Scientific Research

Explaining Perceived Inconsistencies in "Stated Preference" Valuations of Human Life

Philip Thomas, Geoff Vaughan

School of Mathematics, Computer Science and Engineering, City University London, London, UK
Email: pjt3.michaelmas@gmail.com

Abstract

The Relative Utility Pricing model is used to explain the fact that when faced with two "safety packs", the second giving three times the safety benefit of the first, discriminating respondents will place a value on the second pack that is, on average, twice the amount they say they will be prepared to pay for the first. When the safety packs reduce fatal accident frequencies, the "value of a prevented fatality" (VPF) figures deduced from the valuations of the two safety packs must then be significantly different. Such response patterns on the part of respondents were found in a high-profile study carried out on behalf of a number of UK Government Departments. However, the authors of that study considered the responses "aberrant", and dismissed their survey in favour of their later one, which they based on a novel elicitation technique and which led to a VPF that was lower by a factor of between 5 and 10. That method has been shown elsewhere to be invalid, which returns the focus to the original study rejected by its authors. This paper shows that the VPFs produced by the first study are fully explicable and cannot be dismissed if the stated preference approach is to be accepted. However, in view of the difficulties experienced with stated preference techniques in the valuation of life, it is clear that an urgent reappraisal is needed of revealed preference techniques if people's safety is to be safeguarded adequately.

Keywords

Value of a Prevented Fatality, VPF, Relative Utility Pricing, RUP, Stated Preference, Opinion Survey

1. Introduction

Frequently used by governments to set policy, economic measurements can have a great influence on the way in

which we live our lives. If there is a free market in the good, then the measurement task reduces simply to finding the price on which buyer and seller agree. However, there are and always will be instances where no obvious market exists and yet it is still desirable to know what value people place on the amenity so that a cost-benefit analysis may be carried out. This applies in the case of public goods such as clean air or the continued survival of a rare species of plant or animal. In particular, it applies also to the very important valuation of human life, which is essential if rational choices on safety systems are to be made.

"Revealed preferences" may enable an inferential measurement of the value of the non-market good to be made. Such methods conform to John Locke's precept: "I have always thought the actions of men the best interpreters of their thoughts", with the good's value determined through observing what quantifiable benefit has been given up in order to secure a certain quantity of that good or a proxy, or else, if the good is undesirable, what quantifiable burden has been taken on to avoid it. Clearly, the revealed preference technique requires a model of the system, and, as a minimum, that model needs to be transparent so that it can be examined in full detail by a person possessed of the necessary mathematical skills. The J-value method [1]-[3], based on the Life Quality Index [4]-[7], is an example of a revealed preference technique applied to safety analysis and the valuation of human life that conforms to these transparency requirements.

However, attempts may be made to measure the value of a non-market good using opinion polling or market survey techniques. These "stated preference" methods are sometimes referred to as "contingent valuation" (CV) [8]. In the stated preference approach, members of a representative sample of the population considered to have consultation rights are asked more or less directly to state their maximum acceptable price (MAP) for the non-market good. As with other opinion polls and market surveys, it is then assumed that an opinion that is in some sense representative of an entire population can be deduced fairly accurately from the collections of such statements made by people in a random sample.

This paper will provide a new analysis of the stated preference approach used by a research team commissioned by the UK Health and Safety Executive (HSE), the UK Department of the Environment, Transport and the Regions, the UK Home Office and the UK Treasury to investigate how much people were willing to pay for measures to protect human life [9]. The study was tasked with estimating for the UK Government the "value of a prevented fatality" (VPF), the maximum amount that it is notionally reasonable to pay for a safety measure that will reduce by one the expected number of preventable premature deaths in a large population.

The VPF is assumed to be the same for all people in the UK, irrespective of age or gender, an assumption that has been questioned elsewhere (e.g. [1]-[7] [10]-[12]). But whatever problems exist with its scientific and philosophic status, the VPF figure, published annually by the UK Department for Transport, is important [13]. Updated to account for changes in gross domestic product (GDP) per head, it is used extensively in the UK as a reference not only by Government departments but also by the Health and Safety Executive. Moreover, its adoption by the HSE has led to its widespread use in UK industry, road and rail transport, chemical and nuclear, as a standard for judging how much should be spent on protection measures aimed at reducing risks to life.

However the project team commissioned to carry out the research reported problems with their findings, concluding [9] that:

"One can therefore have little confidence that the VOSL ('value of a statistical life', used synonymously with VPF) estimates that emerged from these studies can be used as a reliable basis for public policymaking."
and labelled their study "Caveat Investigator"—"let the investigator beware". The Caveat Investigator team came to believe that people could not be expected to answer accurately on how much they would spend to reduce further an already low risk of a fatal injury. So, in a follow-on study for the same sponsors, the Caveat Investigator team decided to seek an estimate of the VPF indirectly, by asking people to consider a lesser injury first, with which they considered people would be better able to cope. They then interpreted the new survey data using the "two-injury chained method", which depends on a fairly involved application of utility theory [14]. According to Wolff and Orr [15], the resulting figure has become the standard used in the UK:

"It appears that the Carthy study is now the primary source of VPF figures, adjusted for inflation and changes in GDP."

Moreover, despite the fact that it has been many years since the survey was made, a 2011 report for the Department for Transport, with authors in common with the Carthy study, recommended "against any early new full scale WTP (willingness to pay) study" [16].

However, considerable problems exist with the validity of the VPF coming from the Carthy study. The validation of the chained method is regarded by Wolff and Orr [15] as "severely challenging":

"The method needs to make assumptions about the shape of an average person's utility function, connecting their attitudes to small risks with their attitudes to the much larger risks that are presented to the subject. Yet the testing of any such assumption, and hence the validation of the method, presents severe challenges. In order to test the chained method it appears necessary to elicit individual willingness to pay for very small changes in risk so they can be compared with attitudes to larger changes. However, if it were possible to elicit such preferences then the chained method would not be necessary; it was introduced precisely because it has so far proven impossible to elicit such preferences in any reliable way. In short, if it were possible to validate the chained method, it would not be necessary."

In fact a fair test of the validity of the chained method using no more data than was available at the time to Carthy *et al*. has now been carried out [17]. The two-injury chained method is found there to fail multiple tests of its validity, so that, unfortunately, no reliance can be placed on the team's second attempt at producing a VPF figure. This returns the focus to the team's initial attempt, reported in [9]. Were the authors right to dismiss that survey or were the results a valid set of stated preferences?

The Relative Utility Pricing model [18] [19] will be used in this paper to shed light on the question.

2. The Problems Reported in the Caveat Investigator Study

In what the authors subsequently named the "First Phase" of the Caveat Investigator study (hereafter referred to as "Phase 1" for compactness), a number respondents were asked to say how much they would be prepared to pay for a number of safety improvements that would reduce their annual risks of death and injury in road accidents. The Caveat Investigator team attempted to test the sensitivity of their results to the magnitude of the risk reduction by including two reductions in the frequency of a fatal accident, labelled [F1] and [F3]. [F1] gave a 1 in 100,000 reduction in the annual frequency per person of death in a road accident while [F3] gave a 3 in 100,000 reduction in the annual frequency per person of death on the road.

In Phase 1, 83 respondents were divided into two groups: 41 people were assigned to the Bottom Up (BU) group and 42 to the Top Down (TD) group. The TD number was later reduced effectively to 40 through the censoring of the opinions of two people, for the reasons given in the note to **Table 1** of [9]:

"One outlier has been removed from the TD sample (a respondent in the lowest income group category whose first willingness to pay response was £5,000). In addition, one respondent in the TD group failed to give revised CV responses."

Each person, *i*, in the BU group is asked to provide his maximum acceptable price (MAP), $V_i(\text{F1})$, for [F1] first and his MAP, $V_i(\text{F3})$, for [F3] next, and so in the knowledge that the smaller reduction [F1] is an option. Taking the initial letter from "valuation", the term, $V_i(.)$, uses upper case lettering to denote a random variable, in line with the proposition that the respondents and hence their opinions are selected at random.

Conversely each person in the TD group is invited to name his MAP, $V_i(\text{F3})$, for [F3] first and his MAP, $V_i(\text{F1})$, for [F1] second. Thus he is unaware of the smaller, [F1] option when making his initial valuation of [F3]. However it is stated that respondents

"had also been given every opportunity to revise their responses after having been presented with a summary of their answers to the full sequence of CV questions"

with a respondent failing to give to give revised CV responses being excluded from the survey, as shown in an earlier quotation in this Section. Hence the person will be aware of the smaller option, [F1], when he makes his final valuation of [F3].

But the analysts attempting to find the VPF (or VOSL) from the results detected a non-obvious pattern in the responses associated with safety measures, [F1] and [F3], and became concerned at the ratio of the stated values assigned by the respondents to [F1] and [F3]. While they were prepared to concede on economic grounds that the value, $V_i(\text{F3})$, placed by respondent, *i*, on [F3] would not be as high as three times the value, $V_i(\text{F1})$, he placed on [F1], the analysts were not prepared to accept the very much lower values for $V_i(\text{F3})$ that they found.

The phenomenon of $V_i(\text{F3}) \ll 3V_i(\text{F1})$ is observed in both the TD group and the BU group. (The symbol $\ll$ is used here as meaning "significantly less" as defined in **Table 5** and **Table 6**, which show that there are a good many instances where $V_i(\text{F1}) \le V_i(\text{F3}) \le 2V_i(\text{F1})$.) In fact, something similar had been observed in previous studies:

"This inverse relationship between the VOSL and the magnitude of risk reduction is evident in the results of earlier CV studies.[12]

12. For within-study evidence of the effect, see for example Jones-Lee *et al*. (1985) or Jones-Lee *et al*. (1995). To the best of our knowledge the only attempt to explore the relationship between estimates of the VOSL and the magnitude of the risk reduction in CV questions on a *between-study* basis is that reported in Beattie *et al* (1998) where the inverse relationship between the VOSL (taken as the sample mean MRS of wealth for risk for each study) and the magnitude of the risk reduction, δr, is found to be highly significant with

$\log_{10} \text{VOSL} = 4.14 - 0.52\log_{10}\delta r$."

Nevertheless not only the project team but also their sponsors became concerned at the phenomenon observed and

"It was felt by both the research team and the project sponsors that the type of question used in the First Phase study could not be employed with any confidence as a basis for obtaining direct empirical estimates of WTP-based values of safety for use in public policymaking."

As a result a "Second Phase" of the Caveat Investigator study was initiated, hereafter known as "Phase 2". Whereas Phase 1 was concerned with non-fatal injuries as well as deaths, Phase 2 was to be a smaller study, dealing only with fatalities.

In Phase 2 all respondents are asked first to imagine that there is a program of road safety improvements affecting the area where they live, assumed to contain a population of 1 million people living in 400,000 households. 52 respondents, divided into two equal groups, given the identifiers, Version A (VA) and Version F (VF), are asked first to place a value to their household on an improvement [R3] that will prevent a number of deaths on the roads over a specified period in the area in which they live. Then they are asked to value another safety improvement [R1] that will prevent a third of that number.

For VA respondents, [R3] is a reduction of 15 road deaths over 1 year, while [R1] is a reduction of 5 road deaths over 1 year. Meanwhile, for VF respondents, [R3] is a reduction of 75 road deaths over 5 years, while [R1] is a reduction of 25 road deaths over 5 years. The valuations sought are first for [R3] and then for [R1], in both VA and VF, but once again the respondents are given the opportunity to revise their initial responses. Moreover, they are also asked the question:

"In the past, we've found that some people say that preventing 15/75 deaths on the roads are worth three times as much to them as preventing 5/25 deaths on the roads: but other people don't give this answer. Can you say a bit about why you gave the answers you did?"

This question was intended to carry out the analysts' wish

"to focus respondents' attention on the relative magnitudes of the two reductions in road fatalities."
and might, indeed, be regarded as giving a rather clear steer on the analysts' concerns.

But once again it is found that respondents in both VA and VF groups tend to give valuations $V_i(\text{R3}) \ll 3V_i(\text{R1})$ in the same way as $V_i(\text{F3}) \ll 3V_i(\text{F1})$ characterised the responses of the two groups, BU and TD, in Phase 1 of the research.

Following the calculation of the corresponding VPFs in the manner described in Section 11 of this paper, a further concern is raised, namely that each of the VPFs generated (£9.83 M and £4.63 M from Phase 1, £8.70 M and £3.87 M from Phase 2) is much higher than the VPF value then in use (£902,500 in 1997 prices [14]):

"Furthermore, when the VOSL estimates that they had collectively generated were fed back to respondents in the course of the third-stage follow-up discussions, it was widely felt that the figures concerned were excessive.[21]

21. In order to put the VOSL figures into perspective, respondents were given some idea of, for example, the approximate number of secondary school teacher-years or hip replacement operations that could be financed by the sums concerned."

The extent and scope implied by the phrase, "it was widely felt", is not clarified either in the main text or the footnote.

Ultimately the Caveat Investigator team became so concerned over the "aberrant response patterns" that they found, specifically that $V_i(x3) \ll 3V_i(x1)$, $x = \text{F,R}$ for many individuals, that they rejected all the results of their study and decided to start again. But are the response patterns, some signs of which are acknowledged to have been seen before, "aberrant", or are they the sort of thing one should expect when asking people to state their valuations of two "safety packs", with pack 2 being three times the size of pack 1?

3. Results of the Caveat Investigator Study

The safety measures offered and results of the Caveat Investigator study are summarized in **Tables 1-6**.

Table 1. Reductions in annual fatality frequency for Phase 1.

	Reduction in annual frequency of fatality
[F1]	1 in 100,000
[F3]	3 in 100,000

Table 2. Reductions in deaths and annual frequency of fatality for Phase 2.

		Reduction in deaths over 1 year	Reduction in deaths over 5 years	Reduction in annual frequency of fatality
VA	[R1]	5	Not stated	0.5 in 100,000
	[R3]	15	Not stated	1.5 in 100,000
VF	[R1]	Not stated	25	0.5 in 100,000
	[R3]	Not stated	75	1.5 in 100,000

Table 3. Phase 1 mean MAPs, standard deviations (s.d.), standard errors (s.e.) and medians.

	TD $(n = 40)$			BU $(n = 41)$			Total $(n = 81)$		
	Mean (£)	s.d. (s.e.) (£)	Median (£)	Mean (£)	s.d. (s.e.) (£)	Median (£)	Mean (£)	s.d. (s.e.) (£)	Median (£)
$V(F1)$	111.05	182.40 (28.84)	50	85.78	87.85 (13.72)	50	98.26	142.20 (15.80)	50
$V(F3)$	171.33	223.83 (35.39)	100	107.27	112.57 (17.58)	67.5	138.90	178.29 (19.81)	85

Table 4. Phase 2 mean MAPs, standard deviations (s.d.), standard errors (s.e.) and medians.

