2008

High Stakes Behavior with Low Payoffs: Inducing preferences with Holt-Laury gambles

John Dickhaut
Chapman University

Daniel Houser

Jason A. Aimone

Dorina Tila

Cathleen Johnson

High Stakes Behavior with Low Payoffs: Inducing preferences with Holt-Laury gambles

Comments
Working Paper 08-09

High Stakes Behavior with Low Payoffs:
Inducing Preferences with Holt-Laury Gambles

by

John Dickhaut*, Daniel Houser, Jason A. Aimone, Dorina Tila, and Cathleen Johnson**

**Interdisciplinary Center for Economic Science
George Mason University
*Chapman University **University of Arizona**

November 2008

High Stakes Behavior with Low Payoffs:
Inducing Preferences with Holt-Laury Gambles

by

John Dickhaut*, Daniel Houser, Jason A. Aimone, Dorina Tila, and Cathleen Johnson**

Interdisciplinary Center for Economic Science
George Mason University
***Chapman University **University of Arizona**

November 2008

Abstract:

A continuing goal of experiments is to understand risky decisions when the decisions are important. Often a decision's importance is related to the magnitude of the associated monetary stake. Khaneman and Tversky (1979) argue that risky decisions in high stakes environments can be informed using questionnaires with hypothetical choices (since subjects have no incentive to answer questions falsely.) However, results reported by Holt and Laury (2002, henceforth HL), as well as replications by Harrison (2005) suggest that decisions in "high" monetary payoff environments are not well-predicted by questionnaire responses. Thus, a potential implication of the HL results is that studying decisions in high stakes environments requires using high stakes. Here we describe and implement a procedure for studying high-stakes behavior in a low-stakes environment. We use the binary-lottery reward technique (introduced by Berg, et al (1986)) to induce preferences in a way that is consistent with the decisions reported by HL under a variety of stake sizes. The resulting decisions, all of which were made in a low-stakes environment, reflect surprisingly well the noisy choice behavior reported by HL's subjects even in their high-stakes environment. This finding is important because inducing preferences evidently requires substantially less cost than paying people to participate in extremely high-stakes games.

1 Introduction

How should subjects be paid? Experimental economics emphasizes the importance of paying subjects with real dollars, although often small in magnitude (of course theory assumes consumption derives from real dollars.) Thus, when numerous challenges to the classical theory of decision-making (Expected Utility Theory) were generated based on hypothetical decisions (See e.g. Khaneman and Tversky, 1979), the results were, and to some extent still are, a matter of debate.

Khaneman and Tversky (1979) defend hypothetical choices arguing that there is no reason for subjects not to tell the truth. Camerer and Hogarth (1999) argue that choices may involve differential productive effort, which can affect what the experimenter observes. Such observations, whether from hypothetical choices or choices for dollars, may be distortions from a subject's true preferences. Camerer and Hogarth (1999) also assert that no experiments paying subjects based on dollars have overturned anomalies observed from hypothetical choices.

An alternative position has its most outspoken advocate in Harrison. In a sequence of experiments, Harrison (1994) has argued that changing incentives critically alters outcomes of the anomalies literature, including the Allais Paradox and the Preference Reversal Phenomena. His criticism, while predominately directed at anomalies based on hypothetical choices, also generates concerns for experiments using small cash payments.

Holt and Laury (2002) (HL) focus on a sequence of paired comparisons, which are not subject to any known anomalies. Their study contrasts the choices made when decisions are hypothetical versus when the decisions lead to real dollar payments. The conclusions regarding the dollar payments are subject to an order effect which Harrison et. al. (2005) controls for and

finds that the qualitative results of HL remain (although the estimated form of the preference function changes.)

An abbreviated version of the HL design is the following: Subjects in each treatment made ten decisions between two gambles (illustrated in Table 1a.) One gamble was a "safe" gamble with a small payout difference between the two possible payouts, while the other was "riskier" with a higher difference between payoffs. As seen in the table, the seven treatments run included four degrees of increasing actual stake sizes and three corresponding hypothetical high stakes treatments where subjects' instructions asked what they would do in the event they faced the decisions in an actual high stake environment. Subjects received payment based on one randomly selected draw from their ten decisions.

