Risk aversion and incentive effects: Comment

Glenn W. Harrison
*University of Central Florida*

Eric Johnson

Melayne M. McInnes

E. Elisabet Rutström
*University of Central Florida*

---

Find similar works at: [https://stars.library.ucf.edu/facultybib2000](https://stars.library.ucf.edu/facultybib2000)

University of Central Florida Libraries [http://library.ucf.edu](http://library.ucf.edu)

This Article is brought to you for free and open access by the Faculty Bibliography at STARS. It has been accepted for inclusion in Faculty Bibliography 2000s by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

---

**Recommended Citation**

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M.; and Rutström, E. Elisabet, "Risk aversion and incentive effects: Comment" (2005). *Faculty Bibliography 2000s*. 5260.

[https://stars.library.ucf.edu/facultybib2000/5260](https://stars.library.ucf.edu/facultybib2000/5260)
Risk Aversion and Incentive Effects: Comment

By Glenn W. Harrison, Eric Johnson, Melayne M. McInnes, and E. Elisabet Rutström*

Charles A. Holt and Susan K. Laury (2002) develop an experimental design to determine the risk attitude of an individual. They use their observations to argue that increased incentives appear to change risk attitudes, leading to greater risk aversion. Popular utility functions that do not allow for such effects are therefore misspecified. Building on this finding, they estimate a flexible utility function that characterizes their aggregate data well, but that does not assume constant (absolute or relative) risk aversion.

The basic Holt and Laury (2002) (HL) design should become an important tool for the interpretation and design of experiments in which risk attitudes could play a role. However, their most important result—showing the effect of scaling up the stakes of the lottery choice task—is confounded by a possible order effect. An order effect occurs when prior experience with one task affects behavior in a subsequent task. The primary methodological contribution of experiments in economics is to enhance control. Ideally, such control makes the explanatory variables of interest, in this case scale, orthogonal to other explanatory variables such as order, allowing clearer inferences about behavior than one could get from field econometric data. We argue that not controlling for order effects results in a misspecification of utility functions that is as important as that of scale.

The subjects in the HL design were given sequences of three or four tasks, as shown in Table 1. Task #1 involved choices over lotteries with a baseline level of prizes, which we will refer to as the “1× scale.” Task #2 involved hypothetical choices over prizes with a scaled-up level of prizes, either 20×, 50×, or 90×. Task #3 repeated task #2, but with choices that involved real payoffs. Task #4 was a return to the baseline task with real 1× prizes. In some sessions, subjects were not given task #2 or task #3.1

What could one infer from the comparison of measured risk attitudes in the 1× and 20× tasks in HL? Unfortunately, any observed difference could be due either to the scale of the prizes involved or the task order, or some combination of both. Thus, the effect of scale is intrinsically confounded with the possible effect of task order.2 This is a logical flaw in their design, which

1 HL recognize the possibility that wealth effects could also confound the effects of scale in an in-sample design. To handle this they use a clever device: when the subjects proceed from task #1 to task #3, they are asked if they are willing to give up their earnings in task #1 in order to play task #3. Since the stakes are so much higher in task #3, all subjects chose to do so. This means that the subjects face tasks #1 and #3 with no prior earnings from these experiments, although they do have experience with the type of task when facing task #3. No such trick can be applied for task #4, since the subjects would be unlikely to give up their earnings in task #3 in this instance. Thus the responses to task #4 have no controls for income built into the design.

2 HL are aware (p. 1647, fn. 4) of the possible effect of task order, albeit in the context of the real choices in task #3 being affected by the immediately preceding hypothetical task #2 with identical prize values. HL also included task #4 as a check for consistency of order effects for the 1× lottery. Nevertheless, the lack of a significant difference in choices between task #1 and task #4 does not prove conclusively that there are no order effects, since it is possible that subjects gave the same responses due to some sense of obligation or preference for being consistent. Moreover, there are no controls for wealth effects in their task #4. Our design does not suffer from such a confound.

