the mean ranking probabilities is within the range of what we observed in our example of a reasonable preference distribution. The big difference in the second-to-last price auction is not. There are sets of preference distributions that yield the parameters estimated for this market (because the sufficient conditions are met, even if narrowly so), but given the difference between μ_g and μ_h one might question the reasonability of such preference distributions. On the other hand these parameters yield exactly the observed frequencies. The question arises then whether imposing a more plausible set of parameters will reject the model.

To answer that question we maximised the likelihood function subject to the standard constraints and $\mu_h \geq k\mu_g$. Given that the best fit is obtained with μ_h much lower than μ_g, that additional constraint, with k high enough, forces the difference between μ_h and μ_g to be smaller and more in line with what we observed in our simulations of reasonable preference distributions. We then obtained by trial and error the values of k that lead to the rejection of the model at the most commonly used significance levels. The results are in table 13.

The rationale for this test is the same as for the test conducted with the Lichtenstein and Slovic's (1971) error model. If the constraint $\mu_h = k\mu_g$ is imposed on the maximisation of the likelihood function, the model reduces, in effect, to a two parameter model, and the χ^2 has one degree of freedom. If one thinks that a plausible value for k should be, say, no less than 0.72, then $\chi^2 \geq 3.85$ (as 3.85 is the lowest value of χ^2 if the constraint is $\mu_h \geq 0.72\mu_g$) and the probability value will be at most 0.05.

Table 13: Fitting the model with reasonability constraints

Pr	χ^2	constraint	μ_g	μ_h	Cov(g, h)
0.10	2.69	$\mu_h \geq 0.682\mu_g$	0.51	0.34	0.047
0.05	3.85	$\mu_g \geq 0.72\mu_g$	0.49	0.36	0.046
0.01	6.72	$\mu_h \geq 0.804\mu_g$	0.47	0.38	0.042

The model is rejected if one thinks that any of the constraints of table 13 is justifiable. Whether they are is a question to which we cannot give a definitive answer, but given that the model cannot explain both markets, we feel inclined to conclude that the model, if anything, explains the second-price auction.

We fitted the model and applied this testing procedure to several of the most prominent datasets. Table 14 summarises the results. The fifth column shows the probability (Pr) value (assuming one degree of freedom) obtained by fitting the model with the standard constraints and the additional constraints $\mu_h \geq k\mu_g$ and $\mu_h \leq 1 - k(1 - \mu_g)$ (the latter is never binding) with $k = 0.5$. These additional constraints guarantee the existence of a suitable set of preference distributions. $\chi^2 = 0$ in this column indicates that the exact solution is feasible under the additional constraints. The fourth column shows the χ^2 value when the model is fitted only with the standard constraints (thus k is zero in the additional constraints). Again, $\chi^2 = 0$ means that the exact solution is feasible. The last two columns show the values of k in the additional constraint $\mu_h \geq k\mu_g$ needed to reject the model at the stated probability values (under the assumption of one degree of freedom).

Table 14: Fitting the random preference model to some experimental results

Experiment	Incent comp	Sample size	χ^2	Pr	Constraint: k	
			$k = 0$	k=0.5	Pr=0.1	Pr=0.01
Lichtenstein and Slovic, 1971						
Experiment I	No	1038	2.3	<0.01	0.26	0.28
Experiment III	Yes	84	0.0	(χ^2=0)	0.67	0.78
Grether and Plott, 1979						
No incentive	No	245	0.0	(χ^2=0)	0.62	0.68
With incentives	Yes	262	0.0	(χ^2=0)	0.66	0.72
Tversky and Kahneman, 1990, set I	No	1074	0.9	<0.01	0.37	0.39

Notes: k is the k in the additional constraint $\mu_h \geq k\mu_g$; sample sizes are the number of subjects multiplied by 6 pairs of bets; Incent comp stands for incentive compatible. An experiment is incentive compatible if subjects decisions have economic consequences for them.

The exact solution is not feasible in some of the datasets, but the χ^2 statistic is not high enough for our testing procedure to reject the model at the 10% significance level (the critical value at the 10% significance level is 2.71). If the demanding sufficient conditions are imposed, two data sets reject the model at the 1%

significance level. None of these were obtained in incentive compatible experiments. All data sets reject the model at 1% significance level if the constraint $\mu_h \geq 0.78\mu_g$ is imposed, and at 10% if $\mu_h \geq 0.67\mu_g$ is imposed. These constraints look sensible in light of our simulations of reasonable preference distributions. We cannot rule out the possibility that other preference distributions that might be deemed reasonable produce bigger differences between μ_h and μ_g. Still, at the very least, we may conclude that there is no evidence so far that a model of random expected utility may explain preference reversal.

6. Conclusions

This paper started with the observation that while preference reversal has been seen as a non-random deviation from the predictions of most deterministic preference theories, very little is known of what a random deviation in a preference reversal experiment may look like. This paper is an attempt to increase our knowledge in that area. We revisited the early, apparently forgotten, Lichtenstein and Slovic's (1971) error model, henceforth the *error model*, and developed a random preference model of choices and valuations, henceforth the *random preference model*. These models combine stochastic processes with deterministic expected utility theory. Both models predict inconsistencies between choices and valuations, but we feel inclined to conclude that they do not explain preference reversal.

Our random preference model is based on Loomes and Sugden's (1995) theory of random preference. Of the recent research in stochastic models of choice, Loomes and Sugden's theory is the one that can be extended to valuations in an obvious way.

When the same assumption is made in both models about whether individuals are homogeneous or diverse, the random preference model is in part observationally equivalent to a particular case of the error model. This is however no more than a curiosity, as that particular case requires the extreme assumption that all individuals prefer the same bet.

The random preference model, even with the core theory restricted to expected utility, is able to predict decision patterns that have generally been viewed as non-

random deviations from most deterministic preference theories, including expected utility. Many, if not most, economists and psychologists have relied on the difference between rates of standard and non-standard reversals conditional on choice as an indication of the non-random nature of preference reversal. According to the random preference model that measure is not very meaningful: the model predicts no limit to that difference. A more meaningful measure is the difference between the unconditional rates of standard and non-standard reversal (standard and non-standard reversals as proportion of total choices): the model predicts an upper limit to that difference.

Changes in these rates of reversal conditional on choice have also been used to justify various assertions as to the effect on the strength of preference reversal of changes in the experimental design. Other measures, namely unconditional rates of reversal, often lead to different conclusions (see Braga 2003). It is illustrative to revisit one famous such assertion. Grether and Plott (1979) were surprised to find out that the introduction of incentive compatibility had made preference reversal stronger. This finding was based on the big increase in the rate of standard-reversal conditional on choice. The unconditional rate of standard reversal actually decreased slightly. We fitted the random preference model to the authors' datasets. The results are shown in table 14. It took a stricter constraint to reject the model with the incentive-compatible dataset than with the non-incentive-compatible one. On face value this means that the introduction of incentives pushed behaviour in the direction of stochastic expected utility. Given that the rationale for the constraints used to reject the model is based on simulations with a particular type of utility functions we are not willing to make much of this result. We may nevertheless conclude that we have no evidence that incentives made preference reversal stronger.

