

2006s-14

# **Estimates of Own Lethal Risks and Anchoring Effects**

*Olivier Armantier*

---

**Série Scientifique**  
*Scientific Series*

---

**Montréal**  
**Septembre 2006**

© 2006 *Olivier Armantier*. Tous droits réservés. *All rights reserved*. Reproduction partielle permise avec citation du document source, incluant la notice ©.

*Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.*



## **CIRANO**

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du Ministère du Développement économique et régional et de la Recherche, de même que des subventions et mandats obtenus par ses équipes de recherche.

*CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère du Développement économique et régional et de la Recherche, and grants and research mandates obtained by its research teams.*

### **Les partenaires / The Partners**

#### **Partenaire majeur**

Ministère du Développement économique,  
de l'Innovation et de l'Exportation

#### **Entreprises partenaires**

Alcan inc.  
Banque du Canada  
Banque Laurentienne du Canada  
Banque Nationale du Canada  
Banque Royale du Canada  
Bell Canada  
BMO Groupe financier  
Bombardier  
Bourse de Montréal  
Caisse de dépôt et placement du Québec  
Fédération des caisses Desjardins du Québec  
Gaz Métro  
Hydro-Québec  
Pratt & Whitney Canada  
Raymond Chabot Grant Thornton

#### **Autres partenaires gouvernementaux**

Industrie Canada  
Ministère des Finances du Québec  
Ville de Montréal

#### **Partenaires universitaires**

École Polytechnique de Montréal  
HEC Montréal  
McGill University  
Université Concordia  
Université de Montréal  
Université de Sherbrooke  
Université du Québec  
Université du Québec à Montréal  
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

Les cahiers de la série scientifique (CS) visent à rendre accessibles des résultats de recherche effectuée au CIRANO afin de susciter échanges et commentaires. Ces cahiers sont écrits dans le style des publications scientifiques. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

*This paper presents research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.*

# Estimates of Own Lethal Risks and Anchoring Effects\*

Olivier Armantier<sup>†</sup>

## Résumé / Abstract

Ce papier rapporte les résultats d'une expérience qui cherchait à tester si les individus percevaient correctement les risques mortels auxquels ils font face personnellement. Les résultats suggèrent que la perception des sujets est autant biaisée lorsqu'ils font des prédictions pour leur propre groupe d'âge ou pour la population entière (c'est-à-dire que les individus surestiment les risques rares, et sous-estiment les risques les plus communs). Cependant, l'hypothèse que les individus possèdent une meilleure connaissance de leurs propres risques ne peut être écartée complètement puisque les réponses du sondage sur leur propre groupe d'âge sont plus homogènes et mieux ordonnées. Finalement, nous démontrons que l'administration des sondages en série a pu générer des effets d'ancrage (*anchoring effects*) qui pourraient expliquer le fait que nos conclusions diffèrent sensiblement d'une étude précédente.

**Mots clés** : effet d'ancrage, espérance rationnelle, perception des risques, risques pour la santé et la sécurité

*The paper reports on an experiment testing whether agents perceive correctly the lethal risks they face personally. The results suggest that subjects exhibit comparable biases when making predictions for their own-age-cohort, or for the entire population (i.e. agents overestimate rare risks, and under-estimate common risks). The hypothesis that agents have better knowledge of their own risks, however, cannot be dismissed entirely, as responses in the own-age-cohort survey are more homogenous and better ordered. Finally, it is shown that administering surveys in succession can generate anchoring effects, which may explain why our conclusions differ markedly from a previous study.*

*Keywords: anchoring, health and safety hazard, rational expectation, risks perception*

**Codes JEL** : D8, C9

---

\* I would like to thank Soiliou Daw Namoro, Nicolas Treich, and Paul Pezanis-Christou for helpful discussions. I also would like to thank the editor and an anonymous referee for helpful comments. All remaining errors are mine.

Version août 2005, révisée et publiée en 2006 dans *The Journal of Risk and Uncertainty* 32, 1, 37-56.

<sup>†</sup> Université de Montréal, CIRANO, CIREQ, CRT, Département de sciences économiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, H3C 3J7, Canada;  
[olivier.armantier@umontreal.ca](mailto:olivier.armantier@umontreal.ca).

## 1. Introduction

During the last two decades, several experimental studies, both in economics and psychology, have identified a number of systematic biases in agents' perceptions of uncertain events.<sup>1</sup> In particular, experimental subjects have been found to consistently overestimate the risks of death due to low-probability causes, while they underestimate lethal risks from high-probability causes.<sup>2</sup> The presence of such biases may have serious implications for the regulation of risks. Indeed, most cost-benefit analyses of risk reduction projects rely on the public perception of risks, while policies aimed at curbing down risky behavior (e.g. smoking) are often devised around objective probabilities, rather than the population actual risk perception.<sup>3</sup>

Benjamin and Dougan (1997), however, challenged the consensus around the presence of systematic biases in agents appraisals of lethal risks. Instead, they suggested that, although they are typically poorly informed about the population-wide risks, agents may be able to evaluate accurately the lethal risks for their own-age-cohort, since these are the most relevant to them. Benjamin, Dougan and Buschena (2001) (hereafter BDB) conducted an experiment whose outcomes essentially supported this hypothesis. If robust, BDB's result would be reassuring, as it would imply that agents' decisions, and risk regulation policies are not skewed by biased perceptions.

In the present paper, the accuracy of own lethal risk perceptions is re-examined using a methodology slightly different from BDB's. Instead of successively administering two surveys to the same subject (i.e. one for the entire U.S. population, followed by one for his/her own-age-cohort), different sets of subjects are recruited to fill only one survey. In treatment 1, a first group of subjects must predict the number of deaths in the entire U.S. population for a series of risks; in treatment 2, a different set of subjects must evaluate the number of deaths within their own-age-cohort for the same lethal risks; finally, in

---

<sup>1</sup>See e.g. Kahneman, Slovic and Tversky (1982), Slovic (1987), Viscusi (1989), Viscusi (1993), or Camerer (1995).

<sup>2</sup>See e.g. Lichtenstein et al. (1978), Granger et al. (1983), Viscusi, Hakes and Carlin (1997), as well as Hakes and Viscusi (1997, 2004).

<sup>3</sup>See e.g. Viscusi (1998), or Salanié and Treich (2003).

treatment 3, the last group must fill the same survey for a different age-cohort.

The experimental outcomes in these three treatments exhibit the traditional biases. In addition, subjects' predictions appear to be as biased for the risks they face personally, for the risks faced by someone from a different age-cohort, and for the risks faced by the entire population. The predictions pertaining to the own-age-cohort, however, are found to be significantly more homogenous across subjects. Furthermore, the ranking of the own-age-cohort lethal risks appears to be significantly more accurate. In other words, our experiment provides evidence that, although they have biased perceptions, subjects may have better information about their own lethal risks.

These experimental outcomes contrast with those in BDB suggesting that own lethal risks predictions are virtually unbiased. To explain this difference, we conjecture that subjects in BDB's two surveys experiment may have anchored their answers in the second survey, on the answers they gave in the first survey. Three additional treatments are conducted to test this conjecture: treatment 4 is essentially similar to BDB's (i.e. the population-wide survey followed by the own-age-cohort survey); in treatment 5, the order of the surveys is reversed (i.e. the own-age-cohort survey is completed prior to the population-wide survey); after filling the population-wide survey, subjects in treatment 6 are asked to evaluate lethal risks for a different age-group. As explained further, we find evidence supporting the anchoring hypothesis. Indeed, subjects' predictions systematically exhibit the traditional biases in the first survey filled, while estimates for a specific age-group appear virtually unbiased when filled after the population-wide survey (i.e. treatments 4 and 6).

The remainder of the paper is structured as follow: the experimental design is introduced in section 2; the outcomes of the single survey treatments are discussed in section 3; the anchoring conjecture is presented and tested in section 4; finally, section 5 concludes.

## 2. The Experimental Design

The experiment was conducted with volunteer undergraduate students at the State University of New York at Stony Brook. The experiment consists in six experimental sessions, one for each treatment, and each with 35 subjects. All subjects within the same treatment were asked to meet in a large room, where they had to complete independently the survey(s). No time limit was imposed, but most surveys were completed within 20 minutes. Note that subjects were not allowed to participate in more than one session. Like in BDB's experiment, the average age of a subject is roughly 21, and the wide majority of respondents (92%) belongs to one of the 15-19 or 20-24 age brackets.<sup>4</sup>

The basic survey administered to subjects is essentially similar to BDB's.<sup>5</sup> Subjects were asked to evaluate the total number of deaths in 1999 for either the entire U.S. population or for a specific subpopulation due to the causes listed in the first column of Table 1.<sup>6</sup> Beyond the fact that subjects in treatments 1, 2 and 3 only have to fill a single survey, three main differences with BDB's experiment should be noted.