* Harrison: Department of Economics, College of Business Administration, University of Central Florida, Orlando, FL 32816 (e-mail: gharrison@bus.ucf.edu); Johnson: Department of Economics, Moore School of Business, University of South Carolina, Columbia, SC 29208 (e-mail: ejohnson@moore.sc.edu); McInnes: Department of Economics, Moore School of Business, University of South Carolina, Columbia, SC 29208 (e-mail: mcinnes@moore.sc.edu); Rutström: Department of Economics, College of Business Administration, University of Central Florida, Orlando, FL 32816 (e-mail: erutstrom@bus.ucf.edu). We thank the U.S. National Science Foundation for research support under grants NSF/IIS 9817518, NSF/MRI 9871019, and NSF/POWRE 9973669 to Rutström and grant NSF/SES 021429.0 to McInnes.
can be fixed either by assuming away the possibility of order effects or redesigning their experiment and controlling for them. One may argue that the effect of increasing stakes from \( \frac{20}{1000} \) to \( \frac{50}{1000} \) (and \( \frac{90}{1000} \)) represents unconounded evidence in favor of a stake effect,\(^3\) but this assumes that there is no interaction effect between order and stake. If the order effect increases with the stakes, then it is possible that much of the change in preferences between \( \frac{20}{1000} \) and \( \frac{50}{1000} \) is due to order, not scale. Further, if the effect of scale on relative risk aversion is diminishing, the confound will become more pronounced as the stakes are increased.

Fortunately, order effects are easy to control for by design. We undertook a new series of experiments that build closely on the basic design features of HL, but allow an identification of the extent to which the apparent scale effects on risk aversion are actually order effects. We focus on the two real decisions from HL in which there is a change in payoff scale. In our first treatment, the subjects were asked to make the choices given in panel B of Table 1, and then provided the opportunity to give up those earnings in return for the chance to participate in choices with payments scaled up by 10.\(^4\)

Scaling the base payoffs, which have prizes ranging between $0.10 and $3.85, provides responses that span prizes between $1.00 and $38.50. This comfortably covers the range of prizes needed to apply the measures of risk aversion to most experiments. We call this the \( \frac{1}{10} \) treatment. In the second treatment, a different sample of subjects was given only the \( \frac{1}{10} \) task. This allows us to test whether the responses in the \( \frac{1}{10} \) task are affected by the prior experience of having seen the \( \frac{1}{10} \) task.

Fortunately, order effects are easy to control for by design. We undertook a new series of experiments that build closely on the basic design features of HL, but allow an identification of the extent to which the apparent scale effects on risk aversion are actually order effects. We focus on the two real decisions from HL in which there is a change in payoff scale. In our first treatment, the subjects were asked to make the choices given in panel B of Table 1, and then provided the opportunity to give up those earnings in return for the chance to participate in choices with payments scaled up by 10.\(^4\)

Scaling the base payoffs, which have prizes ranging between $0.10 and $3.85, provides responses that span prizes between $1.00 and $38.50. This comfortably covers the range of prizes needed to apply the measures of risk aversion to most experiments. We call this the \( \frac{1}{10} \) treatment. In the second treatment, a different sample of subjects was given only the \( \frac{1}{10} \) task. This allows us to test whether the responses in the \( \frac{1}{10} \) task are affected by the prior experience of having seen the \( \frac{1}{10} \) task. Thus, our design allows us to disentangle the effects of order and scale by conducting a between-subjects analysis of the \( \frac{1}{10} \) responses from the \( \frac{1}{10} \) and \( \frac{1}{10} \times \) treatments for pure order effects, and a within-subjects analysis of

\[^3\] Using nonparametric tests, the only significant scale effect across the \( \frac{20}{1000} \), \( \frac{50}{1000} \), and \( \frac{90}{1000} \) tasks is between \( \frac{20}{1000} \) and the two higher scales. There is no significant effect between \( \frac{50}{1000} \) and \( \frac{90}{1000} \). This may be due to the small sample sizes in the \( \frac{50}{1000} \) and \( \frac{90}{1000} \) treatments.

\[^4\] We chose the \( \frac{1}{10} \) scaling, rather than the \( \frac{20}{1000} \) scaling used by HL, since we were measuring risk aversion for use in later experiments for which \( \frac{1}{10} \) spanned the range of payoffs.
the 1×10× responses for scale effects combined with order effects.

