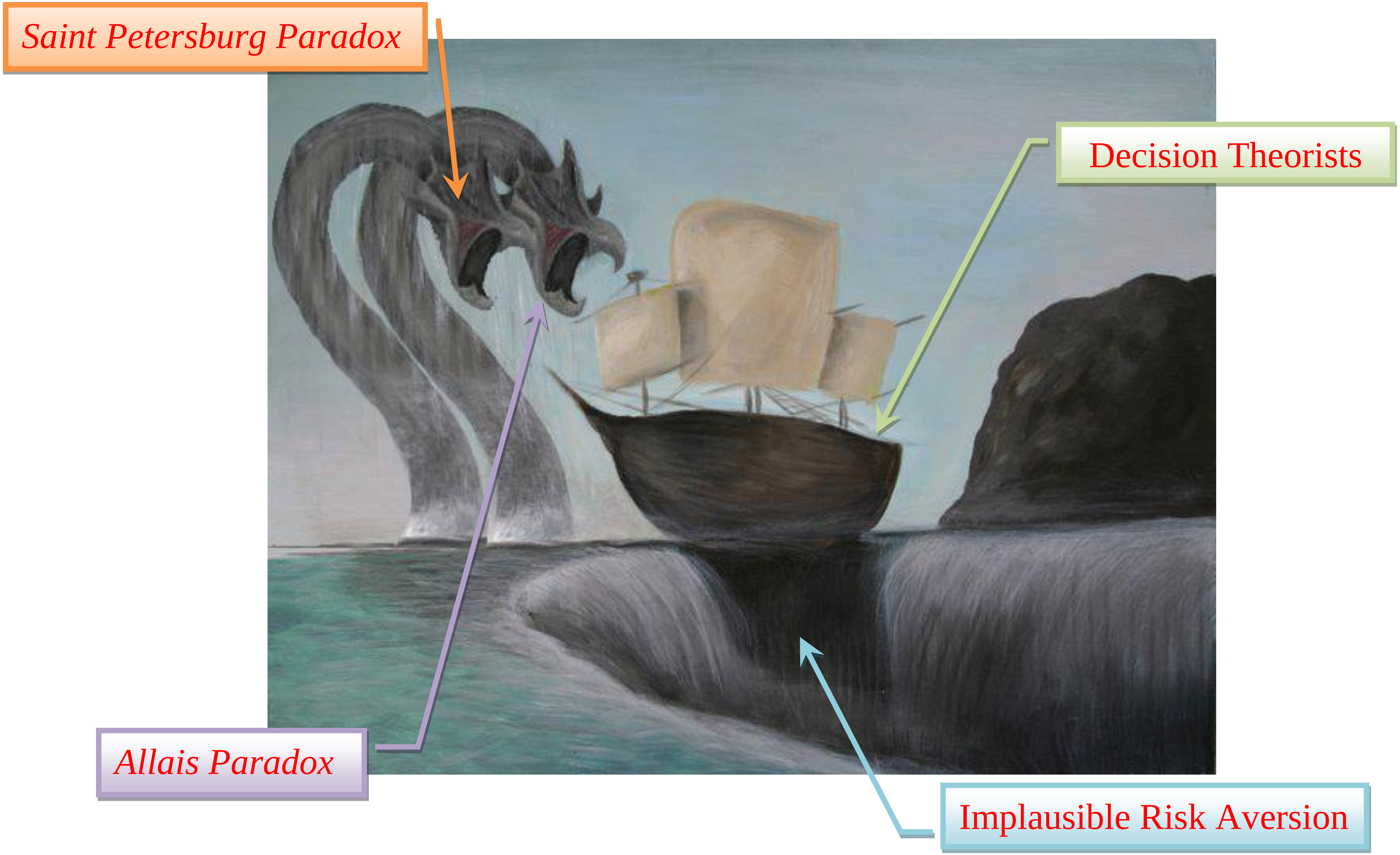
Decision under Risk: Mechanisms, Paradoxes, and Dual Calibrations

James C. Cox

Plenary Lecture: Economic Science Association, Tucson, November 2011

SCYLLA AND CHARYBDIS OF DECISION THEORISTS



Focus on ways in which theories might be implausible:

