Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Comment on “A Model of Probabilistic Choice Satisfying First-Order Stochastic Dominance” by Pavlo Blavatskyy

Graham Loomes, Inmaculada Rodríguez-Puerta, Jose-Luis Pinto-Prades

MANAGEMENT SCIENCE
Vol. 60, No. 5, May 2014, pp. 1346–1350
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2013.1810

Comment on "A Model of Probabilistic Choice Satisfying First-Order Stochastic Dominance" by Pavlo Blavatskyy

Graham Loomes
Behavioural Science, Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom, g.loomes@warwick.ac.uk

Inmaculada Rodríguez-Puerta
Department of Economics, Universidad Pablo de Olavide, 41013 Seville, Spain, irodpue@upo.es

Jose-Luis Pinto-Prades
Yunus Centre for Social Business and Health, Glasgow Caledonian University, Glasgow G4 0BA, United Kingdom, joseluis.pinto@gcu.ac.uk

We present examples of existing evidence that lead us to be cautious about claims made in the original paper [Blavatskyy PR (2011) A model of probabilistic choice satisfying first-order stochastic dominance. *Management Sci.* 57(3):542–548] that the proposed model provides a better fit to experimental data than do existing models. We raise concerns about the accuracy of this and other assertions and about the adequacy of the comparisons made with alternative models in the existing literature.

Keywords: decision analysis; risk; economics; microeconomic behavior; utility; probabilistic choice functions
History: Received June 10, 2012; accepted March 18, 2013, by Rakesh Sarin, decision analysis. Published online in *Articles in Advance* November 4, 2013.

Introduction

Blavatskyy (2011, p. 542) "presents a new model of probabilistic binary choice under risk" that, according to the author, outperforms existing probabilistic models. The main improvement, in relation to existing models, is that the Blavatskyy model satisfies both stochastic dominance and weak stochastic transitivity, which according to the author, are desirable properties of any descriptive choice model because they are rarely violated in the data. Blavatskyy claims that the model can explain "behavioral regularities" such as the common ratio effect or preference reversal. The purpose of the present paper is to examine whether this strong claim can be justified. In this comment, we note that by its very construction the model cannot explain the well-established common consequence effect. Moreover, its ability to explain the common ratio effect is limited: cases that the model does not explain have been frequently observed. Finally, Blavatskyy's abstract claims that the model "provides a better fit to experimental data than do existing models" (p. 542); however, he takes just two data sets and compares the proposed model with a small and selective subset of competing models. This illustrates that the model has *some* power, but it hardly amounts to robust support for the rather sweeping claim made in the abstract. The purpose of the present paper is to draw attention to a body of evidence that predates the Blavatskyy model and that exhibits systematic response patterns that the model does not accommodate.

An Outline of the Model and Some Implications

The Blavatskyy (2011) model, which takes standard expected utility (EU) as its core theory, can be summarized as follows:

(i) For any pair of lotteries X and Y, it is possible to identify a greatest lower bound (GLB), which is defined as the best lottery that is dominated by both X and Y.[1]

(ii) For each lottery in the pair, the key measure is the difference between the EU of that lottery and the EU of the GLB. Denote these differences by X' and Y', respectively.

(iii) The probability that X is chosen rather than Y in a binary choice between the two is expressed as a

[1] It is also possible to identify a least upper bound (LUB), which is the worst lottery that dominates both X and Y. The analysis in Blavatskyy (2011) can equally well be conducted in terms of GLB or LUB, but we shall focus on the analysis in terms of GLB.

function of $\{X', Y'\}$. More specifically, the probability that X is chosen rather than Y is given as

