Quantitative Economics 6 (2015), 825–856 1759-7331/20150825

Response mode and stochastic choice together explain preference reversals

Sean M. Collins
Economics Department, Fordham University

Duncan James
Economics Department, Fordham University

Informed by Grether and Plott (1979) and Cox and Grether (1996), we implement various preference elicitation procedures over a parameter grid. First, we find a lower incidence of preference reversals for probability equivalents from the dual-to-selling version of Becker, Degroot, and Marschak (1964; BDM) than for certainty equivalents from traditional BDM—consistent with conjectures regarding response mode. Second, the Blavatskyy (2009, 2012) model of probabilistic choice can explain the incidence of preference reversals when using probability equivalents. Thus, between response mode (outside the Blavatskyy model) and stochastic choice (as per Blavatskyy), preference reversals in the original certainty equivalent case seem to be explained. We also present estimates for risk and stochasticity parameters; the former are not correlated across mechanisms, but the latter are. Finally, relatively more error-laden behavior (based on *within*-mechanism checks) can be associated with fewer reversals *across* mechanisms. The data make clear, empirically, the logical proposition that reducing reversals requires only a better "match" with binary choice, not necessarily rational behavior at any deeper level.

Keywords. Preference reversals, probabilistic choice, mechanism design, institutions.

JEL classification. C91, D47, D81.

The preference reversal phenomenon has been resistant to solution for decades. Lichtenstein and Slovic (1971, 1973) defined the field. They found that, for a given subject, there is a substantial likelihood that the subject's responses regarding a set of lotteries will differ across mechanisms. Furthermore, those responses will not be rationalizable by a single set of expected utility preferences. There, as in subsequent research,

Sean M. Collins: scollins15@fordham.edu
Duncan James: dujames@fordham.edu
The authors are grateful for helpful comments from the editor, three anonymous referees, Daniel Friedman, Mark Isaac, Maros Servatka, Joel Sobel, Richard Watt, seminar participants at the University of Canterbury, and the Florida State University Experimental Economics Readings Group. The authors acknowledge generous financial assistance from Fordham University's Office of Research in the form of a Faculty Research Grant.

DOI: 10.3982/QE437

the two mechanisms usually juxtaposed are binary (either/or) choice and the selling version of the Becker, Degroot, and Marschak (1964) procedure (hereafter BDM).

The existence of the preference reversal phenomenon was found to be robust to the methodological variations implemented by Grether and Plott (1979). Cox and Grether (1996) succeeded in reducing the incidence of preference reversals, but did so by replacing BDM with various multiperson markets as the valuation task. This changed one aspect of the problem from an individual decision-making setting to a market setting. In this paper, we seek to determine whether one can reduce the incidence of preference reversals while remaining entirely within the realm of individual decision-making. In particular, we set out to see whether using a form of BDM that has as its response mode not currency, but probability, reduces the incidence of preference reversals. That this might be possible is suggested by a discussion in Grether and Plott, based on work by Slovic (1975). There it is conjectured that information processing, as influenced by response mode, might play a role in generating preference reversals (a conjecture that has also received some follow-up study, notably by Tversky, Slovic, and Kahneman (1990)).

In the data reported in this paper, we find that between (a) changing response mode from currency to probabilities and (b) allowing for stochastic choice (as per Blavatskyy (2009, 2012)), one is able to account for the preference reversal phenomenon.

As an adjunct to this, we present parametric estimates from the data generated within the preference reversal experiments, something notably absent from the previous 40+ years of research on the topic. This allows cross-mechanism comparisons for risk preference parameters (analogous to Isaac and James (2000) and Berg, Dickhaut, and McCabe (2005)) and similarly for the parameter governing stochasticity-in-choice (as per Blavatskyy). We find that risk preference parameters are uncorrelated across mechanisms, but that the stochasticity parameters *are* correlated across mechanisms. This further supports the explanation that mechanism-specific mapping of stochastic choice to asymmetric deviations in valuation (relative to risk neutrality) is responsible for generating preference reversals. That is to say, that the operational mechanics of each procedure, notably including response mode, seem to channel stochastic choice by subjects in aggregate patterns characteristic to each procedure.

1. Background

Lichtenstein and Slovic (1971) pioneered the study of preference reversals: the juxtaposition of elicited certainty equivalence responses from the selling version of the BDM procedure with the results of either/or binary choice.[1] The principle is elegantly simple: optimizing agents who prefer lottery A to lottery B in a direct comparison should place a higher valuation on lottery A than lottery B and vice versa. Furthermore, the prescription of standard decision theories (such as expected utility, and many relaxations thereof) is clear: failure to assign valuations in the same ranking as binary expressions of preference allows for cycling in choice; cycling in choice might allow for the agent to be systematically fleeced via a "Dutch book."

[1] There is earlier work on intransitive choice, such as Tversky (1969).

The existence of preference reversals in experimental data tells us there has to be a problem somewhere, whether for the interpretation of such decision theories as being descriptive of human behavior, or for the subjects who might be prone to having Dutch books made against them, or within the complex of mechanisms, procedures, and parameters used in the experiments in question. Follow-up efforts in the literature have sorted through a wide range of experimental design features and alternative descriptive decision theories in an attempt to eliminate the discrepancy between (arguably still prescriptive) decision theory and subject behavior.[2]

Grether and Plott (1979) attempt to eliminate candidate explanations including income effects and hypothetical incentives. Preference reversals continued to be observed nonetheless. Cox and Epstein (1989) set out to see whether a different valuation mechanism might eliminate reversals. They replace BDM with a procedure wherein a given subject states valuations for each of two gambles in a pair; the gamble on which a higher value is placed is then played out by the subject, while the other gamble is bought from the subject by the experimenters for a flat fee. This procedure eliminates the multiple random draws involved in running BDM. But preference reversals are still observed, suggesting that the conjunction of violation of the independence axiom by subjects and use of BDM by experimenters need not be responsible for the occurrence of preference reversals.

Slovic (1975) initiated research investigating the possibility that differences in response mode—the units in which responses are communicated—sets up the conditions for preference reversals to occur. Tversky, Slovic, and Kahneman (1990) present evidence, consistent with that in Cox and Epstein, that changing from BDM to another mechanism that is ordinal in operation does not eliminate preference reversals. The Tversky, Slovic, and Kahneman (1990) procedure uses "prices" only as a way for subjects to indicate preference, not as costs incurred in trade. Specifically, the subject is asked to place a minimum selling price on each of two lotteries; the subject is then asked to express direct preference for one lottery or the other; the subject knows that one of the two preceding tasks will subsequently be selected for actual payment, at which point either the higher "priced" lottery (if the first task is randomly drawn) or the directly preferred lottery (if the second task is randomly drawn) will be played out for money. The authors point out that use of BDM in the presence of violation of the independence axiom cannot explain all preference reversals, as they are still observed when using the mechanism just described. In contrast, both scale compatibility and prominence receive some support as possible contributing factors. *Scale (in)compatibility* refers to the imperfect correspondence between the units in which information is denoted in (a) the attributes of an object being evaluated and (b) the response mode afforded the subject. The application to preference reversals is that asking for currency-denominated certainty equivalents as responses in one mechanism, but in another asking for direct

[2]Seidl (2002) provides a thorough survey of earlier work on the preference reversal phenomenon. He claims that the preference reversal phenomenon has been explained to be caused by three determinants: the elicitation mode of certainty equivalents, intransitivity of preferences, and over- and/or underpricing of lotteries. (He also considers nonlinear probabilities, as these apply to discount reversals.) We address how our findings relate to Seidl's survey in footnote 22.

choice between pie charts with probability-denominated areas may initiate different information-processing sequences in a given subject. *Prominence* refers to the relative salience of attributes of a good or lottery; for instance, probability-area pie charts might make probability more prominent in the sense that probability is considered before currency in an effectively lexicographic manner. The two explanations are not necessarily mutually exclusive.

Another line of inquiry developed around noise-based explanations of preference reversals. These studies also started using probability equivalents (as opposed to just certainty equivalents, as in the original studies). MacCrimmon and Smith (1986), Cubitt, Munro, and Starmer (2004), and Butler and Loomes (2007) all pursue versions of this approach. However, none of them uses probability equivalents elicited in an analogous manner to the way in which certainty equivalents were elicited in the original Lichtenstein and Slovic (1971) study.

New developments in mechanism design and in the modeling of probabilistic choice allow us to revisit this problem with new hope of finding a solution. First, instead of running the original selling version of BDM, which elicits currency-denominated certainty equivalents, one could run a dual version of it. Relative to selling BDM, the dual-to-selling version reverses the direction of trade (cash for lottery instead of lottery for cash) and the response mode (probability instead of cash).[3]

The dual-to-selling version of BDM endows the subject with a cash amount instead of a lottery. Subject response takes the form of a probability that is compared to a random draw. The object obtainable by the subject in place of the initial cash endowment is a lottery with dollar amounts of two end states disclosed at the start of the period, but with probabilities established by the aforementioned draw. The subject exchanges the cash amount for the lottery only if the terms she states are less favorable to her than those established by the random draw (though in the event of an exchange, the random draw becomes the probability used to parameterize the lottery). Thus one can elicit probability equivalents in the same operational manner as traditional elicitation of certainty equivalents (James (2011)).

Second, there is also now available an axiomatically derived model that allows for a particular type of probabilistic choice, can accommodate von Neumann–Morgenstern (VNM) preference structures, and can allow for preference reversals: the stochastic preference model of Blavatskyy (2009, 2012).[4] The Blavatskyy model is constructed so as to be consistent with two main stylized facts regarding empirical studies of decision-making under risk: (a) revealed choice appears probabilistic in nature but (b) seldom violates stochastic dominance.

[3]The dual-to-selling version of BDM differs from the buying version of BDM in that the buying version (a) still has cash as its response mode and (b) does not allow construction of intransitive cycles, as it generates noncomparable wealth positions. The dual-to-selling version both reverses the direction of trade (like the buying version) and switches the response mode to probability. It also preserves the ability to construct intransitive cycles, as with the selling version.

[4]The Blavatskyy (2009, 2012) model is within the class of models originally proposed by Luce (1959). Blavatskyy (2014) also augments the Fechner model to allow for better correspondence with empirical observations concerning preference reversals.

2. Research questions

The overarching concern with preference reversals is that they should not exist, at least in a world of deterministic application of expected utility theory. Given that their existence has proven robust and durable, natural questions then present themselves: Can preference reversals be reduced or eliminated? How do preference reversals happen in the first place? Answering the first question might provide evidence that at least narrows down the possible answers to the second question, and answering that first question might be possible through systematic investigation, and variation, of the operational mechanics of the elicitation procedures.

A possible variation could involve changing the response mode through which subjects express their decisions. Slovic (1975) (and Grether and Plott (1979), citing Slovic) raise the possibility that in asking subjects to (a) nominate currency-denominated certainty equivalents for lotteries and then (b) engage in pairwise choice between pie chart representations of lotteries, wherein the area within each chart is probability-denominated, experimenters might be setting subjects up to fail. The process of formulating and giving a response requires information processing, and different formats for informational input might cue up different, informal "calculational" approaches, which might then reach different conclusions even when the algebraic substance of the input is the same.

A natural progression within the preference reversal literature would then be to alter BDM so as to elicit responses denominated in terms not of currency, but of probability. This might eliminate at least some of the mismatch in media that earlier researchers hypothesized might be creating the conditions for preference reversals. The dual-to-selling version of BDM (James (2011)) is exactly suited to carrying out this perturbation (and will be discussed further in Section 3).

Another possible cause of preference reversals is stochastic choice. It is not necessarily a substitute for, but perhaps rather a complement to, the information processing conjecture just discussed. For example, consider the stochastic choice model of Blavatskyy (2012). It is a system within which the probability that the decision-maker will respond with an incorrect expression of his or her "true preference" varies as a function of the magnitudes of the probabilities and currency amounts associated with lottery objects being compared. This model does not incorporate a role for response mode; such considerations are outside the scope of the Blavatskyy model. As such it may be possible that both stochasticity in expression of preference, as per Blavatskyy, and dependence of human calculation on the format of informational inputs, as per Slovic or Grether and Plott, might each have a role in generating preference reversals.