<Table 1>

The HL results argue that, for a low level of payoffs, there was no distinction between the between hypothetical and real payoffs; however, as stakes increased by factors of 20, 50, and 90 there was a marked divergence between the results with payoffs and the results without payoffs. Figures 1 and 2 portray these results. Note that hypothetical payoffs shift the distributions right.

While providing support for the validity of paying subjects in cash, HL's results cast further doubt on the effectiveness of using small stakes experiments to arrive at insights about how subjects behave in a high stakes environment. To our knowledge the HL experiments are the only risk-elicitation games that use such high dollar stakes[1]. As everyday portfolio and production decisions can easily be of such magnitudes, the generalization from low stakes behavioral experiments appears tenuous.

---

[1] A number of of high-stakes experiments have been studied. For example, Rapoport (2003) reports a high-stakes centipede game and Camerer (2003) reviews various high-stakes ultimatum games.

We here take an important first step towards solving this problem. In particular, we show that one can use low stakes experiments to produce choices that approximate decisions for much higher stakes. We demonstrate that the induction of preferences method (from Malouf and Roth, 1979 and Berg et. al., 1986) can be extended effectively to induce high stakes preferences in environments with low stakes real payments. The method takes an existing preference function (we use the utility function estimates reported by HL) that summarizes choices in a high stakes environment, and then generates low stakes gambles which produce choices consistent with the original preference function (for theory and proofs that formally justify this approach see Berg et. al., 1986).

We induce the preference function inferred by HL, and then compare the choices made with the induced high stakes preference function both to the hypothetical choices of HL (showing the choice distributions are different) and to the real dollar payoff choices of HL (finding the choice decisions indistinguishable.)

2 Background in Inducing

The technique (Malouf and Roth, (1979); Berg et. al. (1986)) is used in choice settings as well as in examining the predictive ability of equilibrium concepts in multiple person settings. The complete story of its validity and potential uses has still not been determined, so that the current study adds to understanding of the technique. Several reviews of the technique exist suggesting that risk aversion can be induced relatively successfully but that there are mixed results for risk preferring and risk neutrality (for other reviews see also Davis and Holt, (1993) Roth (1995), Camerer (2003)).

> **Comment [JAA1]:** Should we cite a few of these studies you (john) refer to here?

Two sets of studies have focused on examining anomalies: those by Harrison (1994) and those by Selton et. al (1999). Harrison generally finds that induction mitigates anomalies, while Selton et al. (1999) argues that with induction the results are no different than paying people in cash (including preference reversal, Allais paradox effects, etc.). Berg, Dickhaut and Rietz (2003) reexamine the Selton et al. (1999) finding, using inducing, and show that unlike Selton et. al (1999), inducing does not refute an expected utility with noise explanation and in fact significantly alters the preference reversal claims of Selton et. al. (1999).

3 Experimental Design

Our goal is to create an experiment that can incentivize high-stakes behavior in a low-stakes environment. To do this we incorporate the inducing procedure of Berg et al. (1986) into the HL experimental design. We study five treatments. The first four coincide with the four treatments in HL: Low payouts, 20x payouts, 50x payouts, and 90x payouts. Often times we would like to examine decisions people make under very high sums, thus we expand the manipulation beyond incentives used by HL. The fifth treatment we run does not occur in HL, but is a treatment comparable to offering 180x the payouts of the HL Low payout treatment. Appendix A contains a transcript of the instructions.