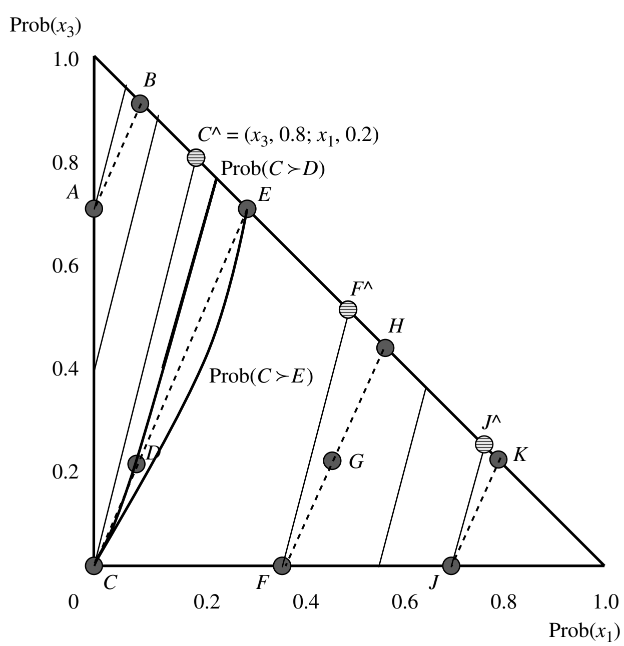
$$\text{Prob}(X \succ Y) = \phi(X')/[\phi(X') + \phi(Y')]. \tag{1}$$

(iv) It follows from (1) that when $\phi(X') = \phi(Y')$—that is, when $\text{EU}(X) = \text{EU}(Y)$—the probability of choosing each option is 0.5. This is independent of the nature of $\phi(\cdot)$ and allows us to reconfigure the Marschak-Machina triangle (Machina 1982), replacing the deterministic indifference curves of standard EU by "50–50 lines"—lines where the probability of choosing any one of the lotteries on the line rather than any other lottery on the same line is 0.5.

In Figure 1, the probability of the highest payoff x_3 is depicted on the vertical axis and the probability of the lowest payoff x_1 is shown on the horizontal axis. Hence, the certainty of the intermediate payoff x_2 is shown by C at the lower left corner of the triangle. Suppose the individual's preferences are such that his EU for this sure payoff is the same as his EU for a lottery offering a 0.8 chance of x_3 and a 0.2 chance of x_1—this lottery is represented by $C^\wedge$ in Figure 1. Then the line connecting C to $C^\wedge$ represents the set of lotteries from which we can draw any pair such that the probability of choosing each lottery in that pair is 0.5.

(v) All other solid lines parallel to the line connecting C with $C^\wedge$ also identify sets of lotteries that have a 0.5 probability of being chosen when paired with any other lottery on that same line. For example, any pair of lotteries on the line connecting F to $F^\wedge$ are equally likely to be chosen in a binary choice between them,

Figure 1 **Diagrammatic Representation in Probability Triangle**



so too any pair of lotteries on the line connecting J to $J^\wedge$.

(vi) For other pairs of lotteries *not* on the same solid line, the probabilities of being chosen will vary. For example, for lotteries C and E in Figure 1, E is stochastically dominated by $C^\wedge$ so that $\text{EU}(E) < \text{EU}(C^\wedge)$, and therefore $\text{EU}(E) < \text{EU}(C)$, with the result that $\text{Prob}(C \succ E) > 0.5$. This is signified diagrammatically by the fact that the dashed line between C and E slopes up from left to right but has a gradient less than that of the solid line connecting C to $C^\wedge$. So in the Blavatskyy (2011) model, we have curves that are composed of all the lotteries that have the same probability of being chosen over one common lottery. In Figure 1 we have a curve composed of all lotteries that have the same probability as E of being chosen over C and another curve composed of all lotteries than have the same probability as D of being chosen over C. The curvature of these curves depends on the specification of $\phi(\cdot)$. For example, if $\phi(x) = e^{\lambda x} - 1$, as Blavatskyy assumes for the purposes of econometric fitting, probability curves are concave below the 50% probability line and convex above this line.

(vii) Therefore, the probability of one lottery on the dashed line between C and E being chosen over another lottery on that same dashed line is likely to vary, depending on the specification of $\phi(\cdot)$. For example, a specification such as $\phi(x) = e^{\lambda x} - 1$ will entail that $\text{Prob}(C \succ E) > \text{Prob}(C \succ D) > 0.5$ because the probability curve that joins C and E will be further from the 50–50 line than the probability curve that joins C and D. Thus, the probability of a more southwesterly lottery being chosen over a more northeasterly lottery on a given dashed line will tend to fall as the distance between them becomes smaller.