A. (generalized) **St. Petersburg Paradox** in which theory seems to imply **acceptances** of large-stakes risks that are empirically implausible; supposed to require **concave utility**

B. **Allais Paradox** in which **seemingly plausible patterns** of risk aversion are supposed to be **inconsistent with linearity** in probabilities

C. **Calibration** of nonlinear utility and probability transformations propositions which show **how seemingly plausible patterns** of small-stakes (resp. same-stakes) risk aversion imply **rejections** of large-stakes (resp. same-stakes) risks **that are implausible**

THE ORIGINAL ST. PETERSBURG PARADOX

The game pays 2^n if the first head appears on flip n.

The probability that the first head appears on flip n is $(1/2)^n = 1/2^n$

The expected value of the game is

$$EV = \sum_{n=1}^{\infty} 2^n \times (1/2^n) = 1 + 1 + \cdots + 1 + \cdots$$

Bernoulli (1738) offered concave utility of payoffs (specifically **log utility**) as a solution; but this only solves his example.

GENERALIZED ST. PETERSBURG PARADOX for EUT

- An example for unbounded u with inverse function u^{-1}.
- Let the prize be $u^{-1}(2^n)$ if the first head appears on flip n.

$$EU = \sum_{n=1}^{\infty} u(u^{-1}(2^n)) \times (1/2^n) = 1 + 1 + \cdots + 1 + \cdots$$

- Arrow (1971), Pratt (1964) and Laffont (1989) define EUT on an unbounded domain and assume bounded utility "in order to avoid the St. Petersburg Paradox".
- <u>But bounded utility can be shown to imply implausible risk aversion</u> (Cox and Sadiraj, 2008).

GENERALIZED ST. PETERSBURG PARADOX for POPULAR THEORIES

Generalized St. Petersburg paradoxes can be constructed for five popular theories of decision under risk if they are defined with unbounded money transformation (utility or value) functions on unbounded domains (Proposition 1, Cox and Sadiraj, 2008).

"Popular Theories": EV, EUT, CPT, RDU, DTEU

SO WHAT?

No one can credibly offer "infinite" prizes.

A bounded domain solves the St. Petersburg Paradox.

Suppose that the maximum prize offered is equal to $\$2^{35}$ = \$34.36 billion (less than the \$40 billion profit of Exxon-Mobil in 2007).

This lottery has expected value of

$$\$36 = \sum_{n=1}^{35} [\$2^n \times (1/2)^n] + \$2^{35} \times (1/2)^{35}$$

It should not be a surprise if someone is not willing to pay a very large amount to play the game.

WHAT HAPPENS IF PEOPLE ARE OFFERED REAL, FINITE ST. PETERSBURG LOTTERIES?