3.1 Stage 1

In Stage 1, subjects make a decision between two lotteries, A and B. This is the only decision the subject makes. Instead of being paid cash from the outcomes of their chosen lottery, as in HL, subjects receive points. Table 2 lays out each of these ten decisions. As shown, in the first decision, there is a 10% (90%) chance of the high (low) number of points. As subjects proceed

through the decisions, the chance increases (decreases) by 10% each decision. After a subject makes their decision (A or B), they roll a ten-sided die, which determines the number of points they receive. A roll of 1 in the first decision means the subject receives the high number of points, a 1 or 2 in the second decision means they receive the high number of points, a 1, 2, or 3 in the third, and so on.

<Table 2>

Table 3 lays out the points that make up the high and low payouts of each lottery (A and B) in the five treatments. The table indicates that the points earned in each of our treatments coincide precisely with the dollars earned in the corresponding HL treatment (third and seventh columns.) For example if a subject would earn 40 dollars in the HL treatment, they would earn 40 points in our treatment.

<Table 3>

3.2 Stage 2

Stage 2 converts the points subjects earn in Stage 1 into monetary cash earnings. Subjects either earn $2.50 for the decision or nothing. The points a subject receives in Stage 1 translate into chances (Bernoulli trial probabilities) to win a $2.50 prize (more points implies more probability of the prize.) Table 3 shows the probability that a subject wins the prize for each of the possible number of points earned in each treatment (fourth and eighth columns.) Subjects roll a 100-sided die on their desk to determine if they win the prize. For example, when a subject has a 2% chance of the prize, if the die lands with sides 1 or 2 facing up, they win the prize, otherwise they do not. Likewise, if a subject has a 73% chance of the prize, with any side between (and including) 1 and 73 facing up, they win the prize.

3.3 Discussion

It is worth emphasizing several features of our design. First, the only difference between the treatments come in the point payouts (and the related prize probabilities.) Secondly, note that while the probability of winning the prize changes between treatments, the value of the prize does not. The prize remains at a constant $2.50 in each of the treatments. Thus, while in HL treatments the amount a subject can win increases dramatically through the treatments, in our treatments a subject never receives the opportunity to earn more than $2.50 for a decision (subjects always participates in a low-stakes experiment.) Finally, with the use of the Berg et al. (1986) methodology, there are no wealth effects in our design.[2] We can therefore pay subjects for each of the ten HL decisions instead of only one randomly selected decision.

3.4 Procedures

Upon arriving to the experiment, the experimenter directs subjects to the appropriate room where they read the instructions in private and listen to the instructions read aloud by the experimenter. The instructions include paid practice to insure subjects understand the procedures. After the instructions, subjects make their first decision between lotteries A and B. Subjects then roll a ten-sided die at their desk (with the monitor watching) to determine the number of points they receive, which is recorded. The subject follows that roll, immediately, with the roll of a 100-sided die to determine if they win the $2.50 prize. Each of the ten HL decisions repeats this procedure. Subjects receive their cash payments immediately prior to leaving the laboratory.

---

[2] Wealth effects are not a problem in this design because we pay subjects in points for their Stage 1 decision. Assuming expected utility maximization, subjects should prefer to choose the point-maximizing option, and this is independent of wealth levels.

3.5 Hypotheses

HL's hypothetical treatments provide data on how subjects make stage 1 decisions in the above design, when faced with hypothetical large payments. In order to demonstrate that high-stakes behavior can be induced in a low-stakes environment, subjects must not approach the induced high-stakes of our design environment as a hypothetical high-stakes environment.

Hypotheses 1 *The distribution of safe choices (choice A) from our choice data based on induced preferences will be statistically distinguishable from the Holt-Laury distribution of safe choices for each hypothetical high stakes treatment.*

Likewise, successfully inducing high stakes subject decision-making behavior requires subjects' decisions in the induced high stakes environment not to significantly differ from subject behavior under actual high stakes. Hypothesis 2 captures this requirement, using the HL data from their actual stakes treatments as a comparison group.

