

# Design III: Implementation

Ryan Oprea

University of California, Santa Barbara

Economics 276a,  
Lectures 4

# Computerization Pro and Con

The earliest experiments were run with pieces of paper, pencils and chalkboards. In the 1970s computerized experiments emerged

Pro:

- No data entry
- Generally much quicker
- Lower staff required
- Better visualization (see below)
- Many experiments are simply infeasible without software.

Con:

- Software development can be expensive and time costly

Most contemporary experiments are run using software in experimental computer labs.

## Existing Packages

Happily there are several good options to prevent you from having to write software from scratch.

By far the most popular is **z-Tree**:

- Contains a full set of built in networking tools.
- Minimal programming required and syntax is very easy to learn.
- Recent versions allow ambitious users to implement surprisingly rich visualizations.

Several other more specialized packages exist

- ConG (Continuous Games) (A very visually rich set of tools for conducting game theory experiments in continuous time);  
EconPort (a set of tools covering common experiment types);  
Veconlab (a set of online tools for running common experiment types)

Also several packages exist for running market experiments

- Marketscape, JMarkets, Edgeworks (coming soon)

Coming soon: Redwood, a set of very flexible, graphical Javascript-based tools.

# Typical Architecture

Most programs for running economics experiments have a predictable set of components:

- **Monitor Program:** Calculates payoffs, realizes random variables and prints data usually to text file. Often includes an interface for the experimenter to use to control the experiment.
- **Client Program:** Generates a user interface providing visual information and a set of widgets (sliders, buttons, text input fields) for subject decision making.
- **Config (or Parameter) Files:** Config files are used by the monitor program to set treatment variables. Usually there is a separate config file for each treatment (at least).

Most programs have a thin-client structure where most of the computation is done by the Monitor program and the client program serves only to generate visualizations and capture subject decisions.

# Visualization and Salience

A premise of most models: subjects know their own preferences intimately.

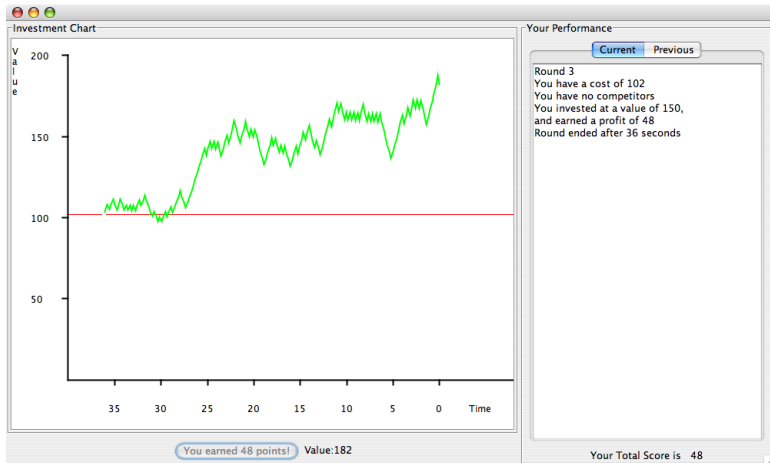
- Experiments are not math tests on people's ability to maximize utility functions.

To fairly test a model you should do everything in your power to make elements assumed in the setup of the model a reality in the experiment (this is an important part of IVT3: Salience)

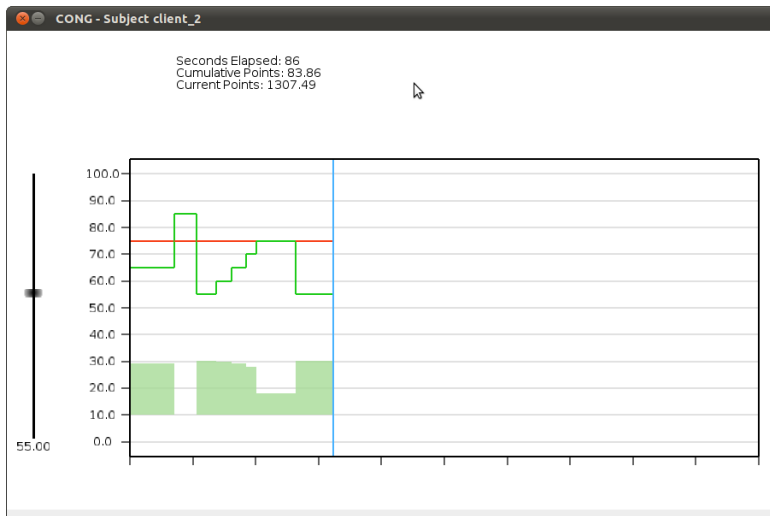
If all subjects were mathematicians who could do calculus in their head, we might just give them mathematical functions. Luckily human beings have equivalents in their visual cortex.