- First, the lethal risks in the present survey are the same as in BDB, except for three small risks, "Poisoning by Solid or Liquid", "Poisoning by Vitamins" and "Small Pox Vaccination" for which no data were available to us. Instead, we have replaced these risks by "Cholera", "Sail Boat Accidents" and "Accidental Poisoning due to Pesticides" which produced approximately the same number of deaths.
- Second, half of the subjects in BDB's experiment were provided with a high anchor (the number of deaths due to motor vehicle accidents), and the other half with a

---

<sup>4</sup>In fact, out of a total of 210 participants, 98 (47%) belong to the 15-19 age-group, 95 (45%) belong to the 20-24 age-group, and 17 (8%) are older than 24. The age distribution in BDB's study is very similar since, out of 50 students, 23 (46%) belonged to the 15-19 age-group, 22 (44%) belonged to the 20-24 age-group, and 5 (10%) were older than 24. Note also that the age distribution is roughly similar across our six treatments.

<sup>5</sup>The instructions given to the experimental subjects are similar to the one found in BDB, and therefore, they are not reproduced here.

<sup>6</sup>Actual 1999 population-wide and per-age-cohort figures may be found on the National Center for Health Statistics web-site at <http://www.cdc.gov/nchs/dataawh/statab/unpubd/mortabs/gmwiki10.htm>.

low anchor (the number of deaths due to electrocution). The object of the present study is not to analyze this form of anchoring. Therefore, following Hakes and Viscusi (2004), subjects in each treatment are only provided with a single anchor (the number of deaths due to motor vehicle accidents).

- Third, each subject in BDB’s experiment received a fixed payment of \$10. In the present experiment, subjects within the same treatment were ranked by order of accuracy for each cause of death. The rankings were then aggregated for each participant across all causes of death. The subject with the best aggregate ranking within his own treatment received \$50. The others did not receive any payment.<sup>7</sup> This payment mechanism offers two potential advantages: first, it is relatively inexpensive to conduct an experiment with several treatments and numerous participants; second, experimental economists generally believe that providing rewards contingent on performance provides subjects with a greater incentive to answer the surveys as accurately as possible.<sup>8</sup>

As we shall see, these three design modifications do not appear to generate any significant treatment effect compared to the experimental outcomes in BDB, and in other comparable studies (e.g. Lichtenstein et al. 1978, Granger et al. 1983, or Hakes and Viscusi 2004).

Let us now turn to the description of the six experimental treatments. The first two columns of Table 2 provides a summary of the design differences between the six different

---

<sup>7</sup>This payment method has been shown to be incentive compatible for subjects with uninformed priors (Ottaviani and Sørensen 2003). In addition, it has been argued that, to elicit beliefs, simple payment mechanisms, such as the one adopted here, should be preferred to proper scoring rules, such as the often used quadratic scoring rule. Indeed, scoring rules are often considerably more complex, and therefore, less transparent to subjects (Hogart 1987, Dufwenberg and Gneezy 2000, or Armantier and Treich 2005).

<sup>8</sup>If the subjects perceive the distribution of the reward as equiprobable among them, then one could question the salience of what would be their expected payoff in that case (i.e. \$1.4). Bolle (1990) however, argues that paying experimental subjects larger amounts of money with low probability is preferable to paying them smaller amounts with certainty, especially when the guaranteed reward is not tied to the subjects’ performance. This conjecture has since then been supported by several experimental studies including Straub and Murnighan (1995), Murnighan and Saxon (1998), Franck and Schulze (2000), Fershtman, Gneezy and Verboven (2005), as well as Armantier (2005).

treatments. The first three treatments consist in a single lethal risk survey for either the entire U.S. population (treatment 1), the subject’s own-age-cohort (treatment 2), or a different age-cohort (treatment 3).<sup>9</sup> The object of this first wave of treatments is to test whether subjects make more accurate predictions for the risks they face personally. In the last three treatments, the same subject has to fill successively two out of the three previous surveys (i.e. the population-wide followed by the own-age-cohort in treatment 4, the own-age-cohort followed by the population-wide in treatment 5, and the population-wide followed by the different-age-cohort in treatment 6). The object of this second wave of treatments is to test whether anchoring may explain why some of the experimental outcomes in the present paper differ from BDB’s.

### 3. Perception of Risks in Single Surveys

We can compare in Table 1 the actual number of fatalities for each cause of death with the geometric mean of the subjects’ responses in each of the six treatments.<sup>10</sup> These statistics suggest that subjects’ perceptions in the first three treatments exhibit the traditional biases. Indeed, small risks, like “measles” or “cataclysmic storms”, are often overestimated, while most common risks, like “suicide” or “hearth diseases”, are grossly underestimated.<sup>11</sup> To confirm statistically the generality of this observation, we consider an empirical methodology consisting in two steps. In step one, we regress the subjects’ individual responses in each survey on the actual number of deaths. In other words, the endogenous variable  $Y_{i,j}$  is defined as subject  $i$  individual response in a given survey for

---

<sup>9</sup>Since subjects are college students the different-age-cohort has been chosen to be the group of people between 40 and 44 years of age.

<sup>10</sup>We chose to present geometric means in Table 1 to reduce the influence of outliers, and to remain consistent with previous studies such as BDB. Note also that, due to space constraints, Table 1 only reports for treatments 2, 4 and 5 the responses of subjects who belonged to the cohort 15-19 or 20-24. Recall however, that these responses represent 92% of our subject pool. In addition, note that the responses of subjects older than 24 will be taken into consideration in the subsequent econometric analysis.

<sup>11</sup>Unlike Hakes and Viscusi (2004), we do not test for each cause of death in each survey whether the true number of fatalities is above or below a 95% confidence interval constructed with the subjects individual responses. Indeed, such a test has very low power here as it relies, depending on the survey considered, on a sample of 14 to 35 observations, including numerous outliers.



the  $j^{th}$  lethal cause, while the exogenous variable  $X_j$  is the corresponding actual number of deaths. In step two, we verify whether subjects' perceptions in a given survey are unbiased by testing the null hypothesis  $H_0 : (c, s) = (0, 1)$ , where  $c$  and  $s$  are respectively the constant and the slope coefficient in the previous regression.

The implementation of this methodology, however, presents three main challenges: i) the samples collected in each treatment include several outliers and high leverage points;<sup>12</sup> ii) there is strong evidence of heteroskedasticity; and iii) subjects' responses are non-negative. In other words, it is highly doubtful that any of the homoskedasticity, zero conditional mean, or normality assumptions apply to the distribution of the error term. As a result, the quality of standard estimation techniques may be altered significantly, and classical testing procedures may be rendered inappropriate in finite samples. For instance, it is well known that in such a context the least-square method produces unstable and inefficient predictions, while asymptotic theory provides poor approximations to the distributions of standard test statistics. The latter is a serious problem here since the conclusions of the present paper will rely heavily on the outcomes of statistical tests. To analyze properly the different samples, we must therefore strive to adopt the most appropriate estimation and testing procedures, as immune as possible from our data deficiencies.

Previous econometric analyses of lethal risk predictions have attempted to address some of the problems listed above by adopting a Log-Log specification, a median regression approach, data aggregation techniques (e.g. using geometric means instead of individual data), and/or traditional standard errors adjustments for heteroskedasticity (e.g. White 1980).<sup>13</sup> Although these techniques are appropriate in numerous applications, they typically fail to resolve simultaneously the three problems mentioned

---

<sup>12</sup>An outlier is a point for which the dependent variable is unusually high or low conditionally on the explanatory variable. In contrast, a high leverage point is an observation with an unusual set of explanatory variables.

<sup>13</sup>Another tempting approach to reduce the effect of outliers is to discard them, and apply standard estimation techniques to the data. This somewhat arbitrary strategy however, is not recommended in general as it yields inconsistent estimates of the standard errors, and biased approximations to the distributions of standard test statistics (see Huber 1981).

previously, and they do not address directly the question of testing (Doksum and Wong 1983, Rasmusen 1989). In contrast, the statistical literature on “Robust Estimation” has provided several techniques over the last three decades yielding more efficient estimates and more powerful tests when the data suffer from heteroskedasticity, non-normality, skewness, heavy-tails and/or outliers.