Hypotheses 2 The *distribution of safe choices (choice A) from our choice data based on induced preferences will be statistically indistinguishable from the Holt-Laury distribution of safe choices for each real stakes treatment.*

5 Results

The experiments took place at the Interdisciplinary Center for Economic Science (ICES) at George Mason University. Subjects were randomly recruited from the George Mason student body. In addition to any amount earned in the experiment, each subject received seven dollars for arriving to the laboratory on time. Subjects spent about 90 minutes in the laboratory.

We report data from 98 subjects in five treatments: 19 in the Low treatment, 20 in the x20 treatment, 17 in the x50, 21 in the x90, and 21 in the x180.[3] We compare the decisions made by these subjects in our induced stakes environments to the decisions made by the subjects in the comparable experimental treatments in HL.

Result 1: *The distribution of safe choices (choice A) from our choice data based on induced preferences is statistically different from the Holt-Laury distribution of safe choices for each of the three HL hypothetical high stakes treatments.*

This result supports Hypothesis 1. The three graphs in Figure 1 show that subjects systematically approach induced preference environments differently than they approach hypothetical high stakes environments. The reported p-values from the Kolmogrov-Smirnov and Kruskal-Wallis tests show significant differences at standard levels for all three treatments. For the Kolmogrov-Smirnov tests: $p=0.046$, $p=0.087$, and $p=0.016$ for the x20, x50, and x90 treatments respectively. For the Kruskal-Wallis: $p=0.032$, $p=0.006$, $p=0.006$ for the x20, x50 and x90 treatments respectively.

Result 2: *The distribution of safe choices (choice A) from our choice data based on induced preferences is statistically indistinguishable from the Holt-Laury distribution of safe choices for each of the four HL treatments.*

This result supports Hypothesis 2. As the four graphs in Figure 2 show, the distribution of safe choices chosen by subjects in each of our induced preference treatments follows that of the distribution of safe choices chosen by subjects in HL. There is no significant difference between

---

[3] These samples are similar to the sample sizes used by HL in their very high stakes treatments (19 and 18 in their x50 and x90 treatments respectively.)

these two distributions in any of the four treatments, at standard levels, neither through use of Kolmogrov-Smirnov 2-tailed tests nor Kruskal-Wallis two-tailed tests.

<Figure 2>

Result 3: *The actual distribution of safe choices with induced preferences of simulated high stakes that are 180 times that of Holt-Laury's low stakes follows the distribution of safe choices predicted by the Holt-Laury power-expo utility function.*

This result supports Hypothesis 3. Figure 3 shows the predicted noisy distribution using HL's power-expo function under an actual 180x stakes environment compared to subjects in our experiment under an induced 180x stakes environment. A chi-squared goodness-of-fit test shows no significant difference between these two distributions (P>0.25).

<Figure 3>

Result 4: *The use of induced high stakes environments in experiments is cheaper than the use of actual high stakes environments.*

Table 4 displays result 4 in greater detail. The first column of the table lists the expected costs of paying one subject for one randomly selected lottery from their ten lottery choices made in a standard actual-stakes HL session. The second column shows the comparable cost when using an induced high stakes environment instead. As seen, the expected cost differences are substantial. As discussed briefly in the design section, in the induced stakes environment, experimenters can pay subjects for all ten lottery decisions without wealth effects concerns. The fourth column in the table lays out the expected cost per subject when paying for each decision. Comparison to the third column demonstrates that even if one could pay for all ten decisions

under actual high stakes without wealth effects, the cost per subject increases rapidly. As a comparison, our induced stakes environment has an expected cost per observation of about 22 dollars in the 180x stakes treatment when paying for all ten decisions, while paying for one decision in an actual 180x stakes environment has an expected cost of over 388 dollars per observation.

<center>**<Table 4>**</center>

6 Discussion

The importance of salient rewards, long emphasized by Vernon Smith and formalized with his "Induced Value Theory" (Smith, 1976), is a defining feature of experimental economics. It separates experimental economics research from much related work on decision making occurring in other social science and business school environments. Unfortunately, this emphasis might be thought to entail an inability to use laboratory investigations to study behavior and decision in very high stakes environments. Here we have argued that, in fact, this is not correct. In particular, we have demonstrated that using the "preference inducement" procedure (Berg et al, 1986) one is able to generate high-stakes behavior within a low-stakes laboratory environment.