(viii) The model's Axiom 4 (Common Consequence Independence) entails that for any given specification of $\phi(\cdot)$, if there are two pairs of lotteries the same distance apart on two different dashed lines with the same gradient, the probabilities of choice will be the same across pairs. For example, $\{A, B\}$, $\{C, D\}$, $\{F, G\}$, and $\{J, K\}$ in Figure 1 lie on parallel dashed lines with the same distance between them, so the model entails $\text{Prob}(A \succ B) = \text{Prob}(C \succ D) = \text{Prob}(F \succ G) = \text{Prob}(J \succ K) > 0.5$.

Implications for the Common Ratio Effect

If we take a specification of $\phi(\cdot)$ that entails $\text{Prob}(C \succ E) > \text{Prob}(C \succ D)$ as in (vii) above and combine it with the property that $\text{Prob}(C \succ D) = \text{Prob}(J \succ K)$ as in (viii) above, we get $\text{Prob}(C \succ E) > \text{Prob}(J \succ K)$. This latter pattern has been frequently observed in experiments yielding what Kahneman and Tversky (1979) called the *common ratio effect* (CRE).

Blavatskyy (2011) claims that the ability of his model to accommodate the CRE is one of its strengths. However, this claim is qualified: his exact words are that the model is "compatible with several behavioral regularities such as *certain types* of the common ratio effect..." (p. 543, emphasis added). In fact, the model can accommodate the CRE only up to a point. If $\text{Prob}(C \succ E) > 0.5$, which is the case in most CRE data, then the kind of $\phi(\cdot)$ needed to produce the effect has the implication that $\text{Prob}(C \succ D) = \text{Prob}(J \succ K) > 0.5$. In other words, under these assumptions, the Blavatskyy model can explain the fact that $\text{Prob}(C \succ E) > \text{Prob}(J \succ K)$, but it does not allow $\text{Prob}(C \succ E) > 0.5$ in conjunction with $\text{Prob}(J \succ K) < 0.5$. Thus, a majority favoring the safer alternative in the $\{C, E\}$ pair may only change to a smaller majority favoring the safer option, but the modal preference cannot reverse to give a majority favoring the riskier option from $\{J, K\}$.

However, in numerous such experiments, the majority preference *does* actually reverse, contrary to Blavatskyy's model. There are many examples of this, including Problems 3 and 4, 3′ and 4′, 5 and 6, 7 and 8, and 7′ and 8′ in Kahneman and Tversky (1979); the O versus L comparison in Chew and Waller (1986); a number of the triangles presented to Group 1 in Loomes and Sugden (1998); the last of the four cases[2] in Table 2 of Bateman et al. (2006); and all three cases in Table 6 and 5 out of 12 cases in Tables 10 and 11 of Bateman et al. (2006).

There is a further implication of the model in relation to CRE that has not been examined much experimentally but for which some evidence exists. In most experimental examinations of the CRE, C is chosen over E by a substantial majority of respondents. But suppose the parameters are set such that E is more likely to be chosen; i.e., $\text{Prob}(E \succ C) > 0.5$. If the specification of $\phi(\cdot)$ is such as to produce the "usual" CRE, the implication will be that $\text{Prob}(E \succ C) > \text{Prob}(K \succ J) > 0.5$—that is to say, $\text{Prob}(K \succ J)$ will be closer to 0.5 than $\text{Prob}(E \succ C)$. The reason is that all functions that produce the usual CRE in the Blavatskyy model have the opposite curvature on the other side of the 50–50 probability line and so here too probabilities get closer to 50–50 when the distance between the lotteries is reduced.

However, in their Experiment 1, Bateman et al. (2006) found three pairs where a clear majority favored the counterpart of E over the counterpart of C and where that majority became even larger in the case of the counterpart of K compared with the counterpart of J—in two of these cases to a highly significant extent.[3] Such patterns are contrary to the model when $\phi(\cdot)$ is specified in the way required to produce a "standard" CRE pattern.

Implications for the Common Consequence Effect and Corollaries

Given Axiom 4 in his paper, Blavatskyy's (2011) model cannot explain the classic form of the Allais paradox (which Kahneman and Tversky 1979 called the common consequence effect (CCE)). This is noted on page 546 by Blavatskyy, who states that the CCE is directly contrary to that axiom. In terms of Figure 1, the classic CCE takes the form of a comparison between $\{C, D\}$ and $\{J, K\}$, where C is typically chosen over D more often than J is chosen over K. But Axiom 4 requires the probabilities to be the same for both pairs.