The theory of robust estimation typically considers three criteria when comparing estimation techniques.<sup>14</sup> First, the notion of “efficiency” measures how well the technique performs relative to least-square estimation when the data are “clean” (i.e. the error term is i.i.d. normal with zero conditional mean). An efficiency level above .95 is typically considered acceptable. Second, the “breakdown point” gives the minimum proportion of outliers which may produce an infinite bias. For instance, the least-square method has a breakdown point of  $1/n$  (where  $n$  is the sample size), meaning that one outlier is sufficient to render the estimate meaningless. In contrast, a robust estimator may reach a breakdown point of 0.5. Finally, the “influence function” is a measure of the influence of high leverage points on the estimates. Unlike least-square methods, robust estimators must have a bounded influence function. Although no robust estimation technique strictly dominates all others, the adjusted M-estimator proposed by Coakley and Hettmansperger (1993), and considered in the present paper, ranks high on all three criteria.<sup>15</sup> In addition, this estimator has been shown to outperform alternative robust estimators in terms of bias and efficiency when the error term is skewed and heavy-tailed, which seems to be the case with our data (see Wilcox 2005). Finally, the adjusted M-

---

<sup>14</sup> “Robust estimation” should not be confused with “robust standard errors” which is a more familiar notion in econometrics (see e.g. Eicker 1967, White 1980, or Newey and West 1987). With the latter, the parameters are estimated using traditional techniques, but their standard errors are estimated without relying on the homoskedasticity and/or no-autocorrelation assumptions. The estimated parameters, standard errors, and test statistics, however, are not protected against additional failures of the classical assumptions such as the presence of outliers, skewness or heavy tails. In contrast, the object of robust estimation, and more generally robust statistics, is to produce estimates, standard errors, and test statistics simultaneously robust to different forms of data deficiencies such as heteroskedasticity, non-normality, and outliers. For an introduction to robust statistics see Wilcox (2005).

<sup>15</sup> Coakley and Hettmansperger’s adjusted M-estimator has an efficiency level above 0.95, a breakdown point of 0.5, and a bounded influence function. In contrast, the popular median regression method, although it reduces the influence of outliers, still has a breakdown point of  $1/n$ , and it is not efficient at protecting against high leverage points (see Wilcox 2005).

estimator has standard properties since it is asymptotically normal with a  $\sqrt{n}$  rate of convergence.<sup>16</sup>

As for the test statistics, the traditional data transformations previously mentioned do not fully resolve the heteroskedasticity, skewness and non-normality problems.<sup>17</sup> This therefore prevents us from using classical asymptotic tests. Instead, we rely on the theory of robust hypothesis testing that has been developed to deal in finite samples with distributions departing from the classical assumptions. The basic strategy behind any robust test consists in three steps: first, one randomly generates  $K$  bootstrap samples from an initial set of robustly estimated parameters; second, the parameters are re-estimated with each bootstrap sample using a robust estimation technique; third, based on the  $K$  bootstrap estimates, one constructs a confidence interval which will be used to accept or reject the null hypothesis. The main difference between robust tests essentially resides in the way the bootstrap samples are generated. The objective of a bootstrap resampling technique is to generate data from a distribution as close as possible from the true (but unknown) data generating process. These simulations are especially difficult in finite samples when the data are highly skewed or heteroskedastic. Horowitz (2000), and Flachaire (2005a,b) show that in such a context a test based on the “Wild Bootstrap” resampling technique introduced by Liu (1988) typically outperforms alternative bootstrap and asymptotic testing procedures, including the heteroskedasticity consistent approach proposed by Eicker (1967) and White (1980). In addition, Horowitz (2000) argues that the wild bootstrap technique is numerically superior as it guarantees that the conditional distribution of the bootstrapped error term has mean zero.

To summarize, the data in the present paper are estimated with the robust estimation method developed by Coakley and Hettmansperger (1993), and the distributions

---

<sup>16</sup>S-PLUS code to implement the Coakley and Hettmansperger’s adjusted M-estimator may be found in Wilcox (2005).

<sup>17</sup>For instance, after a Logarithmic transformation of the data in each of the surveys we collected both the Shapiro-Wilk-Kolmogorov and Jarque-Bera tests strongly reject the normality of the error terms (the  $P$ -values range from  $2.373E^{-7}$  to  $8.330E^{-11}$ ), the White test rejects homoskedasticity (the  $P$ -values range from  $6.374E^{-4}$  to  $2.501E^{-8}$ ), and the null hypothesis of zero skewness is systematically rejected (the  $P$ -values range from  $4.920E^{-5}$  to  $8.212E^{-9}$ ).

of the test statistics are approximated by bootstrap using this robust estimator and simulated samples generated with the wild bootstrap technique developed by Davidson and Flachaire (2001).

The outcomes of the regressions conducted with the data collected in each survey may be found in Table 3.<sup>18</sup> Observe first that subjects' responses seem far from accurate in the first three treatments since the estimated slopes are all close to 0, instead of 1. In fact, a robust test of the unbiased perception hypothesis,  $H_0 : (c, s) = (0, 1)$ , is systematically rejected at the usual significance levels in all three treatments (the  $P$ -values in treatments 1, 2, and 3 are respectively  $2.351E^{-11}$ ,  $8.510E^{-9}$ , and  $3.406E^{-8}$ ). In addition, the constant and slope coefficients are statistically indistinguishable in the first three treatments.<sup>19</sup> In other words, subjects appear to exhibit the same traditional biases when predicting lethal risks for their own-age-cohort, for a different age-group, or for the entire population.

These results are illustrated in Figures 1 to 3 where the geometric means of subjects' responses are plotted along with the regression lines.<sup>20</sup> Note, however, that subjects' predictions appear slightly more homogenous for their own-age-cohort (especially for frequent causes of death), since the aggregate responses seem to bracket the regression

---

<sup>18</sup>To test the robustness of our results, we have also considered alternative specifications and estimation methods such as a robustly estimated individual fixed effect specification, a median regression, and an OLS regression with a Log-Log specification. Although slightly different and often more noisy, these estimation results are globally consistent with the conclusions presented in this paper. In addition, according with Hakes and Viscusi (2004), we find that adding an individual fixed effect does not improve significantly the fit of the model. The alternative estimation results may be found on the author's website at [http://www.sceco.umontreal.ca/liste\\_personnel/armantier/](http://www.sceco.umontreal.ca/liste_personnel/armantier/).

<sup>19</sup>To compare the regression parameters in two different surveys, we adopt the dummy variable alternative to the Chow test. In doing so, we face an additional econometric problem. Indeed, we can see in Table 1 that subjects responses are between 10 and 1,000 time larger in the population-wide survey compared to the surveys for a specific age-group. To deal with this severe scaling and heteroskedasticity problem we must rely once again on robust estimation and testing methods. The resulting  $P$ -values are 0.122 when we merge the samples collected in treatments 1 and 2, 0.190 when we merge the samples collected in treatments 1 and 3, and 0.273 when we merge the samples collected in treatments 2 and 3.

<sup>20</sup>To facilitate the comparison, we have adopted BDB's approach to construct all the figures in the paper. In particular, the own-age-cohort geometric means have been constructed with weights equal to the number of respondents in each group. Note also that the regression lines appear to be non-linear since a logarithmic scale is used.

line more tightly in Figure 2.<sup>21</sup> In other words, although biased, subjects' predictions of their own lethal risks appear somewhat less volatile. To test this hypothesis, the variance of the heteroskedastic error term is modelled as  $\sigma_i^2 = \sigma^2 X_i^\gamma$ , where  $(\sigma, \gamma)$  is estimated consistently from the outcomes of the previous regressions (see Table 3). Using a robust testing procedure, the pair  $(\sigma, \gamma)$  is found to be significantly smaller in treatment 2 than in the other two treatments, thereby confirming that subjects' predictions are more homogenous for the risks they face personally.<sup>22</sup>

This result may be considered of interest in several respects. First, policies aimed at preventing risks, or at correcting the public misperceptions, may be more efficiently implemented when perceptions, although biased, are homogenous. Indeed, a single policy directed toward the representative agent may be devised when the population is homogenous. In contrast, several different (and possibly contradictory) targeted policies may be necessary to influence each type of heterogenous agents. Second, cost-benefit analyses are likely to be more precise when the population surveyed is homogenous, thereby better representing the public true willingness to pay for a risk reduction project. Third, evidence of homogenous perceptions of own risks may be considered reassuring as it may be argued that extreme perceptions by a few may lead to extreme behaviors negatively impacting the entire population. For instance, if a person with a high probability of being contaminated by a virus completely ignores this risk, then the spread of the virus to the population may be facilitated. Likewise, severe overestimation of low risks by some may raise their willingness to pay for insurance, and thereby increase the price paid by others.