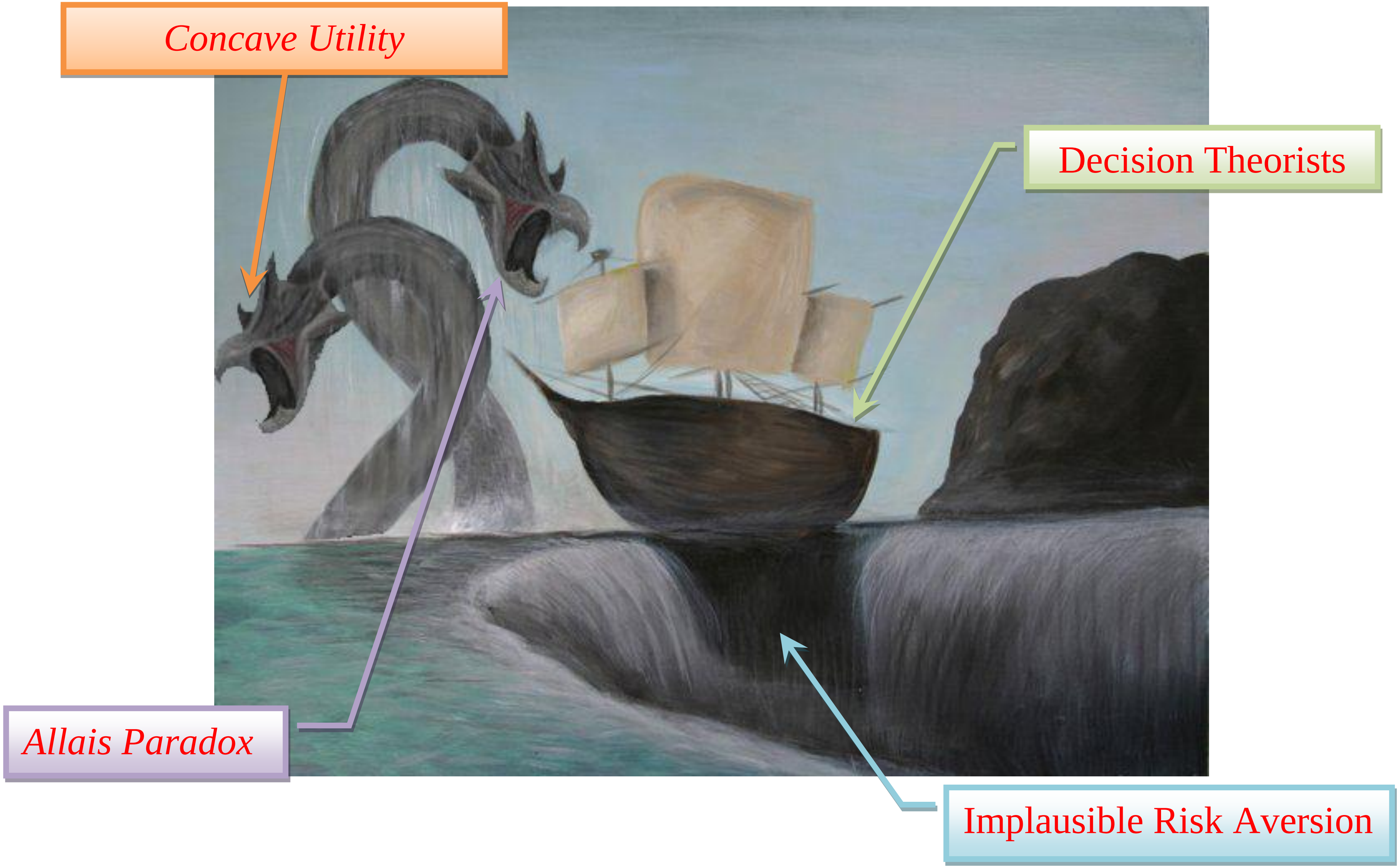
A large majority of subjects reject small-payoff St. Petersburg lotteries (Cox, Sadiraj and Vogt, 2009).

So, maybe we need concave utility after all?

These data and <u>much</u> other data support small-stakes risk aversion.

But, as we shall recall, representing risk aversion with concave utility can produce problems of implausible risk aversion.

DEFLECTED SCYLLA AND CHARYBDIS



ALLAIS PARADOX

First Choice Pair	*Second Choice Pair*
Option 1A: \$1M for sure	Option 2A: 0.11 prob. of \$1M 0.89 prob. of \$0
Option 1B: 0.10 prob. of \$5M 0.89 prob. of \$1M 0.01 prob. of \$0	Option 2B: 0.10 prob. of \$5M 0.90 prob. of \$0

- **Allais reported** that, with hypothetical payoffs, most people choose Option 1A and Option 2B
- Define: P.1 as "Option 1A is preferred"; and Q.1 as "Option 2B is preferred"
- With EUT: P.1 → *not*Q.1 or, equivalently Q.1 → *not*P.1

MOTIVATION FOR NONLINEAR PROBABILITY TRANSFORMATIONS

Data from Much Experimental Literature:

- Allais' (1953) hypothetical experiment
- Many lab experiments with CRE and CCE

have been said to be inconsistent with linearity in probabilities (EUT)

Decades of Research:

- Developed theories with nonlinear probability transformations
- Reported large numbers of experiments said to support such theories

NONLINEARITY IN PROBABILITIES CAN IMPLY IMPLAUSIBLE RISK AVERSION

Cox and Sadiraj (2011) brings some bad news for theories with functionals that are nonlinear in probabilities.