	VA $(n = 26)$			VF $(n = 26)$			Total $(n = 52)$		
	Mean (£)	s.d. (s.e.) (£)	Median (£)	Mean (£)	s.d. (s.e.) (£)	Median (£)	Mean (£)	s.d. (s.e.) (£)	Median (£)
$V(R1)$	79.30	131.30 (25.75)	30	138.33	166.74 (32.70)	90	108.80	151.58 (21.02)	55
$V(R3)$	95.83	131.30 (25.75)	45	196.14	242.41 (47.54)	115	145.14	199.60 (27.68)	77.5

Table 5. Individual MAP ratios for Phase 1, $R_{Fi} = V_i(F3)/V_i(F1)$.

	TD	BU	Total
R_{Fi} incomputable $V(R3) = V(R1) = 0$	6[a]	2[a]	8
R_{Fi} incomputable $V(R1) = 0;\ V(R3) > 0$	3[a]	3[a]	6
$R_{Fi} = 1$	8	15	23
$1 < R_{Fi} \le 2$	14	17	31
$2 < R_{Fi} \le 3$	3	2	5
$R_{Fi} > 3$	6	2	8
Total	40	41	81

[a]9 ratios, R_{Fi}, were listed as "Missing" under TD as well as 5 under BU in **Table 2** of Beattie *et al*. Meanwhile Note 11 of Beattie *et al*. reads: "In addition, 8 respondents gave a zero willingness to pay for both [F1] and [F3] while a further 6 gave a zero response for [F1] but a non-zero response for [F3]. Clearly CV[F3]/CV[F1] ratios cannot be computed for any of these 14 cases." In the absence of more precise information, these 6 respondents have been distributed equally in **Table 5** between TD and BU.

Table 6. Individual MAP ratios for phase 2, $R_{Ri} = V_i(\text{R}3)/V_i(\text{R}1)$.

	VA	VF	Total
$R_{Fi} = 1$	13	9	22
$1 < R_{Fi} \leq 2$	6	13	19
$2 < R_{Fi} \leq 3$	6	3	9
$R_{Fi} > 3$	1	1	2
Total	26	26	52

Table 1 and **Table 2** give descriptions of the reductions in fatal accident frequency and fatal accidents over a period for Phase 1 and Phase 2 that the respondents are asked to value.

Table 3 and **Table 4** give the mean valuations of such reductions. *Note:* the practice used by the Caveat Investigator team has been retained, whereby the monetary values placed on the 5-year safety programmes of VF have been divided by 5 to give annualized amounts.

Table 5 and **Table 6** classify the ratios $R_{Fi} = V_i(\text{F}3)/V_i(\text{F}1)$ and $R_{Ri} = V_i(\text{R}3)/V_i(\text{R}1)$. The ratios are categorized as one of: incomputable, equal to 1, between 1 and 2, between 2 and 3 or above 3.

This paper will now explain these findings and show that what the Caveat Investigator team describes as "aberrant response patterns" are, in fact, fully understandable outputs from an opinion survey aimed at finding the value to the public of safety improvements.

4. Safety Improvements as Commodities

The good offered in Phase 1 of the Caveat Investigator study is described as follows:

"a safety feature fitted to their car which would reduce their own risk of death or injury but would leave the risk to other occupants of the car unchanged. Respondents were also told that if it was to remain effective, the safety feature would have to be renewed every twelve months."

with its merit described simply by a number "1 in 100,000 reduction of annual risk of death" for [F1] and "3 in 100,000 reduction of annual risk of death" for [F3].

The description is bare and factual with no attempt at product differentiation. No allusions are made to brand or to any potential status uplift from having the safety device fitted to the respondent's car. It should be borne in mind that such promotional tactics might well be used by a commercial firm trying to sell its safety wares to a consumer. Thus a commercial vendor might, as a minimum, offer to attach a windscreen sticker to the car to show the enhanced safety performance that the driver has managed to procure.

It may be deduced from the fact that the investigators did not mention any additional benefits or kudos associated with the device that they intended not to cloud the issue in any way, but wanted to convey the safety content of the device as its only benefit. This absence of any attempt at product differentiation makes it reasonable to suppose that the respondents to whom the good was offered would have regarded it as a commodity. The good offered in the Phase 2 of the Caveat Investigator study is:

"a program of road safety improvements affecting the area in which they live, with a population of 1 million people, or approximately 400,000 households."

again with its merit described simply by a number, under VA, the number of deaths averted in one year is 5 deaths for [R1] and 15 deaths for [R3], while under VF, the number of deaths averted in five years is 25 deaths for [R1] and 75 deaths for [R3]. The safety improvement offered is thus a public good with no attempt at product differentiation (difficult in any case for a public good). Hence, for Phase 2 of the Caveat Investigator study too, the good offered to the respondents may be regarded as a commodity.

5. Relative Utility Pricing (RUP) Model for Quantity Discounts

The RUP model, [18]-[20], applies to commodities offered in packs of different sizes and is able to provide an economic explanation for sales promotions such as "buy-one-get-one-free" (BOGOF) and "3 for the price of 2". While the RUP model is not meant to apply to differentiated goods, it has been shown in the previous section

that the hypothetical safety goods being offered to the respondents in the Caveat Investigator surveys are not differentiated goods but commodities offered in different sizes. Using the terminology of pack sizes, the safety improvement of either [F1] and [R1] may be called "safety pack 1", while safety improvements [F3] and [R3] may be termed "safety pack 2". Because the safety packs are offered as commodities, the RUP model should apply to the Caveat Investigator study.

The relevant Equations for the RUP model are now presented for a commodity, B, that is sold simultaneously in packs of different sizes. Let b_j be the amount of commodity, B, contained in pack j: $0 \leq j \leq N$, where $b_j > b_{j-1}$ so that pack N is the largest pack on offer and pack 1 is the entry-level pack, while pack 0 represents the empty pack, equivalent to no purchase: $b_0 = 0$. The "largeness", $L(j)$, of pack j is defined as its size relative to the entry-level pack 1:

$$L(j) = b_j / b_1 \quad \text{for} \ \ 0 \leq j \leq N \tag{1}$$

Thus the largeness of pack j is the number of standard units that pack j contains, where a standard unit comprises the contents of pack 1, namely b_1. Hence $L(0) = 0$ and $L(1) = 1$.

In line with Kahneman and Tversky [21], who have argued that "people normally perceive outcomes as gains and losses [relative to the current asset position], rather than as final states of wealth or welfare", a proposition supported by Thaler [22], the approach is taken in [18] of considering the utility of the contents of the pack. The utility of the empty pack, with contents $b_0 = 0$ will be $u(b_0)$, the utility of the contents, b_1, of pack 1 will be $u(b_1)$, and, in general, the purchase of pack j will result in a utility, $u(b_j)$. Hence the gain in utility resulting from the purchase of pack 1 rather than pack 0 (equivalent to making no purchase) is

$$\Delta u_{10} = u(b_1) - u(b_0) \tag{2}$$

Meanwhile the marginal utility gained, Δu_{nk}, from the purchase of a pack n rather than a smaller pack k is

$$\Delta u_{nk} = u(b_n) - u(b_k) \quad n > k \tag{3}$$

Thus

$$\Delta u_{n0} = u(b_n) - u(b_0) = u(b_n) - u(b_k) + u(b_k) - u(b_0) = \Delta u_{nk} + \Delta u_{k0} \tag{4}$$

As explained in [18], the form of these utility differences, Δu_{hi}, $h > i$, is constrained by the requirement that

$$\Delta u_{n0} = L(n) \Delta u_{10} \quad n > 1 \tag{5}$$

where Equation (5) expresses the condition of "singular homogeneity", as defined and discussed in Appendix A of that paper. Rearrangement of Equation (4) gives:

$$\Delta u_{nk} = \Delta u_{n0} - \Delta u_{k0} \tag{6}$$

and so, after rearrangement and the use of Equation (5):

$$\frac{\Delta u_{nk}}{\Delta u_{k0}} = \frac{\Delta u_{n0}}{\Delta u_{k0}} - 1 = \frac{\Delta u_{n0}}{\Delta u_{10}} \frac{\Delta u_{10}}{\Delta u_{k0}} - 1 = \frac{L(n)}{L(k)} - 1 \tag{7}$$

Considering the case of two packs, the standard economic assumption of a constant price per unit of marginal utility gained implies that $p_{21}/\Delta u_{21} = p_{10}/\Delta u_{10}$ where p_{10} is the price associated with the marginal utility, Δu_{10}, gained by the purchase of pack 1, given that the purchase of pack 0 is possible, while p_{21} is the price associated with the marginal utility, Δu_{21}, gained by the purchase of pack 2, given that the purchase of pack 1 is possible. Hence p_{21} is the price achieved for pack 2 when the option exists to purchase pack 1 for price, p_{10}.

Therefore

$$\frac{p_{21}}{p_{10}} = \frac{\Delta u_{21}}{\Delta u_{10}} = \max\left(1.0, \frac{L(2)}{L(1)} - 1\right) = \max(1.0, L(2) - 1) \tag{8}$$

where the maximum function embodies the general economic assumption that, all other things being equal, less

will not be preferred to more, while the largeness of the entry pack 1 is unity by Equation (1).

Equation (8) expresses a basic consumer preference that underlies the pattern of prices observed for different sizes of packs of diverse goods offered for sale in supermarkets and on the internet [18] [19]. It provides the rationale for the finding that "more than 80% of all promotional activity within supermarkets is a bogof [buy-one-get-one-free] or three-for-two" [23], as can be seen by setting the largeness ratio in Equation (8) to first 2 and then 3.

The ratio of pack sizes, b_2/b_1, is clearly equal to 3 in the case of the safety packs offered in the Caveat Investigator study:

$$\frac{b_2}{b_1} = \frac{3 \times 10^{-5}}{1 \times 10^{-5}} = 3 \quad \text{for Phase 1, both TD and BU}$$

$$\frac{b_2}{b_1} = \frac{15}{5} = 3 \quad \text{for Phase 2, VA} \tag{9}$$

$$\frac{b_2}{b_1} = \frac{75}{25} = 3 \quad \text{for Phase 2, VF}$$

Now $L(2) = 3$ when pack 2 is three times as big as pack 1, and substituting this into Equation (8) gives the ratio of the prices the consumer will find acceptable:

$$\frac{p_{21}}{p_{10}} = 2 \tag{10}$$

In the commercial world, of course, a bargain needs to be struck between vendor and buyer. Thus the valuation of the purchaser may be disputed and resisted by the vendor, who will be prepared to offer quantity discounts only if his variable costs are low enough for each sale to make a non-negative contribution to covering his overall costs. Economic theory suggests that the 3-for-the-price-of-2 sales promotions on the high street need to be consistent with the retailer maximizing his profit, which implies an optimization exercise to find the offered price, as proposed in [18] [19]. But the optimal price is the same as the mean of the MAPs in the target market when the probability distribution for MAP is uniform. Moreover the optimal price and the mean of the MAPs will always be very similar, no matter what probability distribution for MAP is assumed [24] [25]. Hence when a 3-for-the-price-of-2 sales promotion is offered, the mean ratio of the individual's MAPs for packs 1 and 2 will be roughly two.

This suggests that the results described as "aberrant" by the Caveat Investigator team are an instance of 3-for-the-price-of-2 pricing observed frequently on the high street, which, as illustrated in Section 5 and in [18] and [19], is driven by consumer preferences. Given the relative size of safety packs 1 and 2, RUP theory suggests that MAP ratios, $R_{xi} = V_i(x3)/V_i(x1)$, $x = F, R$, can be expected to be significantly less than 3. However the situation is complicated by the fact that the respondents' powers of discrimination are being taxed by the small size of the safety packs on offer.

6. The Limits of Discrimination

Beattie *et al*. recognized that people face problems of discrimination when the risk is already so small that they can distinguish no benefit or utility from reducing it further [9]. Thus the Caveat Investigator team decided before the start of Phase 1 that they would direct their attention to:

"those contexts (such as the roads) in which it was felt that base risks were 'large' enough to render CV questions concerning reductions in risk comprehensible to respondents. By contrast, in other contexts (such as rail or nuclear power generation) it would appear that base risks are so low that CV questions concerning reductions in those risks would involve probabilities that most members of the public would find great difficulty in conceptualising."

Moreover, they concluded after Phase 1 that

"it appeared that a major contributory factor in the relative failure of the direct CV questions was the difficulty experienced by many respondents in dealing with small reductions in already small probabilities. For example, some respondents seemed to regard one probability reduction of 3 in 100,000 and another of 1 in 100,000 as being essentially the same, since both constitute such small absolute numbers."

However, by selecting a population of 1 million people in Phase 2 and then casting [R3] and [R1] as reductions in road deaths of 15 and 5 per year respectively, the Caveat Investigator team are compelling the new groups of respondents to deal with even smaller reductions in annual frequency per person of fatality: 1.5 in 100,000 and 0.5 in 100,000. See **Table 1** and **Table 2**. Hardly surprisingly, a good many of the respondents in Phase 2 are pushed beyond their limit of discrimination:

"Far from attenuating the pronounced embedding effects encountered in the First Phase study, the Second Phase CV question format, if anything, appears to have exacerbated these effects, with 42% of respondents giving identical non-zero CV responses for [R3] and [R1], compared with the 28% of First Phase respondents who gave identical non-zero CV responses for [F3] and [F1]."

In fact the Caveat Investigator team appears to be half expecting this outcome:

"Previous evidence indicates that presenting risk information in this way *in itself* does not eliminate insensitivity to quantity" (italics are those of [9])

but still maintains the hope that

"when used within the three-stage, in-depth study design outlined above, this alternative approach would increase the sensitivity of responses."

In the event, it is clear that the team's hope is not realized. One might speculate that the in-depth nature of the respondents' participation could even have been a causal factor, since the respondents would have had the opportunity to set the raw numbers of deaths prevented in the context of a population of 1,000,000 people. For even a moment's thought on the part of the respondent will transform the 5 deaths averted into a frequency of 5 in a million per person. Similarly, 15 deaths averted is easily translated into 15 in a million per person.

If pushed beyond his limit of discrimination, respondent *i* may suffer from one of 3 types of indiscrimination, in ascending order of severity:

Type 1, where the same non-zero value is placed on the utility of safety packs 1 and 2, as indicated by $V_i(x1) = V_i(x3) > 0$; $x = F, R$.

Type 2, where the utility of safety pack 1 is put at zero by setting $V_i(x1) = 0$ $x = F, R$ but the utility of safety pack 2 is considered greater than zero, so that $V_i(x3) > 0$; $x = F, R$.

Type 3, where the value placed on the utility of both pack 1 and pack 2 is zero: $V_i(x1) = V_i(x3) = 0$; $x = F, R$.

A fourth type, Type 4, is theoretically possible, where the utility of safety pack 1 is valued positively so that $V_i(x1) > 0$, $x = F, R$ but safety pack 2 is valued at zero by putting $V_i(x3) = 0$, $x = F, R$, but this goes against the normal economic principle that more is preferred to less, and does not feature in the results reported for the Caveat Investigator study. Indiscrimination Types 1, 2 and 3 are, however, all present in Phase 1 of the Caveat Investigator study, while Type 1 is present in Phase 2.