Preference reversals are not the only possible violation of standard decision theory that might be observed using the mechanisms employed in preference reversal studies. Observing violation of dominance in a given round is also possible. This could take place when eliciting the certainty equivalent for a degenerate lottery. Across-round ("chained") dominance violations may also be observed when one provides a higher valuation for one lottery than another, despite the former being dominated by the latter. (This is discussed further in Section 3.) Do these other forms of "choice pathology" allow us to further allocate the responsibility for preference reversals?

In the results section, we will address all of these possibilities. We will also present parametric estimates of both risk preference parameters and stochasticity parameters, jointly estimated. These estimates also shed light on the sources—and it seems there are no less than two—of preference reversals.

3. Design

We face a number of constraints in our design. For instance, we are compelled to use the original mechanisms: binary choice and the selling version of BDM (Lichtenstein and Slovic (1971, 1973)). We do also use the dual-to-selling version of BDM (James (2011)) to systematically vary response mode while remaining in the same family of mechanisms as the selling version of BDM.[5] We thus remain entirely within the sphere of individual decision-making mechanisms, rather than using multiperson markets. For purposes of comparability with Grether and Plott (1979), we use what is now often referred to as the *pay one randomly* (POR) payment protocol, wherein a single round from a multiround experiment is chosen for payment.[6] Additionally, our desire to implement a systematic and broad-based set of comparison lotteries leads us to implement a parameter grid in the spirit of Butler and Loomes (2007). We interpret this grid in the context of the probabilistic choice model of Blavatskyy (2009, 2012).

The preference reversal literature has typically paired a p bet (with a relatively higher probability of a lower upstate payoff) and a $ bet (with a relatively lower probability of a higher upstate payoff). Because the grid design we employ contains multiple such pairs, we denote the p bet and the $ bet in a given pair of gambles, i, as Pi and Di, respectively. The grid first presented in Figure 1 graphically illustrates the probabilities and paired outcomes of lotteries with dichotomous outcomes. Each lottery has a $0 down-state outcome. The grid then organizes ordered pairs of probabilities and dollar amounts, which are in turn representative of different gambles (including degenerate gambles, already familiar to economists as certainty equivalents, along the top boundary of the grid).

In traditional preference reversal studies, the numerical elicitation task (BDM) asks the subject to evaluate a gamble located in the interior of the grid, by nominating a dollar amount on the upper boundary of the grid that is acceptable in place of the gamble. This task is then repeated for another such gamble. In the binary choice task, an either/or choice takes place between these two "interior" gambles.

The dual-to-selling BDM mechanism reverses the action over the grid. Using it for probability elicitation in the "down" direction, the subject starts out with a currency amount on the upper boundary of the grid, then nominates probability equivalents that complete the parameterization of lotteries in the grid below.

[5]Bohnet and Zeckhauser (2004) and Bohnet, Greig, Herrmann, and Zeckhauser (2008) present a mechanism with similar incentive properties but different operational mechanics. Andreoni and Sprenger (2011) use a related list mechanism.

[6]The pay one randomly incentive system has been proven by Azrieli, Chambers, and Healy (2012) to be the only incentive compatible payment protocol over multiple rounds, given a hypothesis that any "underlying" preferences are monotonic. A recent study by Cox, Sadiraj, and Schmidt (2014) examines the impact of payment protocol experimentally.

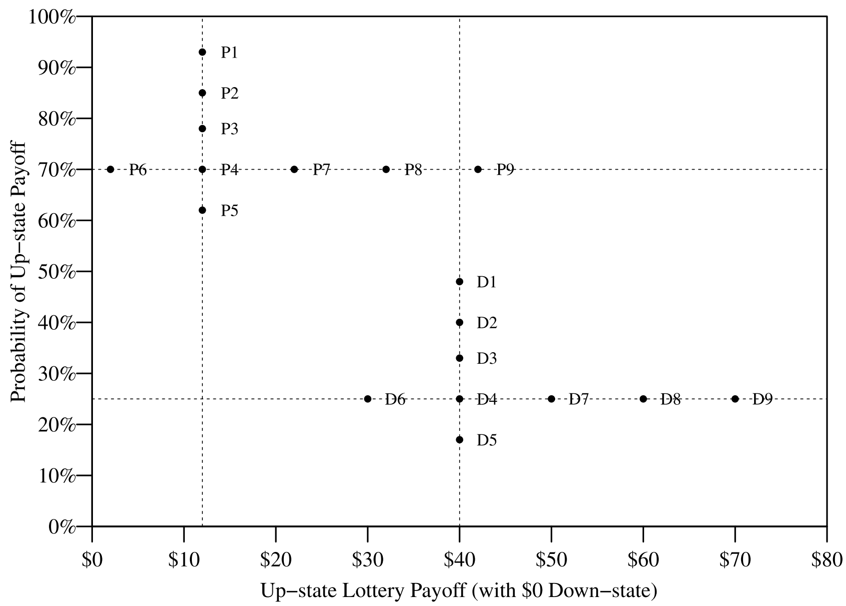

FIGURE 1. Grid of lotteries around the Butler and Loomes (2007) points.

The dual-to-selling procedure works in the following manner. The subject is endowed with a cash amount and is asked to nominate the probability of the high state outcome that would make the latter gamble acceptable in place of that which the subject was endowed with initially. Naturally, the subject is "kept honest" in this endeavor by comparing the probability they nominate to a draw from a uniform distribution over $[0, 1]$, with the subject only making the exchange if the number they nominate is less than the number drawn from the uniform distribution. Suppose the exchange is made; a second draw then takes place to determine the realization of the end states in the gamble. That draw is from a uniform $[0, 1]$, but the mapping of the second draw to the occurrence of either the high or low end state is governed by the realization of the first draw. If the second draw is equal to or less than (resp. greater than) the first draw, the high payoff end state (resp. low state) is realized. This is a dual to the selling version of BDM, the traditional mechanism of the preference reversal literature, with response mode now in probability, rather than currency.

This procedure was used as described above for elicitation of probability equivalents in the down direction in the grid; that is, when the subject was initially endowed with cash. It was also used for elicitation of probability equivalents when the possible exchange was lottery for lottery. These rounds were run as a check/calibration with respect to Butler and Loomes' study, and not as natural pairs or mirror images to the original elicitation task. Instructions detailing both the dual-to-selling version of BDM and the selling version of BDM may be found in Appendix C, available in a supplementary file on the journal website, http://qeconomics.org/supp/437/supplement.pdf. (Appendix B is also available in the supplementary file. Code and data are available in another supplementary file on the journal website, http://qeconomics.org/supp/437/code_and_data.zip.)

It should also be noted that the nature of the dual-to-selling mechanism imposes constraints on the ordering of rounds within an experiment. Since the lotteries in dual-to-selling are not fully specified initially, binary choice involving these lotteries must take place after subjects have specified probability equivalents. This in turn necessitated checks against subjects being able to manipulate this aspect of the design. Subjects did not have any way to know that subsequent binary choices would be related to earlier responses, and binary choice did not involve the exact probability numbers subjects had submitted, but rather lottery probabilities generated by adding a small masking layer to their original answers (drawn from either a discrete uniform distribution over $\{0.01, 0.02\}$ or over $\{-0.01, -0.02\}$). This also eliminated the possibility of indifference between the pair of lotteries presented in the binary choice phase. Note also that Cox and Grether (1996, p. 385) employed endogeneity of a similar kind.[7]

Finally, we should note that earlier work on imprecision and preference reversals by MacCrimmon and Smith (1986) and by Butler and Loomes (2007) explores whether imprecision in preference over certainty equivalents can lead to certainty equivalents that are higher (resp. lower) for the gamble otherwise revealed in binary choice to be dispreferred (resp. preferred). Butler and Loomes (2007) graphically represent the range of valuation responses under the imprecision by cones emanating from the gambles being valued.

The gambles used in Butler and Loomes, for example, are $P4$ and $D4$ (only) in the grid.[8] But might not the incidence of reversals change as one varies the Pi and Di pairs being valued and ranked? The axiomatically derived theory of probabilistic choice proposed by Blavatskyy (2012) can be used to make just such a prediction about the incidence/frequency of occurrence of preference reversals at different locations within the design grid.

In brief, Blavatskyy (2009, 2012) conjectures that an individual chooses some lottery A over B with probability

$$p(A, B) = \frac{\varphi(u(A) - u(A \wedge B))}{\varphi(u(A) - u(A \wedge B)) + \varphi(u(B) - u(A \wedge B))}, \tag{1}$$

[7]The same process of improvement or worsening was applied across all mechanism pairings and rounds. That is to say, it was applied to binary choice even to lotteries that were fully specified at the outset, as with certainty equivalence. The lottery associated with higher upfront cash value in dual-to-selling BDM or the higher certainty equivalent in selling BDM was always improved. The lottery with the higher probability of the reference lottery outcome was improved in the "projecting left" and "projecting right" cases. The other lottery in that binary choice pair always worsened. All numerical computations in this paper accurately reflect these adjustments.

[8]While we used U.S. dollar (USD) denominations that were half as great as Butler and Loomes's (2007) denominations in Australian dollars, the Australian dollar bought a little more than half as much USD at the time Butler and Loomes ran their experiments. We return to this point in footnote 17. However, the primary purpose in varying parameters across the grid is to vary the "relative" distance of the gambles to the boundaries in question, and not to exactly recreate the conditions under which the Butler and Loomes study was conducted.

where $u(\cdot)$ is the von Neumann–Morgenstern expected utility function, $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is a nondecreasing function with $\varphi(0) = 0$, and $A \wedge B$ denotes the greatest lower bound on A and B in terms of weak stochastic dominance. The Blavatskyy model makes distinct predictions for "either/or" binary choice and for valuation elicitation procedures such as BDM and dual-to-selling BDM.

Blavatskyy (2009) offers a means by which this model can allow for preference reversals. In the model, the certainty and probability equivalents of a lottery are random variables. For instance, the certainty equivalent of lottery A is a cumulative distribution function (c.d.f.) $\mathrm{CE}_A = p(x, A)$, where x is the dollar amount of the certainty equivalent (i.e., a degenerate lottery). The observed certainty equivalent is one possible draw from this distribution. The probability that an individual states a higher certainty equivalent for lottery B than lottery A is $p(\mathrm{CE}_B > \mathrm{CE}_A)$. Thus, the probability that one form of preference reversal is observed is $p(A, B) \cdot p(\mathrm{CE}_B > \mathrm{CE}_A)$. A mathematically detailed description of the Blavatskyy model may be found in Appendix A.

An application of the Blavatskyy model to the grid-like design we implement is illustrated in Figure 2. Illustrated in the figure are the lower and upper bounds of the 35% confidence intervals for cash equivalents as predicted by a two-parameter Blavatskyy model fitted for a selected subject.[9] The illustrated case focuses solely on the possible role of stochasticity in certainty equivalents elicited via BDM. The Blavatskyy model also

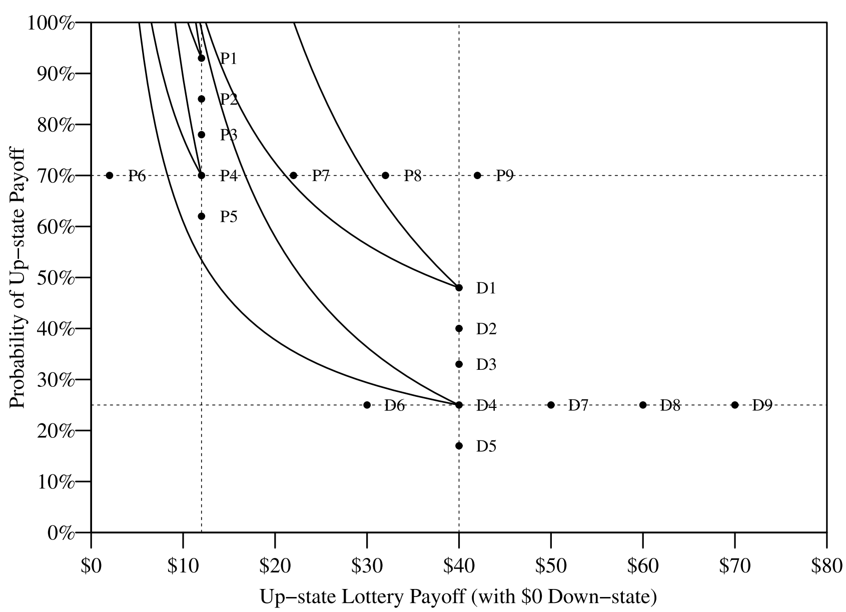

FIGURE 2. Example of Blavatskyy model 35% confidence intervals for gambles $P4$ and $D4$ and gambles $P1$ and $D1$.