For example, consider choices between the pairs of options in Table 1.

Table 1. Varying Probabilities Calibration Pattern

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="6">$0</td><td>$0</td><td>$30</td><td colspan="12">$30</td></tr>
<tr><td>Opt. B</td><td colspan="6">$0</td><td colspan="2">$10</td><td colspan="12">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="7">$0</td><td>$0</td><td>$30</td><td colspan="11">$30</td></tr>
<tr><td>Opt. B</td><td colspan="7">$0</td><td colspan="2">$10</td><td colspan="11">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="8">$0</td><td>$0</td><td>$30</td><td colspan="10">$30</td></tr>
<tr><td>Opt. B</td><td colspan="8">$0</td><td colspan="2">$10</td><td colspan="10">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="9">$0</td><td>$0</td><td>$30</td><td colspan="9">$30</td></tr>
<tr><td>Opt. B</td><td colspan="9">$0</td><td colspan="2">$10</td><td colspan="9">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="11">$0</td><td>$0</td><td>$30</td><td colspan="7">$30</td></tr>
<tr><td>Opt. B</td><td colspan="11">$0</td><td colspan="2">$10</td><td colspan="7">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="12">$0</td><td>$0</td><td>$30</td><td colspan="6">$30</td></tr>
<tr><td>Opt. B</td><td colspan="12">$0</td><td colspan="2">$10</td><td colspan="6">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="13">$0</td><td>$0</td><td>$30</td><td colspan="5">$30</td></tr>
<tr><td>Opt. B</td><td colspan="13">$0</td><td colspan="2">$10</td><td colspan="5">$30</td></tr>
</table>

<table>
<tr><th>Ball</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr>
<tr><td>Opt. A</td><td colspan="14">$0</td><td>$0</td><td>$30</td><td colspan="4">$30</td></tr>
<tr><td>Opt. B</td><td colspan="14">$0</td><td colspan="2">$10</td><td colspan="4">$30</td></tr>
</table>

CALIBRATION OF NONLINEAR PROBABILITY TRANSFORMATIONS

It seems plausible that someone would prefer the three-payoff lottery to the two-payoff lottery in some or all of the pairs of options in Table 1

EUT implies that if an agent prefers the three-outcome lottery in any pair then the agent prefers the three-outcome lottery in all pairs

Suppose that an agent does prefer the three-outcome lottery in every pair of options and that her utility functional is <u>nonlinear</u> in probabilities

THEN

Proposition 1 in Cox and Sadiraj (2011) tells us that **she will also prefer a certain payoff of 3,000 to a 50/50 bet that pays 3 million or 0**

Logical Structure of the **Calibration of Nonlinear Probability Transformation Functions** is to Show that Statements P.2 & Q.2 are Inconsistent:

- P.2 is a seemingly plausible pattern of varying-probabilities, fixed-payoffs preferences over lotteries, as in Table 1
- Q.2 is a plausible preference over large-stakes lotteries as in:

  50/50 bet for 3 million or 0 is preferred to 3,000 for sure

Proposition 1 shows that, with nonlinearity in probs: P.2 $\rightarrow$ *not*Q.2

Of course, by logical equivalence we also have: Q.2 $\rightarrow$ *not*P.2

SO WHAT'S NEW HERE?

- What's new here? Didn't Rabin (2000) already tell us all about this kind of stuff?"

- Well actually: Rabin had nothing to say about the implications of calibration of nonlinear **probability transformations**.

- But the question raises some fundamental issues.