From **Table 5**, 23 people in the TD and BU groups combined suffer from Type 1 indiscrimination, for each of whom the ratio $R_{Fi} = V_i(F3)/V_i(F1)$ is unity. Moreover, 6 respondents suffer from Type 2 indiscrimination and 8 respondents are exhibiting type 3. This makes a total of 37 out of 81 respondents for which data are available who are suffering from a lack of discrimination of some type.

There is, unfortunately not enough information in the Caveat Investigator paper to distribute Types 2 and 3 indiscrimination over the two groups TD and BU with certainty. In its Table 2, the Caveat Investigator team lists 9 ratios, R_{Fi}, under TD as "missing" and 5 under BU. Meanwhile Note 11 of the study reads:

"In addition, 8 respondents gave a zero willingness to pay for both [F1] and [F3] while a further 6 gave a zero response for [F1] but a non-zero response for [F3]. Clearly CV[F3]/CV[F1] ratios cannot be computed for any of these 14 cases."

The 14 incomputable cases translate into the 14 "missing" cases, but no split between TD and BU is given. In the absence of more precise information, and to give indicative figures, it is assumed in **Table 5** that the 6 respondents with Type 2 indiscrimination are distributed equally between TD and BU. This assumption is for illustrative purposes only and has no effect on the calculations or conclusions that follow.

Meanwhile for Phase 2, 22 out of 52 respondents exhibit Type 1 indiscrimination, for whom $R_{Ri} = V_i(R3)/V_i(R1) = 1$. 13 of these are in VA while 9 are in VF. See **Table 6**.

The RUP model presumes that the person assigning a price is able to distinguish the utility of pack 1 from the utility of the empty pack, *viz.* $\Delta u_{10} > 0$. Moreover the RUP model assumes that the person can discriminate between the utilities of pack 2 and pack 1, so that for a present pack-size ratio of 3, as in the present instance, C,

which will be reflected in different prices for the two safety packs. Thus the RUP model should be applied only to those who do not suffer from indiscrimination of any type, and hence to the discriminating 44 out of the 81 respondents in Phase 1 and to the discriminating 30 out of the 52 respondents in Phase 2.

7. Piecewise Uniform Distribution for Discriminating Respondents; Extended Piecewise Uniform Distribution for All Computable MAP Ratios

It is reasonable to suppose that different discriminating respondents will give somewhat different MAP ratios, $R_{xi} = V_i(x3)/V_i(x1)$, $x = F, R$, and the question arises as to how to consolidate these different values into a single, representative figure. If the democratic principle is to be applied so that each person's view is to be accorded the same respect, then the sample mean should be used, as it is the only known measure that satisfies the requirements of Structural View Independence (Thomas, 2014) and can thus absolve the analyst of the charge of bias.

In fact [9] does not list the actual values of the individual MAP ratios, so that reliance must be placed upon the classified data as presented in **Table 5** and **Table 6**. Nevertheless it is possible, making only mild assumptions, to estimate the expected value of the MAP ratio for the discriminating respondents, as well as its variance. Under the presumption that the Caveat Investigator team would have flagged any R_{Fi} or R_{Ri} ratios greater than 4.0 as noteworthy but does not mention any instances, it is assumed first that no respondent has an R_{Fi} or R_{Ri} ratio greater than 4.0. Hence the classifications, $R_{Fi} > 3$ in **Table 5** and $R_{Ri} > 3$ in **Table 6** may be replaced by $3 < R_{Fi} \leq 4$ and $3 < R_{Ri} \leq 4$ respectively. It is now possible to fit a Piecewise Uniform distribution for R_{xi}, $x = F, R$, within each of the intervals, 1 - 2, 2 - 3 and 3 - 4.

Let the number of MAP ratios recorded in a given interval be n_q, where $q = 1$ denotes the interval between 1 and 2, $q = 2$ denotes the interval between 2 and 3 and $q = 3$ denotes the interval between 3 and 4. The response of a discriminating respondent will be captured in one of the three intervals, and, assigning n to the total number of discriminating respondents, it is clear that $n = \sum_{q=1}^{3} n_q$.

The ratios, R_{xi}, may be reclassified by interval, so that there are n_q ratios in the range q to $q+1$

$$q < R_{x,qk} \leq q+1 \quad \begin{array}{c} k = 1, 2, \cdots, n_q \\ q = 1, 2, 3 \\ x = F, R \end{array} \tag{11}$$

Since the distribution is uniform over each interval, the expected value for any observation in the interval is the mid-point:

$$E(R_{x,qk}) = \frac{q + (q+1)}{2} = q + 0.5 \quad \begin{array}{c} k = 1, 2, \cdots, n_q \\ q = 1, 2, 3 \\ x = F, R \end{array} \tag{12}$$

Meanwhile the variance is given by (see e.g. [26])

$$\text{var}(R_{x,qk}) = \frac{((q+1)-q)^2}{12} = \frac{1}{12} \quad \begin{array}{c} k = 1, 2, \cdots, n_q \\ q = 1, 2, 3 \\ x = F, R \end{array} \tag{13}$$

These theoretical values are the true mean and the true variance, given that the distribution within each interval is Uniform. This allows the methods of Appendix B to be used to find the expected value and the variance of the combined distribution that constitutes the Piecewise Uniform Distribution.

While the Piecewise Uniform Distribution applies to those respondents who give MAP ratios greater than unity, it is possible to add in those with Type 1 indiscrimination, whose MAP ratio is 1.0. The methods of Appendix B may be used to include this deterministic component, for which $\mu = 1.0$ and $\sigma = 0.0$, to produce the Extended Piecewise Uniform Distribution, which covers all computable MAP ratios.

8. Expected MAP Ratios for Discriminating Respondents Based on the Piecewise Uniform Distribution

The calculations outlined in Section 7 have been applied to each of the cohorts and the combined cohort for both

Phase 1 and Phase 2. **Table 7** shows the expected values for $E(R_F)$ and $E(R_R)$, while **Table 8** shows the standard errors and standard deviations. **Table 9** contains the 90% confidence intervals, and shows that the RUP-predicted value of 2.0 lies within this interval for each of the cohorts, TD and BU for Phase 1 and VA and VF for Phase 2, as well as for both combined cohorts. The means of MAP ratios for the discriminating respondents lie close to the theoretical value of 2.0 for each of the two cohorts in each of Phases 1 and 2, with the greatest disparity, just over 10%, occurring for the BU cohort in Phase 1 and the VF cohort in Phase 2. The number of respondents in these cases is, of course, low, at 21 and 17 respectively.

The mean for all the discriminating respondents in Phase 1 is 1.98, while the mean for all the discriminating respondents in Phase 2 is 1.93. The 90% confidence interval is 1.77 - 2.18 for Phase 1 and 1.73 - 2.14 for Phase 2. It will be seen that these results conform closely to the prediction of Relative Utility Pricing theory that the MAP ratio will be 2.0.

9. Modeling the MAP Ratio Using a Truncated Lognormal Distribution for All Computable MAP Ratios

As an alternative to the Extended Piecewise Uniform distribution, it is possible also to model all computable MAP ratios using a lognormal distribution, truncated so that the cumulative probabilities below 1.0 and above 4.0 are converted into point values. The probability that the MAP ratio is unity is then the cumulative probability of a MAP ratio up to and including unity, with a similar procedure applied to MAP ratios of 4 and above. When applying the Truncated Lognormal distribution to the combined cohorts of Phase 1 and Phase 2, it is clearly necessary to work with computable MAP ratios. This poses no problems for Phase 2, where the MAP ratio is

Table 7. Expected MAP ratios for discriminating respondents in the combined and split cohorts for Phase 1 and Phase 2. Piecewise Uniform distribution with numbers in each interval equal to those recorded.

	TD $(n = 23)$	BU $(n = 21)$	Total $(n = 44)$
Phase 1			
	2.15	1.79	1.98
	VA $(n = 13)$	VF $(n = 17)$	Total $(n = 30)$
Phase 2			
	2.12	1.79	1.93

Table 8. Standard errors (s.e) and standard deviations (s.d.) for the combined and split cohorts for Phase 1 and Phase 2, discriminating respondents.

Phase 1	TD $(n = 23)$	BU $(n = 21)$	Total $(n = 44)$
s.e.	0.19	0.15	0.13
s.d.	0.91	0.69	0.83
Phase 2	VA $(n = 13)$	VB $(n = 17)$	Total $(n = 30)$
s.e.	0.19	0.16	0.12
s.d.	0.69	0.64	0.68

Table 9. 90% confidence intervals for the combined and split cohorts for Phase 1 and Phase 2, discriminating respondents.

	TD $(n = 23)$	BU $(n = 21)$	Total $(n = 44)$
Phase 1			
	1.84 - 2.47	1.54 - 2.03	1.77 - 2.18
	VA $(n = 13)$	VB $(n = 17)$	Total $(n = 30)$
Phase 2			
	1.80 - 2.43	1.54 - 2.05	1.73 - 2.14

computable for all 52 respondents. But in the case of Phase 1, the requirement for computability of MAP ratios means limiting the matching and comparison process to the 67 computable MAP ratios coming from the 81 respondents.

The mean and the standard deviation may be found by minimizing the chi-squared statistic for the Truncated Lognormal distribution. Since the data for Phase 1 are given in 4 cells, one degree of freedom is left for evaluating the p-value. The optimal values are found as $E(R_F) = 1.569$ and $\sigma_{RF} = 1.132$. See **Table 10**. The associated p-value is 0.095, a figure found after applying the continuity correction of 0.5 necessary when there is just one degree of freedom [27]. Thus when applied to Phase 1, the Truncated Lognormal distribution for R_F satisfies comfortably the requirement conventionally required for goodness of fit, namely that the probability of the observed mismatch between the model-predicted number of ratios shown in **Table 10** and the recorded numbers should be no less than 0.05.

In the case of Phase 2, the smaller number of respondents means that ensuring that the predicted number in each cell is at least 5 requires the final two cells to be combined, leaving just 3 cells against which to make comparisons. This number is insufficient to provide any degrees of freedom for a chi-squared test when an optimal matching exercise is to be used to determine the two distribution parameters. However adopting the Truncated Lognormal distribution from Phase 1 allows its predictions to be tested against Phase 2 data, since there are now two degrees of freedom left over. The resultant *p*-value is then found to be 0.707. This suggests that modeling Phase 2 MAP ratios using the distribution found from Phase 1 is quite satisfactory. See **Table 11**.

The cut-offs at either end of the Truncated Lognormal distribution render difficult if not impossible the task of finding analytically the expected value, $E(R)$, of the MAP ratio for the 44 discriminating respondents of Phase 1. But no such difficulty arises if simulation is used to estimate such a value. Singling out the MAP ratios of those with $R_{Fi} > 1$ in the simulation produces average values for discriminating respondents of:

$$E(R_F) = 1.98 \tag{14}$$

with a standard error of estimation of 0.009 and

Table 10. Fitting a truncated lognormal distribution to the 67 computable ratios, R_{Fi}, from Phase 1: $E(R_F) = 1.569$, $\sigma_{R_F} = 1.132$.

	Actual	Model
$R_{Fi} = 1$	23	23.79
$1 < R_{Fi} \le 2$	31	26.98
$2 < R_{Fi} \le 3$	5	10.03
$R_{Fi} > 3$	8	6.20
	67	67

p-value = 0.095.

Table 11. Testing the truncated lognormal distribution with $E(R_R) = 1.569$, $\sigma_{R_R} = 1.132$ against the 52 ratios, R_{Ri}, from Phase 2.

	Actual	Model
$R_{Ri} = 1$	22	18.47
$1 < R_{Ri} \le 2$	19	20.94
$R_{Ri} > 2$	11	12.60
	52	52

p-value = 0.707.

$$\sigma_{R_F} = 0.70 \tag{15}$$

with a standard error of estimation of 0.01, based on 200 simulations of cohorts of 67. The average number of discriminating respondents is 43.2, which may be used to estimate that the standard error associated with the average value of R_F in such a population is about 0.11.

The figures contained in Equations (14) and (15) are similar to those for the discriminating Phase 1 respondents found when a Piecewise Uniform distribution is used (see **Table 7** and **Table 8**). Likewise, the mean and standard deviation resulting for MAP ratios based on the Truncated Lognormal distribution are close to the parameters found for Phase 2 using the Piecewise Uniform distribution for MAP ratio.

The good match provided for the MAP ratios for Phases 1 and 2 suggests that the Truncated Lognormal distribution should be applicable to any group of individuals asked to value safety pack 1 and a safety pack 2 that is three times bigger when those individuals are being pushed close to their limits of discrimination. The Truncated Lognormal distribution may, indeed, be regarded as a more general model for MAP ratio than the Extended Piecewise Uniform distribution, since it avoids the need to specify the numbers of respondents in each cell. The verification by chi-squared tests that the same Truncated Lognormal distribution is able to represent both Phase 1 and Phase 2 MAP ratios provides further corroboration for the claim of increased generality.

The implication is that the precise distribution of the MAP ratios within each MAP-ratio interval is not important. Specifically, it is not necessary that the distributions should be uniform within each interval, 1 - 2, 2 - 3 and 3 - 4, as assumed in the Piecewise Uniform model. The confirmation of the prediction, based on the RUP model, that the mean of the MAP ratio will be about 2.0 is shown to be insensitive to the precise distribution of MAP ratios within each interval.

10. Validation of the Distributions for MAP Ratio by Embedding Them within a Model for the Complete Opinion Survey

A major objective of the paper has now been achieved, namely a demonstration that the MAP ratios for safety packs 1 and 2, which the authors of the Caveat Investigator study regarded as "aberrant", are fully explicable when it is realized that (i) many respondents are being asked to make judgments beyond their limits for discrimination and (ii) the RUP model will apply to the discriminating respondents.

This section shows that, even though only summary data are available on the respondents' valuations of safety packs 1 and 2, it is still possible to produce useful simulations of both Phase 1 and Phase 2 opinion surveys. This is because the observed valuations, $V(x1)$ and $V(x3)$, $x = F, R$, all obey an approximately lognormal distribution. This is revealed by the close correspondence between the recorded median and the median found from a lognormal distribution in which the recorded mean and the recorded standard deviation are used as parameters. See **Table 12**.

Appendix C details the way in which the Extended Piecewise Uniform distribution can be embedded in a simulation that takes the individuals' valuations of safety pack 1, $V_i(x1)$, $x = F, R$, as inputs and multiplies these by the individuals' MAP ratios, R_{xi}, $x = F, R$, to produce valuations of safety pack 2, $V_i(x3)$, $x = F, R$, as outputs. The responses of the 14 respondents in Phase 1 who suffer from Type 2 or Type 3 indiscrimination need to be removed because their MAP ratio is not computable, and this may be done using the method explained in Appendix A. Conditioning the Phase 1 valuations, $V(F1)$ and $V(F3)$, in this way results in the mean and standard deviations shown in **Table 13** for all computable ratios, R_{Fi}, $i = 1$ to 67. **Table 14** gives the expected values and standard deviations of the computable MAP ratios, R_{xi}, $x = F, R$, for the comb-

Table 12. Medians for the combined cohorts, Phases 1 and 2, based on a lognormal distribution with the same mean and standard deviation.