Harrison and Rutström (2005) propose an alternative approach to evaluate whether agents have better knowledge of the risks familiar to them. Instead of looking at the number of deaths per cause, the idea is to analyze how subjects order risks in terms of

---

<sup>21</sup>Recall that on a graph with a logarithmic scale, what might appear as small differences may in reality correspond to very large differences in absolute terms. This remark is particularly relevant when comparing estimates for frequent causes of death.

<sup>22</sup>The  $P$ -value is 0.203 when we merge the samples collected in treatments 1 and 2, and 0.191 when we merge the samples collected in treatments 2 and 3.

annual fatalities. Indeed, it is possible that although subjects' perceptions for different age-groups are equally biased, their ordering of risks is more accurate for their own-age-cohort. To test this hypothesis, we first rank in descending order the actual number of deaths per risk, as well as each subject's responses. Then, we define the ranking error  $Z_{ij}$  as the difference (in absolute value) between the true rank of lethal risk  $j$ , and the rank inferred from subject  $i$  prediction for risk  $j$ . As a result, we obtain a new sample for each survey composed of integers potentially distributed between 0 (i.e. no ranking error) and 38. The means and standard deviations of  $Z_{ij}$  presented in the last row of Table 3 suggest that the sample of ranking errors has a lower location in treatment 2. This observation is supported by a Mann-Whitney test indicating that the median is significantly lower in treatment 2 than in treatment 1 ( $P$ -value=0.112), and in treatment 3 ( $P$ -value=0.078). In other words, and according with Harrison and Rutström (2005), subjects appear to be more accurate when ranking mortality risks for their own-age-cohort.

In summary, the experimental outcomes in the first three single survey treatments indicate that subjects' beliefs about lethal risks are equally biased for their own-age-group, for a different age-group, and for the entire population. Perceptions of own-risks, however, are found to be more homogenous, and subjects appear to order more accurately risks they face personally. These results are partially at odds with BDB's experiment whose outcomes suggest that own-age-cohort perceptions are virtually unbiased. The object the next section is to explore why our conclusions differ from BDB's.

#### **4. Anchoring Effects in Consecutive Surveys**

BDB's experiment consisted in two surveys that the same subject had to complete successively. In the first survey, the subject was asked to evaluate a series of population-wide lethal risks, while in the second survey, he had to estimate the number of deaths from the same lethal risks within his own-age-cohort. The population-wide survey yielded the traditional biases for low and high causes of death. In contrast with the results presented

in section 3, BDB found the own-age-cohort estimates to be significantly more accurate. In particular, high lethal risks were correctly estimated. Subjects appraisals, however, were not perfectly unbiased since small risks were still slightly overestimated.

Our contention is that the outcomes in BDB's experiment may be explained by the fact that subjects anchored their answers in the second survey, on the answers they gave in the first survey.<sup>23</sup> This hypothesis is plausible since BDB's subjects were first asked an evaluation for the entire population, and then for a subpopulation (their own-age-cohort). If the responses in the first survey act as anchors, then subjects will make a prediction for their own-age-cohort starting from their initial population-wide estimates, but the downward adjustment will be insufficient. As a result, own-age-cohort predictions should be larger than in treatment 2 where subjects, who could not be affected by anchoring since they only had to fill the own-age-cohort survey, exhibited the traditional biases. Anchor theory therefore makes two predictions under BDB's two surveys design: First, the traditional biases should be observed in the first survey (i.e. the population-wide); and second, although the overestimation of rare risks should still be present and potentially amplified, the underestimation of high causes of death should be attenuated and may possibly vanish in the second survey (i.e. the own-age-cohort).<sup>24</sup> This is precisely the outcome observed in BDB's experiment.

To test the anchoring hypothesis, three additional treatments have been conducted:

---

<sup>23</sup>Anchoring has been found to be a robust psychological factor resulting from insufficient adjustments from an initial, arbitrary, and sometimes completely uninformative value (see e.g. Tversky and Kahneman 1974, Camerer 1995, as well as Chapman and Johnson 2002 for surveys). For instance, the outcome of a wheel of fortune has been shown to influence subjects' evaluations of the percentage of African countries in the United Nations (see Tversky and Kahneman 1974).

<sup>24</sup>To illustrate, consider two risks, a common and a rare risk, resulting respectively in 10 and 5,000 annual deaths in a given subject age-cohort. If, like in treatment 2, the subject is only asked for the number of deaths in his own-age-cohort, he might answer 25 and 3,000, thereby exhibiting the traditional overestimation of small risks and underestimation of high risks. If instead, like in BDB's two surveys design, the subject first stated that the number of fatalities in the entire population is 100 and 100,000, then anchor theory predicts that the subject estimates in the second survey for his own-age-group should be close to, but should exceed (25; 3,000). Indeed, according to anchor theory, the subject will start from his population estimates, and then adjusts toward what would otherwise be his own-age-cohort predictions, but insufficiently. For instance, if the subjects estimates in the second survey are (50; 4,500), then observe that the traditional bias almost vanishes for the high risk, while it is slightly amplified for the low risk. In other words, because of anchoring, the subject gives the illusion of nearly accurate perceptions in the second survey.

treatment 4 is identical to BDB’s (i.e. a population-wide survey, followed by an own-age-cohort survey); in treatment 5, the order of the surveys is reversed (i.e. the own-age-cohort survey is completed prior to the population-wide survey); finally, treatment 6 is similar to BDB, except that the second survey asks subjects to evaluate the number of deaths for the 40-44 age group. These treatments, as well as the predicted outcomes under the anchoring hypothesis are summarized in the last three rows of Table 2.

Let us first concentrate on treatment 4. The main objective here is to verify whether the three experimental design modifications described in section 2 (i.e. replacing three lethal risks, providing only the high anchor, and changing the payment mechanism) introduced a treatment effect which could explain why the outcomes in the first three treatments differ from those obtained by BDB. Table 3 suggests that subjects’ perceptions of the population-wide lethal risks are essentially as biased in treatments 1 and 4. In fact, the joint null hypothesis that the pair  $(c, s)$ , the constant and slope coefficients, are equal in the population-wide regressions of treatments 1 and 4 cannot be rejected at the usual significance level ( $P$ -value=0.302). This result confirms, as one may have expected, that predictions in the population-wide survey are similar whether or not the subjects have to fill a second survey.<sup>25</sup> In contrast, Table 3 indicates that the regression outcomes differ when the own-age-cohort survey is filled alone ( $(c, s) = (19.28, 0.20)$  in treatment 2), or after the population-wide survey ( $(c, s) = (29.82, 0.85)$  in treatment 4). In fact, although both constants are statistically insignificant, the slope in the own-age-cohort regression is found to be significantly larger in treatment 4 than in treatment 2 ( $P$ -value=0.618). This result implies that while the lowest risks are estimated with similar precision, the highest risks are further underestimated in treatment 2 compared to treatment 4. In other words, the two surveys design seems to create the illusion that

---

<sup>25</sup>Since no significant difference may be detected between the responses to the population-wide surveys in treatments 1 and 4, the regression model has been re-estimated after merging these two samples. No significant difference has been found compared to the results reported in Table 3. Likewise, no significant difference is found when we pool responses to the population-wide surveys in treatments 1, 4 and 6, and when we pool responses to the own-age-cohort surveys in treatments 2 and 5. The outcomes of these regressions are not presented here, but they are available on the author’s website at [http://www.sceco.umontreal.ca/liste\\_personnel/armantier/](http://www.sceco.umontreal.ca/liste_personnel/armantier/).



subjects have better perceptions of the lethal risks familiar to them. Note, however, that although close, the constant and slope coefficients in treatment 4 are significantly different from  $(0, 1)$ , thereby rejecting the perfect perception hypothesis ( $P$ -value= $2.103E^{-4}$ ).

A comparison of Figures 1 and 4 illustrates the fact that the population-wide estimates are similar in treatments 1 and 4, while a comparison of Figures 2 and 5 illustrates the fact that the own-age-cohort regression line is closer to the diagonal in treatment 4 than in treatment 2. In other words, own-age-cohort predictions for frequent causes of death appear much more accurate in Figure 5 than in Figure 2. Note, however, that respondents in Figure 5 still appear to overestimate small causes of death. As explained previously, these results are consistent with the anchor predictions described in Table 2. In addition, observe that the experimental outcomes in treatment 4 are essentially similar to BDB's. Therefore, the differences between BDB's experiment and the outcomes in the first three treatments of this paper cannot be imputed to the three design modifications described in section 2.