A big part of experiment design (for me anyway) is figuring out how to use visuals to generate high degrees of salience.

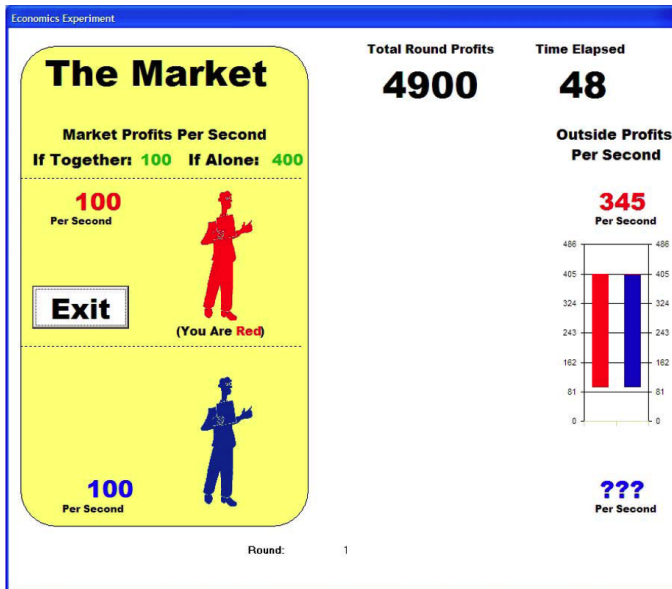
# Real Options Example



# Example: ConG Pricing

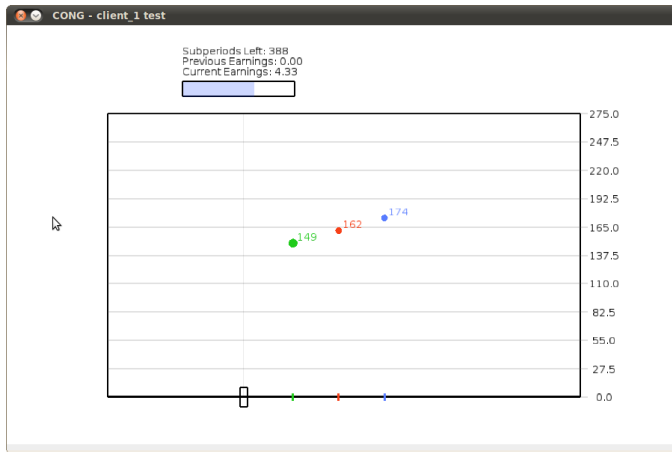


## Example: War of Attrition

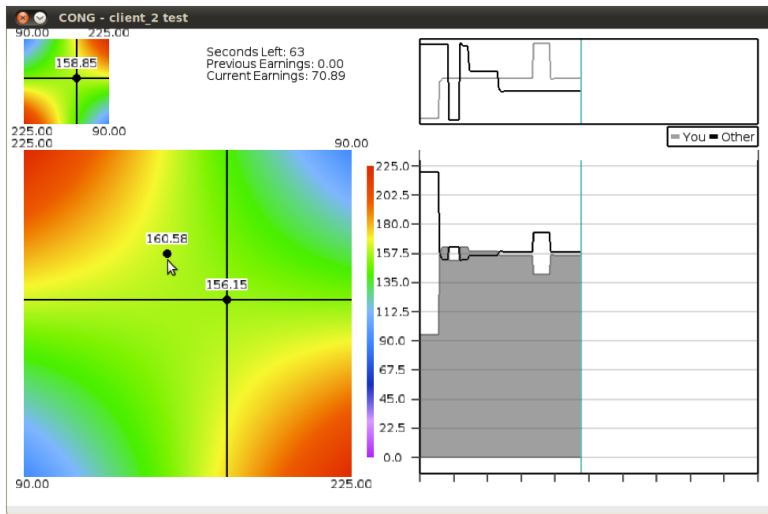




# Example: ConG Bubbles Display



# Example: ConG Heatmaps



# Example: Edgeworks Heatmaps

Bids	Trades Price [ Qty ]	Asks
23.70 [+3.2]	17.91 [2.1]	33.00 [-1.4]
21.81 [+1.4]		41.15 [-1.1]
21.65 [+1.7]		47.14 [-1.0]
19.36 [+1.4]		61.51 [-0.8]

