ESSAYS IN EXPERIMENTAL EXAMINATIONS OF

DECISION MAKING WITH UNCERTAINTY

By

JEFFREY CAMPBELL BIRCHBY

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Economics

Written under the direction of

Barry Sopher

And approved by

____

____

____

____

New Brunswick, New Jersey

May 2013

ABSTRACT OF THE DISSERTATION

ESSAYS IN EXPERIMENTAL EXAMINATIONS OF

DECISION MAKING WITH UNCERTAINTY

By

JEFFREY CAMPBELL BIRCHBY

Dissertation Director:

Barry Sopher

The dissertation examines how individuals behave when facing choices with uncertainty, both within individual and group settings. These questions are probed by examining individuals in experimental situations, with payoffs depending partially upon the choices they make as well as some fundamental uncertainty. In particular, the research seeks to answer, to what extent do individuals consistently evaluate risky gambles, and how do individuals respond to group decisions with either an underlying unknown state and a policy whose payoff depends upon that state, or a coordination task. For both group tasks, the role of communication in overcoming uncertainty is studied.

Within group settings, individuals were tasked with a Median game, a coordination problem with efficiency achieved only at an asymmetrical equilibrium, and an aggregation problem where groups would vote on a risky policy whose success depended upon an underlying state of the world that each member privately received signals about. For the Median game, communication results in large gains compared to the sessions with no communication; an effect that grew as play continued. The types of messages and strategies communicated are analyzed to see which were most effective.

Within a group setting, individuals were given private signals about the probability of payoff-relevant states of the world with payoffs determined by the policy voted for by the group's majority. There was a strict incentive to aggregate information, however this was hindered by the

inclusion of individual payoff biases, creating competing interests and raising a source for disinformation and thus distrust. The efficacy in achieving aggregation of various exogenously given networks was then tested.

Within an individual setting, subjects evaluate scaled lottery questions with consistent aggregate responses that fit well within the established norms of Cumulative Prospect Theory. However, the aggregate data hide considerable individual inconsistency, with the individuals that fit best with a CPT interpretation those that are most variable and for which such a prediction would provide limited if any predictive power.

ACKNOWLEDGEMENTS OF THE DISSERTATION

My thanks to Barry Sopher, for his guidance and advice, as well as to Oriol Carbonell-Nicolau, Gary Gigliotti, and Richard Lau for all of their help.

TABLE OF CONTENTS OF THE DISSERTATION

LIST OF TABLES OF THE DISSERTATION

List of Tables:

LIST OF TABLES OF THE DISSERTATION

List of Tables Continued

LIST OF ILLUSTRATIONS OF THE DISSERTATION

List of Illustrations:

Introduction

Experimental Economics provides a tool for assessing how individuals behave and make decisions. This is useful for when theory is inconclusive, either with no prediction or to distinguish between competing predictions. Following are three such experiments, examining how decision-makers respond when faced with uncertainty.

In the first chapter, the role of communication in achieving efficiency within a group coordination task is studied. Individuals have fundamental uncertainty over the beliefs actions and even payoffs of the other agents. However, it is clear to agents what the optimal group behavior would be. Within similar situations, so long as the "cost" of achieving the higher payoff is born equally, subjects have been able to use communication to coordinate on this equilibrium and achieve higher payoffs for all members. The novel variation introduced here is that the highest group payoff no longer requires unanimous sacrifices, and instead is achievable through costly action of just a majority. The game studied is thus more forgiving than standard variants, although that paradoxically results in worse group performances as self-interested free riding and equity concerns rise sufficiently to weigh down the average payoff.

In a political setting, the bounds of communication and trust are tested with a group voting experiment on whether to adopt a risky policy. Each group member receives a private signal, with communication allowing the subjects to aggregate information and increase their accuracy. The complication to this simple task setting arises with the introduction of a small bias – with either a higher or a lower payoff received by an individual if the group adopts the risky policy. Individuals thus have divergent and imperfectly coinciding interests, creating an incentive to dishonestly report signals as well as to distrust others' reports. Within the range of results theory might predict is an unraveling equilibrium with no communication trusted and no meaningful information communicated. Instead, four exogenously networks are found to have considerable trust and information aggregation – although with stark differences between them – suggesting that the means by which communication occurs are themselves an important input in the communication that occurs.

Finally, individual choices over uncertain payoff lotteries are examined. Competing theories for individual decisions towards risk and probability are well established. Including a zero prize has the effect of focusing behavior and improving consistency. While the aggregate results are found to be highly favorable to a Cumulative Prospect Theory interpretation, examining the individual components of these choices reveals a different story. Individual results are found more consistent with an error smoothing than a consistent inverted S-Shaped probability weighting function, and individual CPT coefficients are highly variable for all but expected payoff maximizers.

Throughout all chapters, the research endeavors to answer how individuals behave regardless of the theoretical predictions. The results are positive but limited for the role of communication in achieving coordination, and even more restrictive of the ability to ‘predict’ the behavior of variable subjects when confronted by risk. However, there are several positive results, in terms of the messages and appeals sent, the network connections of a communication structure, and in describing aggregate behavior and predicting the subject subpopulation that responds consistently to risk.

Chapter 1. Communication, Coordination and Fairness in a Median Game

Chapter 1 Abstract

I explore within a laboratory setting the extent to which communication can increase efficiency within a game where the coordination task cannot be separated from the allocation. Subjects played a median game in which earnings are a multiple of the group's median minus a cost associated with an individual's effort. This structure retains the coordination problem of Pareto ranked equilibrium from the standard minimum effort game, but efficiency is now asymmetric and requires a minority to exert the minimum effort. Communication results in large gains to both effort levels and group medians with the gain relative to a no communication benchmark growing across periods. The effect is most dramatic directly following communication, and dissipates with each repetition.

Subjects were found to be interested both in coordinating to the highest possible median as well as achieving an equitable distribution of gains. Within group communication they suggested unanimous high effort most commonly, but a robust minority advocated a series of rotating free riders. Individuals were most sensitive to receiving a free riding message, with a corresponding increase in their own effort. This effect was not substantial enough to overcompensate for the free riding, with medians negatively impacted by free riding strategies. Efforts were not particularly sensitive to the message type, with two notable exceptions: aggressive, threatening style messages were found to induce higher efforts and trust based messages corresponded with higher efforts across all iterations.

Chapter 1, Section 1 Introduction

Pre-play cheap talk communication has been shown to promote efficiency and enable coordination within a standard minimum effort game framework, as well as to boost generosity within dictator games and voluntary contribution mechanism public good settings. I extend this finding, asking to what extent communication can increase efficiency within a game where the coordination task is not separable from the allocation question. To explore this, experimental subjects played a median game in which the earnings of the subjects were a multiple of the group's median minus a cost associated with an individual's effort. The median game studied here is similar to the standard coordination game, and has sets of Nash equilibrium for each possible effort level that can be ranked by payoff dominance. However, the median game studied herein adds the wrinkle that all individuals exerting maximum effort is no longer stable or efficient. Instead, within the median game, free riding is allocatively efficient and a best response so long as it does not cross the threshold of lowering the group's median. Each equilibrium is composed of a majority of the group exerting the same effort level, and a minority free riding by exerting the minimum possible effort level.

To the best of my knowledge, this game structure is novel to this study, and differs strategically from the weak link game, a voluntary contribution mechanism public goods game, and earlier versions of a median game. Within the minimum game and earlier versions of a median game, payoffs were typically based off a multiple of the group's minimum minus a cost associated with the distance between an individual's effort and the group's minimum. The best response in a standard minimum game is thus to always match the group's minimum, with Nash Equilibrium characterized by sets of identical responses and tiered payoffs. The median game confounds this simple coordination task of achieving the higher group payoff variable with an allocational task of determining who, if anyone, should receive the lower costs associated with free riding. While the optimal response patterns vary significantly, the median game studied herein and the standard minimum effort game, each can be viewed as members of a more generalized class of games with payment based off the n-th lowest input in the group with the median game corresponding to a higher number n.

By reducing the draconian structure of the minimum game, in which the chain is only as strong as the weakest link, the median game can correspond with a variety of situations. One example of a situation which might be represented by the median game structure is a political setting in which politicians must vote on passing a version of an unpopular law that is necessary; examples might include spending cuts or tax increases to combat a growing budget deficit. Each politician individually prefers the strongest possible law be enacted so the problem is fixed as aggressively as possible, but at the same time would benefit from the ability to disclaim responsibility for the costly policy. Public utility may be the dominant concern and be increasing with the steepness with which taxes are raised or government services cut to close a budget gap, but each politician would also prefer if this majority needed to enact such a policy excluded them. Another setting where this incentive structure occurs is a group project or a multiple co-authorship production where, due to a redundancy of inputs, the group's output is not hindered by some limited free riding.

Multiple experiment treatments were run, both including a round of pre-play communication and a baseline of no communication. Communication was found to have a significant and positive effect on the average individual efforts and group medians, resulting in increased payoffs. The effect was most robust in the initial play of the game, with the communication gains relative to the benchmark declining across repetitions within fixed groups. The impact of communication increased over regroupings and new plays of the game. The no communication treatments found significant deterioration in initial effort choices, while the communication treatments had stable initial choices. Communication thus appeared to stabilize expectations and prevent learned pessimism, enabling hope to spring eternally even as its effect was continually eroded across repetitions of play within a fixed group.

Finally, the content and impact of the specific communication is examined. Since communication is found to induce higher effort levels, to what extent can this process be isolated and the communication that is most effective identified? In particular, I examine which arguments correlate with an increased effort when sent and when received. Messages almost always centered on the highest effort level, and were broadly allocated into several categories based upon the appeal used, ranging from explicit appeals for trust to (empty) threats and insults. In addition to explicit effort levels, groups often communicated about a coordinated strategy, with a unanimous high effort plan espoused most often but a robust minority of groups discussed a rotation of low effort individuals. A small minority of subjects pre-announced a strategy of minimum effort and full free riding in each play. Altogether subjects communicated in 93.5% of the times it was possible.

1.2 Experimental Design

The experiment was conducted during the Fall of 2010 at the Gregory Wachtler Experimental Economics Laboratory, part of the Center for Economic Behavior, Institutions and Design at Rutgers University, New Brunswick, NJ. Subjects were undergraduates at Rutgers University and signed up through online recruitment software. The experiment was programmed in Ztree (Fischbacher, 2007), and included 190 subjects across 10 treatment sessions. Each session was either assigned to a no communication treatment, or one with communication via a shared chat box prior to the start of each period. Within the communication treatment, a temporary alias was assigned each period of communication to ensure anonymity and prevent any reputation building effects.

Each period, subjects were grouped into groups of five, and following communication if applicable, they separately answered their effort choices. After all five effort choices were selected, the group's median and the subjects' individual payoff were revealed. Individuals were not presented with the full distribution of effort choices. Within a fixed group, subjects repeated the median game for six iterations, before regrouping and starting the next period. A total of ten periods were done in all but the first session, which ended after six periods.

Each period subjects were presented with one of two possible payoff tables, varying only in the cost of effort. In the first four periods, all individuals shared the same effort cost and payoff matrix, and this was presented as common knowledge. For the last six periods, individuals were assigned one of the two payoff matrices randomly, and were informed that each member of their group had similarly been assigned an effort cost and that effort costs might be dissimilar across group members. A full copy of the instructions can be found in Appendix 1. Final pay was based on cumulative results, and averaged in the range of $24 including a $5 show up fee. Earnings in the median game were parameterized by the formula:

$$Earnings_i = a + b*Median - c_i*Effort_i$$

This was applied with $a = 3$, $b = 1$, and $c_i = .1$ or $.5$, depending upon the period. The payoff tables are displayed in Figure 1 and Figure 2.

Figure 1: Low Effort Cost Payoff Matrix

		Median Value Chosen by Group						
		7	6	5	4	3	2	1
Your Effort Choice	7	9.30	8.30	7.30	6.30	5.30	4.30	3.30
	6	9.40	8.40	7.40	6.40	5.40	4.40	3.40
	5	9.50	8.50	7.50	6.50	5.50	4.50	3.50
	4	9.60	8.60	7.60	6.60	5.60	4.60	3.60
	3	9.70	8.70	7.70	6.70	5.70	4.70	3.70
	2	9.80	8.80	7.80	6.80	5.80	4.80	3.80
	1	9.90	8.90	7.90	6.90	5.90	4.90	3.90

Figure 2: High Effort Cost Payoff Matrix

		Median Value Chosen by Group						
		7	6	5	4	3	2	1
Your Effort Choice	7	6.50	5.50	4.50	3.50	2.50	1.50	0.50
	6	7.00	6.00	5.00	4.00	3.00	2.00	1.00
	5	7.50	6.50	5.50	4.50	3.50	2.50	1.50
	4	8.00	7.00	6.00	5.00	4.00	3.00	2.00
	3	8.50	7.50	6.50	5.50	4.50	3.50	2.50
	2	9.00	8.00	7.00	6.00	5.00	4.00	3.00
	1	9.50	8.50	7.50	6.50	5.50	4.50	3.50

1.3 Results

Communication was found to dramatically raise average effort levels, group medians and ultimately individual payoffs. The effect is most noticeable directly after communicating, with the gains from communication decaying across iterations. Despite that, in each new period communication was found to have a similar effect, with the gains resilient against any learned pessimism. Conversely, treatments without communication had steadily decreasing effort and median levels. As expected, effort cost is a significant factor in subject choices. Efforts were not particularly sensitive to the messages sent, although there were several exceptions to this.

Following is a roadmap of the results. Section 1.3.1 analyzes the impact of having communication possible, without analyzing the actual messages sent. The raw distribution of efforts, cumulatively across all treatments and isolated for the treatments with and without communication are displayed in Tables 1 – 3. Tables 4 – 6 replicate this breakdown for the group medians. After the raw data distributions, generalized least square regressions follow on the individual efforts and medians with Table 7 and Table 8 presenting these findings respectively.

The content of the messages sent is then incorporated into the results in Section 1.3.2, with Table 9 demonstrating the frequency of different message strategies and appeals. Table 10 and Table 11 repeat the GLS analysis on efforts and medians, incorporating the strategies and appeal types. This analysis is restricted to the first iteration following communication to prevent the confounding effect of other group members' actions and past realized medians. Table 12 presents a regression of subjects' messages on their own efforts. Finally Table 13 and Table 14 present the GLS results on efforts and medians including message strategy and appeal types, only after dropping the restriction of examining only the first iteration. These last results should be interpreted carefully, especially in regards to causation, but are informative of the types of communication that ultimately resulted in higher outputs. Section 1.4 presents areas for further research before Section 1.5 concludes.

1.3.1 Results: Raw Communication

The median game studied herein had the effect of focusing behavior into the extremes of either the highest or the lowest possible effort. Table 1 presents the distribution of individual efforts, demonstrating this distinctively bimodal pattern. Across all treatments and iterations, an effort level of 7 or of 1 each occurred individually with a frequency greater than all other efforts combined. This effect was particularly pronounced within the treatment with communication, as shown by Table 2, where 80.10% of all efforts were at one extreme or the other. In contrast, the no communication baseline shows a significant distribution of efforts at all intermediate levels, with a single modal mass at the lowest effort level.

The prevalence of subjects coordinating on the highest level of effort was thus one of the largest distinctions between the experimental sessions that included communication and those that did not. The different distributions of raw efforts depending on the presence of a pre-play communication round contain a pronounced distinction in terms of the frequency of highest possible efforts. The percentage of maximum efforts in the communication treatment was approximately 46%, in contrast to only 18.17% without the possibility to chat before playing the game (see Table 2 and Table 3).

The other salient pattern in the raw effort levels is a significant time series effect across iterations within the same fixed group. The time trend combines a decline in the percentage of high efforts, and a corresponding increase in the percentage of low efforts. The latter occurred similarly in both the communication and the no communication treatments. The percentage of low effort choices rose from 24.81% to 45.96% and from 26.38% to 40.75% from iteration 1 to iteration 6 with and without communication respectively. The percentage of high effort choices in the communication treatment dropped from 58.27% directly following communication to 33.56% in the sixth iteration suggesting that much of the gains from communication diminish rapidly. Without the communication boost to high effort levels in iteration 1, the no communication baseline has a relatively stable level of high efforts, with the percentage dropping only by two percentage points.

Table 1: Distribution of Effort by Iteration: All Response Data

Iteration	Individual Effort Level Frequency						
	7	6	5	4	3	2	1
1	42.01%	7.83%	8.32%	7.99%	4.46%	3.91%	25.49%
2	37.12%	7.55%	8.64%	7.07%	5.38%	5.38%	28.86%
3	33.97%	8.21%	7.88%	7.50%	5.16%	4.78%	32.50%
4	32.28%	7.99%	7.34%	7.23%	5.33%	5.60%	34.24%
5	30.76%	7.17%	7.07%	6.85%	5.11%	5.98%	37.07%
6	27.17%	5.98%	6.79%	6.63%	4.46%	5.27%	43.70%
Total	33.89%	7.45%	7.67%	7.21%	4.98%	5.15%	33.64%

Table 2: Distribution of Effort by Iteration: With Communication

Iteration	Individual Effort Level Frequency						
	7	6	5	4	3	2	1
1	58.27%	6.63%	3.94%	2.79%	1.44%	2.12%	24.81%
2	50.96%	5.77%	6.06%	2.98%	2.31%	2.88%	29.04%
3	47.21%	5.58%	6.15%	3.46%	2.02%	2.60%	32.98%
4	43.94%	5.96%	6.35%	3.94%	2.12%	2.40%	35.29%
5	41.92%	6.15%	5.67%	3.65%	2.69%	3.27%	36.63%
6	33.56%	5.00%	6.06%	3.17%	2.88%	3.37%	45.96%
Total	45.98%	5.85%	5.71%	3.33%	2.24%	2.77%	34.12%

Table 3: Distribution of Effort by Iteration: Without Communication

Iteration	Individual Effort Level Frequency						
	7	6	5	4	3	2	1
1	20.88%	9.38%	14.00%	14.75%	8.38%	6.25%	26.38%
2	19.13%	9.88%	12.00%	12.38%	9.38%	8.63%	28.63%
3	16.75%	11.63%	10.13%	12.75%	9.25%	7.63%	31.88%
4	17.13%	10.63%	8.63%	11.50%	9.50%	9.75%	32.88%
5	16.25%	8.50%	8.88%	11.00%	8.25%	9.50%	37.63%
6	18.88%	7.25%	7.75%	11.13%	6.50%	7.75%	40.75%
Total	18.17%	9.54%	10.23%	12.25%	8.54%	8.25%	33.02%

Group medians demonstrate similar patterns to the individual efforts. Table 4 presents the aggregate results, Table 5 the communication only treatment and Table 6 the no communication treatment. The data shows that there is a glaring juxtaposition between the with communication and the without communication benchmark. The percentage of groups achieving the highest possible median is 48.16% with chat versus 6.98% without. The data again shows a time series deterioration across iterations. Within the communication treatment this can be seen at the extremes; the highest effort declined from 66.35% to 26.92% of all group results while the lowest effort increased from 10.10% to 37.98% of all group results. The without communication medians were fairly uniformly dispersed, yet also show the same increased prevalence of the lowest possible group medians. The without communication results initially had a modal mass at the center of the effort ranges before a spike in the percentage of low effort medians in the last two iterations.