Testing the random preference model against the data was not absolutely conclusive. We fitted the model to seven datasets. Of these, one does not, at naked eye, display preference reversal. Again at naked eye, it displays some bias in the opposite direction, that is, the rate of non-standard reversal conditional on choice exceeds that of standard reversal. The unconditional rates of reversal are virtually the same though. The model cannot be rejected with this data set.

Two datasets reject the model when one imposes a constraint that guarantees, but is not necessary for, the existence of a set of suitable preference distributions for the estimated parameters. These two data sets happen to result from experiments that were not incentive compatible.

Four data sets, two of them from incentive-compatible experiments, reject the model if a further constraint on the parameters is imposed. This constraint looks sensible in light of simulations with a particular type of "reasonable" preference distributions. These distributions were truncated normal probability distributions over sets of utility functions of a particular type. One may not feel entirely confident in rejecting the model on the basis of a particular type of preference distribution.

This points the way to further research. We derived constraints on the parameters of the model from basic statistical principles and the axioms of expected utility. The implication of expected utility to our model was simply that each preference ordering determines the choice of a bet from a pair if and only if it determines that the valuation of that bet is higher than the valuation of the other bet. No other restrictions were imposed on the preference distribution. That is, no restrictions were imposed on the set of preference orderings, except that each preference ordering must be compatible with expected utility, and no restrictions were imposed on the probability distribution over those preference orderings. This is what allows the model to predict a wide variety of decision patterns, including patterns observed in many preference reversal experiments. Many of those preference distributions are nonsensical. In section 4.5 we made simulations with "reasonable" preference distributions, and observed the parameters that they produce, but what is desirable is a normative theory of preference distributions.

A normative theory of preference distributions is necessarily a theory of reasonable preference distributions. An analogy can be made with deterministic preference theory. In the standard theory of preference under certainty, monotonicity and convexity of the indifference curves are properties required on grounds of reasonability. So is independence in the theory of choice under uncertainty. In the same manner it is desirable to require that a preference distribution display normatively appealing properties. Otherwise some predictions of a random

preference model may be dismissed on grounds of relying on absurd preference distributions.

Therefore a generally accepted normative theory of preference distributions is required for a more definitive test of whether the decision patterns observed in preference reversal experiments are compatible with stochastic expected utility.

Appendix A: Regret Theory, Intransitivity, and Valuations

Regret theory assumes that individuals facing a choice between two prospects X and Y compare the outcomes of the prospects in each state of the world. The comparison is evaluated with the function $\Psi(x_i, y_i)$, where x_i, y_i are the outcomes of X and Y in state of the world i, and $\Psi(.)$ has the properties $\Psi(a, a)= 0$ and $\Psi(a, b)= -\Psi(b, a)$. Denoting p_i the probability of state of the world i happening, the individual will prefer X to Y if

$$V(X, Y) = \sum_i p_i \Psi(x_i, y_i) > 0.$$

He will prefer Y to X if $V(X, Y) < 0$, and will be indifferent if $V(X, Y) = 0$. Note that $V(Y, X) = -V(X, Y)$. Such a preference relation may not be transitive in general, but is transitive over the set S of gambles generated by the admissible bids.

Suppose X offers an amount of money x with probability p, and nil with probability $1 - p$. That is, $X= (x, p; 0, 1 - p)$. Let b_j, b_k, and b_l, be any three admissible bids such that $b_j < b_k < b_l$ ($j < k < l$). Making these bids amounts to play the following gambles (see main text):

$$C(b_j,X)= [x,pG(b_j); 0, (1-p)G(b_j); b_{j+1}, G(b_{j+1})-G(b_j);\ldots; b_k, G(b_k)-G(b_{k-1});\ldots; b_N, 1-G(b_{N-1})],$$

$$C(b_k,X)= [x,pG(b_k); 0, (1-p)G(b_k); b_{k+1}, G(b_{k+1})-G(b_k);\ldots; b_l, G(b_l)-G(b_{l-1});\ldots; b_N, 1-G(b_{N-1})],$$

$$C(b_l, X) = [x, pG(b_l); 0, (1-p)G(b_l); b_{l+1}, G(b_{l+1}) - G(b_l);\ldots; b_N, 1 - G(b_{N-1})].$$

For the sake of readability define $C_i = C(b_i, X)$, and $V_{h,i} = V(C_h, C_i)$. Then

$$V_{j,k} = \sum_{i=j+1}^{k} \left[G(b_i) - G(b_{i-1}) \right] \left[p\,\Psi(b_i, x) + (1-p)\Psi(b_i, 0) \right]$$

$$V_{k,l} = \sum_{i=k+1}^{l} \left[G(b_i) - G(b_{i-1}) \right] \left[p\,\Psi(b_i, x) + (1-p)\Psi(b_i, 0) \right]$$

$$V_{j,l} = \sum_{i=j+1}^{l} \left[G(b_i) - G(b_{i-1}) \right] \left[p\,\Psi(b_i, x) + (1-p)\Psi(b_i, 0) \right]$$

Obviously, $V_{j,l} = V_{j,k} + V_{k,l}$. Then if $C_j \geq_p C_k$ and $C_k \geq_p C_l$, then $V_{j,k} \geq 0$ and $V_{k,l} \geq 0$. $V_{j,l} = V_{j,k} + V_{k,l}$, therefore $V_{j,l} \geq 0$, and $C_j \geq_p C_l$. In this example C_i, that is, $C(b_i, X)$ was preference ordered by increasing order of b_i. This is unimportant. For instance, if $C_j \geq_p C_l$ and $C_l \geq_p C_k$, then $V_{j,l} \geq 0$ and $V_{l,k} \geq 0$. $V_{j,k} = V_{j,l} - V_{k,l} = V_{j,l} + V_{l,k}$, therefore $V_{j,k} \geq 0$, and $C_j \geq_p C_k$. It can be easily checked that any other chain of preference (three gambles allow six different preference chains) is also transitive.