On this issue, there is some "special pleading": Blavatskyy claims that although many studies have found evidence of the CCE, these violations are reduced or disappear when lotteries are located *inside* the probability triangle or when the choices involve a gradient of 1. However, there is no mention in the formulation of the model that it applies only to a subset of cases. On the contrary, as already noted, Blavatskyy's abstract claims his proposed model provides a better fit to experimental data than do existing models, but because the great majority of experiments to date have tended to use stimuli where at least one of the lotteries lies on an edge of the triangle, it is strange to introduce the caveat on page 546. Moreover, the great majority of CRE cases involve pairs where *both* lotteries lie on the edge. Part of the claim made for the model is that it accommodates such data, so welcoming "edge" data when they fit with the model but setting them aside when they conflict with the model seems somewhat anomalous.

As noted in (viii) above, the model entails the same choice probabilities for other pairs of lotteries such as $\{A, B\}$, $\{F, G\}$, and $\{G, H\}$, which are on the same gradient and the same distance apart as $\{C, D\}$ and $\{J, K\}$. Fewer studies have examined these other pairs, but there are data that reject this implication of the model. For example, there are a number of such refutations in the patterns reported by Loomes and Sugden (1998), and when testing a different model that shares this

[2] The other three cases also contradict the model but in a somewhat different way, explained below.

[3] In the choice between the certainty of £9 and a 0.50 chance of £25, the choice split was 56:98; in the choice between a 0.20 chance of £9 and a 0.10 chance of £25, the split was 26:128. In the choice between the certainty of £9 and a 0.75 chance of £15, the choice split was 68:85; in the choice between a 0.20 chance of £9 and a 0.15 chance of £15, the split was 23:130. In addition, in the choice between the certainty of £6 and a 0.60 chance of £15, the choice split was 51:102; in the choice between a 0.25 chance of £6 and a 0.15 chance of £15, the split was 38:115.

particular implication, Loomes (2010, Table 2) considered two triangles with different gradients, with each triangle containing a set of three pairs that were all half of the $\{C, E\}$ distance apart and a set of eight pairs that were all a quarter of the $\{C, E\}$ distance apart. These data clearly reject the implications of Axiom 4.

Comparisons and Clarifications

Given the body of evidence that directly challenges Blavatskyy's basic axiom and certain central implications of his proposed model, how is the following statement he makes correct: "Econometric estimation shows that the new model compares favorably with existing models. Its predicted choice patterns often lie significantly closer to actual revealed choices" (Blavatskyy 2011, p. 546)?

In part, the answer may be the particularity of the two data sets to which the models are fitted. However, perhaps a larger part of the explanation lies in the choice of "opponent" models: as any boxing manager knows, a fighter can often build up an impressive record of wins if matched only against a carefully selected subset of opponents.

There are two ways in which other models in the literature differ from the model proposed by Blavatskyy (2011). Any probabilistic model might be thought of as a combination of a "core" theory with a stochastic specification, so different models might vary either in terms of the particular core or in terms of the stochastic specification—or both. His model has an EU core, modified to allow the main arguments to be differences between the EUs of each lottery and the EU of the GLB for that pair. This is then embedded in a stochastic specification of the same general form as the Luce (1959) choice model.

However, given the limitations of the model as far as the CRE and CCE data are concerned, an alternative possibility would be to use a core that is better suited to such data—namely, some form of rank-dependent expected utility (RDEU) model. Alternatives of this kind exist in the literature—for example, see Buschena and Zilberman (2000), a paper that is listed in Table 1 of Blavatskyy (2011)—and it might have been interesting to see how the performance of this kind of model compares with the new model being proposed.

Turning now from the issue of the choice of core model to the issue of the stochastic specification, there is a rather different way of proceeding—the "random preference" (RP) approach (Becker et al. 1963)—which is mentioned but dismissed without (we suggest) adequate consideration. On page 542, Blavatskyy (2011) states that "the main drawback of the random preference model is that it leads to violations of weak stochastic transitivity that are rarely observed in the data..."; on page 546, he goes on to say that "unlike the random preference/utility model, the proposed model does not generate intransitive choice cycles...." Both of these assertions may give false impressions.