- Observation: Rabin and several subsequent authors considered the implications of calibration of concave **payoff transformations**

CALIBRATION OF NONLINEAR *PAYOFF* TRANSFORMATIONS

- Applies to theories w/ nonlinear transformation of money payoffs
- Early results by Hansson (1988)
- Sparked by Rabin (2000)
- Contributions by Neilson (2001), Barberis, et al. (2006), Cox & Sadiraj (2006), Rubinstein (2006), Safra & Segal (2008, 2009)
- All post-Hansson literature until now builds on the varying-payoffs calibration pattern made famous by Rabin (2000)
- Results hold for models defined on terminal wealth or income

Logical Structure of the **Calibration of Nonlinear Payoff Transformation Functions** is to Show that Statements P.3 & Q.3 are Inconsistent:

- P.3 is a seemingly plausible pattern of varying-payoffs, fixed-probabilities, small-stakes risk aversion
- Q.3 is a plausible preference over large-stakes lotteries

- Example of nonlinear payoff transformation calibration:
  - P.3 is the statement that an agent rejects a 50/50 bet with loss of 100 or gain of 110 $\vee w \in [100, 300K]$
  - Q.3 is the statement that an agent with initial wealth of 290K prefers a 50/50 bet with payoffs of 0 or 5 million to getting 10K for sure

Concavity calibration shows for EUTW: P.3 → *not* Q.3

By logical equivalence for EUTW one has: Q.3 → *not* P.3

Subsequent authors used variations on Rabin's varying-payoffs pattern to extend the demonstration of statement P & Q inconsistency to many other theories of decision under risk *that include nonlinear transformation of payoffs*.

The central question: **Is the calibration critique fundamental for evaluating decision theory?**

Addressing this central question **requires addressing questions left open** by previous calibration literature.

QUESTIONS LEFT OPEN BY PREVIOUS LITERATURE

- No implication for nonlinear transformation of probabilities, an alternative way to model risk aversion
- No implication for theories with variable reference points
- No implication for same-stakes domains
- No data that demonstrate empirical relevance
- No insight into what type of theory is immune to calibration critique (could not have without use of duality)

These questions are addressed in Cox & Sadiraj (2008, 2011) and Cox, Sadiraj, Vogt & Dasgupta (2010)

THE DUAL CALIBRATION ANALYSIS OF DECISION THEORY

I. Varying-probabilities, fixed-payoffs calibrations: as in the example in Table 1 above

II. Varying-payoffs, fixed-probabilities calibrations: as in the Rabin pattern used in all previous literature

DUAL PARADOXES FOR PATTERNS OF RISK AVERSION

- **Varying-probabilities**, fixed payoffs patterns (Cox & Sadiraj):
  - conform to expected utility theory
  - imply implausible risk aversion for theories with nonlinear transformation of probabilities

- **Varying-payoffs**, fixed probabilities patterns (Rabin & others):
  - conform to the *dual* theory of expected utility
  - imply implausible risk aversion for theories with nonlinear transformation of payoffs

REFERENCE POINTS

- With **fixed reference point**, theories that transform both probabilities and payoffs (e.g., CPT) are **subject to both** of the dual calibrations

- With **variable reference points**, theories that transform both probabilities and payoffs (e.g., third-generation CPT) are:
  - Immune to varying-payoffs, fixed-probabilities calibration (Wakker, 2005, 2010)
  - Vulnerable to varying-probabilities, fixed-payoffs calibration (Cox & Sadiraj, 2011)

MORE IMPLICATIONS

Theories with nonlinear probability transformations **cannot rationalize** even **same-stakes** risk preferences.

For example, Proposition 1 (Cox and Sadiraj, 2011) tells us that these statements are inconsistent for such theories:

P.1e The three outcome lottery that pays 14 or 4 or 0 is preferred to the two outcome lottery that pays 14 or 0 for all *p* in $\{0.1, 0.2, \ldots, 0.8, 0.9\}$

Q.1e The 50/50 lottery that pays 16.50 or 0 is preferred to a sure payoff of 0.50

EMPIRICAL RELEVANCE

Calibration propositions and the paradoxes they produce are very interesting. But definitive conclusions about theory require empirical support.

Support for what? The propositions are about Statement P and Statement Q inconsistency: P → *not* Q and Q → *not* P

We design experiments to ascertain whether statements P are consistent with subjects' choices because the theory implies *not* P

EXPERIMENTAL DESIGN ISSUES

Here is a testable pattern of varying-payoffs, fixed probabilities lottery preferences that can be calibrated:

P.3a Certain income x is (weakly) preferred to the 50/50 lottery with payoffs $x - \ell$ or $x + g$ for all $x \in [m, M]$

What if we were to use parameters such as ℓ = \$100 , g = \$110, m = \$1K and M = \$350K?