Appendix B: Minimum and Maximum *h* given *g*

1. Minimising *h* Given *g*

We want to solve the following problem:

$$\min\ h = \sum_{i=1}^{m_P} \sum_{j=0}^{m_S} p_{ij} \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p_{kl} \tag{B1}$$

subject to

$$\sum_{i=1}^{m_P} \sum_{j=0}^{i-1} p_{ij} = g\,; \tag{B2a}$$

$$\sum_{i=0}^{m_P} \sum_{j=0}^{m_S} p_{ij} = 1; \tag{B2b}$$

$$p_{ij} \geq 0, \quad i = 0, \ldots, m_P \text{ and } j = 0, \ldots, m_S; \tag{B2c}$$

$$p_{ii} = 0, \quad i = 0, \ldots, m_P. \tag{B2d}$$

The Lagrangian function for this problem is

$$£ = \sum_{i=1}^{m_P}\sum_{j=0}^{m_\$} p_{ij} \sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} - \lambda_g\left(\sum_{i=1}^{m_P}\sum_{j=0}^{i-1} p_{ij} - g\right) - \lambda_1\left(\sum_{i=0}^{m_P}\sum_{j=0}^{m_\$} p_{ij} - 1\right) -$$

$$\left(\lambda'_{11}p_{11} + \ldots + \lambda'_{m_P m_P}p_{m_P m_P}\right) - \left(\left(\lambda_{11}p_{11} + \ldots + \lambda_{1m_\$}p_{1m_\$}\right) + \ldots + \left(\lambda_{m_P 1}p_{m_P 1} + \ldots + \lambda_{m_P m_\$}p_{Pm_\$}\right)\right) \quad \text{(B3)}$$

The Kuhn Tucker conditions for a minimum are

$$\frac{\partial £}{\partial p_{mn}} = 0 \quad \text{for } m = 0,\ldots, m_P \quad \text{and} \quad n = 0,\ldots, m_\$; \quad \text{(B4a)}$$

$$\sum_{i=1}^{m_P}\sum_{j=0}^{i-1} p_{ij} = g\,; \quad \text{(B4b)}$$

$$\sum_{i=0}^{m_P}\sum_{j=0}^{m_\$} p_{ij} = 1; \quad \text{(B4c)}$$

$p_{mm} = 0$ for $m = 0,\ldots, m_P$; (B4d)

$\lambda_{mn}p_{mn} = 0$, for $m = 0,\ldots, m_P$ and $n = 0,\ldots, m_\$$; (B4e)

$p_{mn} \geq 0$, $\lambda_{mn} \geq 0$ for $m = 0,\ldots, m_P$ and $n = 0,\ldots, m_\$$. (B4f)

The $(m_P + 1) \times (m_\$ + 1)$ derivatives (B4a) take one of the following forms:

$$\frac{\partial £}{\partial p_{00}} = \sum_{i=1}^{m_P}\sum_{j=0}^{m_\$} p_{ij} - \lambda_1 - \lambda_{00} - \lambda'_{00} = 0\,; \quad \text{(B5a)}$$

$$\frac{\partial £}{\partial p_{0n}} = \sum_{i=n+1}^{m_P}\sum_{j=0}^{m_\$} p_{ij} - \lambda_1 - \lambda_{0n} = 0 \quad \text{for } n = 1,\ldots, m_P - 1; \quad \text{(B5b)}$$

$$\frac{\partial £}{\partial p_{0n}} = -\lambda_1 - \lambda_{0n} = 0 \quad \text{for } n = m_P,\ldots, m_\$; \quad \text{(B5c)}$$

$$\frac{\partial £}{\partial p_{mn}} = \sum_{k=0}^{m_P}\sum_{l=0}^{m-1} p_{kl} + \sum_{i=n+1}^{m_P}\sum_{j=0}^{m_\$} p_{ij} - \lambda_g - \lambda_1 - \lambda_{mn} = 0 \quad \text{(B5d)}$$

for $m = 1,\ldots, m_P$, $n = 1,\ldots, m-1$;

$$\frac{\partial £}{\partial p_{mn}} = \sum_{k=0}^{m_P}\sum_{l=0}^{m-1} p_{kl} + \sum_{i=n+1}^{m_P}\sum_{j=0}^{m_\$} p_{ij} - \lambda_1 - \lambda_{mn} - \lambda'_{mn} = 0 \quad \text{for } m = 1,\ldots, m_P,\ n = m; \quad \text{(B5e)}$$

$$\frac{\partial £}{\partial p_{mn}} = \sum_{k=0}^{m_P}\sum_{l=0}^{m-1} p_{kl} + \sum_{i=n+1}^{m_P}\sum_{j=0}^{m_\$} p_{ij} - \lambda_1 - \lambda_{mn} = 0 \quad \text{(B5f)}$$

for $m = 1,\ldots, m_P$, $n = m + 1,\ldots, m_P - 1$;

$$\frac{\partial \pounds}{\partial p_{mn}} = \sum_{k=0}^{m_P}\sum_{l=0}^{m-1} p_{kl} - \lambda_1 - \lambda_{mn} = 0 \quad \text{for } m = 1,\ldots, m_P,\ n = m_P,\ldots, m_\$. \tag{B5g}$$

One Solution: the p Probability Distribution*

The preference distribution p^* (B6a-c) ((4-6) in the main text) and the following values of the Lagrange multipliers (B6d-k) satisfy the Kuhn Tucker conditions, that is, (B5a-g) (that is, condition (B4a)) and conditions (B4b-f):

$$\sum_{n=m_P}^{m_\$} p^*_{0n} = 1 - g, \quad \text{for } n = m_P,\ldots, m_\$; \tag{B6a}$$

$$p^*_{m,m-1} = \frac{g}{m_P} \quad \text{for } m = 1,\ldots, m_P; \tag{B6b}$$

$$p^*_{mn} = 0 \quad \text{for all other } (m, n); \tag{B6c}$$

$$\lambda^*_g = \frac{m_P + 1}{m_P} g\,; \tag{B6d}$$

$$\lambda^*_1 = 0; \tag{B6e}$$

$$\lambda^*_{0n} = \frac{m_P - n}{m_P} g \quad \text{for } n = 0,\ldots, m_P - 1; \tag{B6f}$$

$$\lambda^*_{0n} = 0 \quad \text{for } n = m_P,\ldots m_\$; \tag{B6g}$$

$$\lambda^*_{mn} = \frac{m - 1 - n}{m_P} g \quad \text{for } m = 1,\ldots, m_P,\ n = 0,\ldots, m - 1; \tag{B6h}$$

$$\lambda^*_{mn} = \frac{m_P + m - n}{m_P} g \quad \text{for } m = 1,\ldots, m_P,\ n = m,\ldots, m_P - 1; \tag{B6i}$$

$$\lambda^*_{mn} = \frac{m}{m_P} g \quad \text{for } m = 1,\ldots, m_P,\ n = m_P,\ldots, m_\$; \tag{B6j}$$

$$\lambda'^*_{mm} = 0 \quad \text{for } m = 1,\ldots, m_P. \tag{B6k}$$

Equations (B6a-c) merely describe the p^* preference distribution presented in the main text. Of the pairs of optimal bids with positive probability, only the pairs defined by (B6b) rank *P* above *$* (and none ranks *P* and *$* equally). There are m_P such pairs, each with probability g/m_P, thus their probabilities add up to g, and constraint

(B4b) is satisfied. The sum of all the remaining pairs with positive probability is $1-g$ (B6a), thus the probabilities of all pairs add up to 1, and constraint (B4c) is satisfied. Equations (B6a,b), the ones that define the pairs with positive probability, include no pair (m, n) where $n = m$. Thus all these pairs fall under equation (B6c), have zero probability, and constraints (B4d) are satisfied. The corresponding Lagrange multipliers, λ'_{mn}, are all null, meaning that constraints (B4d) would be satisfied even if they had not been imposed. All Lagrange multipliers corresponding to the pairs that may have positive probability are null: the Lagrange multipliers corresponding to pairs in (B6a) are in (B6g); those corresponding to pairs in (B6b) are in (B6h) where $n = m-1$. Thus constraints (B4e) are satisfied. All probabilities and Lagrange multipliers are non-negative, thus constraints (B6f) are satisfied.