Comparing the raw distribution of efforts and medians shows that the efforts alone understate the gains from communication. The communication treatment had a higher occurrence of medians at the upper limit of 7 then efforts, suggesting successful coordination and less wasting of efforts. This is particularly important as it indicates that some of the instances of the low efforts are actually efficient free riding versus median reducing inefficient choices. In contrast, while the no communication treatment had over 18% of efforts at the highest level of 7, slightly less then 7% of all groups there achieved a median of 7. Thus, the instances of high individual effort were more likely to be indicative of wasted effort then emblematic of group coordination and high payoffs.

Table 4: Distribution of Median by Iteration: All Response Data

Iteration	Median Effort Level Frequency						
	7	6	5	4	3	2	1
1	40.76%	9.78%	12.77%	11.68%	7.88%	6.25%	10.87%
2	35.60%	8.15%	14.40%	11.96%	9.78%	8.15%	11.96%
3	31.52%	11.14%	10.87%	10.87%	9.51%	7.88%	18.21%
4	28.53%	10.60%	10.60%	13.04%	8.15%	9.24%	19.84%
5	26.36%	9.24%	11.14%	9.24%	9.24%	11.41%	23.37%
6	18.75%	9.78%	10.05%	10.60%	6.79%	9.24%	34.78%
Total	30.25%	9.78%	11.64%	11.23%	8.56%	8.70%	19.84%

Table 5: Distribution of Median by Iteration: With Communication

Iteration	Median Effort Level Frequency						
	7	6	5	4	3	2	1
1	66.35%	8.65%	6.25%	3.37%	2.40%	2.88%	10.10%
2	58.17%	7.21%	12.50%	4.81%	3.85%	3.85%	9.62%
3	50.00%	9.13%	9.13%	4.81%	3.37%	3.37%	20.19%
4	46.63%	9.62%	11.54%	5.77%	1.92%	2.40%	22.12%
5	40.87%	10.10%	9.13%	6.73%	3.37%	7.21%	22.60%
6	26.92%	9.13%	10.58%	6.25%	2.88%	6.25%	37.98%
Total	48.16%	8.97%	9.86%	5.29%	2.96%	4.33%	20.43%

Table 6: Distribution of Median by Iteration: Without Communication

Iteration	Median Effort Level Frequency						
	7	6	5	4	3	2	1
1	7.50%	11.25%	21.25%	22.50%	15.00%	10.63%	11.88%
2	6.25%	9.38%	16.88%	21.25%	17.50%	13.75%	15.00%
3	7.50%	13.75%	13.13%	18.75%	17.50%	13.75%	15.63%
4	5.00%	11.88%	9.38%	22.50%	16.25%	18.13%	16.88%
5	7.50%	8.13%	13.75%	12.50%	16.88%	16.88%	24.38%
6	8.13%	10.63%	9.38%	16.25%	11.88%	13.13%	30.63%
Total	6.98%	10.83%	13.96%	18.96%	15.83%	14.38%	19.06%

A panel data GLS regression was run on individual effort choices to identify the full magnitude of communication on individual behavior within the Median game. Explanatory variables included the period of the experiment (the round of regrouping, ranging from 1 to 10), the iteration of the experiment (the round of repetition within a fixed group, ranging from 1 to 6), the effort cost which took a binary value of either .1 or .5, a dummy for the presence of an opportunity to communicate, and lastly cross intercept dummies between communication and period, iteration and effort cost. This generalized structure enabled fitting in effect two separate regressions, one for the treatments with communication and one for the treatments without communication by not imposing any structural continuity between the two. Table 7 presents the results of the Generalized Least Squares regression on effort, which was derived from 11,040 individual observations from 190 subjects.

Communication was found to result in an increased effort by .84, nearly a single degree out of the seven available higher and statistically significant at the 1% level. As displayed in Tables 1 – 6, effort choices declined across iterations. The GLS regression attributed a coefficient of -.126 per iteration to effort. For the treatments with communication, this effect was further exaggerated, with an additional coefficient of -.126. Both coefficients were significant at the 1% level. Period also was found to have a negative impact on effort choices, with a raw coefficient of -.102. Communication however was found to have an opposite effect when combined with period, with a dummy slope intercept fitted at .099. Testing that the combined coefficient for the communication treatment is equal to zero returns a Chi Squared Statistic of .10, so the null hypothesis that the period is insignificant in the communication treatment cannot be rejected.

Table 7: Generalized Least Squares Regression on Effort

Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.255	.177	29.73	0.000
Period	-.102	.011	-9.56	0.000
Iteration	-.126	.018	-7.05	0.000
Communication	.840	.233	3.60	0.000
Effort Cost	-2.157	.156	-13.82	0.000
Period w/ Chat	.099	.014	6.89	0.000
Iteration w/Chat	-.126	.024	-5.31	0.000
Cost w/Chat	-.386	.209	-1.85	0.064

n = 11,040 observations on 190 individuals, R^2 = .0856, X^2 = 973.45

While subjects without the opportunity to discuss the experiment slowly became more pessimistic in their actions, subjects in the communication treatments approached each new grouping as optimistic as the last. This optimism reoccurred despite the fact that efforts appeared to quickly disintegrate within each grouping. Finally, effort cost was found significant at the 1% level, with a negative coefficient that corresponds to a .86 difference in average efforts be- tween payoff grids. A communication and effort cost dummy was significant at the 10% level but not at the 5% level, with a marginal reduction in effort when moving from the low to the high cost payoff matrix of .15.

Conducting the same analysis on group medians reveals the same patterns, with an increase in the magnitude of these coefficients. Medians are decreasing in period and iteration by coefficients of -.155 and -.137 respectively, both significant at the 1% level. Communication increases medians by a coefficient of 1.281, a larger effect found than found on individual efforts. Once again this result declines across iterations, here by the steeper -.238 per iteration. Finally, the dummy slope intercept on communication and period again is large enough to offset the no communication decline across periods. The combined coefficient on the period variable for the communication sessions was statistically insignificant from 0, with a Chi Squared (1) Statistic of 0.57.

Table 8: Generalized Least Squares Regression on Median				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	4.88	.275	17.75	0.000
Period	-.102	.011	-3.82	0.000
Iteration	-.126	.018	-4.36	0.000
Communication	1.281	.363	3.53	0.000
Period w/ Chat	.182	.054	3.36	0.001
Iteration w/Chat	-.237	.042	-5.69	0.000

n = 2208 observations on 368 groups, $R^2 = .1616$, $X^2 = 305.90$

The main results found in analyzing the impact of a pre-play round of communication are consistently higher effort and medians directly following communication. These efforts and medians then experience a quick decay across iterations, losing significant after four or five repetitions. However, with each new grouping and pre-play communication round, efforts began again flat with where they had started the previous round. In contrast, the no communication treatments demonstrated carry over effects and thus lowered initial efforts with each subsequent grouping. The inter-temporal stability of first iteration efforts and medians within the communication treatment, despite the seemingly inevitable decline in group medians, suggests that the communication in effect prevented the learning and adapting to this negative pattern. These results were self-fulfilling, with communication subjects rewarded with higher medians and payoffs.

1.3.2 Results: Distribution and Content of Messages

Since the raw possibility of communicating resulted in a dramatic increase in both individual efforts and the group medians, the question naturally follows to what extent can this effect be isolated to identify the individual messages that boost payoffs. Examining the chat transcripts reveals that the individual comments typically fell into one of the broad categories of: numerical messages, group strategy suggestions and one of several different subcategories of appeals for higher efforts. The messages sent by each individual in each period were coded as to whether they fell into any of these categories, with messages often falling into several of them. In addition, the messages individuals received as well as the number of group members they received them from were also identified.

The numerical messages matched the effort range from 7 to 1, with the vast majority for a full effort of 7. Strategies included unanimous full effort by the group, an intricate series of rotating free riders, and occasionally a personal free riding strategy in which the other members of the group would be told the median depended solely on the remaining members since that subject would be choosing the minimum each time. Lastly, participants often tried to talk other members into higher efforts, with the major categories of these appeals identified as threat based, fairness based, coordination based, risk based and trust based. These appeal based categories were not mutually exclusive, and a subject could be coded as using one or more of the message types as applicable. The full distribution of messages sent and their frequency within the communication treatment is displayed in Table 9.

The most prevalent message consisted of indicating the highest level of effort at 7. A full 65.58% of all possible individual communication periods included this. In contrast, the next most prevalent numerical messages were 1 and 6 with frequencies of 6.54% and 4.23% respectively. Indeed, the prevalence of messages of 7 and the near absence of lower effort messages was so common as to make the numerical message sent entirely uninformative within subsequent regressions upon the efforts that individuals took after communication. Strategies most frequently were some version of a plea for unanimous high efforts. However, a robust minority of groups across different sessions had members recommending an equal distribution of free riding gains by rotating turns of low effort. Lastly, a handful of individuals declared themselves permanent free riders, with a strategy to exert low effort each iteration and depend upon the rest of the group to coordinate a higher median.

Examples of messages coded as a unanimous strategy include: "all 7s," "everyone do all 7," and "okay since obviously we cannot collaborate, can we all just do 7?" The rotation strategy included both calls for one and two low effort members, and often detailed the order of low effort based upon the visible aliases identifying group members. The range of messages spanned from "we can alternate who hits 1," to "hey since the median counts, do you guys want to take turns, putting in 1? like D puts in 1 first, n then), then S, then Q, rthen P." Examples of the type of strategy messages that were coded as free riding include: "im always going to go 1, i just told you what i am doing, plan accordingly," and "i always pick 1, so if we get 1, you will know that i am one of them."

Subjects often pleaded to other subjects to adopt their preferred strategy and to pick a high effort level, with the appeals commonly falling into one (or more) of several categories. The most common type of appeal was a coordination based argument, in which subjects argued that the gains from efficiently free riding would be swamped by the increased chance of the group failing to achieve a high median. The typical coordination message was something similar to "everyone thinks that they will be the one to switch to 1 it does not work." The next most common category of appeals was threats, which included cursing and hostile language. Just under 10% of all communication was of this type, with examples such as "yeah all 7's...no one be an a hole and switch it up," which includes the unanimous strategy recommendation, to the blunt "if u press 1 u f*** it up and we all make no money," to "and YOU the one that is going to screw someone over by putting 1 DONT DO IT."

The last message focuses on the effect of a low effort choice as "screwing" the other group members, which was one of the more common word choices. These messages appeared to have a dual connotation, both one of the threatening and hostile nature, and also that there is unfairness in the choice, and therefore were coded as belonging to both categories. The fairness category also included messages directly referencing fairness, such as "and then it's all fair cuz we all make the same." Similarly, trust based appeals were those in which subjects said some variant of trust me, or "we have to trust each other." Lastly, risk based appeals were just under 5% of all communication and typically featured a contrast between the gains from reducing one's effort with the loss if the median fell; "everyone tries to save like .50 cents and we all lose like 5 dollars," or simply "is it worth it to risk that for 0.6?"

Table 9: Communication and Messages Sent

Communication Category	Frequency
Message of 7	65.58%
Message of 6	4.23%
Message of 5	1.83%
Message of 4	1.15%
Message of 3	0.00%
Message of 2	0.48%
Message of 1	6.54%
Unanimous Strategy	29.60%
Rotation Strategy	5.87%
Personal Free Rider Strategy	1.83%
Threat Based Appeal	9.52%
Fairness Based Appeal	8.46%
Coordination Based Appeal	10.19%
Risk Based Appeal	4.62%
Trust Based Appeal	6.63%
Silent	6.54%

The messages received by an individual before the first iteration can be thought of as an input into the effort choice, and are logical variables to include within a regression analysis. Table 10 presents just this, with GLS results of including the strategy messages received on first iteration effort. In addition, Table 11 includes the same GLS analysis run on the group's median in the first iteration following communication. While individual efforts are the direct factors in determining the median, to the extent that the messages sent might be suggestive of these actions this analysis can be informative. However, it is important to note that the median regression results should not be interpreted as causal coefficients. Even with the restrictions of a first iteration following communication, there was still a significant database of 1040 observations on 110 subjects.

Table 10: GLS Regression on 1st Iteration Effort Including Strategies Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.629	.220	25.58	0.000
Period	-.024	.026	-0.92	0.358
Effort Cost	-1.829	.340	-5.38	0.000
Unanimous Strategy	.068	.073	0.93	0.351
Rotation Strategy	.057	.125	0.46	0.648
Free Rider Strategy	.535	.269	1.99	0.046

n = 1040 observations on 110 individuals, $R^2 = .0221$, $X^2 = 35.76$

Table 11: GLS Regression on 1st Iteration Median Including Strategies Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.915	.220	25.58	0.000
Period	.055	.022	2.49	0.013
Effort Cost	-1.720	.308	-5.58	0.000
Unanimous Strategy	.386	.140	2.76	0.006
Rotation Strategy	.607	.264	2.30	0.022
Free Rider Strategy	.028	.457	0.06	0.952

n = 1040 observations on 110 individuals, $R^2 = .0474$, $X^2 = 51.50$

Messages for both a unanimous and a rotation strategy were found to have a statistically insignificant impact on the effort chosen. When confronted with a free rider, subjects did not retaliate with low efforts as well but instead attempted to overcompensate. This response was of a .5 higher effect and was found significant at the 5% level. Medians tell a different story. Here we see that both unanimous and rotation strategy messages are significantly associated with higher medians. Discussion about the group’s strategy is a good sign for the upcoming median even if it did not lead to a statistically significantly improvement in individual efforts. The one exception to this is the free rider strategy, where even though the other group members attempted to over compensate, the median that failed to statistically significantly increase.

Table 12: GLS Regression on 1st Iteration Effort Including Appeals Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.675	.218	26.09	0.000
Period	-.034	.027	-1.28	0.200
Effort Cost	-1.830	.337	-5.43	0.000
Threat	.231	.119	1.95	0.052
Fairness	.125	.123	1.01	0.310
Coordination	-.064	.111	-0.58	0.561
Risk	.166	.164	1.02	0.309
Trust	.038	.116	0.32	0.745

n = 1040 observations on 110 individuals, R^2 = .0249, X^2 = 38.05

Table 13: GLS Regression on 1st Iteration Median Including Appeals Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.954	.154	38.64	0.000
Period	.019	.024	0.78	0.436
Effort Cost	-1.798	.305	-5.89	0.000
Threat	.115	.106	1.08	0.279
Fairness	.271	.108	2.51	0.012
Coordination	.264	.095	2.77	0.006
Risk	.311	.144	2.16	0.031
Trust	.116	.097	1.19	0.234

n = 1040 observations on 110 individuals, R^2 = .0613, X^2 = 67.36

The type of appeals received is included as an explanatory variable in Table 12 and Table 13. The only type of appeal that was found to have significance on the first effort following communication was threats. Threats, which included hostile and derogatory language, were found to increase effort by .23, with a significance level of 5.2%. All other appeal types were not significantly different from zero. Table 13 includes the same analysis on the first iteration median for each individual, with the same caveat holding again that these results should be interpreted carefully. Medians are found to be positively associated with fairness, coordination, and risk based appeals. These variables each had a coefficient between .26 and .31 and were significant at the 5% level. Threats and trust based arguments were positive but not statistically significant.

Since medians show significance for receiving strategies and appeals that are not statistically significant on effort levels, to what extent can this be reconciled? It is possible that the medians are capturing small nudges in the group behavior that the efforts are not sensitive enough to identify. Similarly, they could instead be capturing the effects of an increase in coordination, in which net efforts are constant but efficiency and medians increase. One alternative possibility is that the cheap talk is informative of the actions that the sending agents will take, and the regressions on the medians are just identifying the communication that proxies for higher efforts from the senders. To address this alternative, the messages individuals send was regressed on the subjects' own effort choices. The results are presented in Table 14.