Let us now turn to treatment 5. Table 3 indicates that switching the order in which the surveys were completed has an impact on the subjects' predictions. Indeed, in contrast with treatment 4, the own-age-cohort estimates are now consistent with those obtained in the single survey treatment (i.e. treatment 2). This observation is confirmed statistically by a robust test of the null hypothesis that the constant and slope coefficients are the same in treatments 2 and 5 ( $P$ -value= $0.196$ ). Moreover, although the constants cannot be statistically distinguished, the slope in the population-wide regression is found to be significantly smaller in treatment 5 than in treatment 1 ( $P$ -value= $0.555$ ).

Figures 6 and 7 illustrate the fact that switching the order of the surveys yields the traditional biases both in the population-wide and the own-age-cohort surveys. In addition, a comparison of Figures 1 and 7, indicates a slight increase in the bias for frequent causes of death in the population-wide survey, thereby illustrating the fact the slope is significantly lower in treatment 5 than in treatment 1. These results are once again consistent with the anchor predictions in Table 2. Indeed, subjects' perceptions of lethal risks for their own-age-cohort exhibit the traditional biases when the own-age-

cohort survey is completed first. In addition, if the own-age-group estimates act as anchors in treatment 4, then subjects should adjust insufficiently in the population-wide survey, thereby reinforcing the underestimation of common risks observed in treatment 1. This is precisely what we observed.

Let us now examine the experimental results in treatment 6, in which subjects had to evaluate lethal risks for the entire population, and then for a different age-group. Table 3 indicates that the regressions results are very similar in treatments 6 and 4. In fact, the parameters  $(c, s)$  in the first (respectively second) surveys of treatments 6 and 4 cannot be distinguished statistically.<sup>26</sup> The similitude of the regression lines may be appreciated graphically by comparing Figures 4 and 8, and Figures 5 and 9. The two surveys design in treatments 4 and 6 therefore gives the illusion that respondents are able to provide nearly unbiased estimates for their own, and for a different age-cohort. This result is again compatible with the anchor hypothesis. Indeed, the quality of the estimates in the second survey of treatments 4 and 6 are similar, as subjects appear to anchor on the answers given in the population-wide survey they first filled.

## 5. Conclusion

The presence of systematic biases in the perceptions of lethal risks has generated several debates about the salience of the result, and the origin of the phenomenon. In particular, Benjamin and Dougan (1997) contest the generality of this conjecture. Instead, they suggest that the biases observed in population-wide surveys may be explained by the fact that agents are only aware of their own risks, since these are the most relevant to them. A couple of recent experiments appear to provide some support to this hypothesis (Benjamin et al. 2001, and Harrison and Rutström 2005).

In the present paper, we find that subjects exhibit the traditional systematic biases when they are asked to answer a single survey for either their own-age-cohort, a different

---

<sup>26</sup>The  $P$ -value when we pool the data collected in the first (respectively, second) surveys of treatments 4 and 6 is 0.182 (respectively, 0.203)

age-cohort, or the entire population. The experimental outcomes in the first three treatments, however, also indicate that, although biased, perceptions of own lethal risks are significantly more homogenous across subjects. This result is interesting as it may be argued that agents with homogenous beliefs may be more efficiently informed and regulated. In addition, and according with Harrison and Rutström (2005), it appears that subjects are able to order risks by fatalities more accurately for their own-age-group.

In conclusion, it appears that biased perceptions of lethal risks is a salient and robust phenomenon, but BDB's hypothesis that agents have better information about the risks they face personally cannot be unambiguously dismissed. These conclusions are consistent with Viscusi and O'Connor (1984), as well as Gerking, deHaan and Schulze (1988) who find that workers have slightly more accurate evaluations of the risks related to their own job.

Finally, the paper suggests that the experimental outcomes observed by BDB may be explained by anchor theory. More precisely, we find that, when asked to evaluate the number of deaths for two different populations in two consecutive surveys, respondents anchor their answers in the second survey on the answers they gave in the first survey. As a result, it appears that subjects may have given the illusion of unbiasedness for their own-age-cohort in BDB's study, as they were first asked about population-wide estimates.

## References

- Armantier, Olivier, 2005, "Do Wealth Differences Affect Fairness Consideration?", *International Economic Review*, forthcoming.
- Armantier, Olivier and Nicolas Treich, 2005, "Overbidding in Independent Private Values Auctions and Misperception of Probabilities", Mimeo, Universite de Montreal.
- Benjamin, Daniel K. and William R. Dougan, 1997, "Individuals' Estimates of the Risks of Death: Part I—A Reassessment of the Previous Evidence", *Journal of Risk and Uncertainty*, 15, 115-133.
- Benjamin, Daniel K., Dougan William R. and David Buschena, 2001, "Individuals' Estimates of the Risks of Death: Part II—New Evidence", *Journal of Risk and Uncertainty*, 22, 35-57.
- Bolle, Friedel, 1990, "High Reward Experiments without High Expenditure for the Experimenter?", *Journal of Economic Psychology*, 11(2), 157-167.

Camerer, Colin, 1995, "Individual Decision Making", in: Kagel, J.H. & Roth, A.E. (eds.): "Handbook of Experimental Economics", Princeton, NJ: Princeton University Press, 587-703.

Chapman, Gretchen and Eric J. Johnson, 2002, "Incorporating the Irrelevant: Anchors in Judgments of Belief and Value", In T. Gilovich, D. W. Griffin, D. Kahneman (Eds.), "The Psychology of Intuitive Judgment: Heuristics and Biases", New York: Cambridge University Press.

Coakley, Clint and Thomas Hettmansperger, 1993, "A Bounded Influence, High Breakdown, Efficient Regression Estimator", *Journal of the American Statistical Association*, 88, 872-880.

Davidson, Russel and Emmanuel Flachaire, 2001, "The Wild Bootstrap, Tamed at Last", Mimeo, IER#1000, Queen's University.

Doksum, Kjell and Chi-Wing Wong, 1983, "Statistical Tests Based on Transformed Data", *Journal of the American Statistical Association*, 78, 411-417.

Dufwenberg, Martin and Uri Gneezy, 2000, "Measuring Beliefs in an Experimental Lost Wallet Game", *Games and Economic Behavior* 30, 163-182.

Eicker, Friedhelm, 1967, "Limit Theorems for Regressions with Unequal and dependent Errors", *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1, 59-82.

Fershtman, Chaim, Uri Gneezy and Frank Verboven, 2005, "Discrimination and Nepotism: The Efficiency of the Anonymity Rule", *Journal of Legal Studies*, forthcoming.

Flachaire, Emmanuel, 2005a, "More Efficient Tests Robust to Heteroskedasticity of Unknown Form", *Econometric Reviews*, 24(2), 219-241.

Flachaire, Emmanuel, 2005b, "Bootstrapping Heteroskedastic Regression Models: Wild Bootstrap vs. Pairwise Bootstrap", *Computational Statistics & Data Analysis*, 49(2), 361-376.

Frank, Björn and Gunther Schulze, 2000, "Does Economics Make Citizens Corrupt?", *Journal of Economic Behavior and Organization*, 43, 101-113.

Gerking, Shelby, Mennor deHaan and William Schulze, 1988, "The Marginal Value of Job Safety: A Contingent Valuation Study", *Journal of Risk and Uncertainty*, 1, 185-199.

Granger, Morgan et al., 1983, "On Judging the Frequency of Lethal Events: A Replication", *Risk Analysis*, 3, 11-16.

Hakes, Jahn K. and W. Kip Viscusi, 1997, "Mortality Risk Perceptions: A Bayesian Reassessment", *Journal of Risk and Uncertainty*, 15, 135-150.

Hakes, Jahn K. and W. Kip Viscusi, 2004, "Dead Reckoning: Demographic Determinants of the Accuracy of Mortality Risk Perceptions", *Risk Analysis*, 24(3), 651-664.

Harrison, Glenn W. and Elisabet E. Rutström, 2005, "Eliciting Subjective Beliefs About Mortality Risk Orderings", *Environmental and Resource Economics*, forthcoming.

Hogarth, Robin, 1987, "Judgment and Choice: The Psychology of Decision," 2nd ed. New York:Wiley.