This leaves us with constraints (B4a), or (B5a-g). Note that pairs in (B6a) do not appear in these constraints. In each double sum, for each l there is only one k, such that $p_{lk} > 0$, namely $p_{l+1,l} = g/m_P$, and all the other probabilities are null. Likewise, for each i, $p_{i,i-1} = g/m_P$, and all other probabilities are null. Thus substituting these results and the values of the Lagrange multipliers into equations (B5a-g) yields:

$$\frac{\partial \pounds}{\partial p_{00}} = \sum_{i=1}^{m_P} \frac{g}{m_P} - \frac{m_P - 0}{m_P} g = 0, \quad \text{(from (B6f))}; \tag{B7a}$$

$$\frac{\partial \pounds}{\partial p_{0n}} = \sum_{i=n+1}^{m_P} \frac{g}{m_P} - \frac{m_P - n}{m_P} g = 0 \quad \text{for } n = 1,\ldots, m_P - 1, \quad \text{(from (B6f))}; \tag{B7b}$$

$$\frac{\partial \pounds}{\partial p_{0n}} = \; - \; 0 - 0 = 0 \text{ for } n = m_P,\ldots, m_\$, \quad \text{(from(B6g))}; \tag{B7c}$$

$$\frac{\partial \pounds}{\partial p_{mn}} = \sum_{l=0}^{m-1} \frac{g}{m_P} + \sum_{i=n+1}^{m_P} \frac{g}{m_P} - \frac{m_P + 1}{m_P} g - \frac{m - 1 - n}{m_P} g = 0$$

$$\text{for } m = 1,\ldots, m_P,\ n = 0,\ldots, m-1, \text{ (from (B6d) and (B6h))}. \tag{B7d}$$

Equations (B5e) and (B5f) become the same, since $l'_{mn} = 0$. Thus

$$\frac{\partial \pounds}{\partial p_{mn}} = \sum_{l=0}^{m-1} \frac{g}{m_P} + \sum_{i=n+1}^{m_P} \frac{g}{m_P}_{1} - \frac{m_P + m - n}{m_P} g = 0$$

$$\text{for } m = 1,\ldots, m_P,\ n = m,\ldots, m_P - 1, \text{ (from (B6i))}; \tag{B7e,f}$$

$$\frac{\partial \pounds}{\partial p_{mn}} = \sum_{l=0}^{m-1} \frac{g}{m_{P\ 1}} - \frac{m}{m_P} g = 0 \quad \text{for } m = 1,\ldots, m_P, n = m_P,\ldots, m_\$, \quad \text{(from (B6j))}. \tag{B7g}$$

It may easily be checked that equations (B7a-g) are all true, and therefore conditions (B4a-f) are all satisfied.

Non Quasi-Convexity of h

These conditions are necessary for p^* to be a minimum. If the objective function, h, was strictly quasi convex we would have a guaranty that p^* was a (and only) global minimum. h would be strictly quasi convex if for any two preference distributions p' and p'', and any scalar α such that $0 < \alpha < 1$, $h(\alpha p' + (1 - \alpha)p'') < \max\{h(p'), h(p'')\}$. The following example shows that h is not strictly quasi convex. (It is not quasi convex either. The definition of quasi convexity may be obtained from that of strict quasi convexity by substituting weak inequalities for the strict ones.). Note that p' and p'' in the table below are probability distributions over the listed pairs: their probabilities are non-negative and add up to one.

Table B1: h is not strictly quasi convex

$(b_P, b_\$)$	p'	p''	$0.5p' + 0.5p''$
(0, 4)	0.2	0.2	0.2
(0, 1)	0.4	0.0	0.2
(2, 1)	0.4	0.4	0.4
(2, 3)	0.0	0.4	0.2
$h = (p_{21} + p_{31})(p_{01} + p_{21})$	0.32	0.32	0.36

If h was strictly quasi concave, any preference distribution satisfying the Kuhn Tucker conditions for a maximum would be the global maximum. This is a preference distribution that we would like to find as well. h would be strictly quasi concave if for any two preference distributions p' and p'', and any scalar α such that $0 < \alpha < 1$, $h(\alpha p'$

+ $(1 - \alpha)p'') > \min\{h(p'), h(p'')\}$. The following example shows that h is not strictly quasi concave (and not quasi concave either).

Table B2: h is not strictly quasi-concave

$(b_P, b_\$)$	p'	p''	$0.5p' + 0.5p''$
(0, 4)	0.6	0.6	0.6
(2, 1)	0.4	0.0	0.2
(3, 2)	0.0	0.4	0.2
$h = p_{21}^2 + p_{32}(p_{32} + p_{21})$	0.16	0.16	0.12

As the function h is not strictly quasi convex there may be other minima than p^*. One way to prove that no other minimum is smaller than (8) would be to find all the solutions to the Kuhn-Tucker conditions, and see which one yields the lowest h. The Kuhn-Tucker conditions include an indefinitely large number of non-linear equations and non-negativity constraints over an indefinitely large number of variables. Finding all the solutions to these conditions is not an enterprise one would necessarily like to undertake. We will follow an alternative route. That requires a few definitions.