We focused on the behavior discovered by Holt and Laury (2002). They found that risk attitudes varied systematically with the magnitude of the payoff associated with the decision. Moreover, they estimated a utility function that captured the dependence of choice on size of payoff. In this paper we showed that, in a low-stakes environment, the Berg et al (1986) procedure can be used to generate choices in this risk task that follow the same patterns Holt and Laury found in their high-stakes conditions.

Economically important decisions under risk often occur in high-stakes environments, lending special importance to our study. Also, these risk-decisions such as portfolio choices are lacking strategic elements, which we found attractive for this first-step investigation. In future research, we intend to induce preferences within game and market environments. This can be accomplished within any environment where appropriate data exists to back-out participant preferences (e.g., trader risk-aversion can be inferred from financial market data.) We are confident that doing this will enable preference inducement to become an increasingly valuable procedure for informing high-stakes behavior under novel institutions, and therefore become a key tool for anyone interested in mechanism design.

References

Berg. J.E., L.A. Daley, J.W. Dickhaut, JR. O'Brien (1986) "Controlling Preferences for Lotteries on Units of Experimental Exchange". *Quarterly Journal of Economics*. 101(2) 281-306.

Berg J.E., J.W. Dickhaut, and T. A. Rietz (2003) "Preference Reversals and Induced Risk Preferences: Evidence for Noisy Maximization". *Journal of Risk and Uncertainty*. 27(2) 139-170.

Camerer, C.F. and R. M. Hogarth, (1999) "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework" *Journal of Risk and Uncertainty*, 19, 7-42.

Camerer, C. F. (2003) " Introduction" and " Dictator, Ultimatum, and Trust Games" *Behavioral Game Theory*. Princeton, New Jeresy: Princeton University Press. p. 40-41, 60-62.

Davis, D. and C. Holt (1993) *Experimental Economics* Princeton, New Jersey: Princeton University Press. p 472-476.

Harrison, G. (1994) "Expected Utility Theory and the Experimentalists" *Empirical Economics.* 19, 223-253.

Harrison, G.W., E. Johnson, M. M. McInnes, and E.E. Rutström (2005) "Risk Aversion and Incentive Effects: Comment" *The American Economic Review.* 95(3) 897-901

Holt, C. A. and S. K. Laury. (2002) "Risk Aversion and Incentive Effects." *The American Economic Review*. 92(5) 1644-1655.

Kahneman, D. and A. Tversky (1979) "Prospect Theory: An Analysis of Choice Under Risk." *Econometrica* 47(2) 263-291.

Rapoport A., W. E. Stein, J.E. Parco, and T.E. Nicholas (2003) "Equilibrium Play and Adaptive Learning in a Three-Person Centipede Game." *Games and Economic Behavior.* 43, 239–265.

Roth, A. E. and M.W.K. Malouf. (1979). "Game-Theoretic Models and the Role of Bargaining". *Psychological Review* 86, 574–594.

Roth, A. (1995) "Introduction to Experimental Economics", *The Handbook of Experimental Economics*. Princeton, New Jersey: Princeton University Press. p. 81-83.

Selton R., A. Sadrieh, and K. Abbink (1999) "Money Does Not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse". *Theory and Decision*. 46(3) 213-252.

Smith, V.L. (1976) "Experimental Economics: Induced Value Theory." *The American Economic Review*. 66(2) p. 274-279.