Affordability vs. Credibility with Varying-Payoff Experiments

Those parameters give the concavity calibration real “bite”. But the experiment would be much too expensive!

Suppose the subject always chooses the certain amount $\$x$ and that one of the subject’s decisions is randomly selected for payoff.

Then the expected payoff to a single subject would exceed $175,000.

With a sample size of 30 subjects, the expected payoff to subjects would exceed $5 million.

But why use payoffs denominated in U.S. dollars? After all, concavity calibration is dimension invariant.

Suppose, instead, we were to use dollars divided by 10,000; in that case the example experiment would cost only about $500.

The problem with this "solution" is that if the unit of measure is $1/10,000 then the binary lottery has trivial financial risk ($g-\ell$) of 2.1 cents.

We addressed this problem in two ways:

By running experiments in Calcutta where we could afford to pay significant amounts of rupees

By running experiments in Magdeburg with “contingent euros” using the Magdeburg Casino’s roulette wheels

Varying-Payoffs Exp. Calcutta 30/-20

Row	**Option A** (rupees)	**Option B** (rupees)
1	80 or 130	100
2	980 or 1,030	1,000
3	1,980 or 2,030	2,000
4	3,980 or 4,030	4,000
5	4,980 or 5,030	5,000
6	5,980 or 6,030	6,000

VARYING-PAYOFFS EXPERIMENTS

Calcutta 30/-20: binary lotteries $\{x+30;0.5;x-20\}$; x values from {100, 1K, 2K, 4K, 5K, 6K}; payoffs in rupees.

Calcutta 90/-50: binary lotteries $\{x+90;0.5;x-50\}$; x values from {50, 800, 1.7K, 2.7K, 3.8K, 5K}; payoffs in rupees

Magdeburg 110/100: binary lotteries $\{x+110;0.5;x-100\}$; x values from {3K, 9K, 50K, 70K, 90K, 110K}; payoffs in *contingent* euros

SIGNIFICANCE OF THE RUPEE PAYOFFS

We collected both income and price data. These data reveal that the 50 (= 30 - 20) amount at risk in the Calcutta 30/-20 experiment was the equivalent of:

- A full days pay
- 15 servings of poultry
- 1.5 – 3 moderate quality restaurant meals
- 14 bus tickets

The amount at risk in Calcutta 90/-50 was about 3 times as large.

Power vs. Credibility with Varying-Probability Experiments

The "power" of the calibration increases with the number of sub-intervals of the [0,1] probability interval.

Example: let the high payoff = 3 $\times$ the intermediate payoff then

- P.1 calibration for $p \in \{0.1, 0.2, \cdots, 0.8, 0.9\}$ implies a Q.1 statement that 1,000 for sure is preferred to 50/50 lottery for 33K or 0
- P.1 calibration for $p \in \{0.001, 0.002, \cdots, 0.998, 0.999\}$ implies a Q.1 statement that 1,000 for sure is preferred to 50/50 lottery for 10^{150} or 0

An experiment to test the empirical validity of the P.1 statement for

A. $p \in \{0.1, 0.2, \cdots, 0.8, 0.9\}$ would require the subject to make 9 choices

B. $p \in \{0.001, 0.002, \cdots, 0.998, 0.999\}$ would require the subject to make 999 choices AND "adjacent" choices would involve 0.001 differences in probabilities of the high and low payoffs

In order to have credibility, we use a small number of sub-intervals of [0,1] in our experiments.