The Set of p^+ Probability Distributions

We will define three types of pairs of bids. Pairs where $b_P = 0$ and $b_\$ \geq m_P$, to be called $\$^*$ pairs (because they rank $\$$ above P); pairs where $b_\$ = b_P - 1$, to be called P^* pairs, (because they rank P above $\$$); and all the remaining pairs, to be called O^* pairs. Define p^+ as the set of all probability distributions over $B_P \times B_\$$ that assign zero probability to all O^* pairs. Note that $p^* \in p^+$: all $\$^*$ pairs are in (4) and all P^* pairs are in (5). Consequently all O^* pairs are in (6), and therefore have no probability. However not all p^+ distributions are p^*, as p^+ distributions need not assign the same probability to all P^* pairs, as p^* does.

p^+ probability distributions have the following properties. If p' is a p^+ distribution, g', the probability that P be chosen over $\$$ under p', is the sum of the

probabilities of all P^* pairs, because, of the pairs with positive probability, only P^* pairs rank P above $\$$, and all P^* pairs do that. Thus

$$g' = \sum_{i=1}^{m_P} p'_{i,i-1} \quad \text{for any } p' \in p^+. \tag{B8}$$

A second property is that, under a p^+ distribution, P bid will exceed a $\$$ bid only if both valuations are extracted from P^* pairs (obviously no valuations can be extracted from O^* pairs). Thus if $p' \in p^+$, h', the probability that P be valued above $\$$ under p', defined by (B1) ((3) in the main text), simplifies to:

$$h' = \sum_{i=1}^{m_P} p'_{i,i-1} \sum_{k=1}^{i} p'_{k,k-1} = \sum_{i=1}^{m_P} \sum_{k=1}^{i} p'_{i,i-1} p'_{k,k-1} \quad \text{for any } p' \in p^+. \tag{B9}$$

A last property is that the function h is strictly convex in p^+.

This last property is very useful. The p^* probability distribution, which satisfies the Kuhn-Tucker conditions for a minimum, is a p^+ distribution. Therefore it is also a solution to the Kuhn-Tucker conditions when the domain of the probability distributions is restricted to p^+. As h is strictly convex in p^+, and all the constraints the minimisation of h was subject to are linear, and thus convex, the solution we found, h^*, is the global minimum in p^+. We will show next that for *any* non-p^+ probability distribution p and its implied g and h there is a $p' \in p^+$ such that $g' = g$, and $h' \leq h$. We will also show that either $h' < h$ or $h^* < h'$. Therefore h^* is the global minimum for all possible probability distributions over $B_P \times B_\$$.

Convexity of h in p^+

In this subsection we will deal only with p^+ distributions. Therefore g and h will refer only to the ranking probabilities under p^+ distributions. When the domain of h is restricted to p^+, expression (B9) can be further simplified. Table B3 illustrates how. Its first row lists all the P^* pairs, and so does the first column. Each of the remaining cells shows the probability that pair $(i, i-1)$ be drawn for the valuation of P, and pair $(k, k-1)$ be drawn, independently from $(i, i-1)$, for the valuation of $\$$. That probability is

$p_{i,i-1}p_{k,k-1}$. That is also the probability that the valuation of P is obtained from $(k, k-1)$ and the valuation of \$ is independently obtained from $(i, i-1)$. This is illustrated by the symmetry of the probability matrix.

Table B3: Probabilities of two independently drawn P^* pairs

\$ bid / P bid	(1, **0**)	(2, **1**)	(3, **2**)	. . . . .	(m_P, $\boldsymbol{m_P - 1}$)	Total
(**1**, 0)	p_{10}^2	$p_{10}p_{21}$	$p_{10}p_{32}$	. . . . .	$p_{10}p_{m_P, m_P-1}$	$p_{10}g$
(**2**, 1)	$p_{21}p_{10}$	p_{21}^2	$p_{10}p_{32}$	. . . . .	$p_{21}p_{m_P, m_P-1}$	$p_{21}g$
(**3**, 2)	$p_{32}p_{10}$	$p_{32}p_{21}$	p_{32}^2	. . . . .	$p_{32}p_{m_P, m_P-1}$	$p_{32}g$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
($\boldsymbol{m_P}$, $m_P - 1$)	$p_{m_P, m_P-1}p_{10}$	$p_{m_P, m_P-1}p_{21}$	$p_{m_P, m_P-1}p_{32}$	. . . . .	p_{m_P, m_P-1}^2	$p_{m_P, m_P-1}g$
Total	gp_{10}	gp_{21}	gp_{32}	. . . . .	gp_{m_P, m_P-1}	g^2

Under a p^+ distribution the probabilities of all P^* pairs add up to g. Table B3 contains the probabilities of all possible combinations of two independently drawn P^* pairs. Therefore the sum of the probabilities of all cells in the matrix is g^2. Under a p^+ distribution h is the sum of all the probabilities in the diagonal and lower shaded area (or in the diagonal and upper shaded area, as the matrix is symmetric). As the sum of the probabilities in each of the shaded areas is the same h can be written as

$$h = \frac{g^2 - \sum_{i=1}^{m_P} p_{i,i-1}^2}{2} + \sum_{i=1}^{m_P} p_{i,i-1}^2 = \frac{g^2 + \sum_{i=1}^{m_P} p_{i,i-1}^2}{2}. \tag{B10}$$

The convexity of this function is easy to check. h is strictly convex in p^+ if for any two p', $p'' \in p^+$, with $p' \neq p''$, and any scalar α, $0 < \alpha < 1$, $h(\alpha p' + (1-\alpha)p'') - \alpha h(p') - (1-\alpha)h(p'') < 0$.

$$h(\alpha p' + (1-\alpha)p'') - \alpha h(p') - (1-\alpha)h(p'') =$$

$$= \frac{g^2 + \sum_{i=1}^{m_P} \left(\alpha p'_{i,i-1} + (1-\alpha)p''_{i,i-1}\right)^2}{2} - \alpha \frac{g^2 + \sum_{i=1}^{m_P} p'^2_{i,i-1}}{2} - (1-\alpha)\frac{g^2 + \sum_{i=1}^{m_P} p''^2_{i,i-1}}{2} =$$

$$= \frac{1}{2}\sum_{i=1}^{m_P}\left(\alpha^2 p'^2_{i,i-1} + 2\alpha(1-\alpha)p'_{i,i-1}p''_{i,i-1} + (1-\alpha)^2 p''^2_{i,i-1} - \alpha p'^2_{i,i-1} - (1-\alpha)p''^2_{i,i-1}\right) =$$

$$= \frac{1}{2}\sum_{i=1}^{m_P}\left(\alpha(\alpha-1)p'^2_{i,i-1} + 2\alpha(1-\alpha)p'_{i,i-1}p''_{i,i-1} - (1-\alpha)\alpha p''^2_{i,i-1}\right) =$$

$$= \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^{m_P}\left(p'^2_{i,i-1} - 2p'_{i,i-1}p''_{i,i-1} + p''^2_{i,i-1}\right) =$$

$$= \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^{m_P}\left(p'_{i,i-1} - p''_{i,i-1}\right)^2 < 0$$

because $0 < \alpha < 1$, and $p' \neq p''$, thus for some i, $p'_{i,i-1} - p''_{i,i-1} \neq 0$, and its square is then positive. Therefore h is strictly convex in p^+.

From (B10), noting that $p^*_{i,i-1} = g/m_P$ for $i = 1,\ldots,m_P$ (B6b), we obtain

$$\min h = h^* = \frac{g^2}{2} + \frac{g^2}{2m_P}, \tag{B11}$$

which is the same as (8) in the main text.