We recruited 178 subjects in 11 sessions in October and November 2002 at the University of South Carolina. Of these, 55 participated in the 10× experiments and 123 in the 1×10× experiments.5

Like HL, we find that the majority of subjects choose the “safe” option (lottery A in Table 1) when the probability of the higher payoff is small and then switch over to the “risky” option (lottery B in Table 1) as the difference in expected value increasingly favors the risky option.6

Because the comparison of the 10× responses with and without prior experience is based on a between-subjects comparison, it is important to control for demographics. Differences in the composition of the subjects in the two treatment groups may otherwise confound inferences. To control for demographics, we estimate an ordered probit regression model for the first risky response using the 10× data.7 We include a standard list of sociodemographic characteristics in the model. We also included a binary indicator for tasks that came second (the 10× task in the 1×10× experiments), and dummy variables for each experimental session. Results from estimating the ordered probit model clearly show the importance of demographics and task order.

The ordered probit model allows us to construct Figure 1, which can be directly compared to the figures in HL, and shows the predicted probability of a safe choice for each of the treatments.8 To compare directly to HL, we


6 A small fraction, 16 percent at the 1× scale and 10 percent at the 10× scale, switch more than once. Most of these still choose safe for the first rows in Table 1, and then switch to risky for the last rows with some noise in the middle. Excluding these multiple switchers from the analysis made no difference to our conclusions.

7 The ordered probit specification has three advantages here. First, it recognizes the natural ordering of the ten decisions, which is a central feature of the experimental design. Second, it recognizes that the ten decisions we observe for each individual are not ten independent observations: the probability that an individual chooses a safe choice drops dramatically once a risky choice has been made. The third advantage of the ordered probit is that we can remain agnostic about the functional form of the utility function. We have also undertaken statistical analyses that do not rely on using the first risky choice, and draw the same conclusions about order effects.

8 For each individual and decision, we obtain the predicted probability of a safe choice and then sum over the sample to obtain the cumulative probabilities. We first estimated the model on the 10× data alone. By setting the order dummy equal to 0 or 1, we obtained predictions for
should compare the $10\times$ responses to the $10\times$ responses that follow the $1\times$ responses. The increase in risk aversion in this comparison is due to both order and scale effects. When we compare the $1\times$ responses to the $10\times$ responses with no prior experience, however, the pure scale effect is seen to be significantly smaller.

To measure the economic significance of the scale and order effects, we also estimate an interval regression model under the assumption of constant relative risk aversion (CRRA). This allows us to treat observations of subjects who choose A throughout with appropriate agnosticism, not imposing an arbitrary upper bound on their risk preference parameter. It also allows us to interpret observations of subjects who switch back and forth across several rows as having a wider switching interval, imposing a statistical interpretation of this uncertainty over their risk attitudes. Finally, it allows a direct interpretation of the magnitude of the effects in terms that are familiar to most economists, namely the CRRA coefficient. The interval regression model predicts that the average CRRA coefficient for the $1\times$ scale is 0.37, that it is 0.74 for the $10\times$ scale when there are both order and scale effects present, and that it is 0.57 in the $10\times$ scale when there are only scale effects. The coefficients on dummy variables capturing order and scale effects are each statistically significant, with $p$-values less than 0.05. Hence the order effect in the HL design confounds the inference about scale effects, such that the true scale effect is a little over one-half of the apparent effect when scale and order are confounded.