[9]Subject 19 was selected for the confidence intervals illustrated in Figures 2 and 3 because the subject is "typical"—the subject's two estimated parameters have the lowest Euclidean distance to the median of those estimated within the sample.

predicts a certain probability of choosing either $P4$ or $D4$ in binary choice.[10] Assuming for the moment that stochasticity in binary choice does not contribute to reversals, one would expect to see a lower incidence of reversals in the cross section of results comparing gambles $P1$ and $D1$ than in the cross section of results comparing $P4$ and $D4$, all else constant. In this specific example, the lack of overlap between the confidence intervals associated with $P1$ and $D1$ on the upper boundary signifies that variation in certainty equivalent response is less likely to reverse the ordering of the certainty equivalents than for $P4$ and $D4$, for which the confidence intervals do overlap on the upper boundary.

For other elicitation procedures and types of exchange, the Blavatskyy model makes analogous predictions to those just described. Figure 3 illustrates the four cases we consider in this study: project up, project down, project left, and project right.

Table 1 describes the treatments in the experimental design and details the parameterization of the p-bet and \$-bet pairs.[11] For all treatments, the distance to the "bound-

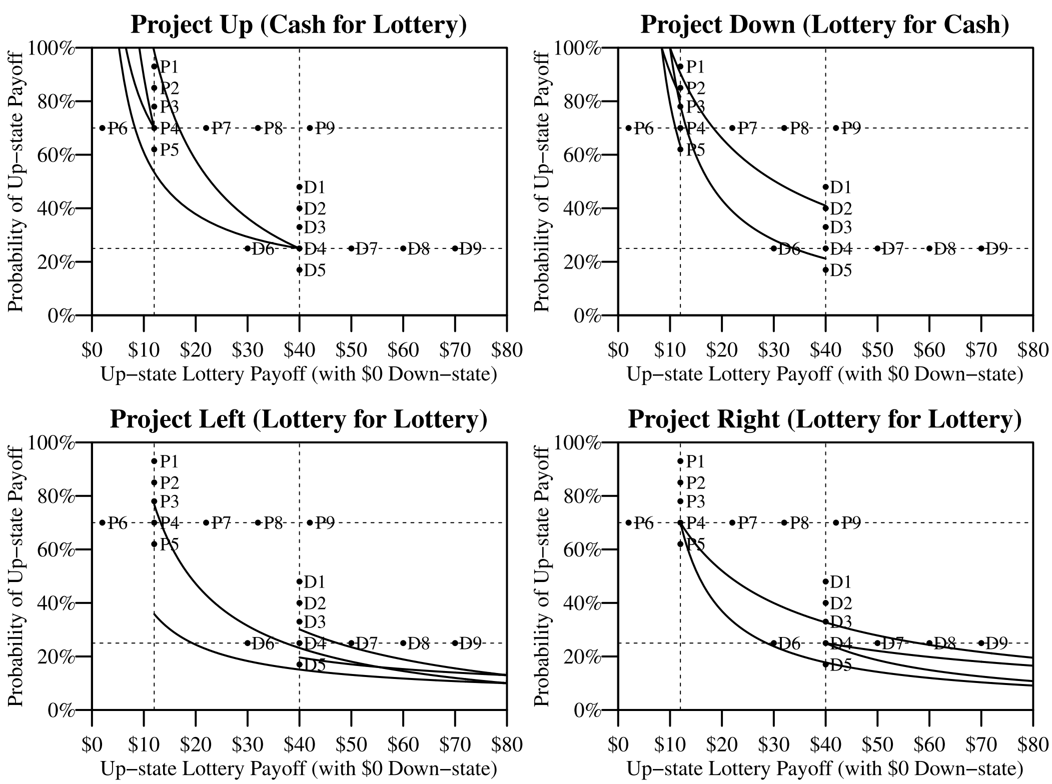

FIGURE 3. Example of Blavatskyy model 35% confidence intervals by treatment for gambles $P4$ and $D4$.

[10] For the gambles illustrated in Figure 2, the fitted model predicts that the subject has an approximately 49% chance of selecting gamble $P4$ over $D4$ and a 26% chance of selecting $P1$ over $D1$ in binary choice.

[11] Experiments were computerized with z-Tree (Fischbacher (2007)). Subjects were invited from a database of volunteers over e-mail using ORSEE (Greiner (2004)) and by manually sent e-mails from the database when ORSEE was unavailable.

TABLE 1. Treatments and parameters defining the p-bet and \$-bet grid.

Treatment Name	Elicitation Mechanism	Elicited Value	p Bets	\$ Bets
Project up (up)	Selling BDM (lottery for cash)	Certainty equivalent (CE)	P1: 0.93 ∘ \$12 *P2: 0.85 ∘ \$12* P3: 0.78 ∘ \$12	D1: 0.48 ∘ \$40 *D2: 0.40 ∘ \$40* D3: 0.33 ∘ \$40
Project down (down)	Dual-to-selling BDM (cash for lottery)	Probability equivalent (PE)	*P4: 0.70 ∘ \$12* P5: 0.62 ∘ \$12	*D4: 0.25 ∘ \$40* D5: 0.17 ∘ \$40
Project left (left)	Dual-to-selling BDM (lottery for lottery)	Probability equivalent (PE)	P6: 0.70 ∘ \$2 *P4: 0.70 ∘ \$12* P7: 0.70 ∘ \$22	D6: 0.25 ∘ \$30 *D4: 0.25 ∘ \$40* D7: 0.25 ∘ \$50
Project right (right)	Dual-to-selling BDM (lottery for lottery)	Probability equivalent (PE)	*P8: 0.70 ∘ \$32* P9: 0.70 ∘ \$42	*D8: 0.25 ∘ \$60* D9: 0.25 ∘ \$70

Note: The p bets (high probability bets) and the \$ bets (high up-state payoff bets) in *italics* were repeated as an ordering control. The down-state outcome of all bets is \$0. The points evaluated in Butler and Loomes (B&L) are $P4$ and $D4$. Certainty equivalents for the vertical p bets ($P6$, $P4$, $P7$, $P8$, and $P9$) were also collected, as were certainty equivalents and probability equivalents for $0 \circ \$12$ and $1 \circ \$12$.

ary" of the grid is varied. For up and down, we collect responses for the vertical pairs of p bets and \$ bets ($P1$–$P5$ and $D1$–$D5$) in the grid. For left and right, we collect responses for the horizontal pairs of gambles ($P6$–$P9$ and $D6$–$D9$).

The remaining details of the design are then as follows: 60 subjects participated in the experiment, each making 90 incentivized decisions producing a total of 5400 observations. A single round was randomly selected for payment, as in Grether and Plott (1979). Of the 90 decisions, each subject made 48 selling/dual-to-selling BDM decisions followed by 25 binary choice decisions and 17 additional selling/dual-to-selling BDM decisions as an ordering control. The first block of 48 was divided into 16 certainty equivalent elicitations (projecting up to the upper boundary, using the selling version of BDM), 12 probability elicitations (projecting down to interior gambles, using the dual-to-selling version of BDM), 10 probability elicitations (projecting left from boundary gambles to interior gambles, using the dual), and 10 more probability elicitations (projecting right from interior gambles to boundary gambles, using the dual).[12] Within blocks, ordering of tasks was varied as a further control.[13]

[12]The reader may wonder why the number of observations varies for each direction/mechanism combination. For all treatments, numerical responses were elicited for 10 grid points ($P1$–$P5$ and $D1$–$D5$ for up and down or $P6$–$P9$ and $D6$–$D9$ for left and right) listed in Table 1. For each of up and down, two dominance checks (detailed in the text) were conducted; these checks relied on degenerate gambles, but were not conducted in left or right, as these latter two treatments did not involve either the upper or lower boundaries of the grid. For up (the "traditional" preference reversal treatment), we provided for violation patterns in certainty equivalents technically possible for certainty equivalent elicitation only by also eliciting $P6$–$P9$, omitting $P4$ as it was elicited previously.

[13]Within the first block, there were a roughly similar number of observations for all six permutations of lottery for cash, cash for lottery, and lottery for lottery. The second block consisted of binary choice. The third block replicated the ordering of the first block and consisted of a subset of the p bets and \$ bets in the first block as an ordering control (as noted in Table 1). We find that there is no statistical difference between the frequency of reversals due to order within the sequence of the treatments by a subject-level Mann–Whitney test on number of reversals ($p \approx 0.3952$ for up and $p \approx 0.09915$ for right, with no in-sample difference for down and left; see Table 3).

"Dominance checks" were employed for each version of BDM. One check was the inclusion of rounds employing degenerate lotteries, for which the response required by the dominant strategy of truthful revelation is known to the experimenter. Note that the Blavatskyy model does *not* allow for violation of dominance in this form: the probabilities assigned by the Blavatskyy model in this case are themselves degenerate, and thus do not allow for any response other than the surplus-maximizing dominant strategy.

Another kind of check was formed by the nature of the grid itself: in the absence of probabilistic choice, but assuming choice respects first-order stochastic dominance, as one descends vertically from Pi to $Pi+1$ or from Di to $Di+1$, the "true" valuation one places on gamble $i+1$ should necessarily be lower than the true valuation placed on gamble i, as both gambles have the same end states but gamble $i+1$ awards less probability mass to the high-state outcome. We denote placing a higher valuation on gamble $i+1$ (or $i+2$, $i+3$, $\dots$) than gamble i in the grid a *chained dominance violation*.

The Blavatskyy model allows for violation of these orderings and can offer predictions as to the frequency of such occurrences. How? It is true that for binary choice between any two lotteries, the Blavatskyy model implies that choice must satisfy first-order stochastic dominance; that is, an individual always prefers A to B if A stochastically dominates B (in which case, $B = A \wedge B$) and vice versa. But for numerical valuations, such as those elicited by BDM, the model *does* allow for a stochastically dominated lottery to be assigned a higher valuation than that assigned to the lottery that dominates it. This is possible because in Blavatskyy's theory of how valuation is arrived at in BDM, the $p(A, B)$ function at the heart of the Blavatskyy model is applied repeatedly; that is, the subject is conjectured to behave as if facing a continuum of binary choices between (a) the lottery being valued and (b) a continuum of certainty dollar amounts. The subject is modeled as (i) making a $p(A, B)$ assessment between each such "pair." Then (ii) a c.d.f. is formed from the continuum of $p(A, B)$ assessments (i.e., assigning probability mass to the events, which in this case are particular dollar amounts for the certainty equivalent). Finally, (iii) a draw is made from this c.d.f. This draw is the subject's response, which is what would ultimately be observed by the experimenter. A certainty equivalent arrived at by this process is the outcome of a process including and reflecting stochastic choice, and two such certainty equivalents may be observed in the reverse order to that suggested by first-order stochastic dominance. Analogous misalignment in numerical responses may also be observed for probability equivalents.

4. Results

4.1 *Preview of findings*

We find that the incidence of preference reversals can be explained by two factors, which are uncovered in two steps. First, switching the numerical elicitation procedure from traditional BDM to dual-to-selling BDM, thus altering response mode, reduces the frequency of reversals. Second, such reversals as remain at that point can then by explained by stochastic choice as per the Blavatskyy model.

Table 2. Proportion of the subject pool exhibiting a given number of reversals.