Table 14: GLS Regression on 1st Iteration Effort Including Own Messages Sent				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.626	.214	26.23	0.000
Period	-.039	.025	-1.59	0.111
Effort Cost	-1.766	.334	-5.28	0.000
Unanimous	.645	.177	3.65	0.000
Rotation	.916	.319	2.87	0.004
Free Rider	-3.083	.602	-5.12	0.000
Threat	.097	.246	0.39	0.694
Fairness	.050	.262	0.19	0.849
Coordination	-.166	.235	-0.71	0.480
Risk	-.055	.356	-0.15	0.877
Trust	.246	.291	0.85	0.398

n = 1040 observations on 110 individuals, $R^2 = .0724$, $X^2 = 81.48$

Unsurprisingly, individuals who claimed they would free ride did so, with the regression fitting an effort drop of 3 for these individuals. While no appeals were statistically significant, both a rotation and a unanimous strategy were correlated with higher efforts. Again, these results should not be interpreted in the standard causal fashion with the messages a subject sends causing them to pick an effort, but instead as a measure of the association between the two actions. With that disclaimer said, this result supports the hypothesis that the increased medians in Table 11 were coming from the effect of high effort individual senders, and not necessarily a product of subjects receiving the strategy messages themselves. However, there is no evidence to support this interpretation for fairness, coordination and risk based appeals.

The raw data painted a compelling picture of communication gains, but with the gains quickly diminishing across iterations. Unfortunately, any analysis across iterations includes a significant confounding effect of the path of the game play. Still, repeating the analysis of the strategy and appeal messages without restricting the domain to the initial iteration following communication provides insight into what type of effect persists and which do not. The full results, taken from 6240 observations on 110 individuals, are presented in Tables 15 – 18.

Table 15: GLS Regression on All Iterations of Effort Including Strategies Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.169	.157	32.88	0.000
Period	-.014	.011	-1.19	0.234
Effort Cost	-2.543	.150	-16.97	0.000
Unanimous Strategy	.065	.033	2.00	0.046
Rotation Strategy	.044	.056	0.78	0.433
Free Rider Strategy	.166	.120	1.38	0.169

n = 6240 observations on 110 individuals, R^2 = .0368, X^2 = 303.50

Table 16: GLS Regression on All Iterations of Median Including Strategies Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.429	.088	61.40	0.000
Period	.000	.012	-.02	0.987
Effort Cost	-2.215	.152	-14.61	0.000
Unanimous Strategy	.208	.032	6.48	0.000
Rotation Strategy	.050	.055	0.91	0.363
Free Rider Strategy	-.667	.118	-5.64	0.000

n = 6240 observations on 110 individuals, R^2 = .0482, X^2 = 296.42

While both a unanimous strategy and a rotation strategy was tied to higher medians when restricting analysis to the first iteration, only the unanimous strategy retains its significance across all iterations. The rotation strategy, while more efficient than everyone exerting the highest effort, is inherently unstable. Across iterations, the instability was significant enough to deteriorate the coordination and make the gains disappear. Similarly, groups initially could almost compensate for a free rider, but as the game was repeated the impact of a free rider increased dramatically. When looking across all iterations, receiving a free rider message did not cause individuals to have effort levels significantly different then zero, while medians declined with a coefficient of -.67 that was significant at the 1% level.

Table 17: GLS Regression on All Iterations of Effort Including Appeals Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.223	.157	33.33	0.000
Period	-.026	.012	-2.20	0.028
Effort Cost	-2.528	.148	-17.03	0.000
Threat	.110	.053	2.10	0.036
Fairness	.036	.054	0.67	0.504
Coordination	.001	.049	0.02	0.980
Risk	-.037	.073	-0.51	0.610
Trust	.220	.052	4.25	0.000

n = 6240 observations on 110 individuals, $R^2 = .0409$, $X^2 = 328.58$

Table 18: GLS Regression on All Iterations of Median Including Appeals Received				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Constant	5.508	.086	64.30	0.000
Period	-.011	.012	-0.93	0.353
Effort Cost	-2.178	.150	-14.48	0.000
Threat	.044	.053	0.84	0.404
Fairness	.095	.054	1.75	0.080
Coordination	.003	.049	0.07	0.945
Risk	.069	.073	0.95	0.341
Trust	.433	.051	8.51	0.000

n = 6240 observations on 110 individuals, $R^2 = .0505$, $X^2 = 306.03$

When the analysis is expanded to include all iterations following communication, trust based appeals emerge as highly influential. Trust based appeals are found to have a .22 coefficient on individual efforts and a .43 coefficient on medians, both significant at the 1% level. Threats are still significant at the individual effort level, but this does not correspond to group success and the coefficient on medians is insignificant. Fairness is significant at the 10% but not the 5% level for medians. Coordination and risk based arguments that were initially significant have lost their significance as the game play progressed.

1.4 Areas for Further Research

The impact of cheap talk communication within a broader context where equity and efficiency are both complementary and competing considerations suggests a wide range of parallel pursuits. Of particular note is that within the current research project all communication was done within an open complete network, thus making all communication public and inherently on an equal level. In contrast, communication often occurs within incomplete or hierarchical networks and the impact of communication and particular message types across network structures may not be universal. As an example, a network with one central communicator who communicates individually with all other group members might lessen the impact of arguments for fairness, while the structure adds salience to trust based appeals.

In addition to the communication structure, the feedback loop from the experiment itself poses a significant avenue of future research. Within the current experiment, subjects were reported their effort choices, as well as the group's median, and finally their payoffs. While these summary statistics provided a meaningful description of the level of overall payoffs achieved by the group, they do not provide the exact distribution of efforts. As a demonstration, consider an individual who chooses an effort level 7, and is returned that the group's median was 5. That subject will be aware that the group's median was less then their effort, and thus any effort above the median was wasted. But the same individual is unaware if the true distribution of actions by the other four agents was (5, 5, 5, 5) or (7, 5, 1, 1), with the first allowing free riding down to 1 while the second requires an effort of at least 5 to maintain the median.

Individual efforts thus would not be expected to converge to a stable outcome, as some level of experimentation could be beneficial. Indeed, this is what was found in the data from iteration to iteration, with medians migrating both down and up to subject experimentation. In contrast, the standard minimum game has efforts converging to the group's minimum and thus minimums are always found to be flat or decreasing across same group repetitions. Introducing additional feedback about the full range of efforts is conjectured to have the effect of eliminating this systematic experimentation, although the welfare implication of this is unclear. Both this and the role of the network structure are examined in a following work.

Finally, to the extent that this research has identified the type of appeals and strategies directly associated with higher effort levels by receiving members, there is an open question as to if this effect can be harnessed. Instead of relying on endogenous appeals that may or may not occur during communication, can subjects be equally effectively primed with an auto-generated message? To the extent that it can, it would be possible to achieve the gains from communication without the very communication it is based upon.

1.5 Conclusion

The median game provided a dynamic setting for testing the impact of communication within a group situation. Treatments without any communication demonstrated a significant decline in efforts and medians across periods. One of the big gains from communication was that it provided a stabilizing effect on the expectations of the subjects, with each period approached with a similar level of optimism and initially high efforts. This effect itself is somewhat surprising, because while communication provided an initial surge in group medians, the effect dissipated the further it was from the communication. The ability to discuss the game with the new group members thus appears to have enabled subjects to convince themselves that this time it would be different, even as it would quickly become not so.

Communication allowed the groups to all identify the higher median as the desirable result, but it also opened the question of whether every member needed to exert that effort, and if not who did. Strategies with everyone exerting high efforts and those with rotating free riders each worked initially, although the more fragile rotation strategy did not succeed across iterations. Similarly, the appeals that focused on fairness and trust were the only ones that had a lasting impact on the medians achieved by the groups. Returning to the political example introduced as a corollary for the median game earlier, these results would correspond with a prediction that, within a group, controversial policies are unlikely to be strategically voted against but rather bought into entirely or not at all. While gains exist that could be captured and evenly distributed by allowing a rotating minority to vote against an unpopular bill, any regime attempting to implement this strategy would be unstable enough to maintain these gains for long.

The efficacy of fairness and trust based appeals, and even the emergence of a complex strategy aimed at equally distributing free rider gains, highlight the importance of these considerations. In contrast, language based upon risk and coordination was found to have no long-term effect. This matches findings in other studies, and suggests the need to incorporate the concept of fairness into economic decision-making and any model explaining human behavior.

Chapter 2: Group Decision Making and Information Aggregation across Networks

Chapter 2 Abstract

In this chapter I study small group decision making in an environment where decision makers receive independent and private signals about the probability of payoff-relevant states of the world, and have a payoff that is contingent upon the voting behavior of a group they are included in. Subjects are assigned randomly to groups of five decision makers, each of which has an incentive to aggregate information and ensure that the payoff maximizing policy is chosen by the group. In later rounds subjects are given individual payoff biases favoring or disfavoring one of the policies, thus creating competing interests and raising a source for disinformation. Subjects can communicate over given network architectures including connected and unconnected networks with free-form messages in computer chat boxes.

Findings include that connected networks are best at promoting full information aggregation, and also lead more often to the socially best policy being adopted. In rounds without payoff biases, communication is generally truthful (in reporting own signals), while in rounds with payoff biases the truthfulness of communication is about 80% overall. Without biases a Full network in which all agents are connected performs best. With biases, a more hierarchical "Star" network performs better then a network with similar number of connections as well as the Full network. Insight into this appears to come from the level of truthful communication across networks – with the Star network's points significantly more honest – and the extent to which the information received by other group members is aggregated into a subject's beliefs.

Chapter 2, Section 1 Introduction

Humans often make decisions in groups, and in the process of making decisions, people will often have to depend upon or make inferences from information that is provided by others. Interesting decision making problems will typically involve some kind of uncertainty, either about the environment in which one is operating or about the preferences of the other individuals involved in the decision making process. While the economic theory of decision making under uncertainty is well-developed, from standard expected utility theory to its many variants, there is no such well developed theory of decision making in groups. In particular, 'simple' group problems can combine communication and network structure sensitivity, information aggregation, coordination and trust within a private signal environment creating complex games. While one can sometimes, in principle, quantify the information available in a given situation, to the extent that information is coming from disparate and, in general, unreliable sources, it is hard to know where to begin to size up the problem to be solved and analyze the resulting behavior.

Communication can result in in a broad range of efficiency in coordination games, particularly when it is common knowledge and visible, see for instance Schotter and Sopher (2006) and Chaudhuri, Schotter and Sopher (2005). But within a general framework, public information can be not just a blessing but also a curse to a group. Increased information can be a negative in instances such as beauty contest style games (see for instance Morris and Shin 2002), although this result itself is sensitive to the parameters, and it can be utility maximizing as well (Angeletos and Pavan 2007). However, instead of public information automatically being provided, we instead are examining information provision within a context that allows selective and untruthful dissemination. The essential problem that we are thus interested in is how and to what extent can or do people trust the information coming from someone else, when the other individual is someone with imperfectly coinciding interests, and to what extent the fundamental patterns of permissible communication affect this.

In order to focus on this aspect of the problem, I study simple situations in which a group shares equally in the total payoff resulting from the group's decision, and in which the group makes a final decision via majority rule—not because I believe majority rule is optimal in any sense, but in order to have a chance of isolating those aspects of the process that I am interested in. I study small group decision making in an environment where decision makers receive independent and private signals about the probability of payoff-relevant states of the world. Two policies, one risky and one safe, are available for adoption. Decision makers first communicate via exogenously given network architectures and then vote to determine which policy to adopt (by majority rule). The group payoff earned is divided equally among the members of the group.

The main goal is to investigate the effectiveness of several network architectures in facilitating full aggregation of information, and to investigate the impact that the introduction of biases (via individual payoff adjustments for one of the available policies) has on the truthfulness of communication and on the "correctness" of the decisions arrived at by groups. Thus, subjects are given the communication networks that people (exogenously) are given to work within a chance to explain variations in the effectiveness of the decision making process, in terms of the ability of the group to reach an objectively better decision. Communication was free-form and unordered, allowing cascading and herding behavior to occur if subjects report projected intentions to vote for or against (Banerjee, 1992 and Bikhchandani, Hirschleifer and Welch 1992) – and allowing full Baysian belief updating if subjects conveyed and believe each other similar to state identification group games such as in Choi, Gale and Kariv (2004) Goyal (2005).

Subjects are assigned randomly to groups of five decision makers in each of 12 rounds of play. In each round one decision making task, as described above, is completed. In each subsequent round subject are randomly reassigned to new groups. Further, subjects receive a new "alias" ID in each round, so there is no possibility of establishing a reputation over rounds of play. In round 1 to 4 there are no payoff biases, but in rounds 5 to 12 there are biases. Biases change from round to round. Network architectures change from round to round. We study 3 connected networks: “Full,” in which each decision maker can communicate with every other decision maker in a group; “Neighbors,” in which decision makers are connected to the person on either side of them in a ring; and “Star,” in which only one decision maker is directly connected to all other in the group, and the rest are directly connected to the center. We also study 1 unconnected network, in which subgroups of 3 and 2 subjects in a group are connected to one other within but not between subgroups. Communication is free-form, via computer chat boxes. Subjects know the payoff biases of other subjects to whom they are connected.

Briefly, I find that the Full network is best at promoting full information aggregation, while Neighbors and Star are close seconds. The “3-2” network does worst. The same ordering of network architectures organizes the data on whether the socially best policy is adopted. In rounds without payoff biases communication is generally truthful (in reporting own signals), while in rounds with payoff biases the truthfulness of communication is about 80% overall. Misrepresentation of signal information occurs mainly in cases where a subject has an incentive to do so; i.e., where a subject would gain from the policy being adopted, even if adoption of the policy would not be socially efficient. Groups manage to adopt the socially efficient outcome remarkably often, in spite of substantial individual biases that frequently provide a majority of the subjects in a group the incentive to deviate from the socially efficient policy and adopt the individually optimal policy.

The rest of the chapter is organized as follows. Section 3.2 contains a discussion of the experimental design and hypotheses stemming from game theoretic, network theoretic, and behavioral considerations. Section 2.3 contains data analysis, including discussion of how they hypotheses perform in the experiment. Section 2.4 concludes.

2.2 Experimental Design and Hypotheses

I am interested in the basic game illustrated in Figure 3. A group needs choose between adopting a risky policy which yields a payoff to each player of 100 in one state of the world, occurring with probability P*, or a payoff of 0 in the other state of the world, occurring with probability 1-P*, or not adopting the policy and taking a sure payoff of 50. P* is uncertain, but each member of the group receives a signal about the true value of P*. Before voting on whether to adopt the policy, members of the group have the opportunity, in general, to communicate with each other, presumably about what the true value of P* is, and what their respective signals might indicate about that question. In terms of preferences, there is only a question of differences in risk attitude to contend with here, as far as making statements about what is optimal for the group. To the extent that all individuals are approximately risk neutral, choosing to adopt the risky policy if P* is thought to be at least .5 should be optimal, since all will share equally in the proceeds.

	P*	1-P*
Adopt Policy	100	0
Do not adopt policy	50	50

Figure 3: Payoff Table for the Basic Group Decision Game

If a group could easily collect everyone's signal and calculate the average, this would be a good estimate of P*, and would serve as the basis for a sensible decision making strategy. In order to explore different "organizational forms," a variety of network communication architectures were imposed to compare the relative effectiveness in reaching optimal decisions. The four architectures used are displayed in Figure 4.

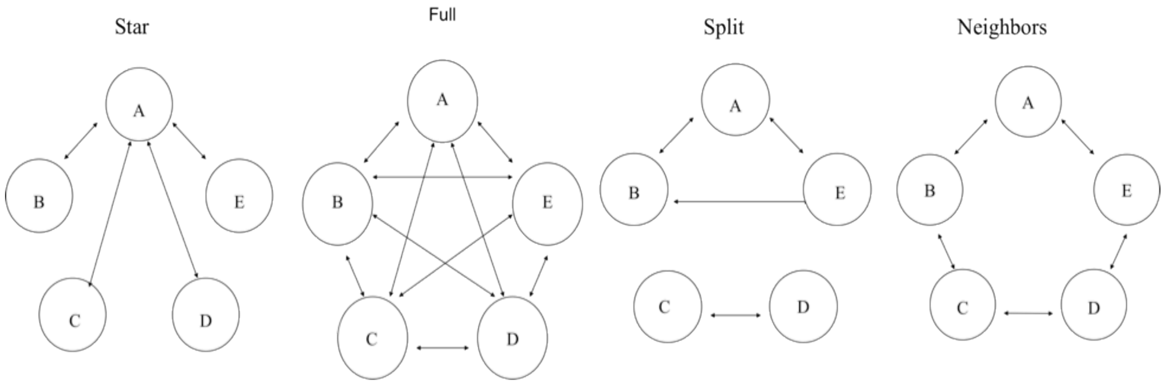


Figure 4: Network Architectures

The Full Network allows every member of the group to communicate with every other member of the group. The Star Network is centralized or hierarchical, evidently giving the one central member the potential for more influence, though in principle the same information that is conveyed in the Full network could be conveyed in the Star (both are connected networks, so everyone can reach everyone else, if not directly then through another person). The Neighbors (also known as Ring) Network is decentralized (but also connected). The Split Network is the only unconnected network. In terms of information aggregation relevant to the basic game in Figure 3, all of the connected networks should, in principle, allow for full aggregation, while the Split network would appear to be at a disadvantage in this respect.

A more challenging problem is to overcome biases that might be present that could lead members of a group to either misrepresent their information or to vote in a way that work against the common good for selfish reasons. By giving an additive bias to the payoff for each player, a basic and stark kind of bias in the decision-making problem was implemented. An individual bias, α (which can be negative or positive), changes, for the individual, the threshold value of P* that makes adopting the policy optimal. For example, α=25 makes the threshold value of P*= 25, and α=-15 makes P*= .65. How exactly does a group deal with these divergent incentives? In order to make the problem as concrete as possible for the subjects in the experiment, each subject observes the biases of the fellow subjects with whom he is directly connected to by a communication link. This in turn, then, induces differences across the different network architectures in the fine structure of information that different people have. For example, in the Full network all biases are known, but in the Star network the central player knows all biases, while the outlying players only know their own and the central player's biases.