We therefore reaffirm the primary conclusion of HL, that risk aversion varies over the income range found in typical experiments. The effect is significantly smaller than they estimate, but the presence of a basic confound in their design does not lead one to reject their qualitative conclusion. Nevertheless, we conclude that order effects are significant and almost as large as scale effects, so that they can lead to misspecifications of utility functions that are as serious as those of scale effects unless properly controlled for.

Practical implications arise for both academic researchers and policy analysts. First, since we find that pure scale effects cause smaller increases in risk aversion than suggested by HL, auction theorists may in many cases be able to continue using the CRRA specification as a local approximation to more general functional forms. The empirical task now is to define better the domains over which CRRA remains a valid approximation, and that is likely to vary with the population sampled and the context of the task. Similarly, in tests of Expected Utility Theory (EUT) certain combinations of risk attitudes and lottery parameter values lead to indifference between lottery pairs, implying that any choice pattern could be consistent with EUT if risk attitudes were uncontrolled. In order to make such tests of EUT operationally meaningful, one therefore has to select the lottery parameters conditionally on the risk attitudes of the respondents, and one has to have relatively precise estimates of risk attitude to do that.

Finally, because the predicted impacts of large-scale policy changes are uncertain, policy analysts must allow for the risk attitudes of households when evaluating welfare changes. Controlling for known systematic effects in risk elicitation, such as order, is easy and removes an important bias. While it is certainly true that other variations in procedures may affect responses in the risk elicitation task, it makes little sense to fail to control for the procedural effects we can easily design away.

We are cautiously optimistic that risk preferences elicited with salient incentives and proper controls for order are robust, particularly when analysts are armed with the flexible utility specification employed by HL to evaluate scale effects. We believe that careful experimental design and

---

the two $10\times$ treatments shown in Figure 1. We then re-estimated the model on the $1\times$ data alone to obtain the predictions for the $1\times$ treatment.

9 See Harrison et al. (2004, 2005) for further discussion of methodological issues in the design and analysis of elicitation designs such as these.

---

10 This issue is discussed in Harrison et al. (2003).

11 In the field of computable general equilibrium models, there has long been a recognition that systematic sensitivity analysis of simulations conditioned on uncertain parameters implies uncertain policy impacts. See Harrison and H. D. Vinod (1992) for example.

12 For example, Harrison, Johnson, McInnes, and Rutström (2005) show that risk preferences appear to be stable over time horizons of several months.
implementation can provide the needed control for most methodological and policy applications.

REFERENCES


25. Astrid Gamba, Elena Manzoni, Luca Stanca. 2017. Social comparison and risk taking behavior. Theory and Decision 82:2, 221-248. [Crossref]


32. Andreas C. Drichoutis, Jayson L. Lusk. 2016. What can multiple price lists really tell us about risk preferences?. Journal of Risk and Uncertainty 53:2-3, 89-106. [Crossref]


34. Antonio Filippin, Paolo Crosetto. 2016. A Reconsideration of Gender Differences in Risk Attitudes. Management Science 62:11, 3138-3160. [Crossref]

35. Julián Pareja Vasseur, Carolina Cadavid Pérez. 2016. Valoración de patentes farmacéuticas a través de opciones reales: equivalentes de certeza y función de utilidad. Contaduría y Administración 61:4, 794-814. [Crossref]


37. Anna Bartczak, Petr Mariel, Susan Chilton, Jürgen Meyerhoff. 2016. The impact of latent risk preferences on valuing the preservation of threatened lynx populations in Poland. Australian Journal of Agricultural and Resource Economics 60:2, 284-306. [Crossref]
38. Marcel Lichters, Claudia Brunnlieb, Gideon Nave, Marko Sarstedt, Bodo Vogt. 2016. The Influence of Serotonin Deficiency on Choice Deferral and the Compromise Effect. *Journal of Marketing Research* 53:2, 183-198. [Crossref]