Number of Reversals		Percent of Subject Pool			
		Project Up	Project Down	Project Left	Project Right
0	(Minimum possible)	0.1167	0.2000	0.1000	0.1833
1		0.2167	0.4500	0.2167	0.3167
2		0.1667	0.1833	0.3333	0.3333
3		0.2333	0.1167	0.2667	0.0833
4		0.1500	0.0500	0.0833	0.0667
5	(Maximum possible)	0.1167	0.0000	0.0000	0.0167

Additionally, we find that this lower incidence of reversals for comparisons of dual-to-selling BDM and binary choice is not necessarily evidence of either greater subject rationality or a better elicitation procedure. This is made clear by the fact that fewer reversals are observed when using the dual-to-selling version of BDM, but more violations of prescribed decision-theoretic behavior are observed internal to this version than internal to the traditional version of BDM.

4.2 *Role of the elicitation procedure and of the response mode inherent in each*

First of all, we find fewer reversals in the project down treatment than in the project up. Table 2 summarizes results across all mechanism/parameterizations, presenting them on a proportion frequency basis.

There are clear differences in prevalence of reversals across treatments. The up treatment (the original Lichtenstein and Slovic preference reversal comparison) exhibits by far the greatest proportion of occurrences, followed by left; right is associated with a still lower incidence of reversals, while down exhibits the lowest. A χ^2 test of homogeneity rejects that the up and down columns of Table 2 are the same ($p \approx 0.0023$); obviously down has more subjects who exhibit few or no reversals.

This finding suggests that changing to a different valuation mechanism with a different response mode has some efficacy in eliminating reversals. As such, it seems that scale incompatibility might be a cause of at least some of the up reversals.[14]

4.3 *Blavatskyy model predictions and fit*

The portion of preference reversals that is not eliminated with a change in response mode appears to be explained by stochasticity-in-choice, as per the Blavatskyy model.

[14] Suppose one were to raise the question, "Is endogenization of the lotteries to be used in binary choice sufficient in itself to change the rate of reversals?" To begin with, it is not obvious a priori that we would expect such to be the case. Also, the similar approach used by Cox and Grether (1996) did not appear to have any such effect. Furthermore, empirically we see that there is not a consistent, across-the-board effect of that sort in our data. Left, using endogenized lotteries, exhibits significantly different ($p < 0.0001$ in a χ^2 test over proportions) and *greater* incidence of reversal than its paired treatment, right, which has initially fully specified lotteries as the basis for binary choice. Down, using endogenized lotteries, exhibits a significantly different ($p \approx 0.0046$) and *lower* rate of reversals than its paired treatment, up.

That is to say, the Blavatskyy model can explain the frequency of preference reversals observed with the dual-to-selling version of BDM, but cannot explain the frequency of reversals observed with the selling version of BDM. Conversely, the response mode can explain the reduction in frequency of preference reversals in changing from certainty equivalents to probability equivalents, but cannot explain the continued, albeit diminished, existence of preference reversals under the latter.

Table 3 presents, for each treatment, for each comparison pair of lotteries, both the raw frequency of reversals and the fitted frequency from the Blavatskyy model. It does so for what have been termed in the literature standard and nonstandard types of reversal.[15] Standard reversals involve selection of the p bet over the \$ bet in binary choice, while nonstandard reversals involve a choice of the \$ bet over the p bet in binary choice (each contravening their respective certainty equivalent orderings). [16] Fitting of the model is done on the individual level and employs joint maximum likelihood estimation using data from the selling version of BDM, the dual-to-selling version of BDM, and binary choice. The estimation procedures are detailed in Appendix A.7.

The results provide constructive, specific feedback concerning the Blavatskyy model. The fitted model works well throughout most of the design space—with the notable exception of the traditional juxtaposition of BDM certainty equivalents and binary choice. In each of treatments left and right, the fitted Blavatskyy model appears to miss in a roughly symmetrical manner, and the overall averages of predicted and actual are close (see the “Total” row). Down, however, shows a systematic tendency by the fitted Blavatskyy model applied in this setting to overpredict standard reversals and in the one possible instance of nonstandard reversal, underprediction. By far the worst performance of the fitted Blavatskyy model in predicting reversals is the up treatment, as the model systematically underpredicts standard reversals and overpredicts nonstandard reversals in this case.[17] This opposite-in-sign performance, compared to that in down, also exhibits far greater magnitudes in terms of the size of the misses. This supports the multicausal (response mode *and* probabilistic choice) explanation of preference reversals, as the “odd pairing out” in terms of excess reversals is the one with the “mismatched units”: BDM, eliciting dollar-denominated certainty equivalents, juxtaposed with probability-area pie charts.

[15]The Blavatskyy model also predicts the incidence of “strong reversals” discussed in Fishburn (1988) and Butler and Loomes (2007). An analysis of strong reversals is presented in Appendix B.3.

[16]In down, it is possible to observe a standard reversal (only) when the upfront value of the \$ bet exceeds that of the p bet; it is possible to observe a nonstandard reversal (only) when the upfront value of the p bet exceeds the upfront value of the \$ bet. In left, a reversal is held to have occurred when, in binary choice, a lottery constructed to be dispreferred to that boundary lottery having a small probability of the high-state outcome (\$80) is chosen over a lottery constructed to be preferred to that boundary lottery having a large probability of the high-state outcome (\$80); as elsewhere, standard reversals involve selection of the p bet over the \$ bet in binary choice, while nonstandard involves a choice of the \$ bet over the p bet in binary choice.

[17]Given the differences in exchange rates (discussed in footnote 8) and other factors, one might argue that the bets $P9$ and $D9$ are closest to the incentives used by Butler and Loomes (2007). For bets $P9$ and $D9$, our data also exhibit more nonstandard than standard reversals in right, consistent with the findings of Butler and Loomes.

TABLE 3. Observed and predicted frequencies of types of reversals by p-bet and \$-bet pair.

		Project Up		Project Down		Project Left		Project Right	
Pair of Bets		Standard	Nonstandard	Standard	Nonstandard	Standard	Nonstandard	Standard	Nonstandard
$P1 \circ D1$ or $P6 \circ D6$	Observed	**0.3167**	**0.0333**	0.1833	–	–	**0.5000**	0.1667	**0.0500**
(Closest)	Predicted	0.1710	0.1796	0.2274			0.7466	0.1283	0.1153
$P2 \circ D2$ or $P7 \circ D7$	Observed	**0.4667**	**0.0667**	0.2167	–	–	**0.4833**	0.1500	0.0833
	Predicted	0.1892	0.1974	0.2537			0.6209	0.1623	0.1322
$P3 \circ D3$ or $P8 \circ D8$	Observed	**0.4500**	**0.0167**	0.2167	–	–	**0.6500**	0.1833	0.1667
	Predicted	0.2004	0.2172	0.3009			0.4734	0.2024	0.1832
$P4 \circ D4$	Observed	**0.4833**	**0.0667**	**0.2333**	–	**0.3500**	–	0.2167	0.1500
(B&L points)	Predicted	0.2278	0.2237	0.3627		0.1930		0.2176	0.2412
$P5 \circ D5$ or $P9 \circ D9$	Observed	**0.5333**	**0.0000**	–	0.5167	0.0333	–	0.0667	**0.3500**
(Farthest)	Predicted	0.2397	0.1973		0.4813	0.0357		0.0727	0.1522
Total	Observed	**0.4500**	**0.0267**	**0.1700**	0.1033	**0.0767**	**0.3267**	0.1567	0.1600
	Predicted	0.2056	0.2031	0.2289	0.0963	0.0457	0.3682	0.1557	0.1648
$P2 \circ D2$ or $P7 \circ D7$ (Order control)	Observed	0.4667	0.0333	0.2167	–	–	0.4833	0.1167	0.1000
$P4 \circ D4$ (Order control)	Observed	0.4667	0.0167	0.2333	–	0.3500	–	0.1333	0.1000
$P4 \circ D4$ B&L (2007)	Observed	0.5018	0.0075	–	–	–	–	0.0524	0.1798

Note: Reported values are observed and predicted frequencies (by the two-parameter fitted Blavatskyy model) of preference reversal by type (standard and nonstandard) as a proportion of total possible reversals pooled across all subjects. Frequencies in boldface are outside of the 95% confidence interval. "Closest" and "farthest" refer to the relative distance to the relevant boundary. Dashes indicate that it not possible to observe a particular type of reversal for the given parameterization, as discussed in footnote 16. Related p-values are reported in Table 12 in Appendix B.

4.4 *Chained dominance violations*

In Section 3 it was noted that while the Blavatskyy model does not admit violations of first-order stochastic dominance in the *binary choice* over two gambles, this does not imply that *two BDM elicitations* of numerical responses must be ordered in a way that is consistent with first-order stochastic dominance. Indeed, the model makes explicit predictions as to the frequency with which numerical responses might be out of alignment in this fashion. Table 4 compares predictions from the fitted model to the violations that occurred in the data.[18]

The raw observed frequency of chained dominance violation is lowest in up, higher in down, and highest in right. Also, we see that by this measure the responses from up are less "irrational" than predicted and right more so. This occurs despite the fact that down and right are the treatments for which the fitted Blavatskyy model does its best job of explaining preference reversals, while for up it does the worst. That is to say, the mechanism in which subjects violate chained dominance *least*, the selling version of BDM, is at the same time part of the most egregious pattern of cross-mechanism inconsistency when paired with binary choice. In contrast, when using the dual-to-selling version of BDM, for which subjects violate chained dominance more often, there is less inconsistency with binary choice. We will return to this seemingly paradoxical point in the conclusion. It turns out to prompt reconsideration of the interpretation of preference reversals.

4.5 *Single-round dominance violations*

Again, we note that a testable implication of the Blavatskyy model is that subjects engaged in binary choice cannot violate first-order stochastic dominance (FOSD). A further nuance to this is that an elicited certainty or probability equivalent may never first-order-stochastically dominate the lottery being valued, nor may it be dominated by that lottery. Thus, certainty equivalents cannot be lower than the low-state or higher than the high-state dollar outcomes, respectively, and analogously for probability equivalents.

When faced with the task of assigning a certainty or probability equivalent to a degenerate lottery, for a given mechanism there is only one response consistent with the Blavatskyy model. For instance, nonviolation of FOSD, applied to a degenerate lottery of \$12 received with probability 1 necessarily maps to a certainty equivalent of \$12 in selling BDM. Analogously, a lottery with end states of \$0 and \$12, and an upfront cash endowment of \$12 necessarily maps to a probability equivalent of 1 in dual-to-selling BDM. In other words, observing a response other than \$12 in the first case, or a probability other than 1 in the second case, is *impossible* in the Blavatskyy model.

[18] Observation of chained dominance violation across p bets or \$ bets is not possible in the left treatment as it is in the other treatments. To change the design parameter environment so as to be able to make this evaluation, one would also have to change the elicitation mechanism in a way that would make it no longer directly comparable in operation to that used in the right treatment.

TABLE 4. Observed and predicted frequencies of chained dominance violations across p-bet and \$-bet pairs.