Table 1

A. Holt - Laury Paired Lottery-Choice Decisions: Low Payout Treatment

| | Gamble A (Safe) | | Gamble B (Risky) | |
	Chance of Receiving 2 Dollars	Chance of Receiving 1.6 Dollars	Chance of Receiving 3.85 Dollars	Chance of Receiving 0.1 Dollars
Decision 1	10%	90%	10%	90%
Decision 2	20%	80%	20%	80%
Decision 3	30%	70%	30%	70%
Decision 4	40%	60%	40%	60%
Decision 5	50%	50%	50%	50%
Decision 6	60%	40%	60%	40%
Decision 7	70%	30%	70%	30%
Decision 8	80%	20%	80%	20%
Decision 9	90%	10%	90%	10%
Decision 10	100%	0%	100%	0%

B. Holt - Laury Treatments

Treatment	Gamble A		Gamble B	
Low	$2.00	$1.60	$3.85	$0.10
x20	$40.00	$32.00	$77.00	$2.00
x50	$100.00	$80.00	$192.50	$5.00
x90	$180.00	$144.00	$346.50	$9.00
Hypothetical x20*	$40.00	$32.00	$77.00	$2.00
Hypothetical x50*	$100.00	$80.00	$192.50	$5.00
Hypothetical x90*	$180.00	$144.00	$346.50	$9.00

*In the hypothetical treatments dollar amounts listed are hypothetical amounts only.

Table 2
DHATJ Paired Lottery-Choice Decisions:
Low Payouts

	Option A		Option B	
	Chance of Receiving 2 Points	**Chance of Receiving 1.6 Points**	**Chance of Receiving 3.85 Points**	**Chance of Receiving 0.1 Points**
Decision 1	10%	90%	10%	90%
Decision 2	20%	80%	20%	80%
Decision 3	30%	70%	30%	70%
Decision 4	40%	60%	40%	60%
Decision 5	50%	50%	50%	50%
Decision 6	60%	40%	60%	40%
Decision 7	70%	30%	70%	30%
Decision 8	80%	20%	80%	20%
Decision 9	90%	10%	90%	10%
Decision 10	100%	0%	100%	0%

Table 3
Holt-Laury and DHATJ Treatment Comparrison

Treatment	Choice	Dollars in HL/ Points in DHATJ	Chance of $2.50 Prize in DHATJ	Treatment	Choice	Dollars in HL/ Points in DHATJ	Chance of $2.50 Prize in DHATJ
Low	A	2	61%	90	A	180	79%
		1.6	52%			144	71%
	B	3.85	97%		B	346.5	99%
		0.1	7%			9	1%
20	A	40	65%	180	A	360	89%
		32	56%			288	83%
	B	77	97%		B	693	100%
		2	2%			18	1%
50	A	100	73%				
		80	64%				
	B	192.5	99%				
		5	1%				

Note that in HL (DHATJ) chance of receiving the high number of Dollars (Points) for choice A or B increases by round starting with 10%, in round 1, and increasing to 100%, in round 10.

Table 4
Cost Comparrison

	Expected Cost per Observation If Paying For 1 of 10 Choices		Expected Cost per Observation If Paying For All 10 Choices	
	HL	**DHATJ**	**HL**	**DHATJ**
Low	$2.41	$1.67	$24.87	$17.06
x20	$47.18	$1.71	$486.18	$17.44
x50	$113.73	$1.85	$1,168.98	$18.75
x90	$204.71	$1.97	$2,104.16	$19.85
x180	$388.17	$2.19	$3,974.67	$21.99

Expected costs calculated by power-expo utility maximizing behavior by all agents