	Reported median	Median based on a lognormal distribution
$V(F1)$	50	55.9
$V(F3)$	85	85.9
$V(R1)$	55	63.4
$V(R3)$	77.5	85.4

Table 13. Phase 1: observed values, conditioned by the removal of zero valuations.

		Mean, £	Standard deviation, £
$V(\text{F1})$	Original	98.26	142.20
	Type 2 and Type 3 respondents removed	118.79	148.44
$V(\text{F3})$	Original	138.90	178.29
	Type 3 respondents removed	154.12	181.50
	Type 2 respondents removed	157.29	184.78

Table 14. Expected values and standard deviations for MAP ratios, R_x, $x = \text{F}, \text{R}$, for Phase 1 and Phase 2, all computable ratios; Extended Piecewise Uniform distribution assumed.

Phase 1	Combined cohort $(n = 67)$	Phase 2	Combined cohort $(n = 52)$
expected value	1.64	expected value	1.54
standard deviation	0.82	standard deviation	0.70

ined cohorts under both Phase 1 and Phase 2.

The mean and standard deviation of the output valuation, $V(x3)$, $x = \text{F}, \text{R}$, may be found either by simulation or analytically. The methods of Appendix E are used in the latter case. **Table 15** shows the mean and standard deviation of $V(\text{F3})$ found analytically and by simulation, and compares these values with the conditioned observed values for $V(\text{F3})$ when the Extended Piecewise Uniform distribution is used to model Phase 1 MAP ratio, R_F. **Table 16** gives equivalent values for the output valuation, $V(\text{R3})$, under Phase 2.

Table 17 and **Table 18** show the average values for mean and standard deviation for $V(x3)$, $x = \text{F}, \text{R}$, based on 200 simulations, and compare these with the conditioned observed values for Phase 1 and the observed values for Phase 2 when the Truncated Lognormal distribution is used to model MAP ratios.

Graphical summaries of the results are provided in **Figures 1-4**, where the models' predictions for the number of respondents giving valuations for safety pack 2 within specified intervals are compared with estimates based on the recorded mean and standard deviation. It is clear from **Figures 1-4** that the models produce good matches to the summary data on observed values, as discussed at greater length in Appendices C and D. Since these models incorporate the probability distributions for MAP ratio that have been described in Sections 7 and 9, this implies a validation of these probability distributions and thus of their predictions for a mean MAP ratio that lies close to the theoretical value of 2.0 for discriminating respondents.

11. Values for the VPFs

For Phase 1, the authors of the Caveat Investigator study explain the calculation of the VPF:

"Suppose that an individual indicates that her willingness to pay for a 3 in 100,000 reduction in the risk of death in a road traffic accident during the forthcoming year is £x. That individual's marginal rate of substitution (MRS) of wealth for risk of death in a road traffic accident would then be well-approximated by £$x \div (3 \times 10^{-5})$. It can be shown that the VOSL is then given by the population mean of these individual MRS."

Since $(1/n)\sum_{i=1}^{n}(V_i(.)/a) = (1/a)\sum_{i=1}^{n}(V_i(.)/n)$ for any $a > 0$, it follows that the VPFs for Phase 1 may be found by dividing the mean of $V(\text{F1})$ by 10^{-5} and the mean of $V(\text{F3})$ by 3×10^{-5}.

Meanwhile for Phase 2 the authors recommend:

"It is most straightforward to proceed by multiplying the mean household willingness to pay by the number of households in the area affected (*i.e.* 400,000) and then dividing the result by the number of fatalities prevented by the safety program concerned."

Hence the VPFs for Phase 2 may be found by multiplying the mean of $V(\text{R1})$ by 400,000 then dividing the result by 5 and by multiplying the mean of $V(\text{R3})$ by 400,000 then dividing the result by 15. (This procedure assumes that the monetary values placed on the 5 year safety programmes of VF have already been divided by 5 to produce annualized equivalents. See note to **Table 3** and **Table 4** in Section 3).

Table 15. Phase 1: mean and standard deviation for $V(\mathrm{F}3)$: $n = 67$; estimation standard errors for simulated average in brackets. Extended Piecewise Uniform distribution for MAP ratio.

	Analytical	Simulation	Conditioned observed
mean	195.0	193.3 (2.4)	157.3
s.d.	262.5	262.4 (8.7)	184.8

Table 16. Phase 2: mean and standard deviation for $V(\mathrm{R}3)$: $n = 52$; estimation standard errors for simulated average in brackets. Extended Piecewise Uniform distribution for MAP ratio.

	Analytical	Simulation	Observed
mean	167.4	168.0 (2.3)	145.1
s.d.	245.3	236.3 (6.6)	199.6

Table 17. Phase 1: mean and standard deviation for $V(\mathrm{F}3)$: $n = 67$; estimation standard errors for simulated average in brackets. Truncated Lognormal distribution for MAP ratio.

	Simulation	Conditioned Observed
mean	194.0 (2.4)	157.3
s.d.	261.7 (7.4)	184.8

Table 18. Phase 2: mean and standard deviation for $V(\mathrm{R}3)$: $n = 52$; estimation standard errors for simulated average in brackets. Truncated Lognormal distribution for MAP ratio.

	Simulation	Observed
mean	176.7 (2.5)	145.1
s.d.	250.9 (7.9)	199.6
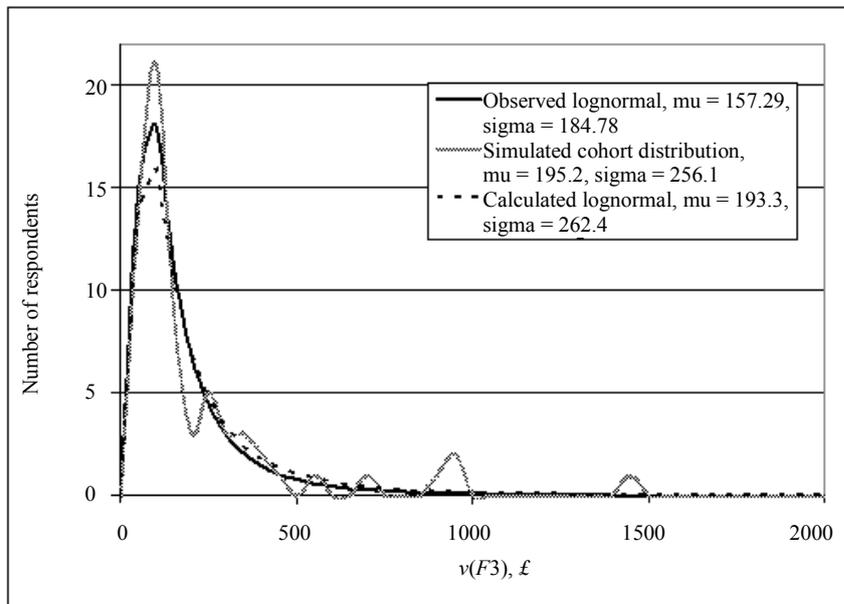
Figure 1. Number of responses in £50 intervals for $V(\mathrm{F}3)$: observed distribution, distribution with parameters from **Table 15** (both assumed lognormal) and distribution from a simulated cohort of 67; Extended Piecewise Uniform distribution for MAP ratio.
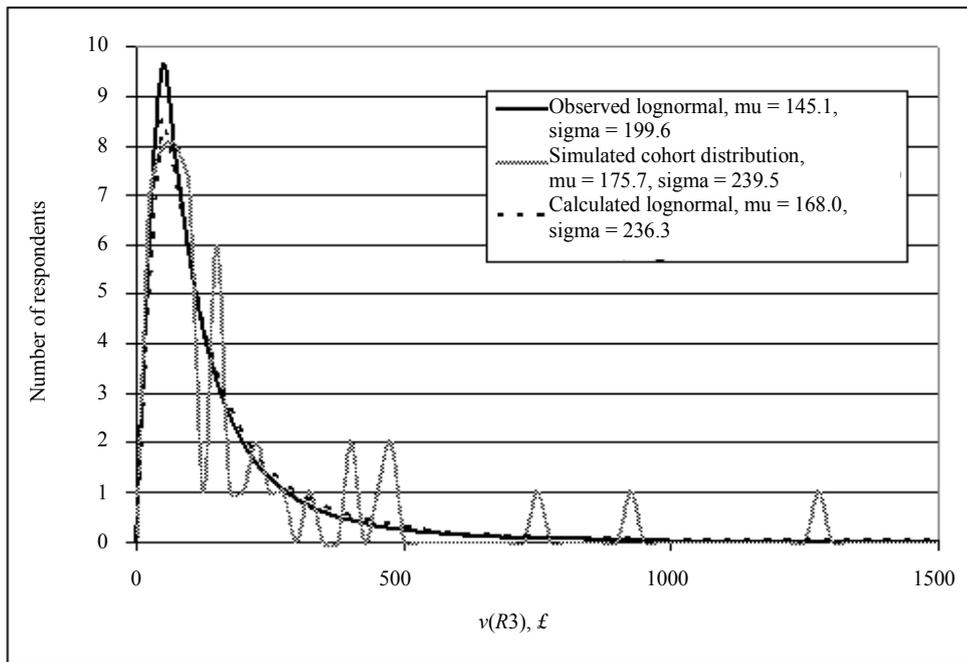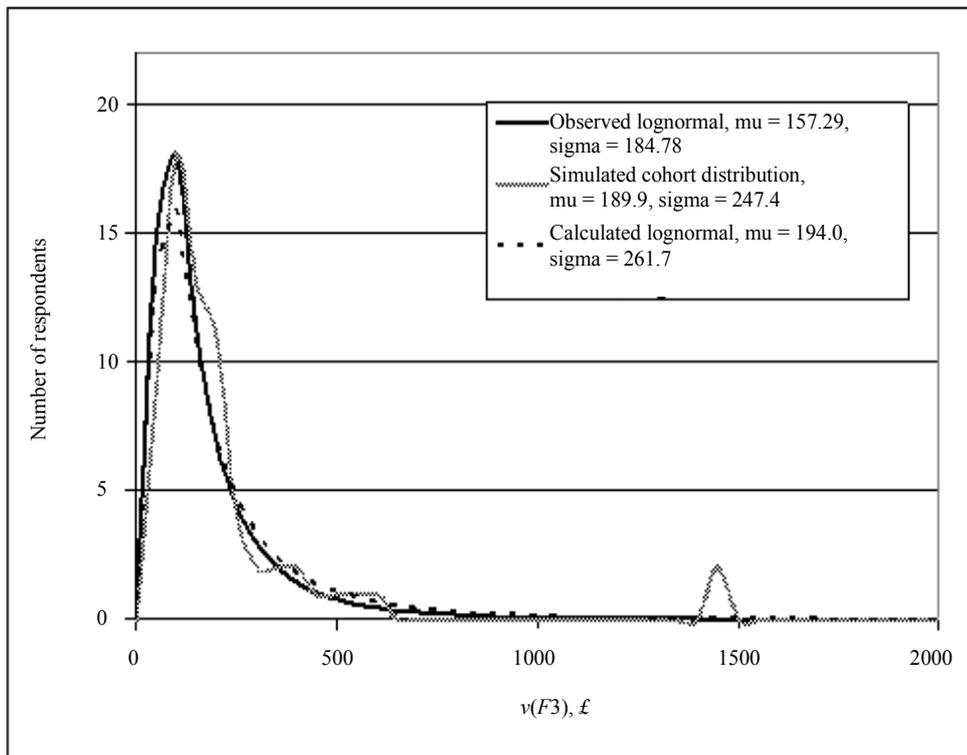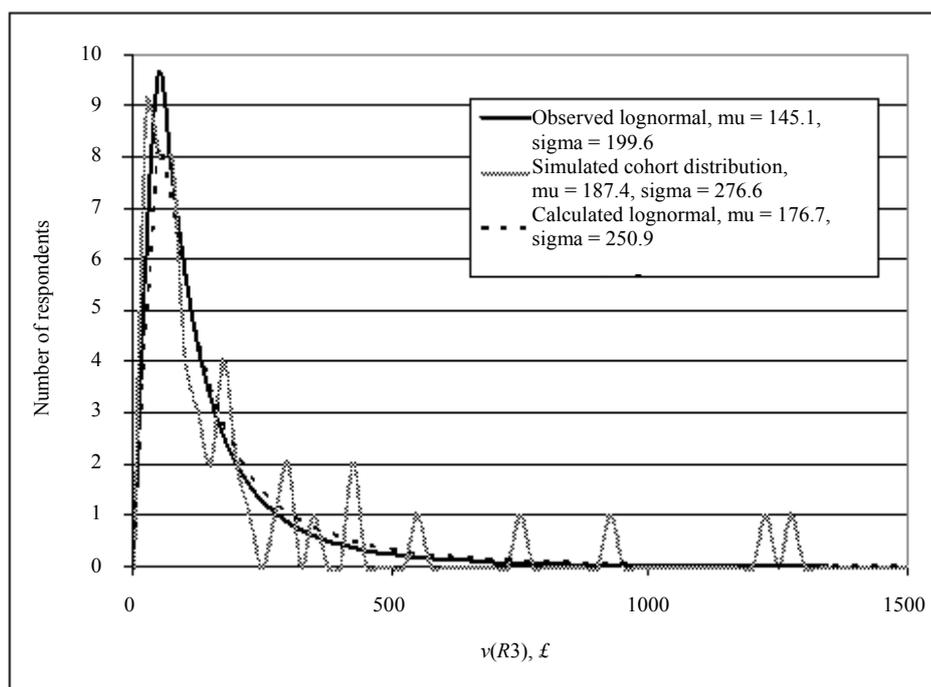
Figure 2. Number of responses in £25 intervals for $V(\text{R3})$: observed distribution, distribution with parameters from **Table 16** (both assumed lognormal) and distribution from a simulated cohort of 52; Extended Piecewise Uniform distribution for MAP ratio.

Figure 3. Number of responses in £50 intervals for $V(\text{F3})$: observed distribution, distribution with parameters from **Table 17** (both assumed lognormal) and distribution from a simulated cohort; Truncated Lognormal distribution for MAP ratio.

Figure 4. Number of responses in £25 intervals for $V(R3)$: observed distribution, distribution with parameters from **Table 18** (both assumed lognormal) and distribution from a simulated cohort; Truncated Lognormal distribution for MAP ratio.

The resulting values for the VPFs are VPF1 = £9.83 M and VPF2 = £4.63 M under Phase 1 while VPF1 = £8.70 M and VPF2 = £3.87 M under Phase 2, as reported by the authors of the Caveat Investigator study. Here VPF1 is derived from safety pack 1 while VPF2 is derived from safety pack 2.

All observed MAP ratios are computable in Phase 2 and so it is possible to derive VPF figures immediately using the simulation models described in Section 10. **Table 19** lists the resulting average values over 200 cohorts each containing 52 respondents: VPF1 = £8.75 M and VPF2 = £4.49 M when the Extended Piecewise Uniform distribution is used, and VPF1 = £8.61 M and VPF2 = £4.68 M when the Truncated Lognormal distribution is used.