Panel A. p Bets			Project Up				Project Down				Project Right			
			$P1$ (FA)	$P2$	$P3$	$P4$ (CT)	$P1$ (FA)	$P2$	$P3$	$P4$ (CT)	$P1$ (FA)	$P2$	$P3$	$P4$ (CT)
$P5$	(FA)	Observed	0.1000	0.1333	0.1451	**0.1167**	**0.2167**	0.1451	0.2833	0.2500	**0.1500**	**0.2500**	**0.5000**	**0.5000**
		Predicted	0.1106	0.1417	0.1977	0.2798	0.1120	0.1421	0.1977	0.2798	0.0377	0.1425	0.2459	0.3639
$P4$		Observed	0.1500	0.1833	0.2833	–	0.2000	0.2167	0.2667	–	**0.1667**	0.2167	0.3500	–
		Predicted	0.2077	0.2714	0.3693		0.2107	0.2740	0.3693		0.0541	0.2070	0.3517	
$P3$		Observed	**0.1333**	**0.2167**	–	–	0.2833	0.2833	–	–	0.1500	0.2167	–	–
		Predicted	0.2947	0.3810			0.2982	0.3831			0.0814	0.3094		
$P2$	(CT)	Observed	**0.2500**	–	–	–	0.3833	–	–	–	0.2000	–	–	–
		Predicted	0.3973				0.3997				0.1412			
Panel B. \$ Bets			Project Up				Project Down				Project Right			
			$D1$ (FA)	$D2$	$D3$	$D4$ (CT)	$D1$ (FA)	$D2$	$D3$	$D4$ (CT)	$D1$ (FA)	$D2$	$D3$	$D4$ (CT)
$D5$	(FA)	Observed	0.1333	**0.1333**	**0.1667**	**0.2167**	0.1667	0.2333	0.2333	**0.1667**	**0.4333**	**0.3167**	**0.3500**	0.3667
		Predicted	0.1709	0.2461	0.3283	0.4061	0.1811	0.2559	0.3356	0.4102	0.0875	0.1259	0.1827	0.2817
$D4$		Observed	0.1833	0.2167	0.3167	–	0.2500	0.3167	0.3667	–	**0.4667**	0.3500	0.3333	–
		Predicted	0.2181	0.3131	0.4124		0.2299	0.3224	0.4169		0.1631	0.2366	0.3411	
$D3$		Observed	0.1833	**0.1833**	–	–	**0.0667**	0.3167	–	–	**0.4500**	0.3667	–	–
		Predicted	0.2734	0.3885			0.2855	0.3950			0.2525	0.3614		
$D2$	(CT)	Observed	**0.2333**	–	–	–	**0.1333**	–	–	–	0.4667	–	–	–
		Predicted	0.3631				0.3724				0.3630			

Note: Reported values are observed and predicted (by the two-parameter fitted Blavatskyy model) frequencies of chained violations as proportion of the total possible such violations pooled across all subjects. Frequencies in boldface are outside of the 95% confidence interval. The acronym FA stands for "further apart" and CT stands for "closer together" on the grid in Figure 1. Chained dominance violations cannot be observed in left, as elaborated upon in footnote 18. Related p-values are reported in Table 12 in Appendix B.

Table 5. Percent of subject pool exhibiting a given number of dominance violations.

	Percent of Subject Pool All Violations		Percent of Subject Pool Excluding Violations by 1¢ or 1%	
Number of Dominance Violations	Project Up	Project Down	Project Up	Project Down
0 (Minimum possible)	0.2333	0.1500	0.3833	0.5000
1	0.5500	0.4667	0.4500	0.2000
2 (Maximum possible)	0.2167	0.3833	0.1667	0.3000

Table 5 reports the frequency with which single-round dominance violations are observed in the rounds of up and down. For the undivided data set, we are unable to reject the hypothesis that the distribution of violations is the same across up and down (with $p \approx 0.1180$ in a χ^2 test for homogeneity). However, for direct revelation mechanisms, it is typical to present a sorting of the data excluding violations of \$0.01 or 1% (Franciosi, Isaac, Pingry, and Reynolds (1993), Isaac and James (2000)). In this case, the distribution of violations in down is statistically different from (and greater than) that of up ($p \approx 0.0112$). Thus, subjects perform no worse, and possibly better, in up rather than down in terms of this particular measure of "rationality."

For completeness, the analogous results for left and right, and for binary choice are as follows. Single round dominance violations were also possible in left and right, which involve the use of the dual-to-selling mechanism in an environment consisting of lottery-for-lottery exchange. The possibilities for violating single-round dominance are different in these cases. In right, FOSD is violated when a subject responds with a higher probability equivalent on the up-state of the lottery on the right boundary of the grid than is present on the up-state of the lottery interior to the grid. In left, responding with a lower probability equivalent on the up-state of the interior lottery likewise constitutes a violation of FOSD.

Unlike say, in down, in left and right, violation of dominance is in principle something that the subject can commit, and the experimenter can infer, in every round. The proportion of responses that violated dominance in left and right are 16% and 34%, respectively.

In binary choice, dominance violations were also possible. Every subject was asked, as an exogenous design parameterization, to make an either/or choice between $P9$ and $D4$ ($P9$ dominates $D4$). Across the entire cross section of subjects, 6.7% of subject responses violated dominance for the choice between $P9$ and $D4$.[19]

[19] Dominance violations in binary choice are observed elsewhere in the preceding literature in similarly frequencies. For instance, Loomes, Moffatt, and Sugden (2002) observe a dominance violation incidence of 1.4% under their parameterization. Note also that within the course of our experiment, instances of dominance violation could be created endogenously and were not limited to the choice between $P9$ and $D4$. This functioned in the following manner: in down and left, subjects respond with probability equivalents; lotteries would then be constructed in a manner informed by these probability equivalent responses. Later, in binary choice, situations in which one lottery dominated the other could thus be parameterized (endogenously) by a subject's earlier probability equivalent responses. This situation occurred 57 and 91 times, for down and left, respectively. In these situations, subjects violated dominance 5.3% of the time in down and 3.3% of the time in left.

Table 6. Intratreatment correlations between reversals, chained violations, and single-round dominance violations.

	Spearman Rank Correlation Coefficient			
	Project Up	Project Down	Project Left	Project Right
Number of reversals and chained dominance violations	−0.0194	0.1279	–	0.1754
Number of reversals and single-round dominance violations	0.0011	0.0471	−0.1551	0.1511
Number of single-round dominance violations and chained dominance violations	0.2586**	0.4098***	–	0.3740***

Note: Reported values are Spearman rank correlation coefficients. The reported coefficients are intratreatment correlations (e.g., the first cell is the correlation between the number of reversals in up and the number of chained dominance violations in up). Chained dominance violations cannot be observed in left, as elaborated upon in footnote 18.
***Significant at the 1% level.
**Significant at the 5% level.

4.6 *Correlations between preference reversals, chained dominance violations, and single-round dominance violations*

Table 6 reports the correlation between the number of preference reversals, number of chained dominance violations, and the number of single-round dominance violations. Looking at correlations between single-round dominance and other features of the data is especially interesting because one is looking at a clear marker of subject mistakes (used also by James (2007), Cason and Plott (2014)) that is outside the purview of the Blavatskyy model. Furthermore, the presence or absence of correlation between violations of single-round dominance, which the Blavatskyy model holds to be impossible, and the phenomena that the Blavatskyy model does attempt to explain might provide stylized facts for the development of the next-generation models of decision-making.

As it turns out, the two different measures of within-mechanism "irrationality"—chained dominance and single-round dominance—are significantly correlated. But these measures of within-mechanism irrationality and that of cross-mechanism *inconsistency*, in the form of preference reversals, are not significantly correlated in cross section. As we shall discuss further in the conclusion, within-mechanism irrationality need not produce a preference reversal across mechanisms as long as the decisions made in the presence of that irrationality are consistent with the decisions made in the other mechanism. Further clues as to how this might come about are presented in the next section, wherein parametric estimates for both risk preference and stochasticity-in-choice are obtained.

4.7 *Parametric estimates and their comparison across mechanisms*

Preference reversals have been studied for over 40 years, but so far no one has concurrently estimated risk preference parameters from such data. We do that and estimate jointly the stochasticity-in-choice parameter allowed by the Blavatskyy model. It turns out that in doing so, key evidence for explaining preference reversals is obtained. Specifically, it appears that the stochasticity-in-choice parameter, α_i, has some predictive power across mechanisms, but at the same time, the risk parameter, ρ_i, does not. This suggests two things. First, subjects with fairly reliable propensities toward stochasticity might be mapped differently by different mechanisms in terms of propensity to take on risk. Second, the correspondence between the opposite mechanics of the two BDM mechanisms and the opposite medians of each distribution of risk parameter estimates (relative to risk neutrality) appears to be consistent with a role for response modes that is outside the scope of the Blavatskyy model. The clash in response modes between the selling version of BDM and binary choice (over probability-area pie charts) is not inconsistent with there being different instances of scale incompatibility for each mechanism, each instance possibly imparting its own asymmetry to responses made with respect to given lotteries. Conversely, the relative lack of conflict between responses from the dual-to-selling version of BDM and those from binary choice seems to result from greater similarity of their respective mappings of subject stochasticity (in ways both within and without the Blavatskyy model) to observable data. This is possibly because the response modes for these two mechanisms are *not* at odds in the same manner as the response modes for the selling version of BDM and for binary choice.

Let us now go into more detail about the process of estimation. So as to make the Blavatskyy (2009, 2012) model estimable, it is necessary to choose a parametric form for the VNM utility and $\varphi(\cdot)$ functions. For the utility function, we employ the constant relative risk aversion (CRRA) specification $u(x) = (x^{1-\rho_i})/(1-\rho_i)$, where ρ_i is the measure of relative risk aversion. An individual is considered risk averse for $\rho_i > 0$, risk neutral for $\rho_i = 0$, and risk loving for $\rho_i < 0$.

For $\varphi(\cdot)$, the function that governs the transformation between the utility of a lottery and its greatest lower bound in terms of weak dominance, there exists no a priori reason to select a particular specification. Thus we employ a power function $\varphi(x) = x^{a_i}$, as suggested by Blavatskyy (2009), with $\alpha_i \geq 0$. Under this specification, α_i may be interpreted as a stochasticity parameter, since it measures the level of stochasticity (or lack thereof) in an individual's choices. If $\alpha_i = 0$, choice is entirely random (with selection of either gamble equally likely), and as $\alpha_i \to \infty$, choice becomes deterministic (the decision maker necessarily choosing the highest VNM-expected-utility gamble).

Figure 4 illustrates the distribution of these parameters when estimated for each mechanism separately, using certainty equivalent, probability equivalent, and binary choice data, respectively. As such, each histogram represents within-mechanism fit rather than out-of-mechanism prediction. As one can see, the shape and location of the distribution of parameter estimates appear to differ from one mechanism to the next. Is this in fact the case? And to what extent are estimates from one mechanism consistent with another, in that behavior in one mechanism can thus be used to predict behavior in another mechanism? We turn next to answer those questions.

In round-robin estimation and prediction, estimates based on data from one mechanism are then tested for their potential to explain the estimates based on data from each of the other mechanisms. Figure 5 illustrates all possible cross-mechanism comparisons of parameter estimates. The three panels on the left plot individual-level constant relative risk aversion (CRRA) parameters compared across mechanisms, while the three panels on the right plot stochasticity parameters. All parameters were generated by max-