Figure 1
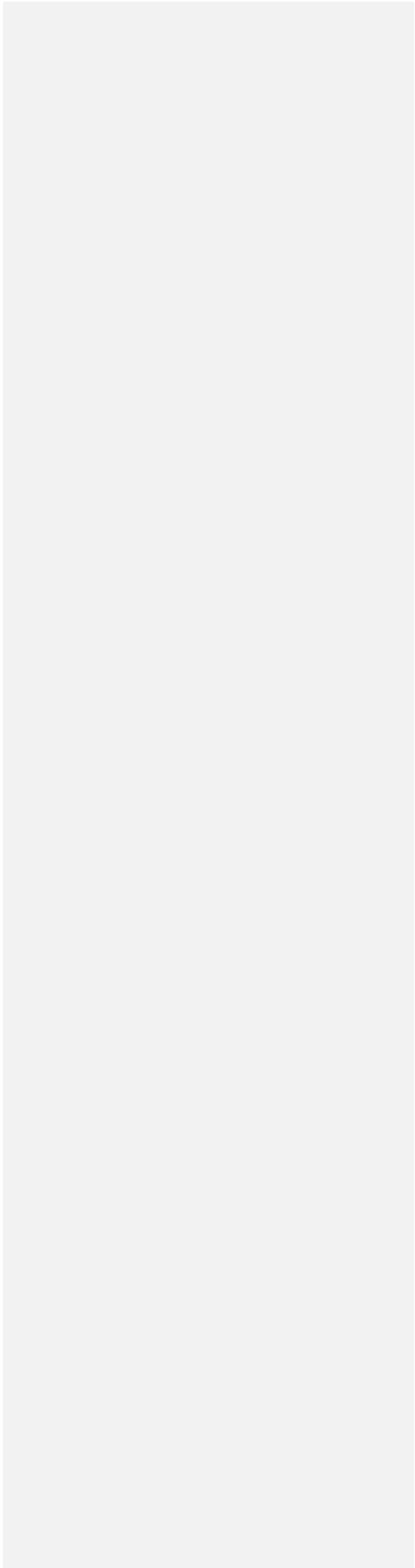HL-Hypothetical vs. DHATJ Comparrison