	P*	1-P*
Adopt Policy	100 + α	α
Do not adopt policy	50	50

Figure 5: Payoff Table for the Basic Group Decision Game with Biases

Before discussing hypotheses, it will be useful to discuss the design and implementation of the games in the laboratory, as the hypotheses will sometimes depend upon surprisingly small details. The games were conducted in the Gregory Wachtler Experimental Economics Laboratory at Rutgers University during 2008 and 2009. Each session consisted of 15 or 20 subjects who played a total of 12 rounds. In each round, subjects played either the basic game from Figure 3, or the game with biases, from Figure 5. The first four rounds were the basic game, and the last eight rounds were the game with biases. Subjects played in each of the four network architectures once during the first four rounds, and in each of the four network architectures twice in the last eight rounds. In each new round, subjects were randomly reassigned to groups of five players. In addition, in each round subjects were given alias IDs that changed from round to round. Thus, although one would inevitably end up with some of the same people in one's group in different periods, it was not at all obvious who was who, due to the changing ID names.

The true value of the probability of success for the risk policy, P*, took on 4 different values: .35, .45, .55 and .65, and all players experienced all of these values over the games that they played. Subjects were not told what the possible underlying values were, but only received signals from a beta distribution centered on the true value for the game they were playing. The games with $P^*<.5$ were, in principle, games in which it would be optimal not to adopt the risky strategy, and vice-versa for games with $P^*>.5$. In practice, games with $P^*=.45$ or $P^*=.55$ could be hard for subjects to judge, as the average signal could be misleading. The effect of the underlying signal distribution was considered in calculating how successful subjects under different network architectures were in achieving the objectively best (at least for the basic game) outcomes. Subjects were not told anything detailed about the underlying distribution, except that the average of the signals for their group would be a good estimate of the true probability of success for the risky policy.

The biases were implemented as a discrete distribution that took on values {-25, -15, -5, 5, 15 and 25} in the experimental currency. The distribution was either "centered," with probabilities {1/14, 1/7, 2/7, 2/7, 1/7, 1/14}, or "extreme," with probabilities {2/7, 1/7, 1/14, 1/14, 1/7, 2/7}, respectively. These were counterbalanced over rounds of play. The underlying distribution of biases was accounted for in assessing the performance of the different network architectures and the behavior of individual subjects. Subjects earned money for every game that they played (that is, the payoffs counted in every game), but each round was as near as possible to an independent observation on a new set of parameters. When appropriate, time trends (i.e., experience in the game) and past earnings are included in the statistical analysis.

Finally, there are a few aspects of the communication technology that may be relevant in the analysis of the experiment. Communication was implemented as computer "chat boxes" on the computer screen. At the beginning of the round, the general nature of the communication architecture was outlined for the subjects (i.e., which kind of network). For every person one was linked to in a particular round, there was a box on the screen indicating the current ID of the person and the bias of the person (if applicable). In addition, in the Full network there was a "broadcast box" which could be used to send a message to all members of the group at once. Also, in the Star network, there was a broadcast box for the central player to send messages to all players in the group. The outside members in the Star network were only connected to the central player, and did not have a broadcast box. All other connections were also one-to-one chat boxes.

Subjects initially read general instructions of the game and questions were answered. Subjects were then formed into groups of five via the computer (the subjects never met face to face and did not know who was in their groups in a particular round), and they were informed of the general structure of communication for that round. Subjects received their initial signals about the value of P*, and were then asked how they would vote, given their initial signal only. The game then moved to the communication stage in which subjects could exchange messages about their signals, about how they were going vote, etc. At the end of the communication stage, subjects then did their final voting, after which the risk policy was either adopted for the group or not. Then the state of the world for the risk policy was realized, based on the true underlying value of P*, and final payoffs for the round were reported to the subjects. All aspects of the experiment were implemented via a program written in Z-Tree, including instructions, preliminary voting, communication, final voting, realization of the state of the world, and reporting of the state and of final payoffs for the round.

Though this analysis is somewhat exploratory, there are some easily agreed on fundamental principles of rational decision-making that can provide guidance as to what might be expected. First of all, in the basic game, with no biases, there is no strategic reason to withhold or misreport one's private information (except for the case already mentioned, risk preferences: e.g., and extremely risk averse individual would appear to have an incentive to under-report whatever signal he or she receives). Under the assumption of risk neutrality, it is always best to opt for the risky policy when the expected P* exceeds .5. When biases are introduced, then there are different incentives, but there are also complexities in the information structure the effects of which could not be predicted. For example, in the Full network, what will be the net effect of everyone knowing everyone else's biases?

One possibility is that everyone will assume that any communication about signals cannot be trusted, and just make a decision based on one's own signal. Paradoxically, then, this would lead to perfect disaggregation of available information, compared to the basic game. Using the same logic, one would not expect either the Star or Neighbors networks to aggregate particularly well, since there is the same problem of how to interpret a message from those with different biases. However, to the extent that one shares a bias (in the same direction) with a neighbor, perhaps one would be more likely to communicate truthfully, and vice-versa. The Split network is interesting in this connection, as it may be more likely for three members together, if they happen to share a bias in the same direction, might have a better chance of doing "the right thing" and adopting the risky policy when it makes sense to do so, than would be the case for the Full network.

On the other hand, though it takes a lot of effort, one could in the Full network communicate individually with only those that you share a bias with (as opposed to using the broadcast box), and in this way a group of three like-biased players might achieve their desired outcome. The answers to these speculations are reported below. Before proceeding to the empirical results, the most speculative aspect of the entire enterprise are worth highlighting. What are the possible larger "social" effects of being in particular network architecture? For example, do some network architectures promote an increased level of trusting behavior on the part of participants?

These are not questions that quantitative social scientists have the training or background to feel confident making predictions about, but it seems at least possible to identify aspects of the different architectures that might have a subtle effect on some participants. For example, the Full network could be seen as fully transparent and a good thing by some (or a risky proposition and/or a place to exploit others by more hard-nosed types). The Star network is centralized, and the central player is the "boss" who controls things. How do the "underlings" on the periphery behave in this context? The Neighbors network is decentralized and may appear more democratic (though more limited than the Full network). Does this lead subjects to be more open and honest? The Split network has insiders and outsiders. The insiders, the connected group of three, can determine the outcome, so the outsiders the connected group of two, really have no role to play. Do the outsiders recognize this fact?

The statistical analysis attempts to quantify everything possible and some things about networks can be quantified. For example, the networks can be compared on the basis of how many others you are directly connected to, and there may be differences within a network (e.g., the Star) in this regard. Beyond the ability to pin down actual observable aspects of the decision-making environment, network dummy variables are included to see if there are identifiable differences in behavior that can be attributed to the network architecture. The hypotheses about information aggregation and resulting adoption decisions are summarized in Figure 6. Behavior at the individual level is also examined independent of network architectures. Hypotheses for individual behavior are summarized in Figure 7.

	Full Network	Star Network	Neighbors Network	Split Network
Basic Game	Maximum efficiency: easy to aggregate all information	Low efficiency: Difficult to aggregate all information	Low efficiency: Difficult to aggregate all information	Minimum efficiency: impossible to aggregate all information
Game with Biases	Lowest efficiency? Increased number of not credible messages	Same problem of credibility as for Full, however less connections and hierarchical structure	Same problem of credibility as for Full, however less connections	Partitioned network, limiting efficiency

Figure 6: Hypotheses Summarized at Network Level

	Across all network Structures:
Basic Game	Easy to aggregate all information, should aggregate fully in all connected networks. Should report own signal truthfully, aggregate information appropriately to update beliefs, and vote for expected payoff-maximizing option.
Game with Biases	Aggregation cannot be assumed. All biases visible, messages not credible. Should distort own information strategically with the incentive to do so increasing in the bias. Should discount others information, vote for expected payoff maximizing choice.

Figure 7. Hypotheses Summarized at Individual Level

2.3 Results

Experimental results follow. In Section 2.3.1, group and individual voting results are presented and benchmarked against the number of correct decisions. Section 2.3.2 follows this up by examining the information aggregation process and specifically when truthful communication occurred. Section 2.3.3 concludes with the determinants of individual beliefs about the expected likelihood of a policy to succeed and he factors of the individual votes presented in aggregate in section 2.3.1. Section 2.4 examines areas for further research before section 2.5 then concludes.

2.3.1 Voting Results

Each subject was asked how he or she would vote prior to communication, and then was given a final vote after communication. Underlying each group decision in turn there was a true underlying probability and an expected probably based upon averaging the signals received. These provide two benchmarks for evaluating the efficacy of communication across the various networks. Against these measures, both the percentage of expected payoff maximizing group decisions and the percentage of expected payoff maximizing individual correct decisions can then be measured. These results in turn are filtered by network structure and by the inclusion or exclusion of biases. Results follow in Table 19 and Table 20.

Due to the nature of aggregation, groups were more likely to already be making the correct decision then the individual members within them. Across all networks, 81.4% of groups were making the correct decision based upon the revealed signals without communication, while the distribution of these signals meant that the number of groups maximizing payoffs based upon the true underlying distribution prior to communication was lower at 74.5%. The introduction of communication increased each of these percentages, climbing to 85.9% based upon average received signal and 76.7% based upon the underlying true probability. Communication had a more dramatic effect on the individual decision level, increasing the number of correct decisions based upon the average signal received from 68% to 80.6%.

Table 19: Summary Statistics on Correctness of Group Preliminary Voting and Final Voting Compared to Average of Signals in Bold, *Compared to Underlying True Value of P* in Italic*

	Basic Games Only		Biased Games Only		All Games	
	Preliminary Vote	Final Vote	Preliminary Vote	Final Vote	Preliminary Vote	Final Vote
All Networks	**84.43%** *74.59%*	**87.70%** *72.95%*	**79.59%** *74.48%*	**84.94%** *78.66%*	**81.44%** *74.52%*	**85.87%** *76.73%*
Full Network	**93.75%** *84.38%*	**100.00%** *84.38%*	**84.38%** *71.88%*	**84.38%** *75.00%*	**87.50%** *70.04%*	**89.58%** *78.13%*
Star Network	**70.00%** *63.33%*	**73.33%** *60.00%*	**81.48%** *74.07%*	**90.74%** *79.63%*	**77.38%** *70.24%*	**84.53%** *72.62%*
Neighbors Network	**81.25%** *68.75%*	**87.50%** *68.75%*	**75.41%** *77.05%*	**83.61%** *78.69%*	**77.42%** *74.19%*	**84.95%** *75.27%*
Split Network	**92.86%** *82.14%*	**89.29%** *78.57%*	**78.33%** *75.00%*	**81.67%** *81.67%*	**82.95%** *77.27%*	**84.09%** *80.68%*

The gains from communication appear to be somewhat robust to the introduction of biases, with the general results showing a lower overall rate but also a lower preliminary rate of votes matching the expected underlying state. The group results across all networks shows an increase from 84.4% to 87.7% within periods with no biases, and an increase from 79.6% to 84.9% in the periods with biases. The more sensitive individual votes move from 68.7% to 82.6% in periods without biases and 67.8% to 79.6% in the periods with biases. There is ample evidence to reject the hypothesis that communication provides no benefit when individual biases are present, with these gains within the range of those found without biases.

Table 20: Summary Statistics on Correctness of Individual Preliminary Voting and Final Voting Compared to Average of Signals in Bold, Compared to Underlying True Value of P* in Italic

	Basic Games Only		Biased Games Only		All Games	
	Preliminary Vote	Final Vote	Preliminary Vote	Final Vote	Preliminary Vote	Final Vote
All Networks	**68.69%** *65.74 %*	**82.62%** *71.80 %*	**67.78%** *66.03 %*	**79.58%** *73.14 %*	**68.09%** *65.93%*	**80.61%** *72.69 %*
Full Network	**70.63%** *67.50%*	**95.00%** *84.38%*	**67.50%** *63.13%*	**81.56%** *72.81 %*	**68.54%** *64.58 %*	**86.04%** *76.67%*
Star Network	**64.67 %** *63.33%*	**72.00%** *62.67%*	**70.00%** *67.04%*	**84.81%** *76.67%*	**68.10%** *65.71 %*	**80.24%** *71.67 %*
Neighbors Network	**64.38%** *60.63 %*	**83.75%** *67.50%*	**67.21%** *67.54%*	**80.00%** *73.11 %*	**66.24%** *65.16%*	**81.29%** *71.18 %*
Split Network	**75.71 %** *72.14%*	**78.57%** *72.14%*	**66.67%** *66.67%*	**72.33 %** *70.33 %*	**69.55 %** *68.41%*	**74.32%** *70.91%*

While the aggregate results finds similar gains from communication with and without individual biases, decomposing these results by network structure reveals dramatic heterogeneity. Without biases, the Full network aggregates information the most based upon the group and individual vote results. Groups voted for the correct policy based upon the average signal in 100% of all votes for the Full network, while the all other networks were under 90%, and the star network was only at 73%. In contrast, when biases were introduced, the star network performed the best 90.7% of group decisions versus nothing above 85% for the others. The results suggest that the optimal structure for communication depends upon the incentive structure involved, with a fully connected network best at a simple aggregation problem while the hierarchical centralized star network performs best with the introduction of individual idiosyncratic interests.

2.3.2 Aggregation and Truthful Communication

The different biases create different cutoff points for risk neutral voters. Subjects thus have two distinct albeit inseparable tasks: aggregating information so they make the right decision for their vote given all available signal information, and ensuring that the majority of the group votes as if following his cutoff rule. This second task is at odds with the first, by truthfully revealing a signal the subject helps the group aggregate information as efficiently. However, by distorting one's own signal one it is possible to change the group's beliefs about the expected value of a policy, and thus possibly conform to the individual's own (biased) voting rule. As an example, consider a group where the distribution of signals is (50%, 55%, 40%, 40%, 40%). The average signal is a 45%, and if there are no biases, one would expect a full-information equilibrium with all five risk neutral individuals voting "No". In contrast, if the individuals have biases assigned as follows (+10, -5, -5, -25, +25) then in the "honest complete information equilibrium" one would expect to see a split vote of (Yes, No, No, No, Yes). And because their cut points differ from the group, both subject 1 and subject 5 would benefit from overstating their signal and swaying at least one other individual to vote with them.

The game is inherently a cheap talk situation, with no credible mechanism for imparting the truthfulness of one's message. If one assumes a limited level of rationality, with strategic communication but naive acceptance of messages, then one would expect to find *all* individuals with biases shading their signals, with the amount of shading depending upon the size of their bias. While there is no pure strategy equilibrium for subjects, adding a second step of rational strategy will just make this worse, as communication from subjects with larger biases, which are publicly observed, are expected to lie more and thus need to lie even further. Viewed in terms of Young (1993), neither of these would be a stable equilibrium of the repeated game, while the complete disintegration of trust and aggregation would be. However, assuming some positive relationship between a conveyed signal and other's beliefs, then incentive to dishonestly communicate one's signal should be increasing in the scale of the individual bias. This forms a testable hypothesis: Does the size of a subject's bias affect his honesty? If there is a disutility associated with dishonest behavior – as found even in payoff enhancing situations such as in Erat and Gneezy (2012) – then one would expect the existence of a bias threshold for lying.

In contrast to this prediction, one would expect that individuals, when communicating *only* to like biased individuals would be more likely to share their true signal. Here the information aggregation problem remains but there is no longer an incentive to alter the effective voting rule of the other subject by supplying them with a misrepresented signal. Thus there is a second testable hypothesis: Individuals communicating with similarly biased individuals will lie with decreased frequency. Within connected networks, this hypothesis relies on a trust among thieves assumption that the similarly biased subjects could be co-opted into a conspiratorial plot; that the subjects could first accurately share signals with each other and then trust each other to mischaracterize them to the remaining group members. Finally, several other variables may affect the rate of truthful communication; the underlying network architecture, the extremeness of the received signal, and the past experiment performance of the subject. In addition I tested for the difference in a 15 person and 20 person experiment session (with group size fixed at five for both).

The results are presented in Table 21 and support the first hypothesis; the incidence of lying was materially affected by the absolute size of the bias of an individual. As the incentive to deviate from truthful behavior increased, the incidence of lying increased, supporting a cost benefit approach to lying or strategically misrepresenting one's signal. This was found significant just above the 5% level. Of the conjectures, the hypothesis that the signal that subjects receive affects the rate of lying was rejected. While this result is compatible with a theory of rational utility maximizers, the result is somewhat counter intuitive: what a subject knows doesn't affect if they lie about what they know. There is a significant time series effect, with the incidence of untruthful communication increasing as play progresses.

The underlying structure of the communication structure and even the experimental session are each also found significant. The null hypothesis that behavior is consistent across all network structures can be accepted for all but one structure – the Star. Here, the deviation can further be isolated to the behavior of the star points, with the rate of lying from the Center statistically indistinguishable from that in other network structures. In contrast, the star points had a decreased rate of untruthful communication significant at the 5% level, and just outside the 1% level. Communication from these subjects had two salient features, that it was inherently one on one and that it was one on one with an individual in a position of more power and possibly knowledge. The behavior in the split network provides a nice counter point for teasing apart these two factors, in that network subjects on the two side each had communication limited to the more intimate on one one but did not have an unequal basis for the relationship. And in contrast, the two side of a split network saw no decreased incidence of lying, suggesting that the hierarchical nature of the network structure is in some way increasing the truthful behavior. To further test this finding, Table 22 presents results of a random effects regression on truthful communication using both the number of connections as well as a dummy for being a star point as variables. A decreased number of connections (depending upon the network, these ranged from 4 to 1) was not statistically significant, however the particular network structure of being a star point was again significant.

Finally, each experiment sessions either had 15 or 20 subjects forming the pool from which the groups of five where anonymously grouped. Despite the lack of permanent identifiers within a session from play to the next, the smaller sessions were found to have increased truthfulness, and on the 1% level. While a full investigation of this is outside the scope of this work, it suggests that the cost of lying is influenced by the size and perceived anonymity of a setting. This result also suggests the delicacy to which small setup details can affect behavior.