Clearly the general trend is captured that VPF2 is significantly lower than VPF1. Moreover, although the average VPF2 values coming from the simulation are above the recorded value of £3.87 M, there is an estimated 21% chance of the simulated VPF2 being at or below £3.87 M (43 out of 200 simulations under the Extended Piecewise Uniform distribution, 42 out of 200 simulations under the Truncated Lognormal). **Figure 5** gives an indication of the distribution of VPF2 values based on use of the Extended Piecewise Uniform distribution.

However, the inclusion of the VPF values for respondents who are unable to discriminate the benefits of one or the other safety pack or else between the two safety packs causes the figures for VPF to be biased low. Conditioning the observed valuations for safety pack [F1] by removing the zero valuations assigned by those with Type 2 or Type 3 indiscrimination increases the VPF1 figure for Phase 1 from £9.83 M to £11.88 M. The methods of Appendix A, based on the schematic of **Figure 6** are applied here to effect this removal.

Taking out the valuations of safety pack [F1] made by the respondents with indiscrimination of Types 2 and 3 has the effect of increasing VPF2 under Phase 1 from £4.63 M to £5.24 M. The valuations of those suffering from Type 1 indiscrimination may now be excluded also. Under the assumption that it is only the valuations of [F3] and [R3] that will be affected by Type 1 indiscrimination, and not the valuations of [F1] and [R1], the effect of these exclusions will be seen on VPF2 but not on VPF1. The conditioned, observed value of VPF2 rises to £5.91 M in Phase 1 and to £4.58 M in Phase 2.

Table 19 shows the full list of observed figures, with and without conditioning, as well as the average values based on 200 simulations of the cohorts. (The Extended Piecewise Uniform distribution is equivalent to the Piecewise Uniform distribution when all respondents are discriminating).

Table 19. VPF values (1997 £M); safety pack 2 offers three times the safety benefit offered by safety pack 1. The simulation values are averages over 200 cohorts of respondents.

		Phase 1 combined cohorts		Phase 2 combined cohorts	
		VPF1	VPF2	VPF1	VPF2
Indiscrimination Types 1, 2 and 3 present	Observation	9.83	4.63	-	-
Indiscrimination Type 1 present	Observation	-	-	8.70	3.87
	Observation conditioned to remove Types 2 and 3	11.88	5.24	8.70	3.87
	Simulation: Extended Piecewise Uniform	11.80	6.42	8.75	4.49
	Simulation: Truncated Lognormal	11.70	6.43	8.61	4.68
All people can discriminate	Conditioned observation	11.88	5.91	8.70	4.58
	Simulation: Extended Piecewise Uniform	11.80	7.71	8.75	5.64
	Simulation: Truncated Lognormal	11.70	7.88	8.61	5.70
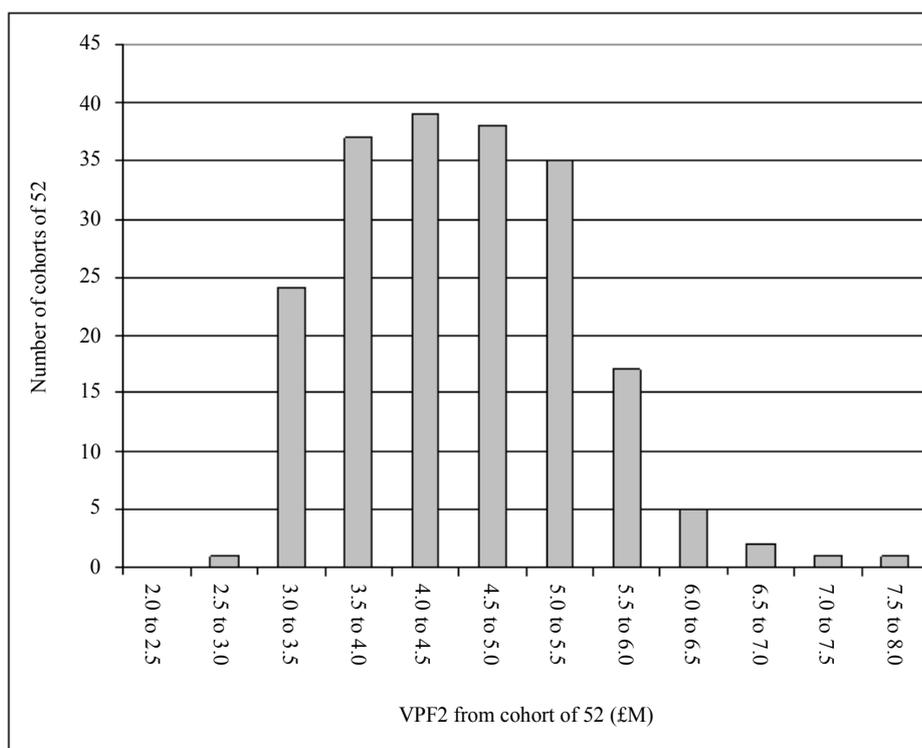
Figure 5. Number of cohorts within intervals of VPF2: 200 simulations using Extended Piecewise Uniform distribution for MAP ratio for Phase 2.

The phenomenon has thus been explained whereby significantly different VPF values will arise when people are asked to value an entry-level safety pack and another that offers a greater improvement in safety. However, it is necessary to remove the effects of indiscrimination if the results are to be generalized to situations where the safety improvements are large enough for everyone to discriminate, as might be the case if safety pack 1 offered a reduction in fatal accident frequency of 10^{-4} rather than 10^{-5} per person per year or lower.

When the effects of indiscrimination are removed, the evidence from the Caveat Investigator surveys is that people will provide valuations of a life-prolonging safety pack 1 that imply an average value of VPF in the range: £8.70 M $\leq$ VPF1 $\leq$ £11.88 M, while their valuations of a second safety pack offering three times the benefits
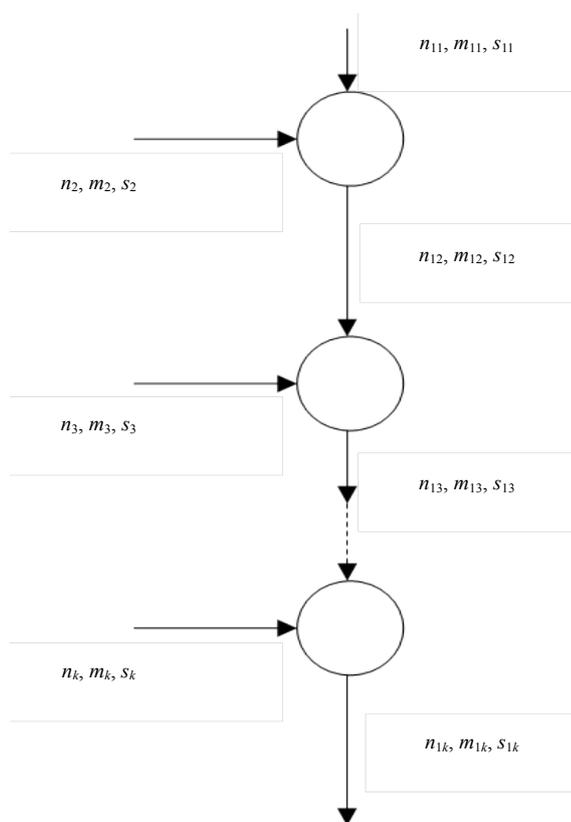
Figure 6. Flow diagram for calculating sample mean and sample standard deviation of combined sample.

will be in the range: £4.58 M ≤ VPF2 ≤ £5.91 M. The above valuations are generally higher than the figures declared by the Caveat Investigator authors.

The conversion from 1997 £s to 2014 £s requires an increase by a factor of 1.622, so that, in today's money the figures would be: £14.11 M ≤ VPF1 ≤ £19.27 M and £7.43 M ≤ VPF2 ≤ £9.59 M, between five times and ten times the value in use in the UK today [13]. Since these figures are fully explicable in terms of human preferences, it must be concluded that, if regulators and others wish to be guided by stated preference methods there is no valid justification for rejecting them.

RUP theory suggests that, when offered two safety packs, the second giving n times the benefit of the first, where $n \geq 2$, the ratio of the MAPs amongst discriminating respondents will be $n-1$, leading to a ratio of $\text{VPF2}/\text{VPF1} = (n-1)/n$. This will have a significant effect on the VPF2 value when n is relatively small, with an individual's VPF2 being expected to be of the order of 20% lower than his VPF1 when $n = 5$, for example, reducing to 10% lower when $n = 10$. In general we may deduce that for discriminating respondents:

$$0.5\text{VPF1} \leq \text{VPF2} \leq \text{VPF1} \quad \text{for} \quad n \geq 2 \tag{16}$$

with the higher VPF, VPF1, being the limiting value in the sense that $\text{VPF2} \to \text{VPF1}$ for high n. For the case where everyone can discriminate, the RUP model predicts a VPF2 for Phase 1 of £7.92 M (two-thirds of the conditioned observed VPF1 of £11.88 M), a value that is very close to the average simulated values given in the last two lines of **Table 19** (£7.71 M and £7.88 M) and comparable with the conditioned observed value of £5.91 M. Similarly for Phase 2, the RUP model predicts a VPF2 of £5.80 M based on the observed value of £8.70 M for VPF1, a figure that is very close to the average simulated values (£5.64 M and £5.70 M) and of a similar size to the conditioned observed value of £4.58 M.

12. Discussion

The Relative Utility Pricing model has shown that what the authors of the Caveat Investigator study described as

"aberrant responses" are actually nothing of the kind. Removing the responses of the people pushed beyond their limits of discrimination and then applying the RUP model makes it clear that the discriminating respondents are valuing different sizes of safety packs in the same way as they would value other commodities offered for sale in supermarkets or on the internet in different pack-sizes. The RUP model predicts that discriminating people, those able both to see a benefit from the smaller safety pack and to discriminate between the two packs, will build a substantial quantity discount into their valuation of a safety pack providing three times the safety benefit offered by the entry-level safety pack. The average ratio of the valuations amongst discriminating respondents is forecast to be 2.0, rather than 3.0 or a value close to it.

This prediction receives robust validation from the data in all of the surveys. Using the Piecewise Uniform distribution to interpret the classified data for MAP ratio, the 90% confidence interval includes the value 2.0 for discriminating respondents in each of the 6 cohorts in the Caveat Investigator study: TD, BU and the combined cohort from Phase 1, as well as VA, VF and the combined cohort from Phase 2.

The standard error reduces and the average valuation ratio approaches closely the theoretical MAP ratio of 2.0 when the larger numbers associated with the combined cohorts are taken into account. Thus the standard error goes down to 0.12 and the average MAP ratio is 1.93 when the Piecewise Uniform distribution is used to interpret the classified data for MAP ratio for Phase 2, which has 30 discriminating respondents. Then for Phase 1, in which there are 44 discriminating respondents, the standard error is 0.13, while the average MAP ratio, at 1.98, has moved to within 1% of the predicted value of 2.0.

An alternative distribution, the Truncated Lognormal, gives a good match to the classified MAP ratio data, as confirmed by the high p-values gained in the associated chi-squared tests. Simulations using this alternative distribution show that the MAP ratio for discriminating respondents has a mean of 1.98 with a standard error of 0.11, thus confirming once more the prediction of the RUP model. Moreover the success of this alternative distribution demonstrates that the confirmation of the RUP prediction is insensitive to the precise distribution of MAP ratios within each classification interval.

Further validation for the Piecewise Uniform and Lognormal distributions for MAP ratio for discriminating correspondents is provided by the simulations of both the Phase 1 survey as a whole and the Phase 2 survey as a whole. The good visual match between the distribution for the simulated valuation of safety pack 2 and the reported distribution, deduced to be approximately lognormal, confirms the two distributions as adequate models for the MAP ratios of the discriminating respondents.

The simulations, broadened out to include respondents with Type 1 indiscrimination, also give an insight to the behavior of the two VPFs, VPF1 and VPF2, that are inevitable when people are asked to state the values they will assign to two safety packs, the larger conferring a low multiple of the benefits of the smaller. Based on an average over all respondents, a simulation of Phase 2 using the Extended Piecewise Uniform distribution for MAP ratio produces VPF1 = £8.75 M and VPF2 = £4.49 M, which may be compared with the recorded results of VPF1 = £8.70 M and VPF2 = £3.87 M. Similar values are produced when the Truncated Lognormal distribution is used. It is clear that the modeling has captured the phenomenon of the two distinct and widely separated VPF modes observed in practice. The average figure for VPF2 from the simulations is within 16% of the observed value and, furthermore, the simulated VPF2 is predicted to be at or below £3.87 M for about 20% of Phase 2 surveys.

The comparison between the simulations and the recorded results for Phase 1 is complicated by the fact that 14 out of the 81 respondents of Phase 1 suffer from the more severe indiscrimination of Types 2 and 3, rendering their MAP ratio incomputable. Removal of the Type 2 and Type 3 responses by the conditioning method of Appendix A leads to conditioned observed VPF values of VPF1 = £11.88 M and VPF2 = £5.24 M, which may be compared with the simulated values of VPF1 = £11.80 M and VPF2 = £6.42 M calculated when the Extended Piecewise distribution for MAP ratio is used, with similar values being produced when the Truncated Lognormal distribution is applied. When all types of indiscrimination are removed, the simulations produce average VPF2 values that are similar to the predictions of RUP theory and of a similar size to the conditioned observed value of VPF2. The phenomenon of two distinct and widely separated VPF modes is correctly predicted.

It is necessary to strip out the effects of indiscrimination to generalize the results, and this leads to higher VPF values. In £s of 2014, VPF1 then lies between £14.11 M and £19.27 M, while VPF2 is between £7.43 M and £9.59 M. RUP theory suggests that VPF2 will converge to the larger VPF1 when safety pack 2 offers safety benefits that are very much higher than those offered by safety pack 1, which asymptotic behavior might mark out the higher VPF1 as having more general applicability.

There would appear to be no valid reason for dismissing the VPFs listed if the decision is taken to follow the stated preference route for valuing human life. Far from being caused by "a number of theoretically irrelevant factors", as claimed by the Caveat Investigator authors, the two VPF figures, VPF1 and VPF2, have been explained fully in terms of economic theory. While the fact that these VPF values are up to an order of magnitude greater than the VPF currently in use might constitute an inconvenience, that in itself can hardly justify a rejection of these values by those committed to stated preference methods.

The difficulty remains, of course, that opinion surveys will tend to provide a significantly different VPF figure depending on the size of the safety benefit. No unique figure for VPF can be expected when stated preference techniques are used.

13. Conclusions

Two distinct and widely separated VPF figures, VPF1 associated with safety pack 1 and VPF2 associated with safety pack 2 are the inevitable outcome of any survey that asks people to value two safety packs, where the larger one confers a safety benefit that is a low multiple of that conferred by the smaller safety pack.

The basic mechanism underlying the two, markedly different VPFs, is Relative Utility Pricing. By the RUP model, people assigning their MAPs to two packs of a commodity, with the second pack being n times larger than the first, will put a valuation on the larger pack that is only $n-1$ times greater than the value they put on the smaller pack.