For Any p There is a $p' \in p^+$: $g' = g$ And $h' \leq h$

For any non-p^+ distribution, p, and the corresponding g and h, there is a $p' \in p^+$ such that g' and h', the ranking probabilities under p', verify $g' = g$ and $h' \leq h$. To prove this we will first obtain two intermediate results. Loosely speaking result 1 states that when the probability of a pair (m, n) is decreased by some amount, and the probability of (m', n), with $m' < m$, is increased by the same amount, h either decreases or stays the same. Result 2 states that the same will happen when some probability is transferred from (m, n) to (m, n'), with $n' > n$.

Result 1. Consider two probability distributions, p and p', over the same pairs of bids, and their implied h and h'. Let the two distributions differ only in that, for some two pairs of bids (m, n) and $(m-1, n)$ with $1 \leq m \leq m_P$, $p_{mn} > 0$, $p'_{mn} = p_{mn} - d$, $0 < d \leq p_{mn}$, and $p'_{m-1,n} = p_{m-1,n} + d$. There must be such pairs unless g is zero, in which case h is also zero. Then $h' \leq h$ regardless of all other probabilities. To prove this we will compute $h' - h$. Note that $\Pr(b_\$ < i)$ is the same in both distributions, that is,

$$\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} = \sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p'_{kl},$$

since $p'_{mn} + p'_{m-1,n} = p_{mn} + p_{m-1,n}$, and all other probabilities are the same in both distributions. Therefore

$$h' - h = \sum_{i=1}^{m_P}\sum_{j=0}^{m_\$} p'_{ij}\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p'_{kl} - \sum_{i=1}^{m_P}\sum_{j=0}^{m_\$} p_{ij}\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} = \sum_{i=1}^{m_P}\sum_{j=0}^{m_\$}(p'_{ij} - p_{ij})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl}.$$

Let

$$\sum_{i=i_0}^{i_1}(.) = 0$$

whenever $i_0 > i_1$. Then

$$h' - h = \sum_{i=1}^{m-2}\sum_{j=0}^{m_\$}(p'_{ij} - p_{ij})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} + \sum_{i=m-2}^{m}\sum_{j=0}^{m_\$}(p'_{ij} - p_{ij})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} + \sum_{i=m+1}^{m_P}\sum_{j=0}^{m_\$}(p'_{ij} - p_{ij})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} =$$

$$= \sum_{i=m-2}^{m}\sum_{j=0}^{n-1}(p'_{ij} - p_{ij})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} + \sum_{i=m-2}^{m}(p'_{in} - p_{in})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} + \sum_{i=m-2}^{m}\sum_{j=n+1}^{m_\$}(p'_{ij} - p_{ij})\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} =$$

$$= (p'_{m-1,n} - p_{m-1,n})\sum_{k=0}^{m_P}\sum_{l=0}^{m-2} p_{kl} + (p'_{m,n} - p_{m,n})\sum_{k=0}^{m_P}\sum_{l=0}^{m-1} p_{kl} =$$

$$= d\sum_{k=0}^{m_P}\sum_{j=0}^{m-2} p_{kl} - d\sum_{k=0}^{m_P}\sum_{l=0}^{m-1} p_{kl} =$$

$$h' - h = -d\sum_{k=0}^{m_P} p_{k,m-1} \le 0. \qquad \text{(B12)}$$

The above expression is the product of d by the probability that P be valued at $m - 1$. The intuition is simple. When the $ bid is $m - 1$, P is valued above $ if the P bid is extracted from pair (m, n) but not if it is extracted from pair $(m - 1, n)$. Thus when a probability d is transferred from pair (m, n) to pair $(m - 1, n)$, h decreases by $d \times \Pr(b_\$ = m - 1)$. Obviously h will stay the same only if $\Pr(b_\$ = m - 1) = 0$.

Suppose now that $2 \le m \le m_P$, and that a third probability distribution, p'', differs from p' only in that $p''_{m-1,n} = p'_{m-1,n} - d$, and $p''_{m-2,n} = p'_{m-2,n} + d$, $0 < d \le p'_{m-1,n}$. From (B12)

$$h'' - h' = -d\sum_{k=0}^{m_P} p_{k,m-2}$$

and

$$h'' - h = -d\sum_{k=0}^{m_P}\sum_{l=m-2}^{m-1} p_{kl}\ .$$

Generally if two probability distributions, p and p', over the same pairs of bids, differ only in that, for some two pairs (m, n) and (m', n) with $m' < m \le m_P$, $p_{mn} > 0$, $p'_{mn} = p_{mn} - d$, and $p'_{m'n} = p_{m'n} + d$, $0 < d < p_{mn}$, then regardless of all other probabilities

$$h' - h = -d\sum_{k=0}^{m_P}\sum_{l=m'}^{m-1} p_{kl} = -d \times \Pr(m' \le b_\$ \le m-1) \le 0. \tag{B13}$$

Note that $h' = h$ only if the probability that $m' \le b_\$ \le m-1$ is zero.

Result 2. Consider again two probability distributions, p and p', over the same pairs of bids, and their implied h and h'. Let p and p' differ only in that, for some two pairs (m, n) and $(m, n+1)$ with $n < m_P$ and $p_{mn} > 0$, $p'_{mn} = p_{mn} - d$, and $p'_{m,n+1} = p_{mn+1} + d$, $0 < d \le p_{mn}$. Again there must be such pairs unless g and h are both zero. Then $h' \le h$ regardless of the common part of the two distributions. We will again compute $h' - h$.

$$h = \sum_{i=1}^{n}\sum_{j=0}^{m_\$} p_{ij}\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} + \sum_{j=0}^{m_\$} p_{n+1,j}\sum_{k=0}^{m_P}\sum_{l=0}^{n} p_{kl} + \sum_{i=n+2}^{m_P}\sum_{j=0}^{m_\$} p_{ij}\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} =$$

Note that

$$\sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p_{kl} = \sum_{k=0}^{m_P}\sum_{l=0}^{i-1} p'_{kl} \quad \text{for } i = 1,\ldots, n, n+2,..., m_P.$$

If $i \le n$, the only probabilities that differ across distributions, those of pairs (m, n) and (m, n+1), are not included in the sums. If $i \ge n + 2$ both probabilities are included, but $p_{mn} + p_{m,n+1} = p'_{mn} + p'_{m,n+1}$. Also the marginal probability distributions of b_P are the same under both p and p', that is,

$$\sum_{j=0}^{m_\$} p_{ij} = \sum_{j=0}^{m_\$} p'_{ij} \quad \text{for any } i,$$

because either $i \ne m$, and none of the two different probabilities are involved, or $i = m$, and both are.