Table 21: Regression for Truthful Communication by Network Type

GLS Random Effects Probit on Indicator for Truth *(assuming a truth-relevant message)*

Variable	Coefficient	Std. Error	z-statistic	Prob>z
Bias Size	-.014	.007	-1.94	.053
Signal Clarity	.006	.005	1.21	.228
Ring	.065	.164	0.39	.693
Two in Split	-.229	.222	-1.03	.303
Three in Split	-.160	.192	-0.83	.407
Star Center	-.067	.322	-0.21	.835
Star Point	.472	.191	2.47	.013
Session Size	-1.600	.382	-4.18	.000
Period	-.064	.021	-3.03	.002
Constant	3.973	.385	10.31	.000

n = 1782 observations on 160 individuals, Log likelihood = -401.39, $X^2 = 53.70$

Table 22: Regression for Truthful Communication by Connections GLS Random Effects Probit on Indicator for Truth *(assuming a truth-relevant message)*				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Bias Size	-.015	.007	-1.97	.048
Signal Clarity	.005	.005	1.05	.292
Connections	.037	.059	0.63	.529
Star Point	.459	.186	2.47	.013
Session Size	-1.590	.380	-4.18	.000
Period	-.064	.021	-3.02	.003
Constant	3.831	.403	9.50	.000

n = 1782 observations on 160 individuals, Log likelihood = -402.46, X^2 = 52.09

The practical matter that subjects communicated entirely with the broadest audience possible hinders testing the second hypothesis. Thus, while subjects had the means to engage in conspiratorial one to one conversation with every other member in the Full network, they instead communicated to the entire group simultaneously in the broadcast box. To handle this complication, both the average bias in a group setting and the network structures where communication was inherently one to one: Star and the two subject side of a Split network. For each methodology employed, the hypothesis that similar biases increase honesty was rejected. Within the Full network, there was no effect of any average bias. Similarly, within the Star network, there was no effect of the average bias of the Star periphery or points, on the rate of lying of the center. Even in the communication between the Star center and Star points, there was no similar bias effect. Examining the mental process by which honesty was increased in the Star points is beyond the scope of this work, but this network architecture is the only one in which an individual might be construed to have someone in a position of power above them, and is broadly compatible with previous findings on the impact of group leaders.

2.3.3 Belief Formation and Individual Voting Decomposition

Examining subjects' beliefs about the expected probability of a policy succeeding provides a second methodology for examining the extent to which information aggregated. Information aggregation should be represented by beliefs updated from an individual's signal to include the signal information received by all connected members of the group. As such, the complete networks – Full, Star and Ring – were used to test the sensitivity of beliefs to the signals received by all other members of the group. To the extent of information aggregation in final beliefs and the hypothesis that information aggregation is increasing in the level of connections and decreasing in the minimum number of connections needed for complete aggregation, one would expect beliefs within the Full network to be most sensitive to the signals received by other members. When comparing the Star and Ring networks, there was no clear prior prediction, with hypothesized opposite effects for the increased number of connections and for the increased minimum distance for information to travel in the Ring network. Across all networks, the introduction of biases should result in a decreased confidence in the information being communicated and thus beliefs more sensitive to their starting signal compared to the signals received by other group members.

Results are presented in Tables 23 – 28 below, including all periods with and without biases as well as separately only those with biases only. Signals were included as demeaned evidence, with a signal of 63.1% thus 13.1% positive evidence. The signals of connected members were then summed to form a connected evidence variable. Thus, if the average signal of the four connected members was 52%, the connected evidence would be recorded as 8% positive evidence, implicitly weighting by the numbers of observers. Finally, the subject's own bias was included. While the bias was a strictly additive payoff and not state relevant, including this allowed a check on if beliefs were 'irrationally' incorporating expected payoffs.

Several interesting results emerge from the beliefs. First, within the Full and the Ring networks subjects are significantly weighting the underlying signal evidence of other group members compared to their own evidence – a coefficient of .773 on others' evidence versus .287 on own evidence in the All network and .480 and ,245. This effect is not significantly reduced by the introduction of biases, with similar results derived from the rounds with biases only. In contrast, the Star network has an opposite effect with a coefficient of .233 on others' signal evidence versus an own signal evidence coefficient of .407. It is important to note that while it appears that beliefs are overcompensating for others evidence, this is based upon the true signal received by the subjects. As such, the 'overweighting' may represent a breakdown in the information being transmitted with only extreme signals communicated, systematic dishonest communication being naively accepted, a discounting of incompatible evidence whereby minority signals in the other's evidence is ignored, or a phenomena of over communicating and over incorporating group evidence.

While the individual payoff biases were paid regardless of whether a policy was successful so long as a policy was voted for, these biases were found to have a positive coefficient significant at the 10% but not the 5% level in both the Full and Ring networks. There thus is limited evidence suggesting that individuals are irrationally incorporating these payoffs into their expectations of the ultimate underlying state. In contrast the Star network showed no such significance, raising the question as to how varying the communication environment but not the underlying payoff structure altered expectations and reported beliefs.

Table 23: Random Effects Regression on expected Probability in a Full Network				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.773	.053	14.46	0.000
Own Evidence	.287	.035	8.30	0.000
Own Bias	.091	.054	1.68	0.094
Constant	50.72	0.84	60.53	0.000

n = 480 observations on 160 individuals, R^2 = .4596, X^2 = 407.29

Table 24: Random Effects Regression on expected Probability in a Full Network with Biases				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.721	.065	11.04	0.000
Own Evidence	.216	.043	5.02	0.000
Own Bias	.081	.054	1.49	0.135
Constant	50.67	1.12	45.06	0.000

n = 320 observations on 160 individuals, R^2 = .3761, X^2 = 212.66

Table 25: Random Effects Regression on expected Probability in a Star Network				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.233	.051	4.53	0.000
Own Evidence	.407	.036	11.16	0.000
Own Bias	-.010	.068	-0.15	0.880
Constant	57.29	1.80	53.77	0.000

n = 428 observations on 160 individuals, R^2 = .2630, X^2 = 151.29

Table 26: Random Effects Regression on expected Probability in a Star Network with Biases				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.237	.069	3.42	0.001
Own Evidence	.357	.046	7.83	0.000
Own Bias	-.050	.071	-0.71	0.477
Constant	55.02	2.41	22.84	0.000

n = 270 observations on 160 individuals, R^2 = .2289, X^2 = 78.98

Table 27: Random Effects Regression on expected Probability in a Ring Network				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.480	.046	10.45	0.000
Own Evidence	.245	.035	6.93	0.000
Own Bias	.097	.056	1.75	0.081
Constant	51.21	.985	52.00	0.000

n = 465 observations on 160 individuals, R^2 = .3134, X^2 = 243.89

Table 28: Random Effects Regression on expected Probability in a Ring Network with Biases				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.443	.058	7.62	0.000
Own Evidence	.234	.045	5.18	0.000
Own Bias	.094	.058	1.61	0.106
Constant	51.28	1.212	42.32	0.000

n = 305 observations on 160 individuals, R^2 = .2665, X^2 = 133.85

While Table 19 and Table 20 present the percentage of groups and individuals voting correctly based upon network structure, the determinants of individual votes can also be regressed just as beliefs were. One would expect that votes would be based upon expectations, but unlike beliefs they would also be expected to include the individual biases. Results are presented for the complete networks separately in Tables 29 – 31 below. As might be expected, the coefficients on signals and other subjects' signal evidence mirror those that constituted the believed probability of the true state, with the Full and Ring network both showing a much greater relative coefficient on the rest of a group members' signals when compared to one's own signal. This raises the question, is beliefs the only underlying variable or does the separate evidence variables have additional significance. To test this, Table 32 includes beliefs as an explanatory variable pooling across all networks. The inclusion of beliefs improves the fit of the regression, but both a subject's own signal as well as the signals received by their group still retains statistical significance at the 1% level.

In both the combined regression including beliefs and the individual network specific regressions, own bias has the expected positive coefficient and significant at the 1% level. However, a subject's own bias is significantly smaller than the combined coefficient of the signal received by a subject and all members of the group. Individuals appear to be voting in a manner that evaluates expected individual profit from an individual specific component as less significant then the same amount of expected profit from an increase in average signals that would translate to the entire group.

Table 29: Random Effects Probit Regression on Voting in a Full Network				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.084	.007	11.32	0.000
Own Evidence	.024	.004	6.45	0.000
Own Bias	.021	.006	3.56	0.000
Constant	.096	.077	1.25	0.21

n = 480 observations on 160 individuals, Log likelihood = -177.82, X^2 = 162.56

Table 30: Random Effects Probit Regression on Voting in a Star Network				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.015	.004	3.57	0.000
Own Evidence	.030	.004	8.34	0.000
Own Bias	.014	.006	2.56	0.010
Constant	.449	.0150	2.99	0.003

n = 428 observations on 160 individuals, Log likelihood = -235.83, X^2 = 76.58

Table 31: Random Effects Probit Regression on Voting in a Ring Network				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Connected Evidence	.043	.005	9.34	0.000
Own Evidence	.023	.003	6.74	0.000
Own Bias	.029	.005	5.34	0.000
Constant	.125	.071	1.76	0.079

n = 465 observations on 160 individuals, Log likelihood = -214.10, X^2 = 140.12

Table 32: Random Effects Probit Regression on Voting in every Network including Belief				
Variable	Coefficient	Std. Error	z-statistic	Prob>z
Belief	.057	.004	15.22	0.000
Connected Evidence	.014	.003	7.02	0.000
Evidence	.017	.002	8.45	0.000
Own Bias	.021	.003	6.73	0.000
Constant	-2.776	.202	-13.76	0.000

n = 1813 observations on 160 individuals, Log likelihood = -712.07, X^2 = 427.80

2.4 Discussion and Areas for Further Research

The full process by which free-form communication is incorporated into beliefs is complex, and beyond the scope of this paper. However, within an aggregation framework the performance of groups with several different communication structures were tested. As these networks were exogenously dictated, and archetypical in nature, a significant task remains in terms of generalizing the findings and finding an encompassing theory of behavior that would be able to make predictions across a new network structure, or make predictions about the optimal network given a distinct task. Further areas for research include, varying the number of group members and the process for selecting a group policy, including two risky options instead of a single risky option and a safe option, and incorporating group identities as befits a political environment. The authors in a subsequent work examine the last two of these.

Additional, exogenous network connections formation with a cost to connecting and communicate is an area ripe area for further study. A generalized theory of individual behavior in group decision making with uncertainty and imperfectly aligned interests requires substantial insights into the interactions between the underlying payoff structure and the network structure and background fundamentals.

2.5 Conclusion

When examining the results a clear pattern of behavior does emerge. The Full network with maximum connections performed best when all interests were perfectly aligned. In contrast, when interests become divergent, and individual payoffs create conflicting goals to the group's information aggregation problem, the group's performance is no longer strictly increasing in the level of connections. Strictly decreasing the number of connections does not appear to be optimal either, but rather the creation of *unequal* connections – those where one individual is in a position of greater connectivity then the other. Within the Star network, communication was more truthful and beliefs did not reflect on overweighting of other members' signals, resulting in the best outcomes when biases were introduced. While the optimal network structure appears to depend upon the level of divergent incentives, there is no predicted optimal structure when these fundamentals are unknown. Towards this end, this paper hopes to have added insights into the type of networks that promote different levels of trust, aggregation and coordination.

Chapter 3: The Consistency of Consistency and Preference Aggregation in Individual Choice with Uncertainty

Chapter 3 Abstract:

Subject errors and inconsistencies can have drastic effects on estimating and testing utility theories in experimental economics. To examine repeated individual choices under uncertainty and the impact of subject variability on the acceptance of decision theory formulation, subjects were asked a sequence of preference questions between compound lotteries including scaled versions of equivalent questions. Subject response data was then used to estimate the level of individual inconsistency, the persistence of inconsistency, and the effects of inconsistency on aggregating individual behavior into a theory of decision-making.

Findings include that significant aggregate response consistency exists across scaled lottery questions, for scaling coefficients in the range of 50% to 200%; random response patterns can be firmly rejected and deviations from consistent responses were unsystematic. On aggregate, individual responses to different base questions demonstrated little interrelationship. Parameter estimates were not significantly predictive, except for a small subsample of consistent subjects that maximized expected value, for questions in which a lottery was compared with a second lottery featuring a 50% chance of a monetary prize, and a 50% chance of receiving zero, and finally for controlling for individual variability.

Significant heterogeneity existed in response consistency, with the level of consistency a persistent individual trait both within and across scaled question variants. Individual variance has a systematic effect on risk aversion and prospect weight coefficients. As variance increases, power coefficients and CPT coefficients decrease. The more variable a subjects responses, the more the coefficient estimates demonstrate mean reversion and thus an inverted S-shaped prospect weighting. Eventually, as variance increases, the effects are reversed and coefficients increase.

Chapter 3, Section 1 Introduction

Experimental studies measured individual choices under uncertainty, attempting to quantify the appropriate decision making process employed by individuals and provide a definitive model. Given a proposed framework, subject response data is used to estimate coefficients on individual risk aversion, loss aversion, and subjective probability. Models of individual choice are then rejected or accepted based upon the individual coefficients estimated, either combined into a representative agent framework or based upon the percentage of the population conforming to a given theory. Yet there is often a fundamental disconnect within the data with average individual coefficients often supporting a cumulative prospect theory model but the underlying observed cumulative prospect theory weights on which this interpretation rests, are highly variable. Coefficient estimates vary widely across individuals, both based upon on stable individual characteristics but also on such transitory phenomenon as mood (Fehr, Epper, Bruhin, and Schubert (2007).

Separately, experimental data often finds that, even in exact repetitions of simple binary choices with significant economic stakes, individuals respond inconsistently. Within a random sample of the adult population of Rwanda, Jacobsen and Petrie (2009) find that despite significant financial stakes "over 50% of the participants making at least one mistake. Importantly, errors are informative." When examining the impact of mood on probability weighting, Fehr, Epper, Bruhin, and Schubert (2007) discard the data of over 15% of their subjects, who demonstrated repeated inconsistency within the same preference solicitation. These subjects on two or more certainty equivalent questions switched from preferring the risky lottery to the certainty equivalent and then back to then preferring the risky lottery as the certainty equivalent increased. Indeed, these examples are more the norm than the exception, with all experimental data revealing inconsistent responses that ideally should be modeled. The overall level of individual inconsistency found in experimental economics might best be summarized by Hey (2005) "if we ask the same questions to the same subjects they typically give different answers, even in situations where they 'should not'."

The impact of fundamental individual inconsistency on decision theory is something that has been noted before, with Hey (2005) and Loomes (2005) both emphasizing the need for a joint determination between errors and a model of individual choice under uncertainty. Earlier work by Carbone and Hey (2000) found that the error structure was highly heterogeneous, concluding that "This message seems to be very clear: for many subjects the Constant Probability story is 'best'; for many others the White Noise story is 'best'. Trying to get one error story as 'best' for all subjects would appear to be seriously misleading." This paper proceeds in that path and explores the measure of individual consistency when faced with repeated similar decision choices, the heterogeneity of individual consistency, and finally the impact of this inconsistency on parameter estimates for decision making models. Continuing in this direction, this study examines the effects of separating the coefficient estimated and the theoretical models of individual choice with the variable responses of experimental subjects that inevitably occur.

To provide some further motivation for the following work, consider the following scenario in which experiment subjects are asked to fill in the certainty equivalent to receiving \$100 with a 90% probability. Risk neutral subjects with a perfectly consistent expected value utility function would reply \$90. For risk neutral subjects that, through whatever functional form or trembling error explanation, have an additive error term (the multiplicative case is similar), the replies would distributed according to the joint distribution of the true CE of \$90 and the error. Yet risk dominance and common sense suggests that a CE above \$100 would be an incompatible answer, effectively truncating all positive errors greater than \$10. For simplicity, assume that the errors are additive and uniform on the range [-20, +20], and that errors that result in an incompatible CE are re-sampled by subjects. This simplified set of assumptions would result in a study with 2/3rds of all subjects demonstrating a probability weighting less than the expected utility value of 90%, and the average coefficient would be 85%. Similarly, with the same error specification 2/3rds of all subjects would demonstrate a probability weighting greater than the expected utility value of 10%, for the certainty equivalent of receiving $100 with a 10% probability.

While driven by asymmetrical error distribution versus a true underlying preference the actual empirical results would be identical to a uniform error distribution within 15 of a Certainty Equivalent of 85. The distinction between a "true" population CE of 90 that, via truncating is expressed as an average CE of 85 and a CE of 85 would then appear little more than philosophical. Average response data would support a CPT interpretation for a representative consumer, and aggregating responses by subject would not necessarily change this conclusion either. Assuming individual errors are identical and independently distributed, 4/9ths of the subject body would fully conform to a standard cumulative prospect theory weighting interpretation (CE less than 90 and CE greater than 10), while another 4/9ths would have one of two coefficients matching a CPT (suggesting one error).