In the case of the Caveat Investigator study, where safety pack 2 confers three times the benefit of safety pack 1, the MAP ratio is predicted to be 2.0, and this figure has been confirmed by the behavior of the discriminating respondents, to whom RUP applies. Their MAP ratios have been found to average between 1.93 and 1.98, with 2.0 included in the 90% confidence intervals.

Simulation models, broadened to include not only discriminating respondents conforming to RUP but also respondents with Type 1 indiscrimination, have reproduced approximately the two very different VPF values found from the views of all the respondents in Phase 2 of the Caveat Investigator study. Similar results were produced for Phase 1, although in that case it was necessary to condition the observations to remove the effect of those who set a zero value on safety pack 1, namely those suffering from Type 2 or Type 3 indiscrimination.

It may be concluded that there is no inconsistency in the VPF results reported in the Caveat Investigator study, despite the fact that they were dismissed by its authors.

The VPF values for general use need to be modified if they are to be applied to situations where the safety benefits are large enough for everyone to discriminate. Stripping out the effects of indiscrimination leads to a VPF1 of between £14.11 M and £19.27 M, and a VPF2 that lies between £7.43 M and £9.59 M (2014 £s).

By RUP theory, the lower VPF2 will converge to the larger VPF1 when safety pack 2 offers safety benefits that are very much higher than those offered by safety pack 1. This asymptotic behavior marks out the higher VPF1 as having more general applicability.

Since the VPF figures observed can be explained fully in terms of economic theory, there is clearly no valid reason for dismissing the VPFs above if the stated preference route is chosen as the vehicle for valuing human life unless the Caveat Investigator study contains a fault or faults yet to be revealed. It is true that the VPF values found in the study are up to an order of magnitude greater than the VPF currently in use in the UK today, but while this might constitute an inconvenience that can hardly justify a rejection of these values by those who are committed to stated preference methods.

As noted in the Introduction, however, there are reasons for viewing stated preference techniques with caution, as there will always be doubts on whether the person's stated valuation will conform to what he will actually be prepared to pay. Moreover, the whole concept of a unique figure for VPF of general applicability has been shown to be incompatible with the use of stated preference techniques. It is thus apparent that there are significant difficulties in the interpretation of surveys to establish a VPF. This has been shown not only here but also by the demonstrated invalidity of the two-injury chained method, which was advanced by the Caveat Investigator team as a notional replacement for their initial opinion surveys.

The problems and confusions associated with the current, stated preference approaches for assessing how much to spend to safeguard human life mean that there is a pressing and unfulfilled need to ensure adequate protection for UK public and workers from industrial and transport hazards. A re-appraisal is urgently required of methodologies, such as J-value, that are based on the surer ground of revealed preferences, and can allow ro-

bust decision making on safety by industry and the regulatory authorities.

Acknowledgements

References

[1]   Thomas, P.J., Stupples, D.W. and Alghaffar, M.A. (2006) The Extent of Regulatory Consensus on Health and Safety Expenditure: Part 1: Development of the J-Value Technique and Evaluation of the Regulators' Recommendations. *Process Safety and Environmental Protection*, **84**, 329-336.

[2]   Thomas, P.J., Stupples, D.W. and Alghaffar, M.A. (2006) The Extent of Regulatory Consensus on Health and Safety Expenditure: Part 2: Applying the J-Value Technique to Case Studies across Industries. *Process Safety and Environmental Protection*, **84**, 337-343.

[3]   Thomas, P.J., Jones, R.D. and Kearns, J.O. (2010) The Trade-Offs Embodied in *J*-Value Safety Analysis. *Process Safety and Environmental Protection*, **88**, 147-167. http://dx.doi.org/10.1016/j.psep.2010.02.001

[4]   Nathwani, J.S., Lind, N.C. and Pandey, M.D. (1997) Affordable Safety by Choice: The Life Quality Method. Institute for Risk Research, University of Waterloo, Waterloo, Ontario.

[5]   Nathwani, J.S., Lind, N.C. and Pandey, M.D. (2008) Engineering Decisions for Life Quality. How Safe Is Safe Enough? Springer, Dordrecht.

[6]   Pandey, M.D. and Nathwani, J.S. (2003) A Conceptual Approach to the Estimation of Societal Willingness-to-Pay for Nuclear Safety Programs. *Nuclear Engineering and Design*, **224**, 65-77. http://dx.doi.org/10.1016/S0029-5493(03)00062-1

[7]   Pandey, M.D., Nathwani, J.S. and Lind, N.C. (2006) The Derivation and Calibration of the Life-Quality Index (LQI) from Economic Principles. *Structural Safety*, **28**, 341-360. http://dx.doi.org/10.1016/j.strusafe.2005.10.001

[8]   Ciriacy-Wantrup, S.V. (1947) Capital Returns from Soil-Conservation Practices. *Journal of Farm Economics*, **29**, 1181-1196.

[9]   Beattie, J., Covey, J., Dolan, P., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Robinson, A. and Spencer, A. (1998) On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1—Caveat Investigator. *Journal of Risk and Uncertainty*, **17**, 5-26. http://dx.doi.org/10.1023/A:1007711416843

[10]   Sunstein, C.R. (2004) Lives, Life-Years, and Willingness to Pay. *Columbia Law Review*, **104**, 205-252. http://dx.doi.org/10.2307/4099352

[11]   Sunstein, C.R. (2004) Valuing Life: A Plea for Disaggregation. *Duke Law Journal*, **54**, 385-445.

[12]   Thomas, P. and Vaughan, G. (2013) All in the Balance: Assessing Schemes to Protect Humans and the Environment. *Nuclear Future*, **9**, 41-51.

[13]   Department for Transport (2013) WebTAG Data Book, WebTAG Table A 4.1.1. https://www.gov.uk/government/publications/webtag-tag-data-book

[14]   Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N. and Spencer, A. (1999) On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 2—The CV/SG "Chained" Approach. *Journal of Risk and Uncertainty*, **17**, 187-214. http://dx.doi.org/10.1023/A:1007782800868

[15]   Wolff, J. and Orr, S. (2009) Cross-Sector Weighting and Valuing of QALYs and VTPFs. A Report for the Inter-Departmental Group for the Valuation of Life and Health. Final Report, 8 July 2009. http://www.ucl.ac.uk/cpjh/docs/IGVLH.pdf

[16]   Spackman, M., Evans, A., Jones-Lee, M., Loomes, G., Holder, S. and Webb, H. (2011) Updating the VPF and VPIs: Phase 1: Final Report: Department for Transport. NERA Economic Consulting. http://assets.dft.gov.uk/publications/pgr-economics-rdg-updatingvpfvpi-pdf/vpivpfreport.pdf

[17]   Thomas, P.J. and Vaughan, G.J. (2014) Testing the Validity of the "Value of a Prevented Fatality" (VPF) Used to As-

sess UK Safety Measures. *Process Safety and Environmental Protection*, in Press.
http://dx.doi.org/10.1016/j.psep.2014.07.0011

[18] Thomas, P. and Chrystal, A. (2013) Explaining the "Buy One Get One Free" Promotion: The Golden Ratio as a Marketing Tool. *American Journal of Industrial and Business Management*, **3**, 655-673.
http://dx.doi.org/10.4236/ajibm.2013.38075

[19] Thomas, P. and Chrystal, A. (2013) Using Relative Utility Pricing to Explain Multibuy Prices in Supermarkets and on the Internet. *American Journal of Industrial and Business Managemen*t, **3**, 687-699.
http://dx.doi.org/10.4236/ajibm.2013.38078

[20] Cookson, C. (2014) Mystery of Supermarket Egg Pricing Cracked. *Financial Times*, UK Edition, London, 4 January, 41. http://www.ft.com/cms/s/2/ec12addc-6c2a-11e3-a216-00144feabdc0.html#axzz38gPHLvPR

[21] Kahneman, D. and Tversky, A. (1979) Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, **47**, 263-292. http://dx.doi.org/10.2307/1914185

[22] Thaler, R.H. (1985) Mental Accounting and Consumer Choice. *Marketing Science*, **4**, 199-214. Republished 2008 in *Marketing Science*, **27**, 15-25.

[23] Wallop, H. (2008) Buy-One-Get-One-Free Offers Are One of the Most Effective Marketing Tools in the Supermarket Industry. *The Daily Telegraph*. http://www.telegraph.co.uk/news/uknews/2263645/

[24] Thomas, P. and Chrystal, A. (2013) Generalized Demand Densities for Retail Price Investigation. *American Journal of Industrial and Business Management*, **3**, 279-294. http://www.scirp.org/journal/ajibm
http://dx.doi.org/10.4236/ajibm.2013.33034

[25] Thomas, P. and Chrystal, A. (2013) Retail Price Optimization from Sparse Demand Data. *American Journal of Industrial and Business Management*, **3**, 295-306. http://www.scirp.org/journal/ajibm
http://dx.doi.org/10.4236/ajibm.2013.33035

[26] Weisstein, E. (2014) Uniform Distribution. *Wolfram MathWorld*.
http://mathworld.wolfram.com/UniformDistribution.html

[27] Dunn, O.J. (1965) Basic Statistics: A Primer for the Biomedical Sciences. John Wiley and Sons, New York.

Appendix A. Characteristics of a Sample Taken from Several Distributions

A.1 Finding the Sample Mean

The scheme for combining the statistics of the records is given diagrammatically in **Figure 6**. Let there be a sample that combines n_{11} records taken from a distribution with an estimated mean, $\hat{\mu}_{11}$, and an estimated variance, $\hat{\sigma}_{11}^2$, and n_2 records taken from a distribution with an estimated mean, $\hat{\mu}_2$, and an estimated variance, $\hat{\sigma}_2^2$. The combined sample will thus have $n_{11} + n_2 = n_{12}$ records.

Let the sample mean of the combined sample be denoted s_{12}^2 and its sample variance, s_{12}^2. The sample mean will be given by:

$$m_{12} = \frac{1}{n_{12}}\left(\sum_{i=1}^{n_{11}} X_i + \sum_{j=1}^{n_2} Y_j\right) = \frac{1}{n_{12}}\left(n_{11} \frac{1}{n_{11}}\sum_{i=1}^{n_{11}} X_i + n_2 \frac{1}{n_2}\sum_{j=1}^{n_2} Y_j\right) \tag{A.1}$$

where X_i; $i = 1, 2, \cdots, n_{11}$ are the sample values taken from the first distribution and Y_j; $j = 1, 2, \cdots, n_2$ are the sample values taken from the second distribution. Now

$$\frac{1}{n_{11}}\sum_{i=1}^{n_{11}} X_i = m_{11} \approx \hat{\mu}_{11} \tag{A.2}$$

where m_{11} is the mean of the values contained in the combined sample but taken from the first distribution. Similarly

$$\frac{1}{n_2}\sum_{j=1}^{n_2} Y_j = m_2 \approx \hat{\mu}_2 \tag{A.3}$$

where m_2 is the mean of the values contained in the combined sample but taken from the second distribution. Hence the sample mean of the combined sample will be

$$m_{12} = \frac{n_{11}}{n_{12}} m_{11} + \frac{n_2}{n_{12}} m_2 \approx \frac{n_{11}}{n_{12}} \hat{\mu}_{11} + \frac{n_2}{n_{12}} \hat{\mu}_2 \tag{A.4}$$

If the sample is taken from three distributions, then simple extensions of Equations (A.1) to (A.4) give the new sample mean of the extended sample, m_{13}, as:

$$m_{13} = \frac{1}{n_{13}}\left(n_{11} m_{11} + n_2 m_2 + n_3 m_3\right) \tag{A.5}$$

where the subscript, 3, denotes the third distribution and the total number of records is now $n_{13} = n_{11} + n_2 + n_3 = n_{12} + n_3$. Equation (A.5) may be recast as:

$$m_{13} = \frac{n_{12}}{n_{13}}\left(\frac{n_{11} m_{11} + n_2 m_2}{n_{12}}\right) + \frac{n_3}{n_{13}} m_3 = \frac{n_{12}}{n_{13}} m_{12} + \frac{n_3}{n_{13}} m_3 \tag{A.6}$$

Equation (A.6) may now be generalized to account for records being taken from k distributions, when the overall sample mean will be m_{1k}, given by:

$$m_{1k} = \frac{n_{1,k-1}}{n_{1k}} m_{1,k-1} + \frac{n_k}{n_{1k}} m_k \quad \text{for} \quad k > 1 \tag{A.7}$$

A.2 The Sample Variance

For a set of n randomly drawn records, Z_i, with sample mean,

$$m = \frac{1}{n}\sum_{i=1}^{n} Z_i \tag{A.8}$$

the unbiased sample variance, s^2, may be written:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Z_i - m\right)^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Z_i^2 - 2mZ_i + m^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}Z_i^2 - 2m\sum_{i=1}^{n}Z_i + \sum_{i=1}^{n}m^2\right)$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n}Z_i^2 - 2nm^2 + nm^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}Z_i^2 - nm^2\right). \tag{A.9}$$

where Equation (A.8) has been used in the penultimate step. Hence, using the notation employed above, the sample variance, s_{12}^2, of a combined sample drawn from two distributions will obey:

$$\left(n_{12} - 1\right)s_{12}^2 = \sum_{i=1}^{n_{11}}X_i^2 + \sum_{j=1}^{n_2}Y_j^2 - n_{12}m_{12}^2 = \sum_{i=1}^{n_{11}}X_i^2 + \sum_{j=1}^{n_2}Y_j^2 - \frac{1}{n_{12}}\left(n_{11}^2 m_{11}^2 + 2n_{11}n_2 m_{11}m_2 + n_2^2 m_2^2\right)$$

$$= \sum_{i=1}^{n_1}X_i^2 - \frac{n_{11}}{n_{12}}n_{11}m_{11}^2 + \sum_{j=1}^{n_2}Y_j^2 - \frac{n_2}{n_{12}}n_2 m_2^2 - 2\frac{n_{11}n_2}{n_{12}}m_{11}m_2 \tag{A.10}$$

$$= \sum_{i=1}^{n_1}X_i^2 - n_{11}m_{11}^2 + n_{11}m_{11}^2 - \frac{n_{11}}{n_{12}}n_{11}m_{11}^2 + \sum_{j=1}^{n_2}Y_j^2 - n_2 m_2^2 + n_2 m_2^2 - \frac{n_2}{n_{12}}n_2 m_2^2 - 2\frac{n_{11}n_2}{n_{12}}m_{11}m_2.$$

where Equation (A.4) has been used in the second step. But we may define the sample variance, s_{11}^2, by:

$$\sum_{i=1}^{n_{11}}X_i^2 - n_{11}m_{11}^2 = \left(n_{11} - 1\right)s_{11}^2 \approx \left(n_{11} - 1\right)\hat{\sigma}_{11}^2 \tag{A.11}$$

and the sample variance, s_2^2, by:

$$\sum_{j=1}^{n_2}Y_j^2 - n_2 m_2^2 = \left(n_2 - 1\right)s_2^2 \approx \left(n_2 - 1\right)\hat{\sigma}_2^2 \tag{A.12}$$