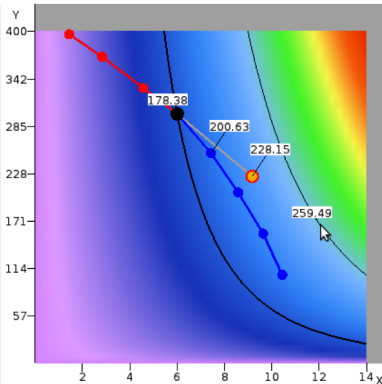
	X	Y
Price:		
+ Qty:	6.00	300.00

**Bid**

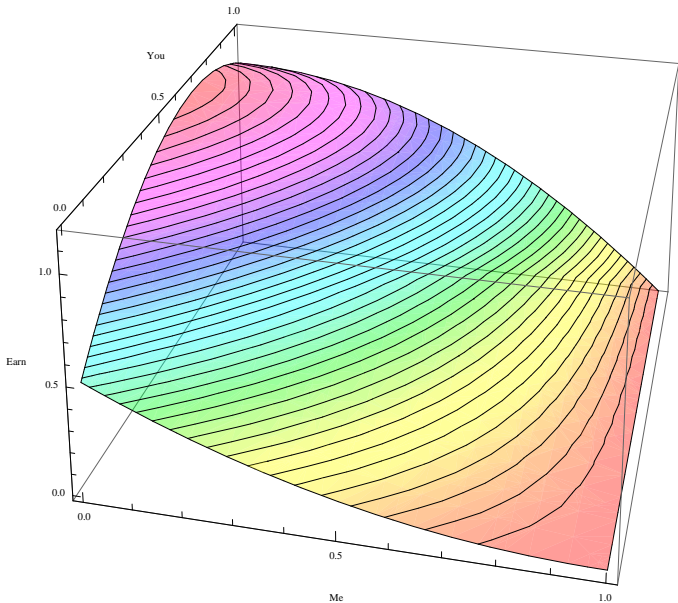
  

	Price:	-Qty:

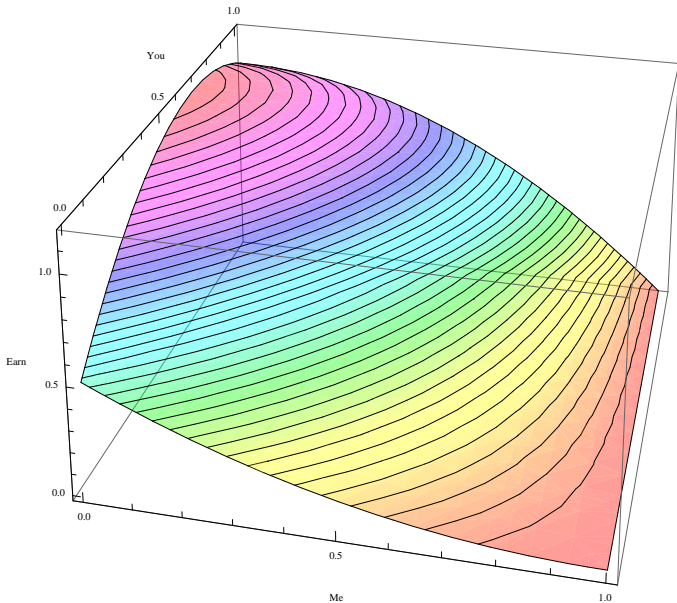
**Ask**



## Dilemma: Continuous Example



## Dilemma: Continuous Example



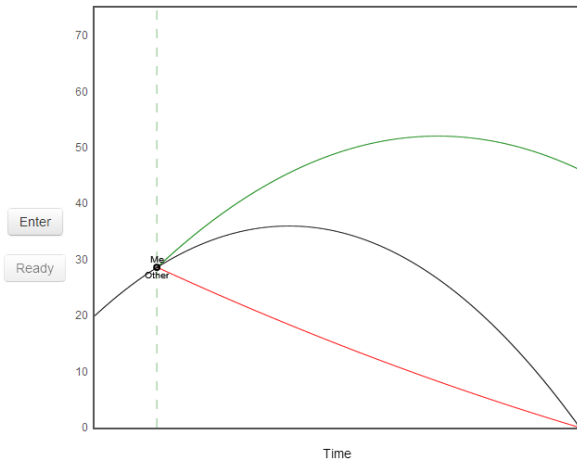
# Continuous Example

Period: 1

Time Remaining: 52

Points: 0

Total Points: 0



Enter

Ready

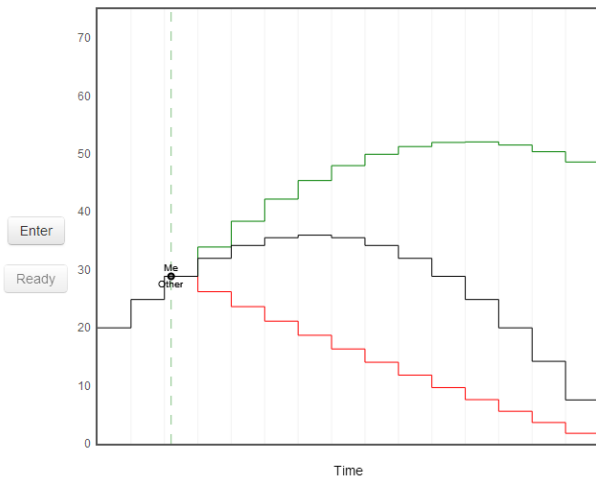
# Continuous Example

Period: 1

Time Remaining: 51

Points: 0

Total Points: 0



# Moves By Nature

Another important issue (raised a bit in the last lecture) is how to realize random variables.

A worry is that both dominance and salience can be compromised if subjects are suspicious of the randomization protocol being used in the experiment.

In complex experiments (e.g. dynamic stochastic decision problems) computer randomization is unavoidable but where possible it is preferable to use:

- Bingo cages
- Urns
- Public coin flips
- Public die rolls



# Elicitation

Fundamentally, experiments measure behavior and often we use this to infer things about

- Preferences
- Beliefs
- Valuation
- Expectation and intentions

An alternative is to create experimental tasks that more directly measure these things, often simply as a way of explaining other behavior.

Experimental economists generally do not simply ask subjects to reveal these things for cheap talk concerns. Instead we use incentive compatible mechanisms.

# Eliciting Values

The Becker-DeGroot-Marschak (1964) procedure measures subjects' willingness to pay (WTP) for an item. Similar to a Vickrey auction.

- Subject submits a bid,  $b$
- Random price  $p$  is drawn
- If  $b > p$ , price is paid and subject receives object.

Can reverse the procedure to elicit willingness to accept (WTA).

Problem: Only works theoretically for risk neutral subjects.

## Eliciting Beliefs

Often it is useful to measure beliefs. Simplest (and most frequent) case is binary: will an event occur. Elicit probability estimate using a *scoring rule*.

Naive (linear) scoring rule elicits probability estimate  $p$  and pay  $s(p) = a + b \sum_{i=1}^n p_i d_i$  where  $n$  is the number of possible events realized and  $d_i$  is a dummy variable valued at 1 if the event occurs and 0 otherwise. Not incentive compatible to reveal true belief.

Quadratic scoring rule:  $s(p) = a - b \sum_{i=1}^n (p_i - d_i)^2$  is incentive compatible.

Scoring rules that are incentive compatible are called proper and there are a number to choose from with slightly different properties (e.g. logarithmic scoring rule, spherical scoring rule)

Problem: Scoring rules often have flat maxima and risk averse subjects may have strong reasons to choose  $p$  near 0 or 1. See Offerman et al (2009) for generalizations that are quite a bit more robust.

# Eliciting Preferences

There are a number of tools in the literature for eliciting risk preferences. I name a few:

- Holt-Laury (2002): Most commonly used. Have subjects choose between ten lottery pairs to get a rough estimate of a population CRRA parameter.
- Choi et al (2007): Use a graphical interface to visualize portfolio choice problems and have subjects make a number of decisions to get relatively strong portrait of individual preferences.
- Wang et al (2010): Propose dynamically optimized sequential experiments (DOSE). These are basically computerized tasks that generate questions based on previous answers and a Bayesian algorithm to significantly reduce the number of questions needed to elicit preferences.