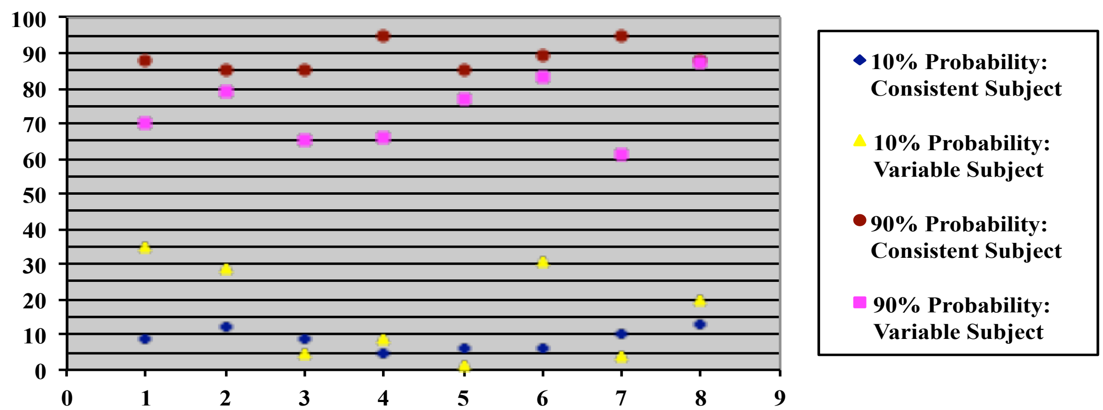
While the aggregate data supports a CPT interpretation, panel data and Bayesian analysis provide a powerful check on this interpretation. If the data generating process were a true CPT model with heterogeneous individual parameterizations, then there would be a (albeit imperfect given subject error) relationship between subject responses across different questions. Deviations from given objective probabilities should noisily reflect an overall attitude about risk, with the direction and size of deviations predictive of responses to different questions. Thus an individual's response underweighting one probability should correspond with a response overweighting another, and vice versa.

In contrast to this prediction, for the hypothetical data generating process described above individual level subject responses to any question will provide no change in the likelihood of a subject demonstrating a CPT consistent parameter than the subject average response rate. This effect would hold not just for different fundamental probability questions (such as the certainty equivalent of a prize occurring with 90% probability or 10% probability) but would occur even if the original question was re-sampled. Conditioning on any subset of a subject's previous responses would result in a predicted response identical to the simple unconditional expectation. Bayesian panel data analysis would thus reject that there is a meaningful individual prospect theory, as subject responses are perfectly uninformative.

The result is that given individual heterogeneous variability is introduced, it is no longer so simple to discerning if aggregate evidence of a cumulative prospect theory is the product of noise, or a true data generating mechanism. As an example, suppose there are two subsets of the population: nearly consistent individuals occurring with probability p and inconsistent individuals occurring with probability 1 – p. Consistent individuals have response errors to certainty equivalent questions that are uniformly distributed on the range [-5, +5] while inconsistent individuals have response errors that are uniformly distributed on the range [-30, +30]. Consistent individuals will thus have observed responses centered around the objective probability, with no problem with censored coefficients for a CE of receiving a prize of $100 with a probability of 90%. In contrast, the inconsistent individuals will be distributed similar to the first example, with risk dominance censoring the upper errors and leading to a distribution biased away from the extreme probabilities.

The effect can be seen in Graph 1, where random responses were drawn according to the above distribution. Average responses and any error minimizing estimates find the variable subject is overweighting the low probability and underweighting the high probability, thus demonstrating classical inverted S shaped subjective probability weights.

Graph I: Random Responses of Consistent and Variable Subjects



Given the simplification of the above data generating process and two population types, individuals demonstrating traditional CPT coefficients (overweighting small probabilities and underweighting large probabilities) will be disproportionately composed of the inconsistent type. This type dependence creates an artificial correlation between responses compatible with traditional CPT parameters. The result is that a standard Bayesian test would no longer reject CPT but would instead find evidence of a correlation across responses: subject responses would be informative in predicting future behavior, since they are informative of the level of variability in subject responses.

The unfortunate result is that individual heteroskedastic error components when combined with risk dominance can cause bias even for panel data Bayseian testing the independence of probability weights. In the above example, heterogeneous variability in the data generating process resulted in mean reversion in probability weights for the most unstable individuals. The presence of one sided long tails caused a centralizing bias to any parameters for prospect weights, with the impact strongest on the same extreme probability values for which the need for prospect theory weighting is stressed. This result does not depend on the simplifications of uniform errors or that subjects can be described as one of two types, but arises as a significant proportion of the population demonstrates preference instability, even in the simplest of settings and with exact question repetitions.

While a prospect theory weighting function would indirectly measure instability for the above example, an alternative would be to estimate a direct measure of the stability of individual preference revelations. A measure of individual consistency will be itself dependent upon the assumed error process. Individual variability can be measured as the probability of a random tremble, a degree of imprecision in selecting based upon the distance between expected payoffs response (such as in a Quantal Response Equilibrium model) or as a perturbation to a fundamental decision making process. In selecting between alternatives, it is worth noting that, just as in the above example, a model should not just be selected based upon maximizing a likelihood function, but that the model should be expected to retain some predictive power and thus perform reasonably out of sample. For any measure of inconsistency to be meaningful, it needs to be itself consistent. Similarly, prospect weights and risk aversion coefficients should themselves be persistent or at least robust to similar situations. If not, then we must either complement these variables with additional variables that will control for situational biases, (regret, loss aversion and level and spacing are approaches in this direction) or relegate them to a status of perhaps being useful for capturing variance but explanatorily insignificant.

A brief outline of the chapter organization follows. In Section 3.2, the experimental design and motivation for using scaled question variants is discussed. Section 3.2 also contains a survey of related literature and the motivation for viewing scaled question variants as comparable. The experimental results are in section 3.3 with subsections 3.3.1 examining consistent response patterns, 3.3.2 looking at near consistent response patterns, 3.3.3 at rank based consistency, and 3.4.4 at consistency in the compound – simple lottery choices. In Section 3.3.5, the distribution of consistency is examined. Section 3.3.6 examines parameterizing the data and the implications of the level of inconsistency found in subject data on expected utility, prospect theory, and loss aversion. Finally Section 3.4 concludes. Appendix 2 includes the subject questionnaire.

3.2. Experimental Design and Scaled Lottery Motivation

The experiment was conducted at Rutgers University in the Gregory Wachtler Experimental Economics Laboratory in the Fall of 2008. The subjects were undergraduate students from the Arts and Sciences and Business who enrolled through an online recruiting system. Subjects were given a series of preference questions between risky lotteries. The lotteries each consisted of two monetary prizes (positive and non-zero except where noted below) each with a mutually exclusive and exhaustive probability of the prize being received. These lotteries were replicated for eleven rows per question, with the first lottery remaining constant while the second lottery featured a single prize that increased per row. The replications thus increased the attractiveness of the second lottery option. Subjects were asked to respond by indicating a switch point, at which point the second lottery was preferred to the stable lottery. Responses of always preferring the left, or immediately preferring the right lottery were allowed, giving subjects a range of twelve possible responses per base choice question. A typical question would look like the following:

Question 1

$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 6.80,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 7.50,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 8.30,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 9.30,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 10.60,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 12.50,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 15,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 18,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 22,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 30,	$9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 40,	$9/10^{th}$ chance at .50

A sequence of twenty of these base choice questions was presented to the subjects. Within the sequence, there were four fundamental left hand side lotteries. Each of these left hand side lotteries was paired with one of two right hand side options. Twelve questions had a compounded lottery featuring two monetary prizes, with the larger monetary prize having a probability that varied per base question, and ranged from 90%, 70% 30% and 10%. The remaining eight questions consisted of a simple 50% probability of a dollar amount and a 50% probability of receiving no monetary prize. The risk neutral objective probability weighting break-even switch point was intentionally varied per question. The base probability choice questions were then presented as unique scaled variants and spaced throughout the survey. Finally, subjects were also asked to allocate 100 tickets between five lottery options.

Subjects were informed that a choice question and row would be selected at random, and subjects would then receive their preferred lottery choice. A second random number would determine the lottery results, with the subjects receiving any monetary prize, in addition to a small show up fee. The average earnings including the show up fee were in the range of $10. Full instructions and questions are included in the Appendix. Prior to the presentation of results, it is worth addressing the extent to which scaled question variants can be treated as essentially identical questions, and thus used as a measure of consistency. Choice questions all started with a static right hand side lottery with a probability p_{LHS} of receiving a prize a and if not receiving prize b with those associated probability (1- p_{LHS}). On the left would be a sequence of lotteries in which the prizes increased, while the probabilities remained constant: probability p_{RHS} of receiving a prize c_i and if not receiving prize d_i with those associated probability (1- p_{RHS}), with $c_{i+1} > c_i$ and $d_{i+1} > d_i$, with i increasing from 1 to 12. Subject (interior) responses framed the lottery of:

$$\mathbf{p_{RHS}(c_{j+1}) + (1- p_{RHS})(d_{j+1}) \geq p_{LHS}(a) + (1- p_{LHS})(b) \geq p_{RHS}(c_{j+1}) + (1- p_{RHS})(d_{j+1})}$$

Given the above formulation, scale indifference requires that:

$$\mathbf{p_{RHS}(c_{j+1}) + (1- p_{RHS})(d_{j+1}) \geq p_{LHS}(a) + (1- p_{LHS})(b) \geq p_{RHS}(c_{j+1}) + (1- p_{RHS})(d_{j+1}) \text{ if and only if}}$$
$$\mathbf{p_{RHS}(k^*c_{j+1}) + (1- p_{RHS})(k^*d_{j+1}) \geq p_{LHS}(k^*a) + (1- p_{LHS})(k^*b) \geq p_{RHS}(k^*c_{j+1}) + (1- p_{RHS})(k^*d_{j+1})}$$
for any scalar k > 0

Expected value maximization, expected utility theory and cumulative prospect theory all support scalar indifference – that individual choices should be consistent across scaled variants of lottery questions. To see this, it is worth noting that the preference interpretations of probability weights are all independent of the relative scale of the prized associated with the probability weight. Furthermore, so long as a power functional form can be attributed to the preferential evaluation of monetary prizes, we can then factor the scalar effect out of a question. In contrast to this, it has been proposed that the prospect weights in CPT should be expanded to include a systematic sensitivity to the scale and also the distance between prizes within a lottery. While full results follow, a few points here are worth noting. The responses found in this study are consistent with subjects demonstrating intentional consistency across scaled questions. This result holds when searching for exact consistency, or loosening to allow minor deviations or even comparing responses by population rank.

Furthermore, to the extent that a scale and level effect occurs, it should be systematic and demonstrate a level of internal consistency. As such, both aggregate and individual data should demonstrate response patterns that vary in opposite directions for opposite scalar effects. If mean responses increases as scale increases ($k > 1$) then the mean response should decrease when the question scale is reduced ($k < 1$). While two of four base questions demonstrated this trend, the other two demonstrated the opposite. Individual level data is even worse for a systematic scalar relationship: while a negative correlation between the changes in subjects' responses would be predicted as scale varies from increasing to decreasing, the exact opposite is found. Spearman rank correlation finds a positive correlation between the changes in responses when scaled is increased and when scale is decreased for all four base questions, with the correlation significant at the 5% level for 3 of the four questions using Sidak adjusted errors.

The study finds significant evidence that repeated scaled lottery questions are treated as the same base question by a significant proportion of the population, regardless the measure of consistency employed. In contrast, no evidence of a systematic relationship, on either the individual or aggregate level, between question responses and increasing or decreasing the scale of a question by a factor of 2. Full results follow.

3.3. Experimental Results

Without imposing a utility framework on the response data, consistency can be measured in several ways. First, the direct percentage of subject responding exactly consistent to a scaled question variant can be tested against the null hypothesis that responses are random. While this level of precision should be expected from a fully calculating and unerring subject, the questions allowed twelve possible answers per question. If subjects are consistent but exhibit a slight trembling when faced with a wide range of responses, then this view of consistency will be too narrow, and ultimately reject the null hypothesis for reasons other than subjects representing inconsistent utility preferences.

To combat this difficulty consistency is also defined to allow for small deviations while capturing if subjects exhibited a degree of choice stability. One such approach is to examine an ε-consistency, only viewing a response as inconsistent if it deviates beyond a preset threshold. As an alternative approach, consistency is also measured by examining the ordinal rank of the subjects' responses. This approach still enables the measuring of an individual's deviation from a previous response without the imposition of an evaluative framework. Instead, deviations are measured by rank in subject response data and thus the results are not dependent upon the supposition of a very system being examined.

An additional advantage of a rank based approach is that it enables direct response comparisons, not just for scaled question variants, but across all questions. If individuals are inconsistent when presented with similar situations, then we should expect this inconsistency to carry over to dissimilar questions. Alternatively, consistency should suggest that, while different fundamental questions should not a priori have the exact same response, individual choices should be generated by some fundamental preference relationship. As such, to the extent that the various questions are not pathological, a consistent ordering across subjects should be expected under any theory of individual choice. To the extent that individuals may be found to respond consistently to separate questions, and yet randomly across different questions, individual choice theories aggregating preferences must be flexible enough to account for this or abandoned.

3.3.1 Perfectly Consistent Response Patterns

The strictest measure of consistency is the percentage of subjects responding identically to scaled question variants. Unwavering subjects should respond identically to every scaled question variant, forming a constituent response triple when the fundamental base lotteries were paired with three different complex lottery scaled variants. A less strict measure of consistency then requiring identically to *every* question variant is the number of consistent response pairs per base question. It is worth noting that consistent question response pairs are not independent, even if subjects' responses are randomly generated: a subject that responds the same to question 1a and 1b, and to 1a and 1c, must necessarily respond consistently to 1b and 1c. The percentage of subjects responding consistently to all three scaled versions of a base question averaged 12.6%. All results presented are after excluding a single subject that responded the same response to every question. If this subject were included, the number would increase to 13.2%. The level of consistency was itself variable, with a ranging from 9.3% to 16.2% of the subject population. In contrast, if subjects were randomizing across all possible choices the expected rate of consistent response triples would be 0.7%, significantly less then found. Similarly, subjects responded consistent paired responses 24.5% of the time, while random responses would be expected to result in this less than 8.3% of the time.

To accurately assess the significance of these results, Monte Carlo analysis was conducted with 5,000 simulations of 153 subjects responding to the question variants. The simulation was done under two scenarios. Monte Carlo 1 used a random response distribution where each response was equally likely. Monte Carlo 2 used the average response distribution per question variant, thus recognizing the question specific response patterns and therefore an increased likelihood of matching responses. The number of subject consistent response patterns was recorded in Table 33 and Table 34 below, along with the 99th percentile for Monte Carlo 1, the 99th percentile for Monte Carlo 2, and the maximum value occurring under Monte Carlo 2 for consistent response triples and pairs respectively.

Table 33: Compound – Compound Lottery Consistent Response Triples

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Consistent Triples	14	4	6	8
Question 2 Consistent Triples	15	4	7	11
Question 3 Consistent Triples	25	4	7	10
Question 4 Consistent Triples	23	4	4	6

Table 34: Compound – Compound Lottery Consistent Response Pairs

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Consistent Pairs	92	53	65	78
Question 2 Consistent Pairs	107	53	67	81
Question 3 Consistent Pairs	134	53	72	85
Question 4 Consistent Pairs	117	53	63	75

For every question variant, the number of consistent response triples was greater than the maximum realization under Monte Carlo simulation, even using the subject response distribution. The hypothesis that the number of individuals demonstrating perfect consistency was derived from random responses across questions can soundly be rejected. Similarly, subject consistent response pairs were also greater than the largest realization under Monte Carlo simulation, and for all four scaled question variants. Again, the hypothesis that uncorrelated individual responses account for the consistency demonstrated can be rejected. The findings are consistent with expected utility and prospect theories prediction that scaled lottery switch questions are identical.

It is worth noting that each of these tests is top tail sensitive, with a small consistent subset of the subject body capable of driving the results. For instance, for question variant 4, the 23 consistent triples accounted for almost 60% of the total consistent response pairs, and exceeded the 99th percentile of the Monte Carlo Analysis for the entire subject body. To check the robustness of rejecting random responses, the same Monte Carlo analysis was conducted for the number of subjects with at least one consistent response pair per question variant (out of a possible three). This test thus examines the probability of a subject not responding consistently to any question pair, or a complete inconsistency. Results support the previous findings and are presented in Table 35. For every question variant, the number of subjects demonstrating at least one consistent response exceeds the maximum Monte Carlo value. Once again the null hypothesis of random subject responses can be rejected.

Table 35: Consistent Response Pairs to a Compound – Compound Lottery

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Subjects with a Consistent Pair	64	49	58	63
Question 2 Subjects with a Consistent Pair	77	49	59	67
Question 3 Subjects with a Consistent Pair	84	49	62	72
Question 4 Subjects with a Consistent Pair	71	49	57	64

While it is not possible to identify an implicit trembling rate without a distributional assumption on responses when subjects tremble, the data suggests that individuals varied from a true response at a rate in excess of 50%. The simplest assumption is to simply rule out the possibility of a consistent response occurring, through multiple trembles. This assumption is a reasonable approximation when we are considering the possibility of a consistent response triple, which would require three of the exact same response trembles out of 12 possible responses. Under this simplifying assumption, the data suggest a 50% tremble rate from the percentage of consistent triples, and a similar 50.5% from the total number of consistent response pairs.

However, if we accept a 50% tremble rate as a basis for analysis, then it makes sense to be concerned with the possibility of a consistent response pair occurring through two matching trembles. Assuming random responses when subjects tremble, the implied tremble rate from consistent responses climbs to 53%. Weighting small trembles more likely then large trembles further increases the implied rate. For example, a 20% chance of coinciding trembles increasing the implied rate to almost 58% and a 25% chance of coinciding trembles increasing the implied rate to 61% from consistent pairs, while the rate implied by consistent triples increases to 52.5%. A third estimate of the implied tremble rate can be obtained from the percentage of subjects demonstrating at least one consistent response pair. Using the percentage of subjects with at least one consistent response pair, the implied tremble rate is 55.3% assuming random responses when trembling, 65.1% with a 20% chance of coinciding trembles and 78.6% with a 25% chance of coinciding trembles. The different tremble rates implied by the size of the top and bottom tails of the distribution have important implications for the conditional distribution of trembling; for the consistency of consistency. This topic will be revisited later, however it is worth noting that consistent responses do not appear independent random events, but instead consistent responses are found to increase the likelihood of further consistent responses.