39. Hazel Bateman, Loretti I. Dobrescu, Ben R. Newell, Andreas Ortmann, Susan Thorp. 2016. As easy as pie: How retirement savers use prescribed investment disclosures. *Journal of Economic Behavior & Organization* 121, 60-76. [Crossref]

40. Johannes G. Jaspersen. 2016. HYPOTHETICAL SURVEYS AND EXPERIMENTAL STUDIES OF INSURANCE DEMAND: A REVIEW. *Journal of Risk and Insurance* 83:1, 217-255. [Crossref]


42. Howard Kunreuther, Erwann Michel-Kerjan. 2015. Demand for fixed-price multi-year contracts: Experimental evidence from insurance decisions. *Journal of Risk and Uncertainty* 51:2, 171-194. [Crossref]

43. Anna Bartczak, Susan Chilton, Jürgen Meyerhoff. 2015. Wildfires in Poland: The impact of risk preferences and loss aversion on environmental choices. *Ecological Economics* 116, 300-309. [Crossref]

44. Andreas C. Drichoutis, Rodolfo M. Nayga. 2015. Do risk and time preferences have biological roots?. *Southern Economic Journal* 82:1, 235-256. [Crossref]


46. P. Crosetto, A. Filippin, J. Heider. 2015. A Study of Outcome Reporting Bias Using Gender Differences in Risk Attitudes. *CESifo Economic Studies* 61:1, 239-262. [Crossref]

47. Uwe Dulleck, Jonas Fooken, Jacob Fell. 2015. Within-Subject Intra- and Inter-Method Consistency of Two Experimental Risk Attitude Elicitation Methods. *German Economic Review* 16:1, 104-121. [Crossref]


49. Constantinos Antoniou, Glenn W. Harrison, Morten I. Lau, Daniel Read. 2015. Subjective Bayesian beliefs. *Journal of Risk and Uncertainty* 50:1, 35-54. [Crossref]


51. Yueming Qiu, Gregory Colson, Carola Grebitus. 2014. Risk preferences and purchase of energy-efficient technologies in the residential sector. *Ecological Economics* 107, 216-229. [Crossref]


56. Giovanna Devetag, Andreas Ortmann. Solving Coordination Problems Experimentally 357-384. [Crossref]
57. Steffen Andersen, Amalia Di Girolamo, Glenn W. Harrison, Morten I. Lau. 2014. Risk and time preferences of entrepreneurs: evidence from a Danish field experiment. *Theory and Decision* 77:3, 341. [Crossref]
58. Charles A. Holt, Susan K. Laury. Assessment and Estimation of Risk Preferences 135-201. [Crossref]
60. Astrid Gamba, Elena Manzioni. 2014. Social Comparison and Risk Taking Behavior. *SSRN Electronic Journal*. [Crossref]
61. Antonio Filippini, Paolo Crosetto. 2014. A Reconsideration of Gender Differences in Risk Attitudes. *SSRN Electronic Journal*. [Crossref]
63. Alexander Harin. 2014. A 'Certain-Uncertain' Inconsistency of the Main Experimental System of Utility and Prospect Theories and Models. *SSRN Electronic Journal*. [Crossref]
64. Constantinos Antoniou, Glenn W. Harrison, Morten I. Lau, Daniel Read. 2014. Information Characteristics and Errors in Expectations: Experimental Evidence. *SSRN Electronic Journal*. [Crossref]
69. Astrid Matthey, Tobias Regner. 2013. On the independence of history: experience spill-overs between experiments. *Theory and Decision* 75:3, 403-419. [Crossref]
72. Christoph Saenger, Matin Qaim, Maximo Torero, Angelino Viceisza. 2013. Contract farming and smallholder incentives to produce high quality: experimental evidence from the Vietnamese dairy sector. *Agricultural Economics* 44:3, 297-308. [Crossref]