Other preferences of course can be elicited with simple tasks (e.g. Dictator games are sometimes used to measure social preferences).

# Eliciting Strategies Instead of Actions

In dynamic games, it is sometimes useful to employ the strategy method (Introduced by Selten, 1967). Instead of having subjects choose their actions in sequence, have subjects effectively program robots to make decisions prior to letting the clock run.

In timing games this might be as simple as a decision time. In more complex strategic problems this might be a full characterization of decisions as a function. Does two things:

- Shows researchers full strategies (i.e. actions off the equilibrium path).
- Avoids nasty censoring problems in timing games.

Question: Does the experience of sequence and time matter? There are a few investigations but I think it is an interesting open question.

# Other

Experimental economists prefer incentive compatible mechanisms but it has become increasingly popular to collect other information:

- Surveys to collect demographic information, income etc.
- IQ tests
- Common psychological batteries (e.g. )

Explorations of how this sort of information organizes and explains incentivized lab behavior is a growing area in experimental economics.

# Running a session

Most experimental sessions follow a predictable pattern:

- 1 Recruit subjects (over-recruit)
- 2 Randomly assign subjects to seats
- 3 Read instructions (possibly administer quiz)
- 4 Run experimental task
- 5 Pay subjects

# Randomization and Blocking for Treatment Assignment

A key inferential advantage of experiments is random assignment of treatment to subject.

- Avoids self selection
- Also avoids some unwitting experimenter effects

Strictly speaking this implies, for between-designs you should randomize

- Assignment of treatments to sessions
- Assignment of treatment to lab (if a multisite project)
- Assignment of treatment to experimenter (if multiple people will be running)

Even better, employ blocking:

- Run all treatments in a randomized sequence and repeat until full design is fulfilled.
- Have all labs and experimenters run an equal number of sessions under each treatment.



## Recruiting to an Experiment

Current best practice in experimental economics is to randomize selection of subjects to experiment in the recruiting process.

Typically subjects are recruited out of classrooms using a script to sign up to an online piece of recruiting software that stores records in a database.

- Current best practice is to populate the database using survey classes from across the curriculum to avoid biases induced by subject of concentration.
- Avoid self selection, friends attending experiment together etc.

Recruiter software then randomly invites subjects by email. Often experiments use filters (or blocking) implemented by the recruiter by:

- Experience levels in other recent experiments
- Demography (race, gender, income or even measured characteristics from previous experiments)
- Academic major

Rare to get full showup among subjects who've signed up; typically over recruit by 25% or so.

## Instructions

Careful instructions are a critical part of experiments – they are important enough to (usually) be included in manuscripts submitted to journals.

Some pointers:

- 1 Instructions should be really clear; the goal is to have subjects completely understand the relationship between actions and payoffs (IVT3,salience).
- 2 Avoid use of overtly normative language.
- 3 Avoid leading language.
- 4 Avoid technical language or terms of art. Explaining the subject's task is not the same as giving a theory seminar.
- 5 Instructions should be as short as possible – subjects' eyes tend to glaze over if instructions are too long.
- 6 Include visuals of the software (or decision sheets) in the instructions.
- 7 Read and reread your instructions with the eyes of a critical reader of your paper. Are you biasing your results?

# How To Deliver

There are a few common ways to deliver instructions

- ① Hand out paper instructions and let subjects read them individually.
- ② Program up instructions with visuals and possibly interactive decision components.
- ③ Hand out paper instructions and read them aloud, possibly with visuals projected on a screen.

Post instructions training

- ① Questions and Answers
- ② Quizzes
- ③ Paid or unpaid practice periods

# Framing

An issue with some controversy attached is how to frame subjects' decisions in the instructions (and software).

Most of the literature has the view things should be antiseptic (don't call subjects firms, don't label actions by their economic interpretation). Sometimes called neutrality.

There is still some worthwhile debate about this:

- Naturalistic framing can seriously improve salience.
- Frames are not neutral but they may be important parts of decision making.
- Market experiments (and other complex experiments) often use non-neutral framing (i.e. prices are prices, values are values, costs are costs).