Therefore

$$\sum_{j=0}^{m_\$} p_{ij} \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p_{kl} = \sum_{j=0}^{m_\$} p'_{ij} \sum_{k=0}^{m_P} \sum_{l=0}^{i-1} p'_{kl} \quad \text{for } i \le n \text{ and } i \ge n+2,$$

and

$$h' - h = \sum_{j=0}^{m_\$} p'_{n+1,j} \sum_{k=0}^{m_P} \sum_{l=0}^{n} p'_{kl} - \sum_{j=0}^{m_\$} p_{n+1,j} \sum_{k=0}^{m_P} \sum_{l=0}^{n} p_{kl} = \sum_{j=0}^{m_\$} p_{n+1,j} \sum_{k=0}^{m_P} \sum_{l=0}^{n} (p'_{kl} - p_{kl})$$

$$= \sum_{j=0}^{m_\$} p_{n+1,j} \left(\sum_{k=0}^{m-1} \sum_{l=0}^{n} (p'_{kl} - p_{kl}) + \sum_{k=m+1}^{m_P} \sum_{l=0}^{n} (p'_{kl} - p_{kl}) + \sum_{l=0}^{n} (p'_{ml} - p_{ml}) \right) =$$

$$= \sum_{j=0}^{m_\$} p_{n+1,j} \sum_{l=0}^{n} (p'_{ml} - p_{ml}) = \sum_{j=0}^{m_\$} p_{n+1,j} (p'_{mn} - p_{mn}) =$$

$$h' - h = -d \sum_{j=0}^{m_\$} p_{n+1,j} = -d \times \Pr(b_P = n+1) \le 0. \qquad \text{(B14)}$$

The interpretation of expression (B14) is similar to that of (B12). When P is valued at n + 1, P is valued above \$ if \$ is valued at n but not if it is valued at n + 1. Thus when a probability d is transferred from (m, n) to (m, n + 1), h decreases by $d \times \Pr(b_P = n + 1)$, that is, it decreases unless $\Pr(b_P = n + 1) = 0$.

This result is generalised as the previous one. If two probability distributions, p and p', over the same pairs of bids differ only in that, for some two pairs (m, n) and

(m, n') with $n \le n'$, $p_{mn} > 0$, $p'_{mn} = p_{mn} - d$, and $p'_{mn'} = p_{mn'} + d$, $0 < d \le p_{mn}$, then regardless of the common part of the two distributions

$$h' - h = -d \sum_{i=n+1}^{n'} \sum_{j=0}^{m_\$} p_{ij} = d \times \Pr\,(n+1 \le b_P \le n') \le 0. \tag{B15}$$

Note that $h' = h$ only if the probability that $n + 1 \le b_P \le n'$ is zero.

Procedures to transform any non-p^+ distribution in a p^+ distribution. Let p be *any* non-p^+ probability distribution over $B_P \times B_\$$. g and h have the usual meanings. Then there exists a probability distribution $p' \in p^+$ such that $g' = g$ and $h' \le h$. If $p \notin p^+$ then there exists at least one O^* pair (m, n) such that $p_{mn} > 0$. Loosely speaking we will find the p' distribution by transferring the probability, if positive, of every O^* pair to pairs that are either P^* or $\$^*$ while preserving the relative ranking of P and $\$$, thus leaving g unchanged. If a pair (m, n), with $p_{mn} > 0$, ranks P above $\$$ and is not a P^* pair, then $n < m - 1$; if it ranks $\$$ above P and is not a $\$$ pair, then $m > 0$ or $n < m_P$ or both. This exhausts all possibilities as we are assuming no indifference.

In the first case, $n < m - 1$, p_{mn} is transferred to the P^* pair $(m, m - 1)$. That is, $p'_{mn} = 0$ and $p'_{m,m-1} = p_{m,m-1} + p_{mn}$. Note that this keeps P ranked above $\$$, and therefore g does not change. From (B15),

$$h' - h = -p_{mn} \sum_{i=n+1}^{m-1} \sum_{j=0}^{m_\$} p_{ij} = -p_{mn} \times \Pr\,(n+1 \le b_P \le m-1) \le 0. \tag{B16}$$

In the second case, $n > m$ and $m > 0$ or $n < m_P$, the probability is transferred first from (m, n) to pair (m, m_P), and then from (m, m_P) to the $\$^*$ pair $(0, m_P)$. This keeps $\$$ ranked above P, and leaves g unchanged. h either decreases or stays unchanged. The first transfer leads, according to (B15), to the following change in h:

$$h'' - h = -p_{mn} \sum_{i=n+1}^{m_P} \sum_{j=0}^{m_\$} p_{ij} = -p_{mn} \times \Pr(n+1 \le b_P \le m_P) \le 0.$$

According to our convention the sum above is null if $n \ge m_P$. The second transfer leads to, according to (B13),

$$h' - h'' = -p_{mn} \sum_{k=0}^{m_P} \sum_{l=0}^{m-1} p_{kl} = -p_{mn} \times \Pr(0 \le b_\$ \le m-1) \le 0.$$

Again this sum is null if $m = 0$. Thus

$$h' - h = h' - h'' + (h'' - h) = -p_{mn} \times [\Pr(n+1 \le b_P \le m_P) + \Pr(0 \le b_\$ \le m-1)] \le 0. \qquad \text{(B17)}$$

After this procedure has been applied to all O^* pairs with positive probability under p the resulting probability distribution, p', is a p^+ distribution, $g' = g$, and $h' \le h$. Now either the resulting p' is p^* or not. If not $h^* < h'$ (that is, $h(p^*) < h(p')$), as we have already shown that p^* is the global, and only, minimum in p^+. If $p' = p^*$, then $h^* = h' < h$. To see why notice that, as shown in condition (B6b) (or (5) in the main text), under p^* the P bid takes all the values from 1 to m_P, each with probability g/m_P, and the \$ bid takes all the values from 0 to $m_P - 1$, each also with probability g/m_P.

Now consider the process of probability transfers from O^* pairs to P^* and $\$^*$ pairs that transformed p into p'. Consider specifically the probability distribution, call it p'', that resulted during that process when there was only one O^* pair with positive probability left. Let that pair be (m, n).

If $m < n$ the change $h' - h''$ brought about by the last probability transfer is given by (B17). $h' - h'' = 0$ only if $\Pr(n+1 \le b_P \le m_P) = 0$, and $\Pr(0 \le b_\$ \le m-1) = 0$. As either $n \le m_P - 1$ or $m > 0$ (or (m, n) would not be a O^* pair), both probabilities are zero only if at least one P^* pair has zero probability, either the pair (1, 0) or the pair $(m_P, m_P - 1)$. The probability of these pairs does not change with the last transfer, as the probability of (m, n) is transferred to $(0, m_P)$ (because $m < n$). Thus after the last probability transfer either $h' > h''$ or the resulting p^+ probability will not be p^*, and $h^* < h'$.