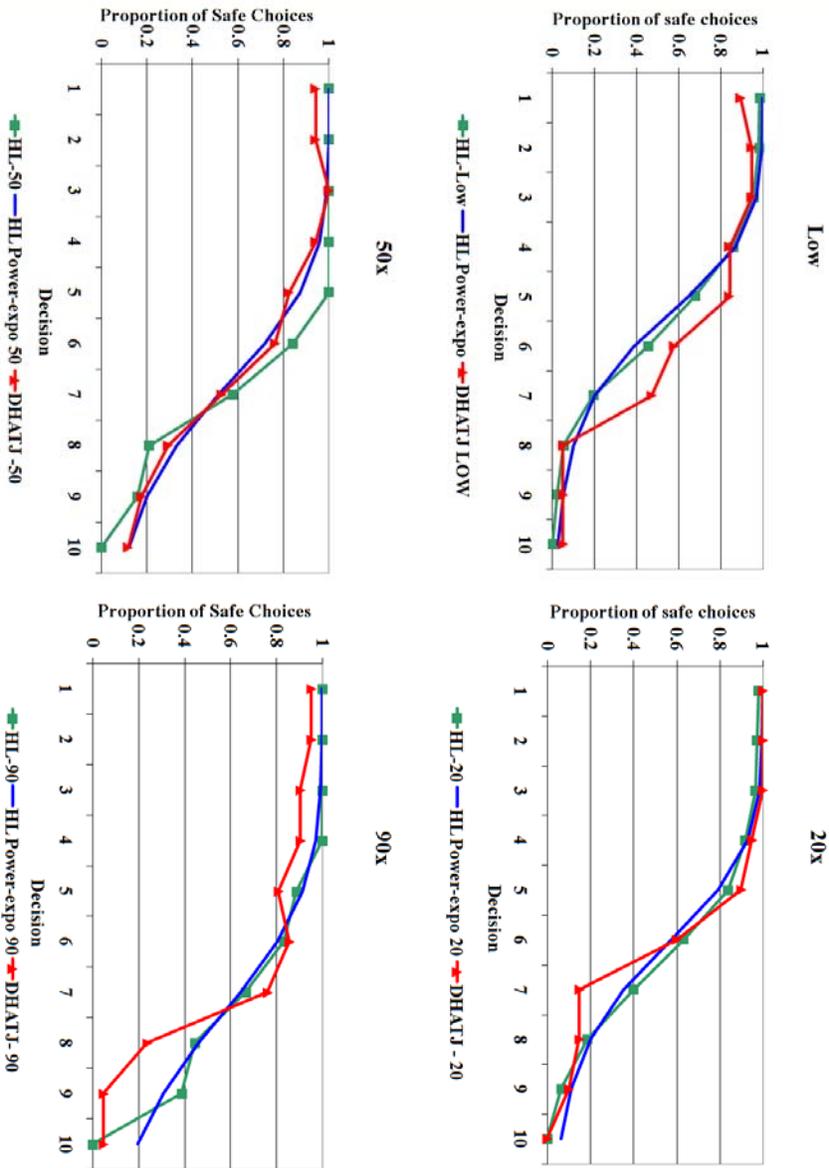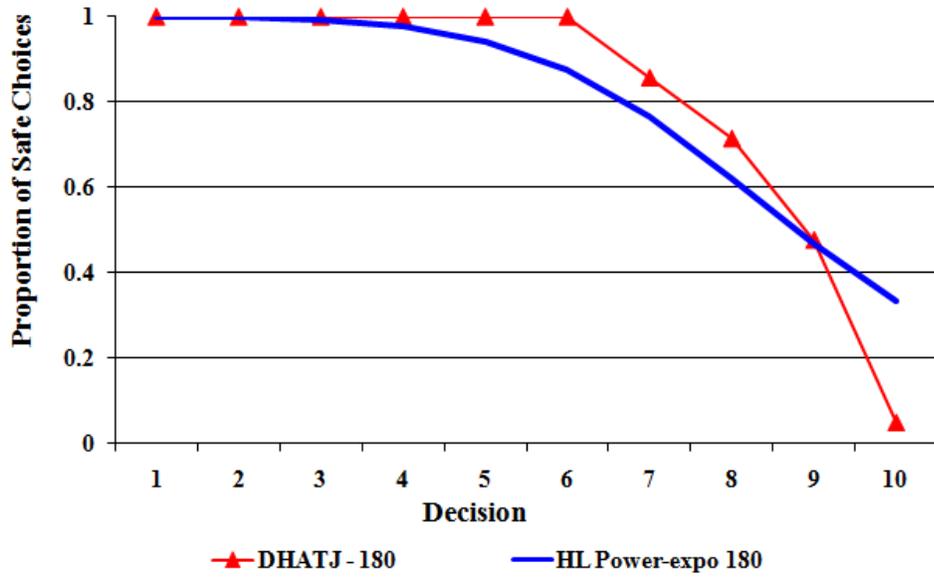**Figure 2**
HL vs. DHATJ Comparrison

Figure 3
DHATJ Induced High Stakes 180x HL

Economic Science Institute Working Papers

2008

08-09 Stecher, J., Shields, T. and Dickhaut, J. Generating Ambiguity in the Laboratory.

08-08 Stecher, J., Lunawat, R., Pronin, K. and Dickhaut, J. Decision Making and Trade without Probabilities.

08-07 Dickhaut, J., Lungu, O., Smith, V., Xin, B. and Rustichini, A. A Neuronal Mechanism of Choice.

08-06 Anctil, R., Dickhaut, J., Johnson, K., and Kanodia, C. Does Information Transparency Decrease Coordination Failure?

08-05 Tila, D. and Porter, D. Group Prediction in Information Markets With and Without Trading Information and Price Manipulation Incentives.

08-04 Caginalp, G., Hao, L., Porter, D. and Smith, V. Asset Market Reactions to News: An Experimental Study.

08-03 Thomas, C. and Wilson, B. Horizontal Product Differentiation in Auctions and Multilateral Negotiations.

08-02 Oprea, R., Wilson, B. and Zillante, A. War of Attrition: Evidence from a Laboratory Experiment on Market Exit.

08-01 Oprea, R., Porter, D., Hibbert, C., Hanson, R. and Tila, D. Can Manipulators Mislead Prediction Market Observers?