- Horowitz, Joel, 2000, "The Bootstrap", in "Handbook of Econometrics", Volume 5, J. J. Heckman and E. E. Leamer (eds), Elsevier Science.
- Huber, Peter, 1981, "Robust Statistics", Wiley, NY.
- Kahnemann, Daniel, Slovic Paul and Amos Tversky, 1982, "Judgement Under Uncertainty: Heuristics and Biases", Cambridge: Cambridge University Press.
- Lichtenstein, Sarah, Paul Slovic, Baruch Fischhoff, Mark Layman and Barbara Combs, 1978, "Judged Frequency of Lethal Events", *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 551-578.
- Liu, Regina, 1988, "Bootstrap Procedure under some non-i.i.d. Models", *Annals of Statistics*, 16, 1696-1708.
- Murnighan, Keith and Michael Saxon, 1998, "Ultimatum Bargaining by Children and Adults", *Journal of Economic Psychology*, 19, 415-445.
- Newey, Whitney and Kenneth West, 1987, "Hypotheses Testing with Efficient Method of Moments Testing", *International Economic Review*, 28, 777-787.
- Ottaviani, Marco and Peter Sørensen, 2003, "Forecasting and Rank-Order Contests", Mimeo, London School of Economics.
- Salanié, François and Nicolas Treich, 2003, "Regulating an Agent with Different Beliefs", Mimeo, Université de Toulouse.
- Rasmussen, Jens, 1989, "Data Transformation, Type I Error Rate and Power", *British Journal of Mathematical and Statistical Psychology*, 42, 203-213.
- Slovic, Paul, 1987, "Perception of Risk", *Science*, 236, 280-285.
- Slovic, Paul, Baruch Fischhoff and Sarah Lichtenstein, 1982, "Facts versus Fears: Understanding Perceived Risk", in D. Kahneman, P. Slovic, and A. Tversky, eds., "Judgement Under Uncertainty: Heuristics and Biases", Cambridge: Cambridge University Press.
- Straub, Paul and Keith Murnighan, 1995, "An Experimental Investigation of Ultimatum Games: Information, Fairness, Expectations, and Lowest Acceptable Offers", *Journal of Economic Behavior and Organization*, 27(3), 345-364.
- Tversky, Amos and Daniel Kahneman, 1974, "Judgments under Uncertainty: Heuristics and Biases", *Science*, 185, 1124-1131.
- Viscusi, W. Kip and Charles O'Connor, 1984, "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?", *American Economic Review*, 74, 942-956.
- Viscusi, W. Kip, 1989, "Prospective Reference Theory: Toward an Explanation of the Paradoxes", *Journal of Risk and Uncertainty*, 2(3), 235-264.
- Viscusi, W. Kip, 1993, "The Value of Risks to Life and Health", *Journal of Economic Literature*, 31, 191-246.
- Viscusi, W. Kip, 1998, "Rational Risk Policy", Oxford: Oxford University Press.
- Viscusi, W. Kip, Jahn K. Hakes and Alan Carlin, 1997, "Measures of Mortality Risks", *Journal of Risk and Uncertainty*, 14, 213-233.
- White, Halbert, 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, 817-838.

Wilcox, Rand, 2005, "Introduction to Robust Estimation and Hypothesis Testing",  
San Diego, CA: Academic Press.

**Table 1**  
**Number of Deaths per Cause and Age in the US in 1999**

Cause of Death	Actual Number of Deaths in 1999				Geometric Means of Subjects' Responses in Treatment											
	Total	15-19	20-24	40-44	1	2	3	4			5			6		
					Total	15-19	20-24	40-44	Total	15-19	20-24	15-19	20-24	Total	Total	40-44
Small Pox	0	0	0	0	80.9	4.7	9.7	70.8	103.1	21.1	15.5	6.5	2.2	68.5	183.0	36.7
Cholera*	1	0	0	0	122.3	11.2	7.7	26.9	240.9	38.9	24.9	5.3	13.8	33.7	139.9	33.0
Measles	2	0	0	0	126.8	33.0	16.1	6.7	223.3	42.1	19.0	30.9	29.0	113.5	207.3	30.1
Botulism	4	0	0	0	524.0	57.7	146.0	24.9	524.3	73.9	177.2	99.0	70.1	309.8	391.3	48.5
Sail Boat Accident*	6	0	1	0	432.0	39.3	9.9	19.1	584.6	80.7	27.3	51.6	14.6	865.7	1,136.9	92.8
Fireworks	7	1	0	0	381.6	7.1	21.0	16.1	225.1	30.2	27.1	10.5	12.6	81.4	95.4	9.4
Accidental Poisoning due to pesticides*	12	1	1	0	460.9	22.6	7.7	10.0	600.6	60.7	37.7	27.9	19.3	133.6	417.6	53.4
Flood	15	0	2	1	772.3	8.2	12.9	11.5	946.9	47.4	30.8	17.6	18.9	334.8	700.8	74.8
Contact with Venomous Animals	17	0	1	4	321.5	11.7	7.0	81.9	484.1	35.6	39.6	15.2	13.4	252.8	417.0	79.9
Syphilis	33	0	0	0	155.9	4.3	4.9	17.3	247.2	18.3	22.9	6.3	2.9	177.8	453.5	94.5
Lightning	64	4	6	7	510.2	7.7	14.0	15.7	731.0	21.5	48.0	13.5	24.0	323.4	619.6	24.7
Cataclysmic Storm	129	2	4	17	1,686.4	25.0	21.2	43.4	1,029.2	52.2	80.4	57.4	33.1	579.9	1,444.1	219.5
Nonvenomous Animals	148	1	3	7	274.3	14.9	25.0	28.4	398.8	68.4	42.8	38.0	28.1	203.8	288.2	54.5
Collision between Car and Train (car Occupants)	160	22	23	12	1,062.8	13.6	10.6	103.6	407.0	44.9	36.9	13.0	22.6	287.9	751.6	153.1
Polio (Sequelae)	263	0	0	0	2,434.7	11.0	14.1	19.7	1,624.6	56.4	86.3	6.6	19.8	950.5	1,248.3	125.2
Appendicitis	331	2	3	10	884.8	9.6	34.0	19.8	951.6	99.7	79.4	17.3	33.0	762.8	1,230.1	57.3
Pregnancy and Childbirth	406	46	51	60	1,723.3	107.3	120.1	32.4	945.2	31.5	201.7	57.7	95.1	435.9	1,942.5	90.0
Electrocution	437	32	45	60	2,964.0	86.1	211.6	76.1	1,742.0	228.0	120.0	116.3	198.7	1,978.4	1,532.3	216.0
Exposure to Excessive Cold	598	5	7	32	563.7	3.4	31.1	16.7	393.7	26.4	94.0	3.8	34.6	162.8	329.0	62.7
Firearms Accidents	824	111	140	74	6,344.7	75.5	34.5	178.5	11,754.8	216.1	308.7	225.8	40.9	5,942.3	8,761.9	354.6
Tuberculosis	1,116	1	10	36	3,567.0	12.9	25.7	37.8	1,129.9	12.2	112.3	10.8	49.9	1,210.9	1,608.4	61.2
Smoke, Fire and Flames	3,348	97	104	246	2,254.4	26.5	32.5	165.4	4,170.4	80.5	97.9	54.2	43.7	1,178.4	3,166.1	436.7
Accidental Drowning	3,529	289	358	250	3,253.8	22.4	70.2	58.8	2,080.8	185.8	240.8	160.4	42.5	1,261.4	1,661.8	184.8
Asthma	4,172	78	72	201	1,745.6	96.1	69.8	344.8	2,385.8	165.1	226.0	156.5	163.9	1,476.1	1,984.6	317.7
Hepatitis	4,853	3	11	518	1,971.5	6.9	28.6	48.9	3,885.2	49.6	35.2	17.5	120.3	6,528.3	2,642.8	167.8
Stomach Cancer	12,711	5	8	316	9,476.8	27.7	44.4	253.0	6,907.4	96.3	51.9	16.0	77.2	4,082.0	6,952.1	438.9
Accidental Falls	13,162	79	163	358	1,546.3	21.1	28.5	35.1	719.3	46.5	114.3	13.0	33.1	319.1	467.9	86.3
Homicide	16,889	1,889	3,109	1,672	19,537.0	290.7	643.2	573.3	25,369.9	1,811.1	2,200.4	821.4	462.4	8,186.9	22,381.1	1,522.1
Emphysema	17,787	0	1	78	9,687.1	12.3	15.4	36.0	5,227.5	18.5	55.8	7.2	33.8	5,696.5	6,201.4	266.8
Leukemia	21,014	200	282	422	8,578.1	19.7	108.0	47.7	10,597.9	58.8	124.9	92.3	141.6	6,528.3	7,791.2	316.0
Suicide	29,199	895	3,006	3,825	4,097.9	104.7	76.5	326.9	2,538.8	368.0	426.3	115.3	190.7	1,640.7	3,531.8	1,587.5
Breast Cancer	41,528	3	10	1,755	19,084.5	45.0	74.8	231.3	26,951.6	123.2	487.0	27.0	56.0	14,169.7	16,262.3	810.0
Diabetes	59,873	37	78	945	7,037.4	35.6	51.4	414.4	8,476.9	97.2	197.9	64.4	186.4	9,016.7	8,271.8	597.8
All Accident	97,860	5,265	8,391	7,128	31,806.3	753.4	678.4	1,119.3	42,809.9	5,201.1	3,240.7	694.5	1,751.8	27,286.5	45,830.1	3,529.2
Lung and Bronchus Cancer	152,063	6	14	2,054	30,282.6	56.8	77.7	555.2	46,124.4	91.4	116.9	45.9	82.1	21,271.8	33,212.8	3,289.5
Stroke	199,450	25	35	2,314	19,048.1	16.2	106.4	769.2	42,867.7	693.9	843.4	40.3	95.2	23,258.2	24,035.7	1,429.0
All Cancer	549,838	452	560	2,063	93,854.4	221.6	156.1	824.7	73,400.3	667.1	760.0	257.3	208.9	58,131.6	53,450.6	3,377.9
Hearth Diseases	725,192	453	616	8,212	51,850.3	194.2	234.4	2,157.3	68,957.3	840.0	950.6	188.7	267.3	28,408.2	35,567.4	6,773.3
All Diseases	2,293,539	8,513	8,487	45,324	247,028.3	3,368.1	5,707.4	9,380.6	328,083.9	8,079.8	14,369.3	3,913.7	3,875.1	159,968.1	210,497.8	39,760.7