3.3.2. Epsilon Consistent Response Patterns

An alternative loosening of consistency is to look for the number of consistent response patterns, allowing for small nearby deviations. When faced with a choice between two compound lotteries, it is reasonable that subjects might approximate the value of the alternatives, and thus demonstrate small deviations. Defining near consistency as a response within a two response range of another (out of 12 total possible responses), the above analysis and Monte Carlo simulation is repeated below. Table 36 reports the number of near consistent triples, Table 37 reports the number of near consistent response pairs and Table 38 reports the number of subjects with at least one near consistent response pair to the compound-compound lottery choice questions.

Table 36: Compound – Compound Lottery Epsilon 2 Consistent Response Triples

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Epsilon 2 Consistent Triples	47	27	38	43
Question 2 Epsilon 2 Consistent Triples	44	27	28	36
Question 3 Epsilon 2 Consistent Triples	53	27	25	31
Question 4 Epsilon 2 Consistent Triples	71	27	37	41

Table 37: Compound – Compound Lottery Epsilon 2 Consistent Response Pairs

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Consistent Pairs	235	198	225	243
Question 2 Consistent Pairs	225	198	193	207
Question 3 Consistent Pairs	240	198	180	194
Question 4 Consistent Pairs	295	198	223	234

Allowing for small deviations considerably increases the level of consistency seen. While only 12.6% of the subject body responses were perfectly consistent to three scaled variants of the same compound-compound lottery choice question, this number rises dramatically to over 35% when allowing small deviations up to two responses. The number of near consistent triples is greater than the maximum Monte Carlo simulation value for every scaled question variant (Table 36), so once again we can reject the hypothesis that the number of individuals demonstrating consistency was derived from random responses across questions. Furthermore, near consistency occurs more often than not, with 54.2% of all response pairs near consistent (Table 37). Every scaled question variant is above the 99th percentile for Monte Carlo 1 and Monte Carlo 2. Furthermore, three of four scaled question variants had more near consistent response pairs then the maximum Monte Carlo realization. It is worth noting that 64.8% of all near consistent responses are from near consistent response triples, and just as with a stricter definition, subject consistency it is a top driven phenomena.

To address the robustness of near consistency, the bottom of the distribution was tested employing Monte Carlo analysis for the number of subjects with at least one near consistent response pair. Since near consistency allows any response within a range of two, the possibility of three responses constrained to a range of 12, and with none within two of each other is rapidly diminishing number. Results are in Table 34 and generally support the rejection of the null hypothesis of random subject responses, albeit not as conclusively as the top of the distribution.

Three of four scaled question variants having more subjects demonstrate one near consistent response pair than the 99th percentile for Monte Carlo 1. Two of the scaled question variants were significant at the 99th percentile for Monte Carlo 2, one was not significant at the 99th but was at the 95th percentile. The last scaled question variant was not significant at the 90th percentile for either Monte Carlo.

Table 38: Subjects with Epsilon 2 Consistent Response Pair to a Compound – Compound Lottery

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Subjects with a Consistent Pair	137	134	138	142
Question 2 Subjects with a Consistent Pair	136	134	132	138
Question 3 Subjects with a Consistent Pair	136	134	129	137
Question 4 Subjects with a Consistent Pair	127	134	139	144

By increasing the definition of consistency, there has been a corresponding decrease in the implied (large) tremble rate. Based on near consistent triples with zero probability attached to three near consistent large trembles, the implied tremble rate is reduced to a still robust 29%. Again, examining near consistent response pairs and making distributional assumptions about responses when trembling can increase the implied tremble rate. Although loosening the view of consistent action reduced the tremble rate, the rate remains sizable as most trembles are not small deviations, but represent large differences in revealed preferences.

3.3.3 Rank Based Consistent Response Patterns

While perfect and near consistent response pairs and triples provide important insights into the extent that subject responses demonstrate stable underlying preferences, they do have two drawbacks. First, consistent and near consistent measures are limited in evaluating the distance between responses to a simple binary variable, consistent or trembles. This eliminates significant data, and creates a dependence on the threshold for measuring consistency. The second major problem with examining the consistent response rate is that it is incapable of comparing consistency across dissimilar questions, when there is no reason to expect specific responses a priori.

Rank based tests provide a solution to both of these problems. Comparing subject rank within subject responses measures the degree of variability in subject responses. As such, it allows a comparison of similar question variants and can be used to examining the extent to which scaled question variants are treated as a single fundamental choice problem without imposing a theory of individual choice. Using Spearman's rank correlation coefficients, Table 39, eight out of twelve coefficients for scaled question variants are positively significant at the 5% level, and six of these are significant at the 1% level using Sidak-adjusted errors. The scaled question variants have an average Spearman's rank correlation of .2985, with a lack of consistency on one scaled question variant resulted in two of the four rejections. While not as compelling as the consistent response pattern data, Spearman's rank correlation results support the belief that scaled question variants are comparable choice questions.

Table 39: Spearman Rank Coefficients for Compound – Compound Scaled Question Variants

	1a	1b	1c	2a	2b	2c	3a	3b	3c	4a	4b	4c
1a	1.00											
1b	**.3378** ** **(.0013)**	1.00										
1c	**.1040 (1.000)**	**.1237 (.9999)**	1.00									
2a	.0351 (1.000)	-.0128 (1.000)	-.1162 (1.000)	1.00								
2b	.1567 (.9726)	-.0272 (1.000)	-.2107 (.4468)	**.4353** ** **(.0000)**	1.00							
2c	.0860 (1.000)	-.0637 (1.000)	.0228 (1.000)	**.3864** ** **(.0001)**	**.1426 (.9955)**	1.00						
3a	.0370 (1.000)	.0347 (1.000)	-.0919 (1.000)	.3421** (.0010)	.3110** (.0060)	.0853 (1.000)	1.00					
3b	-.1277 (.9997)	-.0494 (1.000)	.0666 (1.000)	.1714 (.8987)	.3199** (.0037)	-.0934 (1.000)	**.1699 (.9095)**	1.00				
3c	-.2147 (.3464)	-.0580 (1.000)	-.1498 (.9878)	.1907 (.7033)	.1344 (.9989)	.0856 (1.000)	**.2796** * **(.0303)**	**.2953** * **(.0138)**	1.00			
4a	.2692* (.0492)	.3184** (.0040)	.1118 (1.000)	-.0214 (1.000)	.0877 (1.000)	.0502 (1.000)	.0120 (1.000)	-.1378 (.9979)	-.2604 (.0731)	1.00		
4b	.3343** (.0016)	.2936* (.0151)	.0445 (1.000)	.1111 (1.000)	.0210 (1.000)	.1314 (.9994)	-.0629 (1.000)	-.2173 (.3703)	-.2820* (.0269)	**.4530** ** **(.0000)**	1.00	
4c	.1678 (.9231)	.1222 (.9999)	.1262 (.9998)	.0281 (1.000)	.0298 (1.000)	.1056 (1.000)	-.0306 (1.000)	-.0681 (1.000)	-.2503 (.1123)	**.4233** ** **(.0000)**	**.4312** ** **(.0000)**	1.00

Sidak-adjusted significance level is displayed in parenthesis following the coefficient. Scaled question variant correlations are bolded. Coefficients significant at the 5% level have an asterisk, and coefficients significant at the 1% level have a double asterisk.

In contrast, when comparing different base questions, there appears to be very little consistent relationship to the distributional rank of subject responses. Out of a total of fifty-four cross question variant correlations, only eight were significant at the 5% level, and just five of these were significant at 1%, average (Table 35). Indeed the average Spearman's rank correlation when comparing questions that are not scaled question variants was effectively zero, at -.0002. The lack of rank consistent results for questions that aren't scaled variants, poses an interesting dilemma discussed in further detail in Section 6; to what extent can an aggregate theory of individual choice under uncertainty exist if there is no consistent pattern of responses to lotteries with different base probabilities and prize amounts. Central to expected utility is the idea of a constant exponential coefficient on monetary prizes and to prospect theory, that there exists a relationship between the prospect weights for different objective probabilities. This needs to be reconciled with the extent to which subject data across question variants appears random, or else these coefficients appear to be randomly assigned based on the specific situation and not a fundamental or meaningful component of an individual.

3.3.4 Compound-Simple Lottery Consistent Response Patterns

In addition to the twelve compound lottery preference questions representing four scaled variants, subjects were also asked eight compound – simple lottery preference choices. The fixed left hand side lottery was a scaled variant of one of the four base questions previously discussed, while the right hand side was now a lottery consisting of a 50% probability of a prize amount and a 50% probability of receiving no prize amount. While the compound – simple lottery contained positive probability of a zero payoff if the right hand simple lottery was chosen, the overall risk level of the question was consistent with the compound – compound lottery choice questions. Even in a strict max min approach, the difference between lotteries questions was minimal. The compound – compound lotteries had a lowest possibility prize of $0.25, which occurred for two separate questions, with an attached probability of 30% in one question variant, and 90% in another question variant; comparing with $0.00 occurring with probability 50% in the compound – simple lotteries.

Despite the fundamental similarity of the preference questions, results for these eight choice questions are included separated here because they are markedly distinct then the above results. Consistent responses and near consistent response pairs are presented in Table 40 and Table 41. Consistent response and near consistent response pairs for every fundamental lottery variant were above the maximum Monte Carlo realization, regardless of parameterization. Over

35% of all response pairs were fully consistent, and 63.7% were nearly consistent. Assuming no coinciding trembles, the implied error rate is close to 40% and 20% respectively. While allowing for matching trembles increases these rates, the compound – simple lottery reveals far more preference stability within scaled question variants than the same compound lottery paired with another compound lottery. It is also worth noting that over half of all errors remain even after increasing consistency to include responses within two of the original response.

Table 40: Compound - Simple Lottery Consistent Response Pairs

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Consistent Pairs	60	21	43	49
Question 2 Consistent Pairs	57	21	45	55
Question 3 Consistent Pairs	44	21	30	41
Question 4 Consistent Pairs	54	21	29	36

Table 41: Compound – Simple Lottery Epsilon 2 Consistent Response Pairs

	Subject Response Data	Monte Carlo 1, 99th Percentile	Monte Carlo 2, 99th Percentile	Monte Carlo 2, Max Value
Question 1 Consistent Pairs	91	72	81	88
Question 2 Consistent Pairs	97	72	86	92
Question 3 Consistent Pairs	97	72	74	81
Question 4 Consistent Pairs	105	72	82	92

Spearman results for the compound – simple lottery preference questions are presented in Table 42, and are directly analogous with the results in Table 39. All four scaled question variant pairs were significant at the 1% level, with an average correlation coefficient of .4474; showing significantly more preference stability when compared against an average coefficient of .2985 for the compound – compound questions. The largest distinction between the compound – compound and the compound – simple lottery preference questions appears not just when comparing scaled question variants, but when comparing across different base questions. While few of the cross base question correlations were significant with the compound – compound lottery questions, for the compound – simple questions 23 out of 24 total cross question correlations are significantly positive at the 5% level, and 21 of 24 at the 1% level. The average correlation coefficient across different base questions was .4060.

Table 42: Spearman Rank Coefficients for Compound – Simple Scaled Question Variants

	1d	1e	2d	2e	3d	3e	4d	4e
1d	1.00							
1e	**.4299** (.0000)**	1.00						
2d	.4907** (.0000)	.2292 (.1187)	1.00					
2e	.2761* (.0160)	.3420** (.0005)	**.4071** (.0000)**	1.00				
3d	.5248** (.0000)	.3771** (.0000)	.5329** (.0000)	.2803* (.0131)	1.00			
3e	.4390** (.0000)	.3898** (.0000)	.3386** (.0006)	.4062** (.0000)	**.4288** (.0000)**	1.00		
4d	.5220** (.0000)	.3267** (.0011)	.4253** (.0000)	.3289** (.0010)	.5080** (.0000)	.4062** (.0000)	1.00	
4e	.3950** (.0000)	.4272** (.0000)	. 3430** (.0004)	.4003** (.0000)	.4429** (.0000)	.5506** (.0000)	**.5237** (.0000)**	1.00

Sidak-adjusted significance level is displayed in parenthesis following the coefficient. Scaled question variant correlations are bolded. Coefficients significant at the 5% level have an asterisk, and coefficients significant at the 1% level have a double asterisk.

While scaled questions are still more likely to receive similar responses than different base questions, the magnitude of this effect is dwarfed by the presence of a simple lottery alternative featuring a probability of receiving a zero dollar prize. As measured by rank based tests, the individual consistency in the compound – simple lottery questions was even greater than the scaled question consistency for the compound – compound questions. This is in sharp contrast to different compound – compound base lottery questions, where no consistent aggregate behavior emerged. While it is difficult to detangle if this consistency emerges as a result of increased calculating ability or because of risk aversion related to receiving zero, it is worth noting that over 73% of consistent responses occurred with a stated preference of always receiving the compound lottery. This choice represents a drop of up to $4.5, and a reduction of 22-30% in expected value for the last row depending on the question variant.

3.3.5 Consistency of Consistency

While any measure of consistency has as its ultimate goal to state a precision for model estimates we might expect in response to a new study, we must first start by examining the raw responses of the subjects. To the extent to which raw responses are variable, the ensuing coefficient estimates will inevitable be the same. This effect is felt within a single coefficient model such as expected utility with a power function, but is complicated and magnified by two parameter models such as PPT and the addition of a loss or regret coefficient. Two (or more) parameter models must rely on multiple responses to separate questions to calibrate coefficients. While additional variables inevitable add explanatory value to noisy data, the result is that a small set of questions with underlying individual variance can create a large and highly variable span of parameters. A full discussion of this impact can be found in Birchby and Sopher 2007, unpublished.

While less than 25% of all response pairs were perfectly consistent for any scaled question variants, the probability of a third consistent scaled variant increased dramatically in the subjects with a consistent response. Taking any scaled question pairing, subjects with a consistent response had a third consistent response 51% of the time, or more than doubling the average rate of consistency. Restricting the analysis to the question pairs with the smallest number of consistent responses, yields a consistent response triple in 77 out of 121 possible subjects or a statistically significant jump in consistency to over 63%.

A similar pattern is repeated when examining near consistency. Here 54.2% of all response pairs were near consistent for any scaled question variants, and the probability of a third near consistent scaled variant response increased to 64.8% when examining subjects with a single response near consistent response. Once again when restricting the analysis to the question pairs with the smallest number of near consistent responses, the rate of consistency conditional on having a previous consistent response jumps to 70%. While the appropriate formulation may be debated, consistency and near consistency within a scaled question variant increase the chance of additional consistency, suggesting a stable individual characteristic worthwhile to estimate.

Previous sections have established that the within question consistency occurs far more than any random response pattern would predict, but that there seemed to be little to no relationship between responses across different base questions (as demonstrated in Table 39). The goal for any decision-making theory is to achieve predictive power not just for a single question, but also across questions. Towards this end, a natural question is: to what extent is consistency persistent not just within a scaled question variant, but also across different base questions? Looking at the distribution of consistent and near consistent response triples in Table 43, there is a pattern of consistency suggesting an individual persistence and dependence. An independent binary variable measuring consistent response triples would be expected to have a dramatic fall off as the number of consistent triples increased from 2 out of 4, to 3 and then 4out of 4. In contrast, the data in Table 43 have a clearly thick tail, with little difference in the count between 2, 3, and 4. Similarly, over two thirds of all subjects were not even consistent for a single response triple; a mass disproportionate with the general rate of consistency.

Table 43: Distribution of Consistent and Near Consistent Response Triples

	0 Response Triples	1 Response Triple	2 Response Triples	3 Response Triples	4 Response Triples
Fully Consistent	103 Subjects	34 Subjects	8 Subjects	5 Subjects	3 Subjects
Near Consistent	35 Subjects	58 Subjects	30 Subjects	23 Subjects	7 Subjects

Given the above intuition, it is not surprising that both top and bottom distribution tests reject that consistent response triples are independently drawn from a binary distribution – confirming that consistency is a persistent individual trait. Using the observed frequency of consistent triples for a binary test of the probability of having conditional distribution given in

Table 43 the odds of at least 103 subjects not having a single consistent triple are 1.4%, while the odds of at least 3 subjects having all four triples consistent is 0.001%. At the 98% and 99% power, independence of consistency is rejected for fully consistent response triples across different base questions. The power with which the null hypothesis of independent consistency can be rejected is similar for the top of the distribution even when using the considerably lessened standards of near consistency. The experimental near consistency rate associates a probability of less than 1% for at least 7 subjects to have 4 near consistent responses triples. The bottom of the distribution loses some of its power when employing near consistency: a one-sided test rejecting independence at the 10% level, but not at the 5% level. Against a two-sided alternative, independence can still be rejected even here at the 5% level.

The distribution of consistent question pairs and near consistent question pairs for the scaled question variants are presented in Table 44. Once again the upper end of the distributions resoundingly rejects independence of consistent question pairs, at the 99% level. However, two very different patterns of results emerge. Similar to the consistent response triples, the fully consistent question pairs are distributed with a median (and mode) at 2 consistent pairs and a long persistent tail. In contrast, the near perfect response pairs are distributed far more evenly, with a center at 6 near consistent response pairs, and declining tails in both directions. Shapiro-Francia and Shapiro-Wilk tests of the normality of the response pair distributions reject the normality for consistent response pairs at the 1% level, but fail to reject the null for near consistent responses (Table 45 and Table 46). The broad measure of responses within two for individual questions is essentially a random variable, with the noise of inconsistent respondents drowning the signal of consistent individuals that are captured by the more strict measures of full consistency or near consistency across all three versions of a scaled question.