Therefore:

$$\left(n_{12} - 1\right)s_{12}^2 = \left(n_{11} - 1\right)s_{11}^2 + \left(n_2 - 1\right)s_2^2 + n_{11}m_{11}^2\left(\frac{n_{12} - n_{11}}{n_{12}}\right) + n_2 m_2^2\left(\frac{n_{12} - n_2}{n_{12}}\right) - 2\frac{n_{11}n_2}{n_{12}}m_{11}m_2$$

$$= \left(n_{11} - 1\right)s_{11}^2 + \left(n_{12} - n_{11} - 1\right)s_2^2 + n_{11}\left(\frac{n_{12} - n_{11}}{n_{12}}\right)m_{11}^2 - 2n_{11}\left(\frac{n_{12} - n_{11}}{n_{12}}\right)m_{11}m_2 + n_{11}\left(\frac{n_{12} - n_{11}}{n_{12}}\right)m_2^2. \tag{A.13}$$

Noting that $n_{11} + n_2 = n_{12}$

$$\left(n_{12} - 1\right)s_{12}^2 = \left(n_{11} - 1\right)s_{11}^2 + \left(n_2 - 1\right)s_2^2 + \frac{n_{11}n_2}{n_{12}}\left(m_{11}^2 - 2m_{11}m_2 + m_2^2\right) \tag{A.14}$$

Thus the sample variance, s_{12}^2, of a combined sample taken from two distributions is:

$$s_{12}^2 = \frac{n_{11} - 1}{n_{12} - 1}s_{11}^2 + \frac{n_2 - 1}{n_{12} - 1}s_2^2 + \frac{n_{11}n_2}{n_{12}\left(n_{12} - 1\right)}\left(m_{11} - m_2\right)^2 \tag{A.15}$$

For the general case where the sample is drawn from k distributions, Equation (A.15) becomes:

$$s_{1k}^2 = \frac{n_{1,k-1} - 1}{n_{1k} - 1}s_{1,k-1}^2 + \frac{n_k - 1}{n_{1k} - 1}s_k^2 + \frac{n_{1,k-1}n_k}{n_{1k}\left(n_{1k} - 1\right)}\left(m_{1,k-1} - m_k\right)^2 \quad k > 1 \tag{A.16}$$

Equations (A.7) and (A.16) may be applied successively with $k = 2$ then $k = 3$ and so on until all the distributions have been included. See **Figure 6**.

A.3 Removing a Component Distribution from the Combined Sample

The starting point for the treatment above is that the characteristics of the component distributions are known and the aim is to characterize the combined sample. There may be occasions, however, when the characteristics of the combined sample are known, at least approximately, as well as those of one of the component distributions. The following treatment applies when estimates of mean and variance are available for the combined sample and for distribution 2 and it is necessary to estimate the sample mean, m_{11}, and sample variance, s_{11}^2, of

the distribution identified with the subscript, "11", in **Figure 6**.

Rearrangement of Equation (A.4) gives:

$$m_{11} = \frac{n_{12}m_{12} - n_2 m_2}{n_{11}} \tag{A.17}$$

Meanwhile, from Equation (A.15):

$$s_{11}^2 = \frac{n_{12}-1}{n_{11}-1}s_{12}^2 - \frac{n_2-1}{n_{11}-1}s_2^2 - \frac{n_{11}n_2}{n_{12}(n_{11}-1)}(m_{11}-m_2)^2 \tag{A.18}$$

The calculation relies on the use of the estimated means and variances:

$$m_2 \approx \hat{\mu}_2; \quad s_2^2 \approx \hat{\sigma}_2^2; \quad m_{12} \approx \hat{\mu}_{12}; \quad \text{and} \quad s_{12}^2 \approx \hat{\sigma}_{12}^2 \tag{A.19}$$

Appendix B. Characteristics of a Cohort Involving Several Distributions for Each of Which the Population Mean and Population Variance Are Known Exactly

A method that parallels the development of Appendix A is presented for the case where the population mean and variance are known rather than estimates. An alternative method based on the same principles is presented in Section B.3.

B.1 Finding the Mean of the Combined Cohort

The general outline of **Figure 6** is followed once again, but now with μ and σ replacing m and s respectively. Let there be a cohort that combines n_{11} records, $X^{(11)}$, taken from a distribution with a true mean, μ_{11}, and a true variance, σ_{11}^2, and n_2 records, $X^{(2)}$, taken from a distribution, with a true mean, μ_2, and a true variance, σ_2^2. The combined cohort will thus have $n_{11} + n_2 = n_{12}$ records, $X^{(12)}$, with a population mean, μ_{12}, and population variance, σ_{12}^2.

If a record, $X^{(12)}$, is chosen at random from the combined cohort, its expected value will depend on whether the record comes originally from the "11" distribution or from the "2" distribution. Define a random variable, W, as an indicator variable that takes the values:

$$W = \begin{cases} 1 \text{ if the "11" distribution is the original source} \\ 2 \text{ if the "2" distribution is the original source} \end{cases} \tag{B.1}$$

It follows that

$$E\left(X^{(12)}\big|W=1\right) = E\left(X^{(11)}\right) \tag{B.2}$$

$$E\left(X^{(12)}\big|W=2\right) = E\left(X^{(2)}\right) \tag{B.3}$$

The expected value of a record in the combined cohort, $E\left(X^{(12)}\right)$ is then:

$$E\left(X^{(12)}\right) = E\left(E\left(X^{(12)}\big|W\right)\right) = P(W=1)\times E\left(X^{(12)}\big|W=1\right) + P(W=2)\times E\left(X^{(12)}\big|W=2\right) \tag{B.4}$$

The probability of choosing a record originally from distribution "11" is $P(W=1) = n_{11}/n_{12}$, while the probability of choosing a record originally from distribution "2" is $P(W=2) = n_2/n_{12}$. On substituting these values into Equation (B.4) and also using Equations (B.2) and (B.3), it is found that:

$$E\left(X^{(12)}\right) = \frac{n_{11}}{n_{12}}E\left(X^{(11)}\right) + \frac{n_2}{n_{12}}E\left(X^{(2)}\right) \tag{B.5}$$

or, in more compact notation:

$$\mu_{12} = \frac{1}{n_{12}}\left(n_{11}\mu_{11} + n_2\mu_2\right) \tag{B.6}$$

If the cohort is taken from three distributions, then simple extensions of Equations (B.1) to (B.5) give the new

population mean of the combined cohort, μ_{13}, as:

$$\mu_{13} = \frac{1}{n_{12}}\left(n_{11}\mu_{11} + n_2\mu_2 + n_3\mu_3\right) \tag{B.7}$$

where the subscript, 3, denotes the third distribution and the total number of records is now $n_{13} = n_{11} + n_2 + n_3 = n_{12} + n_3$. Equation (B.7) may be recast as:

$$\mu_{13} = \frac{n_{12}}{n_{13}}\mu_{12} + \frac{n_3}{n_{13}}\mu_3 \tag{B.8}$$

Equation (B.8) may now be generalized to account for records being taken from k distributions, when the true mean will be μ_{1k}, given by:

$$\mu_{1k} = \frac{n_{1,k-1}}{n_{1k}}\mu_{1,k-1} + \frac{n_k}{n_{1k}}\mu_k \quad \text{for} \quad k > 1 \tag{B.9}$$

B.2 Population Variance for the Combined Cohort

The population variance, σ_{12}^2, may be written:

$$\sigma_{12}^2 = E\left(\left(X^{(12)}\right)^2\right) - \left(E\left(X^{(12)}\right)\right)^2 = E\left(\left(X^{(12)}\right)^2\right) - \mu_{12}^2 \tag{B.10}$$

The principles outlined in Section B.1 can be applied not only to the expectation of $X^{(12)}$ but also to its square, $\left(X^{(12)}\right)^2$. Hence, by direct analogy with Equation (B.5):

$$E\left(\left(X^{(12)}\right)^2\right) = \frac{n_{11}}{n_{12}}E\left(\left(X^{(11)}\right)^2\right) + \frac{n_2}{n_{12}}E\left(\left(X^{(2)}\right)^2\right) \tag{B.11}$$

Substituting from Equation (B.11) into Equation (B.10) gives:

$$\sigma_{12}^2 = \frac{n_{11}}{n_{12}}E\left(\left(X^{(11)}\right)^2\right) + \frac{n_2}{n_{12}}E\left(\left(X^{(2)}\right)^2\right) - \mu_{12}^2 \tag{B.12}$$

which may then be developed as follows:

$$
\begin{aligned}
n_{12}\sigma_{12}^2 &= n_{11}\left(E\left(\left(X^{(11)}\right)^2\right) - \left(E\left(X^{(11)}\right)\right)^2\right) + n_2\left(E\left(\left(X^{(2)}\right)^2\right) - \left(E\left(X^{(2)}\right)\right)^2\right) + n_{11}\left(E\left(X^{(11)}\right)\right)^2 \\
&\quad + n_2\left(E\left(X^{(2)}\right)\right)^2 - n_{12}\mu_{12}^2 \\
&= n_{11}\sigma_{11}^2 + n_2\sigma_2^2 + n_{11}\mu_{11}^2 + n_2\mu_2^2 - \frac{1}{n_{12}}\left(n_{11}^2\mu_{11}^2 + 2n_{11}n_2\mu_{11}\mu_2 + n_2^2\mu_2^2\right) \\
&= n_{11}\sigma_{11}^2 + n_2\sigma_2^2 + n_{11}\left(1 - \frac{n_{11}}{n_{12}}\right)\mu_{11}^2 - \frac{2n_{11}n_2\mu_{11}\mu_2}{n_{12}} + n_2\left(1 - \frac{n_2}{n_{12}}\right)\mu_2^2 \\
&= n_{11}\sigma_{11}^2 + n_2\sigma_2^2 + n_{11}\left(\frac{n_{12} - n_{11}}{n_{12}}\right)\mu_{11}^2 - \frac{2n_{11}n_2\mu_{11}\mu_2}{n_{12}} + n_2\left(\frac{n_{12} - n_2}{n_{12}}\right)\mu_2^2 .
\end{aligned}
\tag{B.13}
$$

where Equation (B.6) has been used in the second line. But, since $n_{11} + n_2 = n_{12}$:

$$n_{12}\sigma_{12}^2 = n_{11}\sigma_{11}^2 + n_2\sigma_2^2 + \frac{n_{11}n_2}{n_{12}}\left(\mu_{11}^2 - 2\mu_{11}\mu_2 + \mu_2^2\right) = n_{11}\sigma_{11}^2 + n_2\sigma_2^2 + \frac{n_{11}n_2}{n_{12}}\left(\mu_{11} - \mu_2\right)^2 \tag{B.14}$$

or

$$\sigma_{12}^2 = \frac{n_{11}}{n_{12}}\sigma_{11}^2 + \frac{n_2}{n_{12}}\sigma_2^2 + \frac{n_{11}n_2}{n_{12}^2}\left(\mu_{11} - \mu_2\right)^2 \tag{B.15}$$

Generalizing, the effect on the population of selecting from k distributions is:

$$\sigma_{1k}^2 = \frac{n_{1,k-1}}{n_{1k}} \sigma_{1k}^2 + \frac{n_k}{n_{1k}} \sigma_k^2 + \frac{n_{1,k-1} n_k}{n_{1k}^2} \left(\mu_{1,k-1} - \mu_k \right)^2 \tag{B.16}$$

B.3 Alternative Form

An alternative continuation is possible after Equation (B.12), which may be written:

$$\sigma_{12}^2 = \frac{n_{11}}{n_{12}} M_{11}^{(2)} + \frac{n_2}{n_{12}} M_2^{(2)} - \mu_{12}^2 \tag{B.17}$$

where $M_{11}^{(2)}$, $M_{i2}^{(2)}$ are the second moments about the origin for the distributions identified by the subscripts, "11" and "2". Changing the notation so that the distributions to be combined are numbered $1, 2, 3, \cdots, k, \cdots, N$, Equation (B.17) may be generalized to give the population variance when all N distributions are combined:

$$\sigma_N^2 = \frac{1}{n_N} \sum_{k=1}^{N} n_k M_k^{(2)} - \mu_N^2 \tag{B.18}$$

where

$$\mu_N = \frac{1}{n_N} \sum_{k=1}^{N} n_k \mu_k \tag{B.19}$$

Appendix C. Validation of the Results of the Extended Piecewise Uniform Distribution Model

The validation exercise is applied to combined cohorts from each of Phase 1 and Phase 2, giving the benefit of working with larger numbers. This avoids also the problem, explained in the note to **Table 5**, of lack of information on the distribution of the zero responses between the TD and BU cohorts.

In spite of the limited data available, corroboration for the results of the Extended Piecewise Uniform distribution for the MAP ratio may be achieved by embedding this distribution within an input-output model, where the individual's valuation, $V_i(x1)$, $x = F, R$, is the input and the valuation, $V_i(x3)$, $x = F, R$, is the output produced by multiplying by the individual's MAP ratio, R_{xi}:

$$V_i(x3) = R_{xi} V_i(x1) \quad x = F, R \tag{C.1}$$

A lognormal model for $V_i(x1)$, $x = F, R$ may be used because of the approximately lognormal relationship found to exist in the observed valuations of [F1] and [R1], as revealed by the correspondence between the recorded median and that resulting from a lognormal distribution with the recorded mean and standard deviation as parameters. As shown in **Table 12**, $V_i(x3)$, $x = F, R$ may be considered to obey a lognormal distribution by the same token.

The useful application of Equation (C.1) requires, of course, that the ratio, $R_{xi} = V_i(x3) / V_i(x1)$, $x = F, R$, should exist. This requirement is satisfied for all the respondents in Phase 2, but in Phase 1 there are 14 cases of indiscrimination of either Type 2 or Type 3, where the respondent values $V_i(x1)$ as zero, and assigns $V_i(x3)$ either a positive value (Type 2 indiscrimination: 6 respondents) or zero (Type 3 indiscrimination: 8 respondents). The MAP ratio, R_{Fi} is rendered indeterminate in these 14 instances, making it necessary to filter out the effects of these 14 respondents to produce conditioned observed distributions for $V(F1)$ and $V(F3)$.

C.1 Removing the Zero Responses from the Phase 1 Valuation Distributions (Types 2 and 3 Indiscrimination)

The 14 zero responses inherent in the observed distribution for $V(F1)$ may be removed easily, since the mean and standard deviation of the set of zero valuations are both zero. Applying the method of Section A.3 of Appendix A, the resultant mean of the 67 remaining valuations, $V(F1)$, rises by about £20 to £118.79, while the standard deviation for those 67 valuations of [F1] is now £148.44. The 8 zero responses inherent in the observed distribution for $V(F3)$ may be treated in the same way and this results in a mean of £154.12 for the remaining 73 valuations of the safety pack, [F3], and a standard deviation of £181.50.