# Payments

After the experiment finishes, you pay your subjects. A few pointers:

- Choose a payoff rate that matches an average subject payment that pays off average subject opportunity cost (most sessions these days total \$12-\$18 per hour)
- Typical to calibrate to equilibrium payoffs
- Virtually all experiments include a “show up fee” (focal amount is \$5) guaranteed for subjects who come to the experiment.
- Pay your subjects in private.
- Don't tell your subjects who earned the most (they will ask).
- Don't tell your subjects what the aim of the study was (for reputation reasons) but offer to provide finished paper in  $N$  months after data collection is complete.

# What to Pay On?

**Pay All:** Typically subjects are paid on earnings summed over periods. Some problems:

- ① **Portfolio Effects:** High risk decisions made in each period have much smaller risk over all periods.
- ② **Wealth Effects:** Satisficing subjects may check out or otherwise systematically change behavior after earning a threshold amount.
- ③ **In General:** Choices must have no effect on optima in other periods.

**Pay For Random Period:** Healy et al. argue you should pay on one randomly selected period

- ① If subjects reduce compound lotteries, then assumes subjects are EU.
- ② If subjects don't reduce compound lotteries works under the assumption of event-wise monotonicity.
- ③ See Cox et al. (2014) and Harrison and Swarthout (2014) for some critical points regarding EU assumption and some experiments. See Freeman et al. (2014) for some cautions regarding ambiguity.

## Other Payment Issues

**Infinitely Repeated Games:** As we said earlier, can study infinitely repeated game using indefinitely repeated game with a termination hazard instead of a discount rate. As Charness and Genicot (2009), Fischer (2011), Chandrasekhar and Xandi (2011), Frechette et al. (2011), Sherstyuk et al. (2011) have pointed out:

- 1 If subjects are risk neutral you can pay people the sum of their earnings.
- 2 Otherwise, should pay subjects based only on the last period before the game expires.

**Bankruptcy:** Many experiments have scope for subjects to lose money.

- 1 Can't pay subjects negative amounts so there is implicit limited liability often not included in the motivating model.
- 2 Typical solution: give subjects enough initial money to keep them solvent
- 3 Beware: this can create serious dominance issues.

# Piloting

Piloting is an important part of the design process and almost all experimentalists do it:

- ① Get baseline data for a power analysis.
- ② Learn about the feasibility of a within-design
- ③ Discover side effects of your implementation method you didn't know about.
- ④ Weed out gross design flaws.
- ⑤ Force yourself to think through your data analysis, hypotheses etc. before committing to a final design.

Important to keep piloting from becoming a form of data mining!



# Running an experiment

Most experimental projects also follow a predictable pattern:

- ➊ Initial Design
- ➋ Human subjects / ethics approval
- ➌ Software development /testing
- ➍ Pilot experiments
- ➎ Main data collection ( “Production runs” )
- ➏ Initial analysis
- ➐ Diagnostic treatments
- ➑ Final analysis and writeup

## Bibliography

- Azrieli, Y. C. Chambers and P. Healy. (2014) "Incentives in Experiments: A Theoretical Analysis."
- Becker, G.M., DeGroot, M.H., Marschak, J. (1964) "Measuring utility by a single-response sequential method" Behavioral Science 9(4): 226-232.
- Chandrasekhar, A.G., Xandri, J.P. (2011) "A Note on Payments in Experiments of Infinitely Repeated Games with Discounting." Massachusetts Institute of Technology working paper.
- Charness, G., Genicot, G., 2009. Informal risk sharing in an infinite-horizon experiment. Economic Journal 119, 796812.
- Choi, S. R. Fisman, S. Kariv, D. Gale. (2007) "Consistency and Heterogeneity of Individual Behavior under Uncertainty" American Economic Review, 97(5), pp. 1921-1938.
- Cox, J. C., V. Sadiraj, and U. Schmidt (2014). Paradoxes and mechanisms for choice under risk. Experimental Economics. 1-36
- Fischer, G., February 2011. Contract structure, risk sharing and investment choice, LSE STICERD Research Paper No. EOPP023.
- Frechette, G., Schotter, A., Yuksel, S., 2011. Implementing infinitely repeated games in the laboratory, New York University Working Paper.
- Freeman, D., Y. Halevy, T. Kneeland. (2014) "Eliciting Risk Preferences using Choice Lists."
- Harrison, G.W. and J. T. Swarthout (2014, Oct.). Experimental Payment Protocols and the Bipolar Behaviorist. Theory and Decisions 77(3), 423-438. Special Issue: Decision Making under Risk and Uncertainty: Festschrift in