Varying-Probabilities Exp. Atlanta 14/4

Row	Option A		Option B		
	Payoff 14	Payoff 0	Payoff 14	Payoff 4	Payoff 0
1	1/10	9/10	0/10	2/10	8/10
2	2/10	8/10	1/10	2/10	7/10
3	3/10	7/10	2/10	2/10	6/10
4	4/10	6/10	3/10	2/10	5/10
5	5/10	5/10	4/10	2/10	4/10
6	6/10	4/10	5/10	2/10	3/10
7	7/10	3/10	6/10	2/10	2/10
8	8/10	2/10	7/10	2/10	1/10
9	9/10	1/10	8/10	2/10	0/10

VARYING-PROBABILITIES EXPERIMENTS

Magdeburg 40/10: y = 40 euros, x = 10 euros

Atlanta 40/10: y = 40 dollars, x = 10 dollars

Atlanta 14/4: y = 14 dollars, x = 4 dollars

Calcutta 400/80: y = 400 rupees, x = 80 rupees

TEST RESULTS FOR VARYING-PROBABILITIES DATA

Calibration patterns are exhibited by these percentages of subjects (Wald 90% confidence intervals):

- Calcutta 400/80 experiment: 72% – 74%
- Atlanta 40/10 experiment: 56% – 63%
- Magdeburg 40/10 experiment: 37% – 41%
- Atlanta 14/4 experiment: 74% – 90%

TEST RESULTS FOR VARYING-PAYOFFS DATA

Calibration patterns are exhibited by these percentages of subjects (Wald 90% confidence intervals):

- Calcutta 90/-50 experiment: 80% – 82%
- Calcutta 30/-20 experiment: 36% – 48%
- Magdeburg 110/-100 experiment: 50% – 56%

CONCLUSIONS FROM CALIBRATION ANALYSIS

- The dual calibration critique is a fundamental theoretical criticism of decision theory
- There is significant support for the empirical relevance of the critique
- The two dual propositions and their corollaries tell us what types of models are **not** subject to calibration critique:
    - Linear in Probabilities and Variable Reference Point OR
    - Linear in Probabilities and Concave & Convex Segments (Markowitz, 1952)

CHECKING OUT SCYLLA

- Theorists trying to avoid the Scylla of the Allais Paradox have been sucked into the Charybdis of implausible risk aversion.

- Maybe it is time to revisit the question of whether Scylla is really so scary after all.

CONTENT FROM COX, SADIRAJ & SCHMIDT, 2010

Experiment 2

- We use a crossed design of:

  6 elicitation mechanisms X 5 lottery pairs

We Ask:

- What are the properties of the mechanisms?
- Which paradoxes can be observed with which mechanisms?

MECHANISMS

- We use 5 mechanisms for multiple decision treatments:
  - Pay One Randomly (POR) at the end (also called RLIM)
  - Pay All Sequentially (PAS) during the experiment
  - Pay All Independently (PAI) at the end
  - Pay All Correlated (PAC) at the end (a new mechanism)
  - PAC/n (alternative version of the new mechanism)
- We also use a One Task (OT), between-subjects design for a single choice treatment

POR is incentive compatible if the independence axiom holds. PAC and PAC/n are incentive compatible if the dual ind. axiom holds

PARADOXES

- We consider four paradoxes for theories of decision under risk:
  - Common Ratio Effect (CRE)
  - Common Consequence Effect (CCE)
  - Dual Common Ratio Effect (DCRE)
  - Dual Common Consequence Effect (DCCE)
- We use lotteries with larger differences between EVs than in typical tests for CRE and CCE

QUESTIONS ABOUT PARADOXES

We ask whether we observe:

- CRE with POR or OT
- CCE with POR or OT
- DCRE with PAC, PAC/n, or OT
- DCCE with PAC, PAC/n or OT

CONCLUSIONS ABOUT CRE AND CCE

- No significant CRE or CCE with OT data
- No significant CRE with POR data
- CCE with POR data not significant at 5%

WE CONCLUDE THAT SYLLLA MAY NOT BE A THREAT TO PLAUSIBILITY OF THEORY

- Empirical support for CRE and CCE ("Allais Paradox") is questionable: hence the need for nonlinearity in probabilities is *not* compelling

- St. Petersburg Paradox does not occur on bounded domain: do you feel a need to include "infinite" payoffs in the theory you apply?