If $m > n$, then $m > n - 1$ (or it would not be a O^* pair) and the change in h is given by (B16). That change is null only if $\Pr(n+1 \le b_P \le m-1) = 0$. This requires that at least one P^* pair, the pair $(m-1, m-2)$ have probability zero. This pair does not receive any probability in the last transfer, the pair $(m, m-1)$ does. Therefore after the last probability transfer either $h' < h''$ or the resulting p^+ distribution will not be p^*, and again $h^* < h'$.

A different p^+ distribution could have been obtained without changing g by transferring the probability of (m, n) to other pairs. That does not invalidate our

argument. The procedures we followed transform any non-p^+ distribution into $p' \in p^+$ without changing g. Then we showed that either $h' < h$ or p' is not p^* and therefore $h^* < h'$.

2. Maximising *h* given g

Rather than maximise $h = \Pr(b_P > b_\$)$ given g we will maximise $\Pr(b_P \geq b_\$)$ also given g. This is a much simpler task, and the difference is negligible (only because we are ruling out indifference). To maximise $\Pr(b_P \geq b_\$)$ given that $\Pr(b_{Pi} > b_{\$i}) = g$ is the same as to minimise $1 - \Pr(b_P \geq b_\$)$, or minimise $\Pr(b_\$ > b_P)$ given that $\Pr(b_{\$i} > b_{Pi}) = 1 - g$. This is the problem we have just solved, but formulated in terms of the probabilities that \$ be ranked above *P*. If in conditions (B6a) to (B6c) p^*_{ij} is interpreted as the probability that the optimal \$ bid be i and the optimal *P* bid be j, contrary to what we have been doing, then p^* will be the distribution that minimises $\Pr(b_\$ > b_P)$ given that $\Pr(b_{\$i} > b_{Pi}) = g$. Condition (B6a) would be valid for $j = m_P$ only, as this is the maximum admissible bid for *P*. If we then swap the indices we will have the same distribution, and the probabilities will regain the usual interpretation. Finally, substituting $1 - g$ for g we will have the probability distribution that minimises $\Pr(b_\$ > b_P)$, or maximises $\Pr(b_p \geq b_\$)$, given that $\Pr(b_{\$i} > b_{Pi}) = 1 - g$ or $\Pr(b_{Pi} > b_{\$i}) = g$, and g too regains the usual interpretation. The resulting probability distribution, to be called $p^\#$, is shown in (B18a) to (B18c).

$$p^\#_{m_P,0} = g, \tag{B18a}$$

$$p^\#_{m-1,m} = \frac{1-g}{m_P} \quad \text{for } m = 1,\ldots, m_P, \tag{B18b}$$

$$p^\#_{mn} = 0 \quad \text{for all other } (m, n). \tag{B18c}$$

The maximum value of $\Pr(b_P \geq b_\$)$ may be computed by substituting the above distribution in the function below, which results from extending expression (B1) to include the probability that $\Pr(b_p = b_\$)$.

$$h^{\#} = \sum_{i=0}^{m_P} \sum_{j=0}^{m_\$} p_{ij}^{\#} \sum_{k=0}^{m_P} \sum_{l=0}^{i} p_{kl}^{\#} .$$

We can compute it in a less cumbersome way though. h^*, given by expression (B11), is also the minimum value of $\Pr(b_\$ > b_P)$ given that $\Pr(b_{\$i} > b_{Pi}) = g$. If $\Pr(b_{\$i} > b_{Pi}) = 1 - g$ (that is, $\Pr(b_{Pi} > b_{\$i}) = g$), then

$$\min \Pr(b_{\$i} > b_{Pi}) = \frac{(1-g)^2}{2} + \frac{(1-g)^2}{2m_P} .$$

The maximum of $\Pr(b_P \geq b_\$)$ given that $\Pr(b_{Pi} > b_{\$i}) = g$ is then

$$h^{\#} = 1 - \frac{(1-g)^2}{2} + \frac{(1-g)^2}{2m_P} = g + \frac{1-g^2}{2} - \frac{1-g}{2m_P} . \qquad \text{(B19)}$$

References

Becker, G. M., M. H. DeGroot, and J. Marshak (1963), "Stochastic Models of Choice Behaviour," *Behavioural Science* 8, 41-55**.**

———, ———, **and** ——— **(1964)**, "Measuring Utility by a Single-Response Sequential Method," *Behavioural Science* 9, 226-32**.**

Braga, Jacinto (2003), *Investigating Preference Reversal with New Methods, Models, and Markets*, mimeo, University of Nottingham.

Braga, Jacinto, and Chris Starmer (2001), *Does Market Experience Eliminate Preference Reversal?* mimeo, University of Nottingham.

Carbone, Enrica and John Hey (2000), "Which Error Story is Best?" *Journal of Risk and Uncertainty* 20, 161-76.

Cox, James C., and David M. Grether (1996), "The Preference Reversal Phenomenon: Response Mode, Markets and Incentives," *Economic Theory* 7, 381-405.

Cubitt, Robin, Alistair Munro and Chris Starmer (forthcoming), *Preference Reversals: An Experimental Investigation of Economic and Psychological Hypotheses,* The Economic Journal.

Grether, David M. and Charles R. Plott (1979), "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review* 69, 623-38.

Harless, David H. and Colin F. Camerer (1994), "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica* 62, 1251-89.

Harrison, Glenn W. (1994), "Expected Utility Theory and the Experimentalists," *Empirical Economics* 19, 223-53.

Hey, John D. and Chris Orme (1994), "Investigating Generalisations of Expected Utility Theory Using Experimental Data," *Econometrica* 62 1291-326.

Holt, Charles (1986), "Preference Reversals and the Independence Axiom," *American Economic Review* 76, 508-15.

Karni, Edi, and Zvi Safra (1987), "'Preference Reversals' and the Observability of Preferences by Experimental Methods," *Econometrica* 55, 675-85.

Lichtenstein, Sarah and Paul Slovic (1971), "Reversals of Preference Between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, 89, 46-55.

Loomes, Graham, Peter Moffatt, and Robert Sugden (2002), "A Microeconometric Test of Alternative Stochastic Theories of Risky Choice," *Journal of risk and Uncertainty* 24, 103-30.

Loomes, Graham and Robert Sugden (1983), "A Rationale for Preference Reversal," *American Economic Review* 73, 428-32.

———, **and** ——— **(1995)**, "Incorporating a stochastic element into decision theories," *European Economic Review* 39, 641-48.

———, **and** ——— **(1998)**, "Testing Different Stochastic Specifications of Risky Choice," *Economica* 65, 581-98.

Quiggin, John (1982), "A Theory of anticipated Utility," *Journal of Economic Behaviour and Organisation*," 3, 323-43.

Sugden, Robert (forthcoming), "Reference-dependent subjective expected utility," *Journal of Economic Theory*.

Tversky, Amos, Paul Slovic, and Daniel Kahneman (1990), "The Causes of Preference Reversal," *American Economic Review* 80, 206-17.