\* indicates a cause of death different from the ones used by Benjamin et al. (2001).

**Table 2**  
**Description of the Different Treatments**

	Survey		Expected Outcome under	
	First	Second	BDB's Hypothesis	Anchor Hypothesis
Treatment 1	Population-Wide	—	Traditional Biases	Traditional Biases
Treatment 2	Own-Age-Cohort	—	Unbiased	Traditional Biases
Treatment 3	Different-Age-Cohort	—	Traditional Biases	Traditional Biases
Treatment 4	Population-Wide	Own-Age-Cohort	First survey biased, second survey unbiased	First survey biased, second survey only biased for small risks
Treatment 5	Own-Age-Cohort	Population-Wide	First survey unbiased, second survey biased	First survey biased, second survey more biased for high risks
Treatment 6	Population-Wide	Different-Age-Cohort	Both surveys biased	First survey biased, second survey only biased for small risks

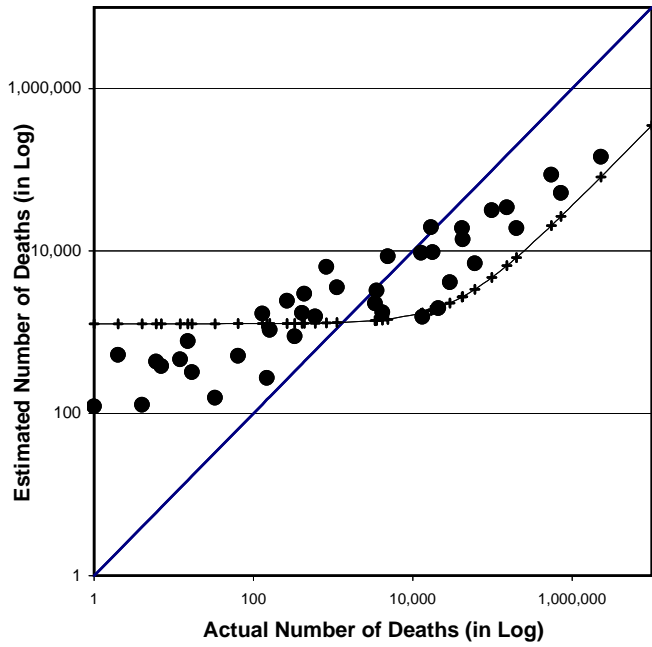
**Table 3**  
**Robust Estimation Results**

	Treatment 1	Treatment 2	Treatment 3	Treatment 4		Treatment 5		Treatment 6	
	Population-Wide	Own-Age-Cohort	40-44 Age-cohort	Population-Wide	Own-Age-Cohort	Own-Age-Cohort	Population-Wide	Population-Wide	40-44 Age-cohort
Constant	1,268.216* (386.736)	19.283 (13.625)	42.008 (46.303)	1,760.803* (413.231)	29.820 (19.832)	25.672 (11.400)	967.230* (308.289)	1,517.104* (426.372)	41.436 (36.303)
Slope	0.035* (0.002)	0.204* (0.007)	0.167* (0.028)	0.039* (0.003)	0.847* (0.019)	0.221* (0.010)	0.022* (0.001)	0.034* (0.003)	0.815* (0.035)
R <sup>2</sup>	0.342	0.471	0.356	0.297	0.518	0.381	0.416	0.301	0.441
$\sigma$	6.742* (1.644)	1.553* (0.426)	5.007* (1.132)	7.218* (1.435)	1.780* (0.504)	1.468* (0.444)	7.882* (1.916)	6.520* (1.482)	4.816* (1.108)
$\gamma$	1.260* (0.125)	0.524* (0.066)	0.892* (0.090)	1.114* (0.106)	0.567 (0.079)	0.541* (0.052)	1.058* (0.151)	1.337* (0.120)	0.851* (0.083)
Ranking Errors*	9.326 (7.675)	6.540 (5.334)	7.911 (5.906)	8.639 (8.004)	8.529 (6.270)	6.986 (5.929)	8.870 (8.473)	8.977 (7.570)	8.146 (6.552)

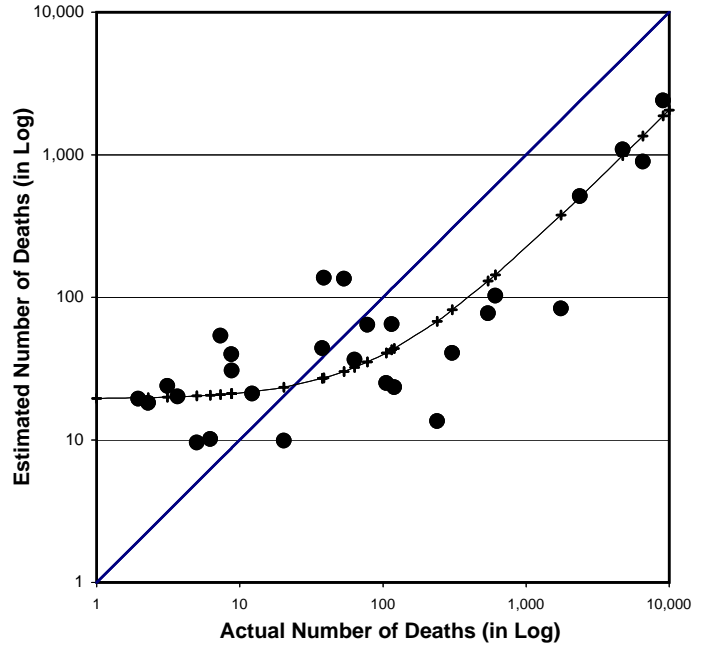
\* The first number refers to the mean of the ranking errors, while the number in parenthesis refers to the standard deviation of the ranking errors. The sample size in each treatment is 1,365. The symbol \* denotes an estimated parameter significantly different from 0 at a 5% significance level. The numbers in parenthesis refer to standard deviations.



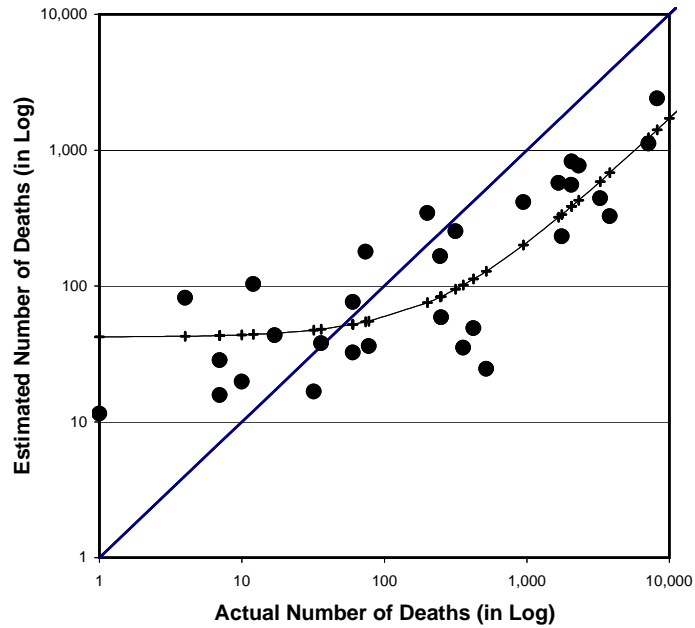
**FIGURE 1**  
**Population-Wide Estimates**  
**Treatment 1**  
(Single Survey Treatment)