It is necessary also to remove from the distribution for $V(F3)$ the effect of the valuations from the six res-

pondents suffering from Type 2 indiscrimination. It is assumed first that the value placed on safety pack 1, $V(\text{F1})$, by respondents with Type 1 indiscrimination will be generally similar to those of the discriminating respondents, and hence that their distribution for $V(\text{F1})$ will have a mean and a standard deviation the same as those just found for the 67 respondents just discussed, namely £118.79 and £148.44 respectively. The additional assumption is now made that the distribution for $V(\text{F3})$, for those with Type 2 indiscrimination matches the distribution for $V(\text{F3})$ of respondents suffering from the less severe Type 1 indiscrimination. Since those with Type 1 indiscrimination place the same value on safety pack 2 as for safety pack 1, so that $V_i(\text{F3}) = V_i(\text{F1})$, it follows that the distribution for $V(\text{F3})$ amongst those with Type 2 indiscrimination will have a mean and standard deviation of £118.79 and £148.44 respectively. Removing the effect of the six valuations with these characteristic parameters produces a final distribution for the valuation of [F3] among the 67 respondents who are discriminating or suffer from Type 1 indiscrimination that is characterized by a mean of £157.29 and a standard deviation of £184.78. The results are summarized in **Table 13**.

The parameters for the conditioned, observed distribution $V(\text{F3})$ are, in fact, relatively insensitive to the parameters for $V(\text{F3})$ assumed for those with Type 2 indiscrimination. As an example, reducing by a factor of two the mean and standard deviation of the valuation of [F3] on the part of the 6 respondents with Type 2 indiscrimination increases the conditioned observed values to £162.60 and £186.09 respectively.

C.2 Modeling the Output Distribution for *V*(F3) in Phase 1

A single mean and variance for the MAP ratio, R_F, may be found under the assumption of an Extended Piecewise Uniform distribution for MAP ratio using the methods of Appendix B. See **Table 14**. The action of the Extended Piecewise Uniform distribution for MAP ratio on the input distribution for $V(\text{F1})$ is then modeled using Equations (E.3) and (E.9) from Appendix E to calculate analytically the parameters of the output distribution, $V(\text{F3})$.

If the valuation, $V(\text{F1})$, and the MAP ratio, R_F, are assumed to be independent then the correlation coefficient, ρ, will be zero, and this produces predictions of £195.0 and £262.5 for the mean and standard deviation of $V(\text{F3})$. An exact match with the conditioned observed mean for $V(\text{F3})$, £157.29, is produced when $\rho = -0.31$, while the standard deviation is reduced to £248.0. The negative value of ρ suggests that there may be a negative correlation between input valuation, $V(\text{F1})$, and MAP ratio, but that it is very weak: $\rho^2 = 0.096$. Hence models based on independence between $V(\text{F1})$ and R_F are likely to give a reasonable representation of reality.

As an alternative, it is possible to simulate numerically the multiplication of sets of two random variables, assumed independent. The assumption is made that the conditioned input valuation, $V(\text{F1})$, with parameters given in **Table 13**, follows a lognormal distribution in the same way as the original distribution for $V(\text{F1})$. The valuations, $V(\text{F1})$, of 200 cohorts of 67 people are then simulated, with each being multiplied by a realization of one of the components of the Extended Piecewise Uniform Distribution for MAP ratio. 67 output valuations, $V(\text{F3})$, are then produced for each cohort. Cohort averages are then found and these are averaged further over the 200 simulations.

The results derived under the assumption that $V(\text{F1})$ and R_F are independent are shown in **Table 15**, where they are compared with the conditioned observed values for the 67 respondents who are either discriminating or suffer from Type 1 indiscrimination. **Figure 1** illustrates that there is a very close match between the calculations and the distribution for $V(\text{F3})$ based on the conditioned observed figures in terms of the numbers found in valuation intervals of £50. Shown are (i) the conditioned observed distribution for $V(\text{F3})$ among discriminating and Type 1 respondents, assumed lognormal, (ii) one realization of 67 people chosen to have a sample mean and sample standard deviation each close to the figures of **Table 15**, and (iii) the calculated distribution based on the simulation average figures from **Table 15**.

Justification for taking the distribution of $V(\text{F3})$ after conditioning to be lognormal is provided by the facts that (i) the distribution for the unconditioned version is approximately lognormal, (ii) the distribution for the corresponding valuation, $V(\text{R3})$, in Phase 2 (where no conditioning is required) is approximately lognormal and (iii) it is clear from **Figure 1** that the simulation's predicted output distribution for $V(\text{F3})$ is approximately lognormal.

A further way of judging the proximity of the results to the observation is found by carrying out 1,000 simulations of cohorts of 67 people when it is assumed that the distribution for $V(\text{F3})$ is lognormal with a mean of £193.3 and a standard deviation of £262.4, the simulation averages from **Table 15**. It is found that the sample

mean and sample standard deviation with this distribution both lie below the conditioned observed values, £157.29, and £184.78, on 93 occasions out of 1,000. This implies that the true values for the mean and standard deviation might be those suggested by the model but the values recorded in an experiment could still be as low as or lower than those recorded in the Caveat Investigator study with a probability of about 9.3%. This single-sided probability is well above the double-sided probability of between 1% and 5% below which the credibility of a model is conventionally regarded as suspect.

This suggests that there are no grounds for regarding the model incorporating the Extended Piecewise Uniform distribution for MAP ratio as inadequate for Phase 1. This provides validation for the finding that the average MAP ratio for discriminating Phase 1 respondents is close to 2.0.

C.3 Modeling the Output Distribution for V(F3) in Phase 2

The absence of Types 2 and 3 indiscrimination in Phase 2 means that it is possible to use the observed means and standard deviations listed in **Table 4** for both $V(\mathrm{R1})$ and $V(\mathrm{R3})$ without the conditioning required in Phase 1. The modeling then follows the pattern laid out in Section C.2.

A single mean and variance for the MAP ratio, R_R, is found under the assumption of an Extended Piecewise Uniform distribution for MAP ratio using the methods of Appendix B. See **Table 14**. The action of the Extended Piecewise Uniform distribution for MAP ratio on the input distribution for $V(\mathrm{R1})$ is then modeled using Equations (E.3) and (E.9) from Appendix E to calculate analytically the parameters of the output distribution, $V(\mathrm{R3})$.

An exact match with the observed mean for $V(\mathrm{R3})$, £145.1, is produced when $\rho = -0.21$, suggesting that there may be a negative correlation between input valuation, $V(\mathrm{R1})$, and MAP ratio. However such a correlation is very weak: $\rho^2 = 0.044$. It is concluded that models assuming independence between $V(\mathrm{R1})$ and R_R are likely to be adequate.

Table 16 shows the predicted parameters for $V(\mathrm{R3})$ from the analytic and the simulation methods under the assumption of independence between $V(\mathrm{R1})$ and R_R, and compares these with the observed parameters. The closeness of the match is illustrated by **Figure 2**, which compares the numbers in valuation intervals of £25 for (i) the observed distribution for discriminating and Type 1 respondents, (ii) one realization of 52 people chosen to have a sample mean and sample variance each close to the figures of **Table 16** and (iii) for the calculated distribution based on simulation averages from **Table 16**.

A further way of judging the proximity of the results to the observation is found by carrying out 1,000 simulations of cohorts of 52 people when it is assumed that the distribution for $V(\mathrm{R3})$ is lognormal with a mean of £168 and a standard deviation of £236.3, the simulation averages given in **Table 16**. It is found that the sample mean with this distribution lies below the observed values of sample mean, £145.14, and sample standard deviation, £199.6, on 218 occasions out of 1,000. This implies that the true values for the mean and standard deviation could be those suggested by the model but the values recorded in an experiment could still be as low as or lower than those recorded in the Caveat Investigator study with a probability of about 21.8%, substantially above the level at which the credibility of a model is conventionally regarded as suspect.

Again no good grounds have been found for regarding the model incorporating the Extended Piecewise Uniform distribution for MAP ratio as inadequate for Phase 2. This provides validation for the finding that the average MAP ratio for discriminating respondents in Phase 2 is close to 2.0.

Appendix D. Additional Validation of the Results of the Truncated Lognormal Distribution Model

Simulation may be used to find the distribution for $V(\mathrm{F3})$ from the distribution for $V(\mathrm{F1})$ when the Truncated Lognormal distribution is used to represent the MAP ratio.

D.1 Phase 1

Table 17 compares the distribution parameters found by simulating the decisions of 200 cohorts of 67 respondents using the Truncated Lognormal distribution to represent the behavior of the MAP ratio, R_F. The parameters are very similar to those of produced using the Extended Piecewise Uniform distribution, and to those observed. See **Figure 3**.

Based on a thousand simulations of groups of 67, where $V(\mathrm{F3})$ is assumed to follow a lognormal distribu-

tion with a mean of 194 and a standard deviation of 261.7, (see **Table 17**), both the sample mean and sample standard deviation will fall below the reported values of 157.3 and 184.8 in 95 cases in 1,000, implying that the true values for the mean and standard deviation could be those suggested by the model but the values recorded in an experiment could still be as low as or lower than those recorded in the Caveat Investigator study with a probability of about 9.5%, well above the level at or below which the credibility of a model is conventionally regarded as suspect.

The good match suggests that the model incorporating the Truncated Lognormal distribution for MAP ratio provides a reasonable model for Phase 1, thus providing further validation for the finding that the average MAP ratio for discriminating respondents in Phase 1 is close to 2.0.

D.2 Phase 2

Table 18 compares the distribution parameters found by simulating the decisions of 200 groups of 52. The results are reasonably close, both to the results based on an Extended Piecewise Uniform distribution for MAP ratio and to the observed data. See **Figure 4**.

Based on a thousand simulations of groups of 52, where the valuation of [R3] is assumed to follow a lognormal distribution characterized by a mean of 176.7 and a standard deviation of 250.9 (see **Table 18**), the sample mean and sample standard deviation of the realized sample of 52 will each be less than the reported values of 145.14 and 199.60 in 156 cases in 1,000, implying that the true values for the mean and standard deviation could be those suggested by the model but the values recorded in an experiment could still be as low as or lower than those recorded in the Caveat Investigator study with a probability of about 15.6%, again well above the probability at which the credibility of a model is conventionally called into question.

Once again the model is found not to be inconsistent with the observations, providing validation for the Truncated Lognormal distribution for MAP ratio, and, in particular, the finding that the average MAP ratio for discriminating respondents in Phase 2 is close to 2.0.

Appendix E. Mean and Standard Deviation of the Product of Two Random Variables

In general, if $Z = h(X,Y)$ is a function of two random variables X and Y with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, then a first-order Taylor series expansion about the point, (μ_X, μ_Y) gives:

$$Z = h(X,Y) \approx h(\mu_X, \mu_Y) + (X - \mu_X)\frac{\partial h(\mu_X, \mu_Y)}{\partial x} + (Y - \mu_Y)\frac{\partial h(\mu_X, \mu_Y)}{\partial y} \tag{E.1}$$

Hence

$$\text{var}(Z) = \sigma_Z^2 \approx \sigma_X^2 \left(\frac{\partial h(\mu_X, \mu_Y)}{\partial x}\right)^2 + \sigma_Y^2 \left(\frac{\partial h(\mu_X, \mu_Y)}{\partial y}\right)^2 + 2\sigma_{XY}\frac{\partial h(\mu_X, \mu_Y)}{\partial x}\frac{\partial h(\mu_X, \mu_Y)}{\partial y} \tag{E.2}$$

where $\sigma_{XY} = E((X - \mu_X)(Y - \mu_Y))$ is the covariance of X and Y, or

$$\text{var}(Z) \approx \sigma_X^2 \left(\frac{\partial h(\mu_X, \mu_Y)}{\partial x}\right)^2 + \sigma_Y^2 \left(\frac{\partial h(\mu_X, \mu_Y)}{\partial y}\right)^2 + 2\rho\sigma_X\sigma_Y\frac{\partial h(\mu_X, \mu_Y)}{\partial x}\frac{\partial h(\mu_X, \mu_Y)}{\partial y} \tag{E.3}$$

where ρ is the correlation coefficient relating the covariance to the standard deviations: $\sigma_{XY} = \rho\sigma_X\sigma_X$.

A second order Taylor series may be invoked in order to determine the mean:

$$Z = h(X,Y) \approx h(\mu_X, \mu_Y) + (X - \mu_X)\frac{\partial h(\mu_X, \mu_Y)}{\partial x} + (Y - \mu_Y)\frac{\partial h(\mu_X, \mu_Y)}{\partial y} + \frac{1}{2}(X - \mu_X)^2 \frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x^2}$$
$$+ \frac{1}{2}(Y - \mu_Y)^2 \frac{\partial^2 h(\mu_X, \mu_Y)}{\partial y^2} + \frac{1}{2}(X - \mu_X)(Y - \mu_Y)\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x \partial y}. \tag{E.4}$$

Applying the expectation operator produces:

$$E(Z) \approx h(\mu_X, \mu_Y) + \frac{1}{2}\left(\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x^2}\sigma_X^2 + 2\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x \partial y}\sigma_{XY} + \frac{\partial^2 h(\mu_X, \mu_Y)}{\partial y^2}\sigma_Y^2 \right)$$

$$= h(\mu_X, \mu_Y) + \frac{1}{2}\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x^2}\sigma_X^2 + \frac{1}{2}\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial y^2}\sigma_Y^2 + \rho\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x \partial y}\sigma_X\sigma_Y. \tag{E.5}$$

or

$$E(Z) \approx h(\mu_X, \mu_Y) + \frac{1}{2}\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x^2}\sigma_X^2 + \frac{1}{2}\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial y^2}\sigma_Y^2 + \rho\frac{\partial^2 h(\mu_X, \mu_Y)}{\partial x \partial y}\sigma_X\sigma_Y \tag{E.6}$$

In the case of a product, where $Z = XY$, then $h = xy$ and

$$\frac{\partial h}{\partial x} = y; \quad \frac{\partial^2 h}{\partial x^2} = 0; \quad \frac{\partial h}{\partial y} = x; \quad \frac{\partial^2 h}{\partial y^2} = 0; \quad \frac{\partial^2 h}{\partial x \partial y} = 1. \tag{E.7}$$

Evaluating those derivatives at (μ_X, μ_Y) and putting the results into Equation (E.3) gives:

$$\mathrm{var}(Z) \approx \mu_Y^2\sigma_X^2 + \mu_X^2\sigma_Y^2 + \rho\mu_X\mu_Y\sigma_X\sigma_Y \tag{E.8}$$

while applying the derivatives and the definition of $h(x, y)$ to Equation (E.6) gives:

$$E(Z) = \mu_X\mu_Y + \rho\sigma_X\sigma_Y \tag{E.9}$$

The equality is exact as written in Equation (E.9), rather than merely approximately true when the function, $h(X, Y)$, denotes a product. This follows from expanding the definition of covariance:

$$\mathrm{Cov}(X, Y) = \rho\sigma_X\sigma_Y = E\left((X - \mu_X)(Y - \mu_Y)\right) = E\left(XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y\right)$$

$$= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X\mu_Y = E(XY) - 2\mu_X\mu_Y + \mu_X\mu_Y = E(Z) - \mu_X\mu_Y. \tag{E.10}$$

from which Equation (E.9) follows.