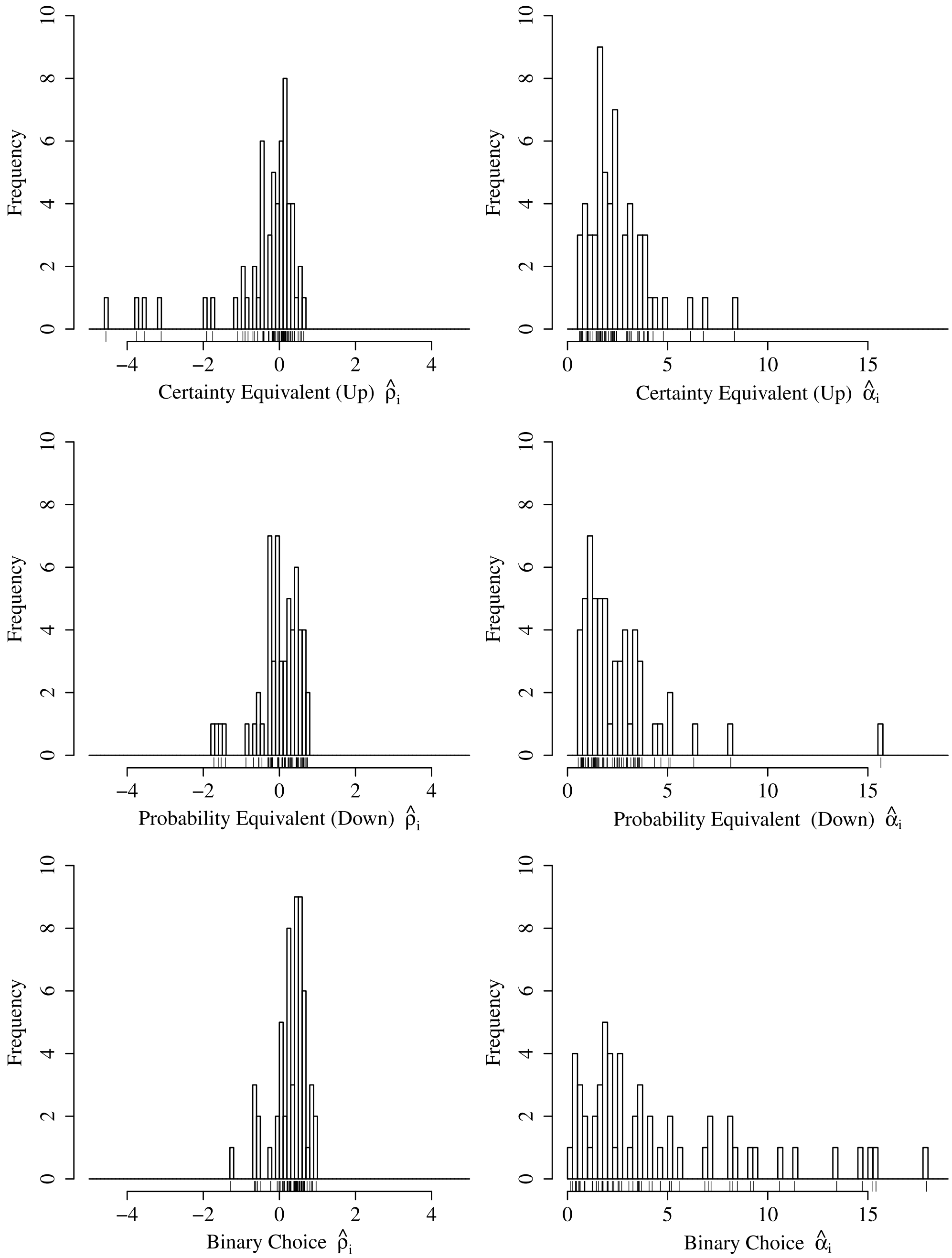

FIGURE 4. Histogram of estimated risk preference (ρ_i) and stochasticity (α_i) parameters.

imum likelihood (ML) estimation of the Blavatskyy model, with each subject's parameter estimate for a mechanism generated using data from that mechanism only. Identical numerical estimates across mechanisms (i.e., preservation of cardinality) would produce a data plot that follows an upward-sloping 45 degree line. Similarly ordered, but

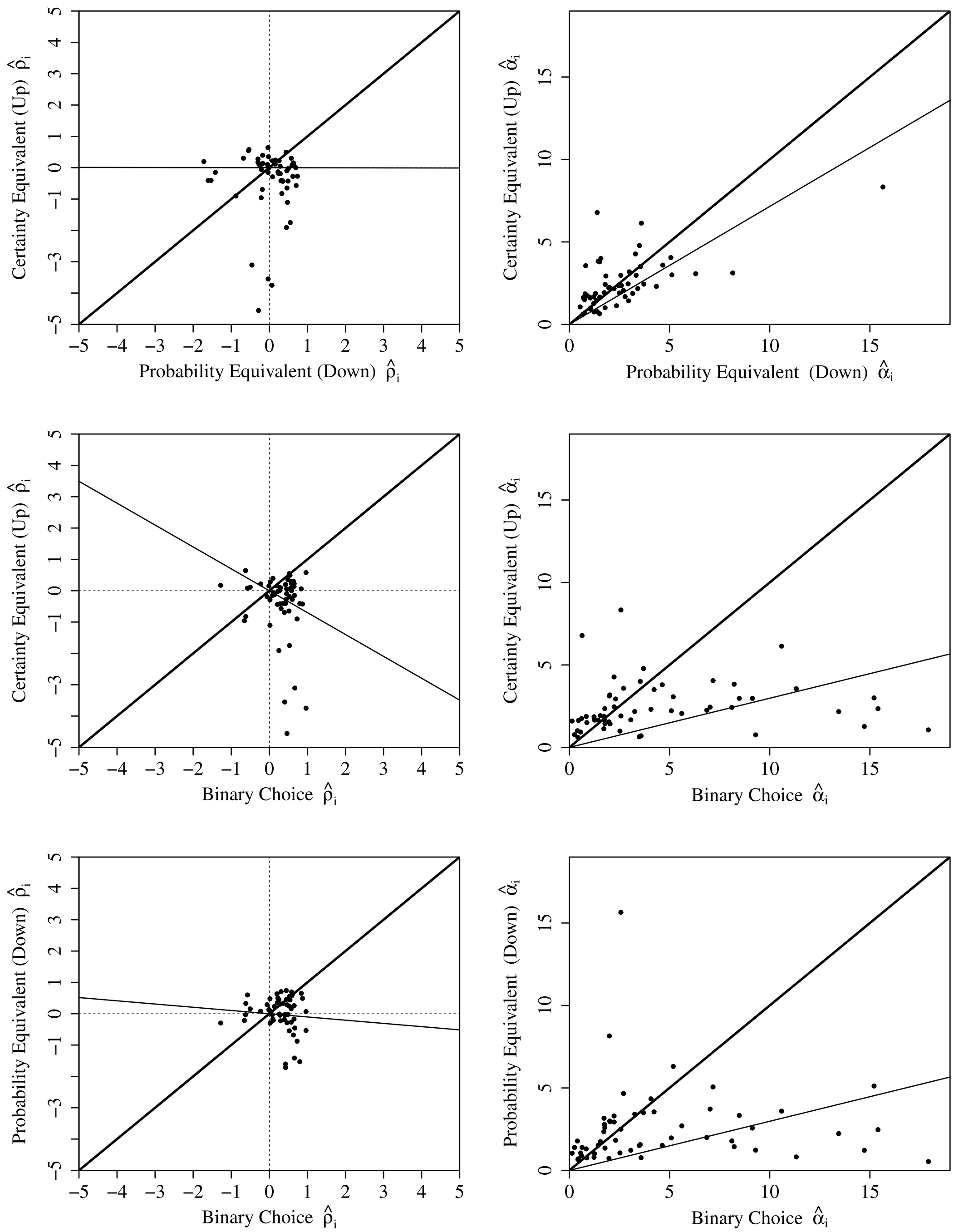

FIGURE 5. Cross sections of estimated risk preference (ρ_i) and stochasticity (α_i) parameters. *Note*: Both the (thicker) 45 degree line and the (thinner) line of best fit, constrained to a zero intercept, are illustrated. The latter is for the purpose of illustration of the trend and is not intended for statistical interpretation.

TABLE 7. The p-values of bootstrapped statistical tests on the estimated Blavatskyy parameters.

	Kolmogorov–Smirnov Tests		Spearman Correlation Tests	
Treatment	PE	BC	PE	BC
	Estimated Blavatskyy risk aversion parameters ($\hat{\rho}_i$)			
CE	< 0.0001	< 0.0001	0.8865	0.5822
PE		< 0.0001		0.6824
	Estimated Blavatskyy stochasticity parameters ($\hat{\alpha}_i$)			
CE	0.2664	0.0860	< 0.0001	0.0013
PE		0.0605		0.0035

Note: Parametric bootstraps employed 9999 replications and used the estimated parameters from the rows in generating data from the null hypothesis. The details of the bootstrap procedure and alternative approaches are described in Appendix B.2.

not numerically identical, estimates across mechanisms (i.e., preservation of ordinality) would produce a data plot that follow an upward-sloping pattern, but not necessarily the 45 degree line.

Table 7 reports the results of six Kolmogorov–Smirnov and six Spearman correlation coefficient tests on the distributions of estimates of ρ_i and α_i. The reported correlations are between the estimated parameters, but the calculated p-values have been bootstrapped for the bias otherwise inherent in conducting tests on fitted parameters. We made the following findings.

1. Individual risk preference parameter estimates from one mechanism do not in general carry over in even ordinal terms to those of other mechanisms. That is to say, knowing a subject's risk aversion parameter in one mechanism will not help one predict that subject's risk aversion parameter in any other of these mechanisms. (This might be thought of as a kind of preference reversal in itself.)

2. Mechanism-level regularities regarding the distributions of ρ_i estimates do exist, however. Changes in the overall location and/or shape of the distributions of $\hat{\rho}_i$ may be predictable based on a change of mechanism in use.

3. Estimates of the α_i parameter *do* display some cross-mechanism predictive ability at the individual level. Specifically, the ordering across individuals of the α_i estimates largely carries over between dual-to-selling BDM, selling BDM, and binary choice.

Mechanisms matter for the overall pattern of results, but at the same time individual subjects' risk parameter estimates do not have any discernible ability to predict across mechanisms.[20] As noted, however, the α_i parameter estimates appear to offer some predictive power across mechanisms. This may indicate that relative individual propensi-

[20]Consider estimates of the risk aversion parameters, ρ_i. First, for all pairwise comparisons of distributions of $\hat{\rho}_i$ from treatments, according to the Kolmogorov–Smirnov test we must reject the null hypothesis that the parameters were drawn from the same distribution. Thus, it would appear that different mechanisms systematically alter the distributions of estimated ρ_i parameters. Second, given a null hypothesis that the true Spearman rank coefficient is equal to zero in any pairwise treatment, we are unable to reject (Spearman rank coefficient test) that the true Spearman rank coefficient is equal to zero in any pairwise treatment comparison. (Indeed, the within-sample coefficients are all small and negative, as reported in Table 8.)

Table 8. Spearman rank correlation coefficient estimates.

	Risk Aversion Parameters ($\hat{\rho}_i$)		Stochasticity Parameters ($\hat{\alpha}_i$)	
Treatment	PE	BC	PE	BC
CE	−0.1593	−0.0297	0.6366	0.3841
PE		−0.0621		0.3479

Note: Unadjusted Spearman rank correlation coefficient estimates are reported. Bootstrapped measures of bias and standard error are reported in Table 16 in Appendix B.

ties to stochasticity in choice might have some stability across mechanisms, though the magnitude of stochasticity in choice may depend upon the mechanism.[21]

This fits well with the literature on the role of cognitive capacity in risky choice. There the finding is that subject mistakes, a function of cognitive ability, generate much of what appears to be violation of risk neutrality at small stakes (Benjamin, Brown, and Shapiro (2013)). Cognitive ability would be carried from one mechanism to another by the subject, though the kind of mistakes one can make (and the magnitude with which one can make them) will differ from one mechanism to the next. Analogous findings, on iterations of reasoning in games (Carpenter, Graham, and Wolf (2013)) and on time preferences (Frederick (2005)) also exist.

Overall, these results might offer further support for the response mode explanation for preference reversals. The same stochastic choice process on the part of a subject might be mapped to responses differently across mechanisms. Conditions for this could be put in place by different response modes in each mechanism. This difference in response modes could lead to different characteristic mistakes across mechanisms.

Binary choice and the dual-to-selling version of BDM might just happen to map stochastic decision-making to a more overlapping range of observed responses, on average across subjects, and thus reduce the incidence of preference reversals. This may be aided by the commonality in units across probability-area pie charts and probability equivalents. This would also be consistent with the observation that the frequency of standard and nonstandard reversals are here observed to be symmetric in dual-to-selling BDM, while standard reversals outnumber nonstandard reversals in BDM.

That this happens need not be a function of whether the mechanisms in question are necessarily "good" or "bad" in the sense of not fostering any other choice pathologies, for example, violation of dominance. Indeed, the dual-to-selling version of BDM is worse than the selling version of BDM on several such scores, and yet it is in greater agreement with the results coming from binary choice. All that is needed for greater agreement across mechanisms might be for them to channel stochastic choice in a sufficiently similar range.

[21] First, we cannot reject the null hypothesis that the distributions of α_i estimates are the same for up and down, but not between binary choice and the others (Kolmogorov–Smirnov test). Thus, at the aggregate level we find that the distributions of $\hat{\alpha}_i$ display similar shape and location across mechanisms within the BDM family, but not between binary choice and either BDM family mechanism. Second, the ordering of individuals' α_i estimates are significantly and positively correlated between *all* mechanisms (Spearman rank coefficient test).