**FIGURE 2**  
**Own-Age-Cohort Estimates**  
**Treatment 2**  
(Single Survey Treatment)

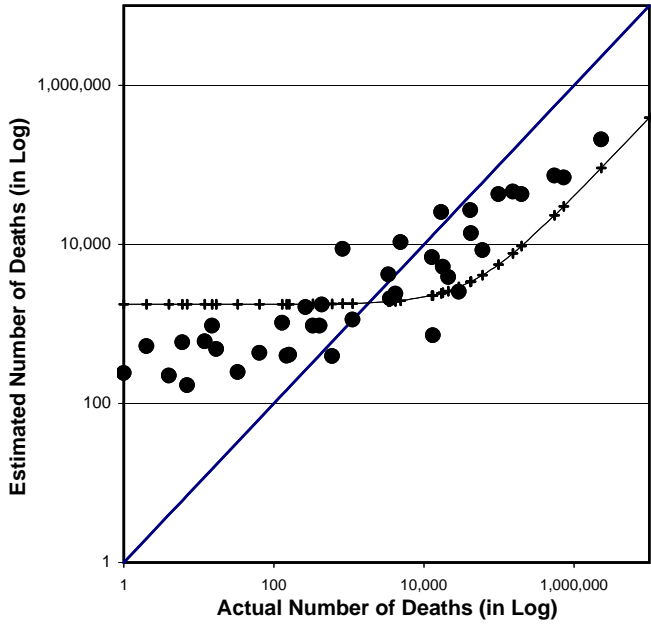


**FIGURE 3**  
**40-44 Age-Cohort Estimates**  
**Treatment 3**  
(Single Survey Treatment)



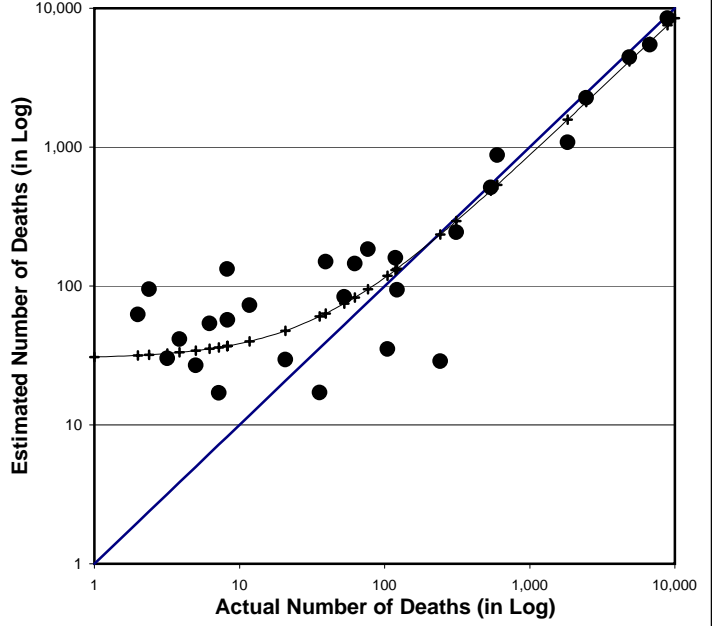
**FIGURE 4**  
**Population-Wide Estimates**  
**Treatment 4**

(Population-Wide Survey then Own-Age-Cohort)



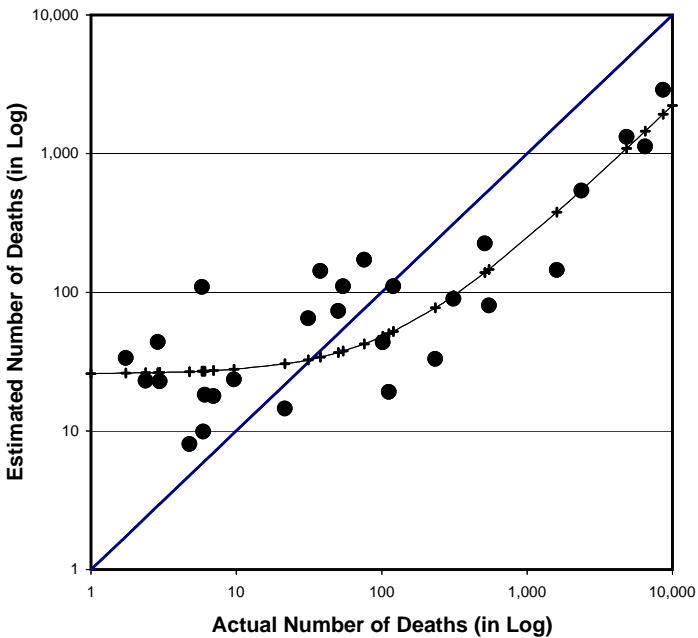
**FIGURE 5**  
**Own-Age-Cohort Estimates**  
**Treatment 4**

(Population-Wide Survey then Own-Age-Cohort)



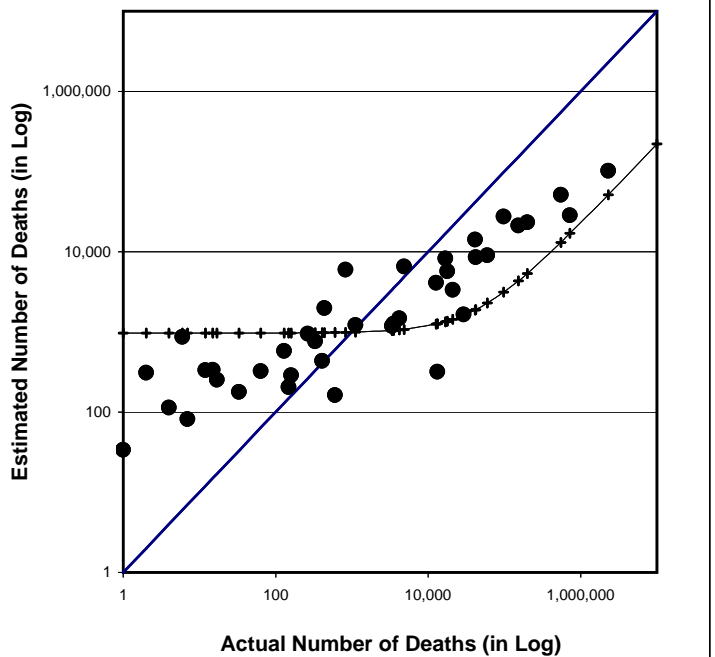
**FIGURE 6**  
**Own-Age-Cohort Estimates**  
**Treatment 5**

(Own-Age-Cohort Survey then population-wide)



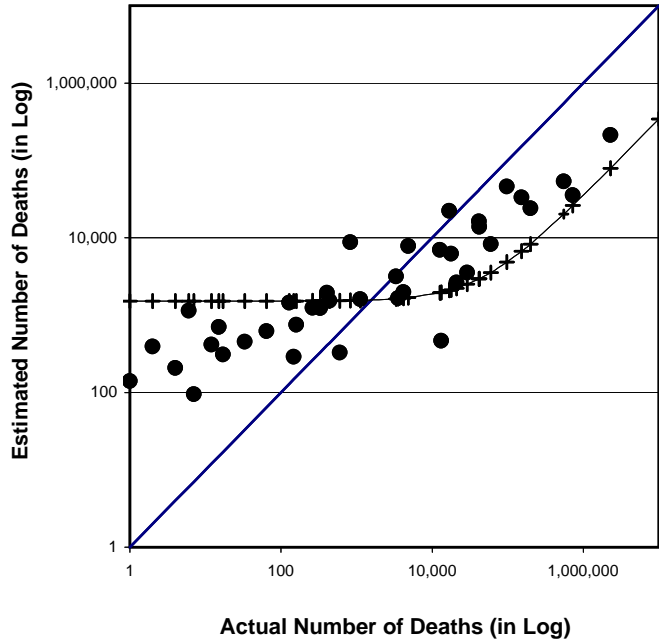
**FIGURE 7**  
**Population-Wide Estimates**  
**Treatment 5**

(Own-Age-Cohort Survey then Population-Wide)



**FIGURE 8**  
**Population-Wide Estimates**  
**Treatment 6**

(Population-Wide Survey then 40-44 Age-Cohort)



**FIGURE 9**  
**40-44 Age-Cohort Estimates**  
**Treatment 6**

(Population-Wide Survey then 40-44 Age-Cohort)