First, it is an error to say that the proposed model does not generate choice cycles. *Any* model of probabilistic choice, including this one, that allows some possibility of either alternative in a binary choice being chosen and that also allows these possibilities to be independent between different binary choices is liable to generate at least some—and conceivably, quite a few—choice cycles. Second, it is not true that an RP specification *necessarily* results in violations of weak stochastic transitivity (WST): whether it may or may not do so depends on the specification of the core distribution of preferences and on the nature of the lotteries being presented.[4]

However, by ruling out *all* RP specifications, Blavatskyy (2011) eliminates a competing model that has had at least some success in the past. Loomes et al. (2002) took an EU core and an RDEU core and embedded each in a simple Fechner specification as well as in an RP specification and applied these models to the data set from Loomes and Sugden (1998). Although Loomes et al. (2002) were guarded about extrapolating too much from a single study, the indication was that, for the data set in question, an RP formulation of a simple form of RDEU performed well, although there were signs that respondents' experience might move their behavior in the direction of an RP formulation with an EU core. They also acknowledged that the low rate of violations of dominance may have favored RP over the simple Fechner specification, in which case the Blavatskyy model, being better suited to this low rate, should provide stiffer competition against both of the RP specifications in Loomes et al. (2002). Unfortunately, by ruling RP out on a priori grounds, Blavatskyy leaves this question unanswered. In the absence of any comparison with existing RP formulations with some established credentials, it seems premature to claim so confidently that the proposed model "provides a better fit to experimental data than do existing models."

[4] For example, if we assume that an individual's preferences consist of a set of von Neumann–Morgenstern utility functions exhibiting constant relative risk attitude with some unimodal beta distribution over the risk attitude coefficients, and if we consider nondegenerate lotteries with the same minimum payoff that are mean-preserving spreads relative to some sure amount, there will be no violations of WST. Actually, such lotteries are not unlike those used in many experiments, but the point of this example is not so much to argue for a particular specification as to demonstrate that RP does not always and necessarily lead to violations of WST.

Concluding Remarks

We have shown that there is a considerable body of evidence that the model proposed in Blavatskyy (2011) cannot explain. Such direct evidence may be a more telling basis for evaluating the model than an econometric fitting exercise that is based on just two data sets and that involves comparisons with just a few alternative models, with other potential competitors omitted from consideration. We share with Blavatskyy the need to develop probabilistic models that respect first-order stochastic dominance. However, we suggest that such models should aspire to account for the *full* range of well-attested empirical regularities. The success of such an enterprise requires us to be alert to the limitations and strengths of new models as well as to evaluate competing models on the basis of an open competition with all potentially viable alternative candidates.

Acknowledgments

Graham Loomes acknowledges support from the UK Economic and Social Research Council [Award RES-051-27-0248]. Jose-Luis Pinto-Prades acknowledges support from the Ministerio de Economía y Competitividad (Spanish Government) [Project ECO2010-22041-C02-01] and the Junta de Andalucía [Project 2009 P09.SEJ.4992].

References

Bateman IJ, Day BH, Loomes G, Sugden R (2006) Ranking versus choice in the elicitation of preferences. Working paper, University of East Anglia, Norfolk, UK. http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/loomes/workingpapers/rankchoice.pdf.

Becker GM, DeGroot MH, Marschak J (1963) Stochastic models of choice behavior. *Behavioral Sci.* 8(1):41–55.

Blavatskyy PR (2011) A model of probabilistic choice satisfying first-order stochastic dominance. *Management Sci.* 57(3):542–548.

Buschena D, Zilberman D (2000) Generalized expected utility, heteroscedastic error, and path dependence in risky choice. *J. Risk Uncertainty* 20(1):67–88.

Chew SH, Waller WS (1986) Empirical tests of weighted utility theory. *J. Math. Psych.* 30(1):55–72.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.

Loomes G (2010) Modeling choice and valuation in decision experiments. *Psych. Rev.* 117(3):902–924.

Loomes G, Sugden R (1998) Testing different stochastic specifications of risky choice. *Economica* 65(260):581–598.

Loomes G, Moffatt PG, Sugden R (2002) A microeconometric test of alternative stochastic theories of risky choice. *J. Risk Uncertainty* 24(2):103–130.

Luce RD (1959) *Individual Choice Behavior* (John Wiley & Sons, New York).

Machina M (1982) "Expected utility" analysis without the independence axiom. *Econometrica* 50(2):277–323.