5. Discussion

Overall, switching from certainty equivalent elicitation to probability equivalent elicitation reduces the frequency of preference reversals in raw data and also reduces them to a level that is predicted by the (fitted) Blavatskyy model—the latter suggesting that a portion of reversals are due to general stochasticity in choice and would be present even given a response mode that is less prone to generating preference reversals.[22]

Our systematic collection of data is summarized as follows:

- When comparing the up (certainty equivalent) and down (probability equivalent) treatments, we see that the dual-to-selling version of BDM usually reduces the frequency of reversals (see Table 2 in Section 4.2).
- At the same time, however, up (certainty equivalent) data usually show a lower rate of chained dominance violations than do down (probability equivalent) data (see Table 4 in Section 4.4).
- There are also interesting regularities viewed in terms of the performance of the fitted Blavatskyy model. It predicts reversals between binary choice and a range of numerical elicitation alternatives very well in right and down, and poorly in up (see Table 3 in Section 4.3).
- Behavior not permitted by the Blavatskyy model is observed. Violations of dominance in a single round are observed in selling BDM, dual-to-selling BDM, and binary choice. This type of violation is at least as prevalent in dual-to-selling BDM as in selling BDM.

How might one put all of this together? Of particular note is the seeming incongruity between the data on internal-to-mechanism rationality checks and the data on across-mechanism preference reversal. Consider the following scenario: the up treatment (using the selling version of BDM) generates chained dominance violations (a) at a low rate within the raw data and (b) at a rate that is less than that which would be predicted by the fitted Blavatskyy model. Indeed, subject behavior is less often in violation of chained dominance and of single-round dominance in the up treatment than in the down treatment (the latter using the dual-to-selling version of BDM), yet the *greatest* frequency of preference reversals in raw data and the *greatest* fitted-Blavatskyy-model underprediction of preference reversals takes place for the juxtaposition of certainty equivalents and

[22] In the context of Seidl's (2002) analysis (previously mentioned in footnote 2), we note the following. First, we find support for the claim that at least one specific aspect of the elicitation mode—response mode within the BDM mechanism—contributes to reversals. Second we find that a model of probabilistic choice (Blavatskyy (2012)) is one way in which *realized* subject responses can exhibit intransitivity that can explain the general prevalence of reversals once the response mode is accounted for. Third, we do not find evidence of systematic over- and/or underpricing of lotteries manifested in the types of reversals observed across all combinations of mechanism and environment (i.e., we observe that standard reversals predominate in up, but a roughly symmetric tendency toward standard and nonstandard reversal predominate in right). In light of this, we believe that we have been able to tie together, in a single study, an explanation of preference reversals that may need only two factors instead of the at least three (not necessarily mutually consistent) factors compiled by Seidl in a survey of many papers.

binary choice data (the up pairing), and the lowest frequency of reversals for the pairing of probability equivalents and binary choice (the down pairing).

This suggests that behavior in one mechanism that might variously be described as irrational, mistaken, or costly according to one or another measure, such as chained or single-round dominance violations, is not necessarily incapable of being consistent with choices from another mechanism. Conversely, while consistency in responses across mechanisms is necessary for a finding of rationality, it may not be sufficient.

Preference reversals might conceivably be reduced, or perhaps even eliminated, by using numerical evaluations with nothing more to recommend them that they express any underlying irrational behavior by subjects in a manner congruent to that in binary choice. This is not inconsistent with the notion (Grether and Plott (1979, 1975)) that asking subjects (a) for dollar valuations of gambles and (b) to choose between probability-area pie charts is in a sense setting subjects up to fail. The mistakes they make, or the heuristics they employ, could well fall in different directions across the two response media: one dollar denominated and the other arguably probability denominated. The superior performance of down in terms of reduced preference reversals might really be more of a superior "match," and the basis for this match might yet turn out to be in shared *ir*rationality. Here, the notion of "constructed preferences" (Lichtenstein and Slovic (2006)) might be a good conceptual basis for exploring what subjects might go through so as to respond in a task: they must establish where their interests lie and how to craft a response that best serves those interests in highly, if subtly, structured tasks. Put this way, it seems natural that various tasks, each with its own structure, might each generate their own respective scatters of responses. Some pairs of these scatters might overlap more than others.[23]

Overall, between switching the response mode from certainty equivalent elicitation to probability equivalent elicitation (physically changing the possibilities for scale incompatibility) and then allowing for stochastic choice as per the Blavatskyy model, one arrives at a situation in which the frequency of preference reversals is no longer at odds with model predictions. That this occurs in those two steps suggests a role for response mode *and* stochastic choice in generating the preference reversals observed in traditional settings, as in Lichtenstein and Slovic (1971, 1973) and Grether and Plott (1979).

[23]Focusing on the pairing that reduces reversals, one possibility might be that binary choice and dual-to-selling might have more in common, in terms of channeling error-prone subjects, than would seem to be the case based solely on within-mechanism dominance violations. Binary choice *does* exhibit very infrequent dominance violations internally. But this makes it all the more strange that it accords best with the alternative mechanism exhibiting the *most* frequent dominance violations internally. Maybe a way to come to grips with this is to consider the possibility that in binary choice the error-proneness of the subjects may not just disappear. Rather, it might manifest itself—be channeled by the mechanism—in other, less obvious ways than dominance violations. One possibility is that it manifests itself instead as more pronounced apparent risk aversion. A wider overlap between dual-to-selling and binary choice in this regard would help reduce reversals, even while the prevalence of dominance violations in dual-to-selling seems not to matter for purposes of not contradicting binary choice.

Appendix A: Details and estimation of the Blavatskyy model

A.1 *Probabilistic choice in the Blavatskyy model*

As noted in Section 3, the Blavatskyy (2009, 2012) model predicts that an individual chooses some lottery A over B with probability $p(A, B)$, where

$$p(A, B) = \frac{\varphi(u(A) - u(A \wedge B))}{\varphi(u(A) - u(A \wedge B)) + \varphi(u(B) - u(A \wedge B))}, \tag{2}$$

where $u(\cdot)$ is the von Neumann–Morgenstern expected utility function, $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is a nondecreasing function with $\varphi(0) = 0$, and $A \wedge B$ denotes the greatest lower bound on A and B in terms of weak stochastic dominance. That is, $A \wedge B$ is a lottery that is dominated both by lottery A and by lottery B for which there is no other lottery dominated both by A and B that stochastically dominates $A \wedge B$.

Preference reversals involve two decision tasks: a binary choice between lotteries and a valuation task. The valuation task may elicit either a certainty or a probability equivalent; in this paper, these are elicited by the selling and the dual-to-selling versions of BDM, respectively. In the Blavatskyy model, the certainty equivalent (CE) of lottery A is a random variable. The observed CE is thus one possible draw from a distribution. That cumulative distribution function is

$$\mathrm{CE}_A(x) = p(x, A) = 1 - p(A, x), \tag{3}$$

where x is the dollar amount of the CE, a degenerate lottery. Similarly, the observed probability equivalent (PE) of a lottery A is a draw from the c.d.f.

$$\mathrm{PE}_A(q) = p(Q, A) = 1 - p(A, Q), \tag{4}$$

where q is the probability of a benchmark dollar amount ($80) in lottery $Q(q)$.

The valuation tasks thus defined, its is possible to calculate the probability of a preference reversal. A standard preference reversal occurs when the p bet is chosen over the $ bet in binary choice, but the $ bet is valued higher than the p bet. Alternatively, a nonstandard preference reversal occurs when the $ bet is chosen over the p bet in binary choice, but the p bet is valued higher than the $ bet.

A.2 *Preference reversals in project up*

Let lottery P denote the p bet and lottery D denote the $ bet. The probability that an individual selects P over D in binary choice is $p(P, D)$. The probability that the same individual states a higher certainty equivalent for D than P is

$$p(\mathrm{CE}_D > \mathrm{CE}_P) = \int \mathrm{CE}_P(x)\, \mathrm{d}\mathrm{CE}_D(x). \tag{5}$$

Thus, the probability that standard reversal occurs (in the project up treatment) is $p(P, D) \cdot p(\mathrm{CE}_D > \mathrm{CE}_P)$. The probability that the same individual states a higher certainty equivalent for P than D is $p(\mathrm{CE}_P > \mathrm{CE}_D) = 1 - p(\mathrm{CE}_D > \mathrm{CE}_P)$. The probability of a nonstandard reversal in up is thus $p(D, P) \cdot p(\mathrm{CE}_P > \mathrm{CE}_D)$.

A.3 *Preference reversals in project right*

The probability that an individual states a higher probability equivalent for P than D is

$$p(\mathrm{PE}_P > \mathrm{PE}_D) = \int \mathrm{PE}_D(q)\, \mathrm{dPE}_P(q) \tag{6}$$

for a particular benchmark lottery. The probability that the individual states a higher probability equivalent for D than P is $p(\mathrm{PE}_D > \mathrm{PE}_P) = 1 - p(\mathrm{PE}_P > \mathrm{PE}_D)$ for the same benchmark lottery. Therefore, a standard reversal occurs in the project right treatment with probability $p(P, D) \cdot p(\mathrm{PE}_D > \mathrm{PE}_P)$ and a nonstandard reversal occurs with probability $p(D, P) \cdot p(\mathrm{PE}_P > \mathrm{PE}_D)$.

A.4 *Preference reversals in project down*

In the project down treatment, a probability equivalent is elicited for a two cash amounts (degenerate lotteries), and then a binary choice is made between the p bet and the \$ bet constructed from those elicited valuations (with up-states of \$12 and \$40, respectively). The decision regarding the cash amounts is arbitrary, so we selected the risk-neutral certainty equivalents for a particular p bet and \$ bet pair, denoted as $\overline{\mathrm{CE}}_P$ and $\overline{\mathrm{CE}}_D$, respectively. The dominance relationship between two lotteries is exogenously determined: if $\overline{\mathrm{CE}}_D > \overline{\mathrm{CE}}_P$, a standard reversal is said to occur if a subject selects the p bet over the \$ bet constructed with the observed probability equivalents of the respective lotteries; if $\overline{\mathrm{CE}}_D > \overline{\mathrm{CE}}_P$, a nonstandard reversal occurs if a subject selects the \$ bet over the p bet. (It is not possible to observe a nonstandard reversal if $\overline{\mathrm{CE}}_D > \overline{\mathrm{CE}}_P$ or a standard reversal if $\overline{\mathrm{CE}}_P > \overline{\mathrm{CE}}_D$.) Since $p(P, D)$ is a function of two independent continuous random variables, PE_P and PE_D, in expectation the probability that an individual will choose the p bet over the \$ bet in PE is

$$E\big[p(P, D)\big] = \iint p\big(P(q_P), D(q_P)\big)\mathrm{PE}'_P(q_P)\mathrm{PE}'_D(q_D)\, \mathrm{d}q_P\, \mathrm{d}q_D, \tag{7}$$

where PE'_P and PE'_D are the probability density functions (p.d.f.s) of the p bet and \$ bet probability equivalents, respectively. The probability in expectation that an individual will choose the \$ bet over the p bet is simply $1 - E[p(P, D)]$.

A.5 *Preference reversals in project left*

The project left treatment uses the same general method for determining whether a reversal has occurred as the PE treatment. For the left treatment in our design, the higher outcomes of the alternative lotteries (\$12 for the p bet, \$40 for the \$ bet) are both lower than the higher outcome of the initial lottery (\$80), and it is trivial to show that $\mathrm{PE}(1) < 1$ for both the p bet and the \$ bet. Consequently, the probability equivalents in left are truncated at the top of the distribution, with the associated p.d.f.

$$\mathrm{PE}'_A(q|\tilde{q} \leq 1) = \frac{\mathcal{PE}'_A(q)}{\mathrm{PE}_A(1)}, \tag{8}$$

where $\mathcal{PE}'_A(q)$ takes the value $\mathrm{PE}'_A(q)$ for $0 < q \leq 1$ and 0 everywhere else. Using the truncated distribution, the approach outlined for calculating the expected probability of reversal in down may then be applied to left.