CHARYBDIS DOES NOT EXIST FOR

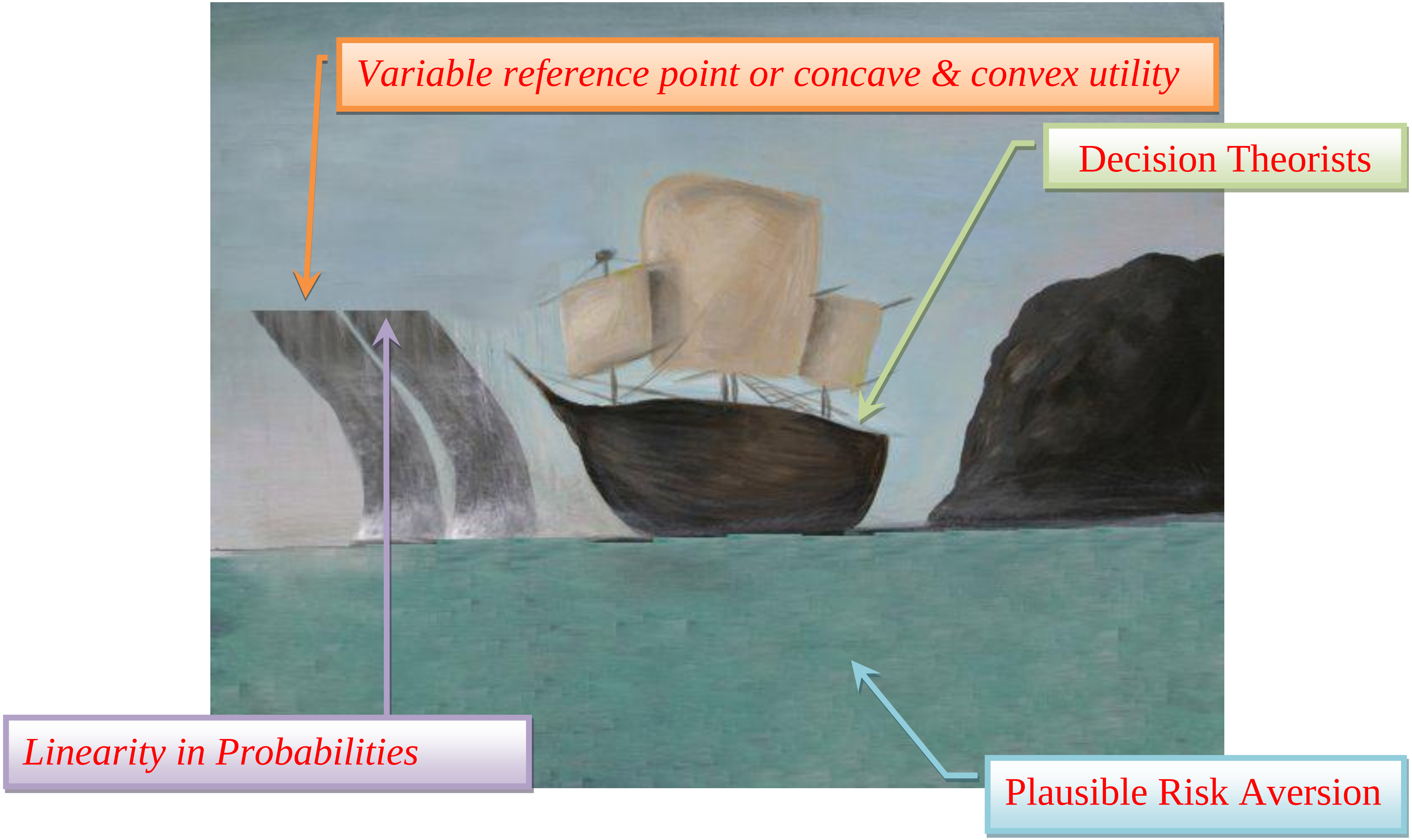
A theory with functional that is

- Linear in probabilities
- Has utility of payoff with variable reference point

The Markowitz (1952) model that is

- Linear in probabilities
- Has utility of payoffs with concave and convex segments

SMOOTH SAILING



Decision under Risk: Mechanisms, Paradoxes, and Dual Calibrations

James C. Cox

EXperimental Economics CENter
Andrew Young School at Georgia State University

Plenary Lecture: Economic Science Association, Tucson, November 2011