81. Constantinos Antoniou, Glenn W. Harrison, Morten I. Lau, Daniel Read. 2012. Subjective Bayesian Beliefs. *SSRN Electronic Journal*. [Crossref]

82. Eric Yongchen Chow. 2012. The Economics of Somatic Responses. *SSRN Electronic Journal*. [Crossref]


84. Louis Lévy-Garboua, Hela Maafi, David Masclet, Antoine Terracol. 2011. Risk aversion and framing effects. *Experimental Economics*. [Crossref]


88. Sujoy Chakravarty, Glenn W. Harrison, Ernan E. Haruvy, E. Elisabet Rutström. 2011. Are You Risk Averse over Other People’s Money?. *Southern Economic Journal* 77:4, 901-913. [Crossref]


91. QUANG NGUYEN. 2010. How nurture can shape preferences: an experimental study on risk preferences of Vietnamese fishers. *Environment and Development Economics* 15:5, 609-631. [Crossref]


93. Larry L. Lawson, Catherine L. Lawson. 2010. Video Game-Based Methodology for Business Research. *Simulation & Gaming* 41:3, 360-373. [Crossref]


98. Vernon L. Smith. Experimental Economics 75-98. [Crossref]
100. Byron W. Keating, Ali M. Quazi, Anton Kriz. 2009. Financial risk and its impact on new purchasing behavior in the online retail setting. Electronic Markets. [Crossref]
107. Katarina Kalovcova, Andreas Ortmann. 2009. Understanding the Plott-Wit-Yang Paradox. SSRN Electronic Journal. [Crossref]
109. Glenn W. Harrison, Sebastian Moritz, Richard Pibernik. 2009. How Does the Risk Attitude of a Purchasing Manager Affect the Selection of Suppliers?. SSRN Electronic Journal. [Crossref]
111. Andreas Ortmann. 2008. PROSPECTING NEUROECONOMICS. Economics and Philosophy 24:3, 431-448. [Crossref]
113. Steffen Andersen, Glenn W. Harrison, Morten I. Lau, E. Elisabet Rutström. 2008. LOST IN STATE SPACE: ARE PREFERENCES STABLE?*. International Economic Review 49:3, 1091-1112. [Crossref]
115. Catherine C. Eckel, Philip J. Grossman. Chapter 113 Men, Women and Risk Aversion: Experimental Evidence 1061-1073. [Crossref]
116. Vernon L. Smith. Experimental Economics 1-19. [Crossref]
120. Peter P. Wakker, Veronika Köbberling, Christiane Schwieren. 2007. Prospect-theory’s Diminishing Sensitivity Versus Economics’ Intrinsic Utility of Money: How the Introduction of the Euro can be Used to Disentangle the Two Empirically. *Theory and Decision* 63:3, 205-231. [Crossref]
122. Yoram Halevy. 2007. Ellsberg Revisited: An Experimental Study. *Econometrica* 75:2, 503-536. [Crossref]
123. Steffen Andersen, Glenn W. Harrison, Morten I. Lau, E. Elisabet Rutström. 2007. Preference Heterogeneity in Experiments: Comparing the Field and Laboratory. *SSRN Electronic Journal*. [Crossref]
131. Frank Heinemann. Measuring risk aversion and the wealth effect 293-313. [Crossref]
132. Susan K. Laury, Charles A. Holt. Further reflections on the reflection effect 405-440. [Crossref]
133. Nathaniel T. Wilcox. Stochastic models for binary discrete choice under risk: a critical primer and econometric comparison 197-292. [Crossref]
134. Steffen Andersen, Glenn W. Harrison, Morten I. Lau, E. Elisabet Rutström. Risk aversion in game shows 359-404. [Crossref]
135. Glenn W. Harrison, E. Elisabet Rutström. Risk Aversion in the Laboratory 41-196. [Crossref]