In the project left treatment, probability equivalents of two "initial" lotteries are used to construct the p bet and the \$ bet (with up-states of \$12 and \$40, respectively), and then a binary choice is made between them. As in down, there is no reason a priori to select a particular initial lottery, so we chose the risk-neutral probability equivalent of an \$80 up-state for the p bet and \$ bet, denoted as $\overline{\mathrm{PE}}_P$ and $\overline{\mathrm{PE}}_D$ respectively. Then, if $\overline{\mathrm{PE}}_D > \overline{\mathrm{PE}}_P$, it is possible to observe a standard reversal, and if $\overline{\mathrm{PE}}_P > \overline{\mathrm{PE}}_D$, it is possible to observe a nonstandard reversal. In expectation, the probability that an individual will choose the p bet over the \$ bet in left is

$$E\big[p(P, D)\big] = \iint p\big(P(q_P), D(q_P)\big)\mathrm{PE}'_P(q_P|\widetilde{q_P})\mathrm{PE}'_D(q_D|\widetilde{q_D})\,\mathrm{d}q_P\,\mathrm{d}q_D, \tag{9}$$

and the expected probability that an individual will choose the \$ bet over the p bet is simply $1 - E[p(P, D)]$.

A.6 *Chained dominance violations*

In Section 3, it is also noted that it is possible for lottery A to stochastically dominate lottery B, but for an individual to report a valuation for B that exceeds that of A. In the up treatment, this occurs with probability $p(\mathrm{CE}_B > \mathrm{CE}_A)$ for any such pair of lotteries. Similarly, in the down and right treatments, this occurs with probability $p(\mathrm{PE}_B > \mathrm{PE}_A)$.

A.7 *Maximum likelihood estimation of Blavatskyy model parameters*

As discussed in Section 4.7, to make the Blavatskyy model tractable, parametric specifications must be chosen for $u(\cdot)$ and $\varphi(\cdot)$. Following the suggestion of Blavatskyy (2009), we impose neoclassical CRRA utility such that $u_i(x) = (x^{1-\rho_i})/(1-\rho_i)$ and a power form of the curvature function such that $\varphi_i(x) = x^{\alpha_i}$.

The parameters to be estimated for subject i are $\{\alpha_i, \rho_i\}$. Define all variables as in Appendix A.1 and, further, let ζ_{ij} take on the value 1 if, for lottery pair i, the subject selects lottery A_i in binary choice, and the value 0 if B_i is selected. Then, for each subject i, there are N_{BC} binary choice observations $\{\zeta_{ij}, A_{ij}, B_{ij}\}_{j=1}^{N_{\mathrm{BC}}}$, N_{CE} certainty equivalent observations $\{x_{ji}, A_{ij}\}_{j=1}^{N_{\mathrm{CE}}}$, and N_{PE} probability equivalent observations $\{Q_{ji}, A_{ij}\}_{j=1}^{N_{\mathrm{PE}}}$. The joint log likelihood function is then

$$\begin{aligned}\ln\mathcal{L}(\rho_i, \alpha_i) = {} & \sum_{j=1}^{N_{\mathrm{BC}}}\big(\zeta_{ij}\ln p(A_{ij}, B_{ij}|\rho_i, \alpha_i) \\ & + (1-\zeta_{ij})\ln\big(1 - p(A_{ij}, B_{ij}|\rho_i, \alpha_i)\big)\big) \\ & + \sum_{j=1}^{N_{\mathrm{CE}}}\ln p'(x_{ij}, A_{ij}|\rho_i, \alpha_i) + \sum_{j=1}^{N_{\mathrm{PE}}}\ln p'(Q_{ij}, A_j|\rho_i, \alpha_i),\end{aligned} \tag{10}$$

where $p'(\cdot)$ are the p.d.f.s of the c.d.f.s from which the certainty and probability equivalents are drawn. Likelihood maximization was conducted using the L-BFGS-B algorithm of Byrd, Lu, Nocedal, and Zhu (1995).[24]

The likelihood function does not admit data that have a probability of 0 or 1. In our case, this implies that dominance checks and subject responses that violate dominance cannot be used in estimation. In up and down, the data are generally well behaved, but in left and right, a large portion of the data exhibits dominance violations inadmissible by the Blavatskyy model (as discussed in Section 4.4). The Blavatskyy model does not allow for such violations, and so we must conclude that some unknown process determines whether a dominance violation has occurred (as discussed in Section 5). We cannot assume out-of-hand that this unknown process is orthogonal to the data-generating process in the Blavatskyy model, so we elected not to include any data from the left and right treatments in the estimation of the subject-level parameters.[25]

References

Andreoni, J. and C. Sprenger (2011), "Uncertainty equivalents: Testing the limits of the independence axiom." Working paper, NBER. [830]

Azrieli, Y., C. P. Chambers, and P. J. Healy (2012), "Incentives in experiments: A theoretical analysis." Working paper. [830]

Becker, G. M., M. H. Degroot, and J. Marschak (1964), "Measuring utility by a single-response sequential method." *Behavioral Science*, 9 (3), 226–232. [825, 826]

Benjamin, D. J., S. A. Brown, and J. M. Shapiro (2013), "Who is 'behavioral'? Cognitive ability and anomalous preferences." *Journal of the European Economic Association*, 11 (6), 1231–1255. [848]

Berg, J., J. Dickhaut, and K. McCabe (2005), "Risk preference instability across institutions: A dilemma." *Proceedings of the National Academy of Sciences of the United States of America*, 102 (11), 4209–4214. [826]

Blavatskyy, P. R. (2009), "Preference reversals and probabilistic decisions." *Journal of Risk and Uncertainty*, 39 (3), 237–250. [825, 826, 828, 830, 832, 833, 844, 851, 853]

Blavatskyy, P. R. (2012), "Probabilistic choice and stochastic dominance." *Economic Theory*, 50 (1), 59–83. [825, 826, 828, 829, 830, 832, 844, 849, 851]

[24] Maximization was conducted in R 2.15.2 using the optim package. Although the concavity of the likelihood function has not been established, the limited dimensionality of the problem also admitted a thorough grid search over the likely parameter space that produced the same results as the L-BFGS-B algorithm. The grid search was conducted over the interval $(0, 25]$ for α_i and $[-5, 1)$ for ρ_i in intervals of 0.01 for both. A large number of estimations were also produced from simulated data across this parameter space using the functions defined in Appendix A.1 to verify the integrity of the ML estimation procedure.

[25] Because there exists no a priori way of dealing with data that cannot be used in estimation, we chose to exclude any observations in the binary choice, up, and the down data for which the probability was 0 or 1 in the reported results. To test the robustness of this decision, we also "bumped" these data back in (e.g., a PE response of zero was changed to 0.01) and reran all routines conducted in the paper; there was little change in the estimates and, relevantly, no change in the inferences to be drawn from the data.

Blavatskyy, P. R. (2014), "Stronger utility." *Theory and Decision*, 76 (2), 265–286. [828]

Bohnet, I., F. Greig, B. Herrmann, and R. Zeckhauser (2008), "Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States." *American Economic Review*, 98 (1), 294–310. [830]

Bohnet, I. and R. Zeckhauser (2004), "Trust, risk and betrayal." *Journal of Economic Behavior and Organization*, 55 (4), 467–484. [830]

Butler, D. J. and G. C. Loomes (2007), "Imprecision as an account of the preference reversal phenomenon." *American Economic Review*, 97 (1), 277–297. [828, 830, 831, 832, 838]

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu (1995), "A limited memory algorithm for bound constrained optimization." *SIAM Journal on Scientific Computing*, 16 (5), 1190–1208. [854]

Carpenter, J., M. Graham, and J. Wolf (2013), "Cognitive ability and strategic sophistication." *Games and Economic Behavior*, 80, 115–130. [848]

Cason, T. N. and C. R. Plott (2014), "Misconceptions and game form recognition: Challenges to theories of revealed preference and framing." *Journal of Political Economy*, 122 (6), 1235–1270. [843]

Cox, J. C. and S. Epstein (1989), "Preference reversals without the independence axiom." *American Economic Review*, 79 (3), 408–426. [827]

Cox, J. C. and D. M. Grether (1996), "The preference reversal phenomenon: Response mode, markets and incentives." *Economic Theory*, 7 (3), 381–405. [825, 826, 832, 837]

Cox, J. C., V. Sadiraj, and U. Schmidt (2014), "Alternative payoff mechanisms for choice under risk." *International Advances in Economic Research*, 20 (2), 239–240. [830]

Cubitt, R., A. Munro, and C. Starmer (2004), "Testing explanations of preference reversal." *The Economic Journal*, 114 (497), 709–726. [828]

Fischbacher, U. (2007), "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10 (2), 171–178. [834]

Fishburn, P. C. (1988), *Nonlinear Preference and Utility Theory*. Johns Hopkins University Press, Baltimore, Maryland. [838]

Franciosi, R., R. M. Isaac, D. E. Pingry, and S. S. Reynolds (1993), "An experimental investigation of the Hahn–Noll revenue neutral auction for emissions licenses." *Journal of Environmental Economics and Management*, 24 (1), 1–24. [842]

Frederick, S. (2005), "Cognitive reflection and decision making." *Journal of Economic Perspectives*, 19 (4), 25–42. [848]

Greiner, B. (2004), "An online recruitment system for economic experiments." In *Forschung und Wissenschaftliches Rechnen 2003* (K. Kremer and V. Macho, eds.), 79–93, Ges. für Wiss. Datenverarbeitung, Göttingen. [834]

Grether, D. M. and C. R. Plott (1979), "Economic theory of choice and the preference reversal phenomenon." *American Economic Review*, 69 (4), 623–638. [825, 826, 827, 829, 830, 835, 850]

Isaac, R. M. and D. James (2000), "Just who are you calling risk averse?" *Journal of Risk and Uncertainty*, 20 (2), 177–187. [826, 842]

James, D. (2007), "Stability of risk preference parameter estimates within the Becker–DeGroot–Marschak procedure." *Experimental Economics*, 10 (2), 123–141. [843]

James, D. (2011), "Incentive compatible elicitation procedures." In *19th International Congress on Modelling and Simulation, Perth, Australia*, 1421–1427, Modelling and Simulation Society of Australia and New Zealand, Canberra, Australia. [828, 829, 830]

Lichtenstein, S. and P. Slovic (1971), "Reversals of preference between bids and choices in gambling decisions." *Journal of Experimental Psychology*, 89 (1), 46–55. [825, 826, 828, 830, 850]

Lichtenstein, S. and P. Slovic (1973), "Response-induced reversals of preference in gambling: An extended replication in Las Vegas." *Journal of Experimental Psychology*, 101 (1), 16–20. [825, 830, 850]

Lichtenstein, S. and P. Slovic (2006), *The Construction of Preference*. Cambridge University Press, Cambridge. [850]

Loomes, G., P. G. Moffatt, and R. Sugden (2002), "A microeconometric test of alternative stochastic theories of risky choice." *Journal of Risk and Uncertainty*, 24 (2), 103–130. [842]

Luce, R. D. (1959), *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York. [828]

MacCrimmon, K. and M. Smith (1986), "Imprecise equivalences: Preference reversals in money and probability." Working paper, University of British Columbia. [828, 832]

Seidl, C. (2002), "Preference reversal." *Journal of Economic Surveys*, 16 (5), 621–655. [827, 849]

Slovic, P. (1975), "Choice between equally valued alternatives." *Journal of Experimental Psychology: Human Perception and Performance*, 1 (3), 280–287. [826, 827, 829, 850]

Tversky, A. (1969), "Intransitivity of preferences." *Psychological Review*, 76 (1), 31–48. [826]

Tversky, A., P. Slovic, and D. Kahneman (1990), "The causes of preference reversal." *American Economic Review*, 80 (1), 204–217. [826, 827]

Submitted February, 2014. Final version accepted December, 2014.