Table 44: Distribution of Consistent and Near Consistent Response Pairs

	0 Pairs	1 Pair	2 Pairs	3 Pairs	4 Pairs	5 Pairs	6 Pairs	7 Pairs	8 Pairs	9 Pairs	10 Pairs	11 Pairs	12 Pairs
Fully Consistent	15	35	38	18	16	10	6	6	1	1	4	0	3
Near Consistent	0	1	7	10	18	22	24	18	22	6	13	6	7

Table 45: Shapiro-Francia W' test for Normal Data Consistent and Near Consistent Response Pairs

	Observations	W'	V'	Z	Prob > z
Pairs: Consistent	153	.88500	14.772	5.322	.00001
Pairs: Near Consistent	153	.99077	1.185	0.356	.36090

Table 46: Shapiro-Wilk W' test for Normal Data Consistent and Near Consistent Response Pairs

	Observations	W'	V'	Z	Prob > z
Pairs: Consistent	153	.87137	15.223	6.179	.00000
Pairs: Near Consistent	153	.99187	0.962	-0.087	.53476

3.4 Implications and Areas for Further Research

While much of the above analysis has endeavored to answer the question of how consistent are individuals without imposing a model of decision making, examining response data within a given model provides the structure needed to answer the degree of variance in the individual data. Subject responses were used to derive implied coefficients for both a power function utility formulation, and separately a cumulative prospect theory weighting function. While the level of individual variability in the face of repeated scaled lottery questions suggests the presence of significant noise in any coefficient estimates, the aggregate data is consistent with similar studies, with most subjects demonstrating some risk aversion and an inverted S-shaped set of prospect weights. For an expected utility power coefficient framework, the individual coefficients consistent with responses were estimated. Individual questions had a broad range of possible implied coefficients, with the average range from .40 to over 2.1. The questions were intentionally not centered on risk neutrality, so the median response possible translated into a power coefficient ranging from .45 to 1.2. Similarly, a Prelec CPT weighting function coefficient was estimated.

Individual average coefficients for the power function and the CPT probability weights varied significantly both across and within individuals. The mean and median power function coefficient based on the compound-compound lottery questions was 1.015 and .98 respectively. Including the compound-simple lottery questions reduced the mean and median coefficients to .944 and .934. Similarly, the average CPT weighting coefficient for compound-compound lottery questions had a mean value of .775 and a median value of .777. Here, the inclusion of the compound-simple questions again reducing the mean and median coefficient, to .664 and .625.

While individual variability will still remain with expanded parameterization of choice theory from expect utility to a CPT framework, to what extent can prospect weights reconcile cross-question variance in individual responses? To best answer this question, the analysis could be restricted to only subject data that is consistent within scaled questions. The alternative is to use all the data, and risk mapping average responses that may not be economically meaningful. Requiring full response consistency for at least two question variants reduces the data pool to 16 of 153 subjects who were fully consistent for at least 2 question variants. For these subjects, 8 were for 2 triples, 5 were for 3 triples, and 3 were for all 4 triples. Since a prospect weighting and value function require two responses to jointly parameterize, this approach is reduced to examining the 8 subjects with consistent responses to more than two scaled base questions. The difficulty entailed in this, is that these are just the individuals who are most likely to follow a strict algorithm when selecting responses; and indeed 5 of 8 subjects demonstrated coefficients across every question consistent with risk neutral expected value maximization.

When including all subject responses, the joint distribution of individual coefficients and the underlying individual variance of these coefficients is revealing. Table 47 displays the joint distribution of standard deviation of individual power coefficient estimates and the average coefficient of subjects within the given range of variability. Similarly, Table 48 provides this analysis for the Prelec weighting function for a CPT approach.

Table 47: Joint Distribution of Power Coefficient Estimates and Standard Deviations

Standard Deviation of Power Coefficient	Subject Count from Compound Questions	Average Power Coefficient	Subject Count from All Questions	Average Power Coefficient
0 to .1	4	1.01	5	1.00
.1 to .2	4	.99	3	.99
.2 to .3	7	.93	11	.88
.3 to .4	12	.85	33	.86
.4 to .5	28	.90	40	.87
.5 to .6	34	.96	25	.94
.6 to .7	29	.99	9	1.03
.7 to .8	12	1.06	10	1.05
.8 to .9	4	1.36	9	1.12
.9 to 1.0	11	1.31	5	1.22
Greater than 1.0	8	1.45	3	1.29

Table 48: Joint Distribution of CPT Estimates and Standard Deviations

Standard Deviation of Weighting Coefficient	Subject Count from Compound Questions	Average Weighting Coefficient	Subject Count from All Questions	Average Weighting Coefficient
0 to .1	8	1.00	7	1.00
.1 to .2	7	.92	4	.99
.2 to .3	16	.84	5	.84
.3 to .35	18	.79	2	.78
.35 to .4	27	.71	4	.68
.4 to .45	29	.70	16	.78
.45 to .5	23	.77	10	.63
.5 to .55	14	.76	14	.60
.55 to .6	6	.80	22	.57
.6 to .65	5	.81	21	.55
.65 to .7	-	-	16	.48
.7 to .75	-	-	10	.57
Greater than .75	-	-	22	.79

An interesting pattern emerges, as the most stable individuals correspond with expected utility risk neutral maximization. After this, an increase in variability corresponds with a steady decline in the estimated power coefficient. Finally, as variability continues to increase, the estimated power coefficient increases substantially. Similarly, when interpreted as subjective prospect weights, the more variable a subject is, the lower the Prelec coefficient and the greater the effect of an inverted S-Shaped set of probabilities the subject's responses generate on average. Thus, as predicted in the introduction, more variable subjects are parameterized by a set of prospect weights with mean reversion, overweighting low probabilities and underweighting high probabilities. Given the nonexistent correlation between responses to different base questions and especially in light of the significant correlation to scaled question variants), there is little alternative but to interpret this as the effect of the variance in individual responses and not the expression of a stable systematic individual approach to uncertainty.

While consistency does emerge in the compound – simple lottery response choices, this stability cannot be attributed to the simplified prospect weights but rather appears to clearly be a risk aversion based phenomena. The average standard deviation of CPT coefficient across the compound-compound lottery questions was 0.38, this figure rises dramatically to 0.67 (0.57 including all). In contrast, the power coefficient standard deviation drops from 0.58 to 0.17 when comparing the compound-compound and compound-simple lottery questions (0.49 including all). Finally, the effects of different base questions can be seen in the response data. The base question sensitivity of the data can be controlled for by adjusting the mean response of each question to equal the subject's average overall coefficient. This approach significantly reduces variance, with the overall remaining standard deviation reduced to slightly over half its original value. Factoring out the effects of a given scaled question variant has some success in decreasing the overall variability of subjects responses, as displayed in Table 49.

Table 49: Joint Distribution of Power Coefficient Estimates and Standard Deviations Controlling for Questions

Standard Deviation of Power Coefficient	Subject Count	Average Power Coefficient	Subject Count with Question Adjusted	Average Power Coefficient
0 to .1	4	1.01	10	.99
.1 to .2	4	.99	26	.86
.2 to .3	7	.93	35	.93
.3 to .4	12	.85	31	.98
.4 to .5	28	.90	25	1.02
.5 to .6	34	.96	15	1.28
.6 to .7	29	.99	6	1.27
.7 to .8	12	1.06	2	1.46
.8 to .9	4	1.36	3	1.64
.9 to 1.0	11	1.31	-	-
Greater than 1.0	8	1.45	-	-

3.6 Conclusion

Significant individual variance exists within responses to lottery preference questions. The variance is not restricted to small trembles, with half of all response inconsistencies based on large deviations. Subjects are significantly more consistent within a scaled lottery framework than across different lottery preference questions, and true consistency across scaled lotteries cannot be rejected. The level of consistency itself appears a consistent individual trait, with consistent responses predictive of future consistent responses, and inconsistent responses increasing the likelihood of further inconsistency. Fitting the response data with a Prelec weighting function and a power coefficient demonstrates the significant underlying variability, both across and within subject. A small subsample of subjects demonstrate significant consistency and coefficients consistent with expected value maximizing. As individual variance increases, coefficient estimates find a greater level of risk aversion and subjective probability weights in the standard inverted S-shape. These weights are a priori expected given the variability of the underlying responses, and as such it is unclear what economic interpretation to give the coefficients besides capturing this variability.

Consistency increased dramatically when the lottery choice preference included a simple lottery option, with the possibility of receiving a zero monetary prize. Subject responses created a sharp distinction between this compound-simple lottery question and one with lower prizes ranging as low as 25 or 50 cents. While there was little to no aggregate consistency across different compound-compound lottery choice questions, this effect dissipated with significant cross question correlation occurring once a zero prize was introduced. The consistency reflected a fundamental risk aversion, with a significant subsample of subjects always rejecting a higher expected value lottery that includes a zero monetary prize, even as the expected value increases significantly over the alternative risky lottery. Describing these preferences is more successful with a power function approach than with CPT prospect weights: fitting these ‘stable’ preferences with variable subjective prospect weights actually increases the variability of the coefficients estimates. While the power function has significantly more success here, it would appear that the displayed preferences are representative of a heuristic rule to avoid a zero monetary payoff than a temporary reduction in the coefficient of risk aversion. Modeling an individual zero dollar disutility similarly yields limited success.

While significant individual variance in lottery preference questions exists, there are several positive claims that can be made. Individuals are consistent across scaled lottery preferences. Consistency is a consistent individual trait. The presence of a zero prize greatly increases consistency. Yet without the zero prize, responses across different base questions are inconsistent, and question dependent for all but the most consistent, expected value maximizers. Lastly, an inverted S-Shaped set of probability weights emerges as the dominant pattern of variable subjects, but it is impossible to attribute this pattern to an underlying preference and not the variance itself.

Conclusion

In the preceding chapters I studied individuals making decisions with uncertainty both in individual and in groups settings, with a particular focus on coordination. I explored this through a series of experiments where subjects were asked to coordinate with themselves to respond consistently in evaluating financial gambles and with others to share information and maximize group performance. It was of particular interest to tease apart the significance of the types of messages sent and the structure of the networks of communication used by the subjects to achieve coordination and information aggregation.

These chapters ultimately addressed the application of experimental methods to theories and the resulting findings showed where theoretical predictions held true, where they were violated by subject behavior, and what happens when we as economists have no clear predications about actions. This set of research brings us one step further understanding decision-making behavior and how subjects react to uncertainty.

Appendix 1: Experiment Instructions

Introduction

This is an experiment about decision making. Funding for this experiment has been provided by the National Science Foundation and Rutgers University. You will be paid $5 for your participation plus an additional amount which depends upon the decisions that you make and upon random luck. Please read these instructions carefully so that you understand how your decisions help to determine your earnings. The currency in this experiment is the Experimental Currency Unit, or ECU. At the end of the experiment your earnings in ECUs will be converted into U.S. Dollars at the rate of $1 per 100 ECUs, and you will be paid this amount in cash before you leave the experiment.

General Instructions

In the experiment you will be making a series of decisions about whether or not to adopt a particular *policy*. In each of several decision-making rounds you will be organized into groups of five people. Each group will decide whether to adopt the policy by *majority vote*. That is, if three or more people in your group vote to adopt the policy, then it will be adopted. A policy will specify a monetary payment to be made to each individual member of a group. The payment may differ, depending upon the policy turns out to be *successful* or *unsuccessful*. Every member of a group which has adopted a policy will receive monetary payoffs according to the policy, regardless of whether a particular member of the group voted for the policy or not. If a group does not vote to adopt a policy, then each member of the group will receive a fixed monetary payment for that round. Your payoffs in a particular round will depend only upon the decisions made by members of your group and upon random factors specific to your group.

Payoffs

The table below shows an example of how the payments in ECUs depend upon whether or not the policy is adopted, and on whether or not, if adopted, the policy is successful.

	Policy Unsuccessful	Policy Successful
Policy Adopted	0 ECUs	100 ECUs
Policy NOT Adopted	50 ECUs	50 ECUs

Notice how the payment to each member of a group varies if the policy is adopted: 0 ECUs if the policy is unsuccessful, 100 ECUs if it is successful. Also notice that the payment if the policy is not adopted is always 50 ECUs.

Probability, or Chance, that the Policy is Successful

An important fact that determines whether the policy will be successful is the *Probability*, or *Chance*, that the policy is successful. We will refer to this probability by the letter *p,* and we will express it as a number which may take on any value from 0 to 1. A value of p=0 corresponds to a 0% chance that the policy is successful, and a value of p=1 corresponds to a 100% chance of success. In general, the value of p will be neither 0 nor 1, but rather some value strictly between 0 and 1. Clearly, the larger is the value of p, the better are the chances that the policy will be successful and yield a payoff of 100 ECUs. Specifically, if p is larger than ½, then you can expect to earn more if the policy is adopted than you will earn if it is not adopted.

Signals about the true value of p

Before your group votes on whether to adopt the policy, each member of your group will receive an informative *signal* about the true value of p. The true value of p has been determined in advance by a random procedure. The signals that the members of your group receive about the true value of p are distributed about the true value of p in such a way that the average of the signals should be a good predictor of the true value of p. That is, although the average of the 5 signals for your group will not necessarily equal the true value of p, the average of a much larger number of signals would tend to equal the true value, and would be more and more accurate as the number of signals that we average over became larger.

Payoff Adjustments

In addition to the payments you receive that depend upon whether the policy is adopted and upon whether the policy is successful or not, there will be other Payoff Adjustments that only depend upon whether the policy is adopted, and not upon whether the policy is successful. Your Payoff Adjustment may be either positive or negative, and thus will either increase both possible payments associated with the policy or it will decrease both of these possible payments. A payoff adjustment is equally likely to be 25, 10, 0, -10 or -25 ECUs for each member of each group, independent of one another. Thus, denoting your Payoff Adjustment by the letter A, the full payoff table is as below:

	Policy Unsuccessful	Policy Successful
Policy Adopted	0 + A ECUs	100 + A ECUs
Policy NOT Adopted	50 ECUs	50 ECUs

Communicating with other Members of your Group

In each round, after you have received your signal but before you vote, you may also have the opportunity to communicate with some, but possibly not all, members of your group. In each round you will be informed of which members of your group you may communicate with. You will also be told the overall structure of who can communicate with whom in each round.

Notice that each member of your group will have received a signal about the true value of p. You are free to communicate about your signal, about how you intend to vote, or anything else that you think would be useful to communicate prior to voting on whether to adopt the policy. The only restriction is that you may not reveal your identity in your messages.

Voting

After a communication period of 3 minutes you will then be asked to decide on whether to adopt the policy by voting to adopt or not to adopt the policy. If your group votes to adopt the policy, then the true value of p for your group for that round will be used to randomly determine whether the policy is successful or not. Each member will then receive the payment specified depending on whether policy is successful or not. This payment may also include an adjustment payoff, as described above. If your group votes not to adopt the policy, then you will each receive the fixed payment specified for that round for not adopting the policy.

At the end of each round your you will be shown how your group voted, what your earnings for that round are, and what your cumulative earnings over the course of the experiment are. At the start of each round you will be given information specific to that round: the communication structure, your adjustment payoffs, etc. In each round the groupings of participants into groups of 5 will be determined randomly, so that you will, in general, be in a different group in each round. You will be assigned a randomly determined ID number in each round of the experiment, so you will not be able to tell who is in your group by their ID number. At the end of the experiment you will be paid your earnings in cash.

We will now answer any questions that you have about these instructions.

Appendix 2: Experiment Instructions

You are about to undertake an experiment in decision making. Following there will be 20 decision questions, in which you are asked to express your preference, between two possible prizes. The prizes are expressed as the form of a lottery, and will have 2 prizes each with an individual probability assigned to it. The prizes are mutually exclusive, and the total probability sums to one, so if the first prize is not achieved, the second prize will occur. The questions are of the following form:

3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 6.80,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 7.50,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 8.30,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 9.30,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 10.60,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 12.50,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 15,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 18,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 22,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 30,	9/10th chance at .50
3/10th chance at 4, 7/10th chance at 1:	1/10th chance at 40,	9/10th chance at .50

The lottery on the left remains constant, while the lottery on the right has a prize increase, as each row goes down. The only necessary response from you will be a **line**, with rows above the line symbolizing you prefer the lottery to the left, and rows below the line symbolizing you prefer the lottery to the right.

Your response will directly affect your pay. One of the 20 questions will be selected randomly, and then a row for that question will be selected randomly. The lottery you responded that you preferred will then be played, with a random draw determining which of the possible monetary prizes are awarded.

In addition, there are 2 final questions with directions given before them.

Question 1

$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 6.80, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 7.50, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 8.30, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 9.30, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 10.60, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 12.50, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 15, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 18, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 22, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 30, $9/10^{th}$ chance at .50
$3/10^{th}$ chance at 4, $7/10^{th}$ chance at 1:	$1/10^{th}$ chance at 40, $9/10^{th}$ chance at .50

Question 2

$9/10^{th}$ chance at 4, $1/10^{th}$ chance at 3	$7/10^{th}$ chance at 5.40, $3/10^{th}$ chance at .50
$9/10^{th}$ chance at 4, $1/10^{th}$ chance at 3	$7/10^{th}$ chance at 5.60, $3/10^{th}$ chance at .50
$9/10^{th}$ chance at 4, $1/10^{th}$ chance at 3	$7/10^{th}$ chance at 5.80, $3/10^{th}$ chance at .50

9/10th chance at 4, 1/10th chance at 3	7/10th chance at 6.00,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 6.20,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 6.50,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 6.80,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 7.20,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 7.70,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 8.30,	3/10th chance at .50
9/10th chance at 4, 1/10th chance at 3	7/10th chance at 9.00,	3/10th chance at .50

Question 3

1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 7.50,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 7.75,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 8.00,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 8.25,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 8.50,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 8.75,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 9.00,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 9.25,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 9.50,	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 9.75	3/10th chance at 4
1/10th chance at 25, 9/10th chance at 5:	7/10th chance at 10.00	3/10th chance at 4

Question 4

7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 15,	9/10th chance at 2

7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 20,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 25,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 30,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 35,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 40,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 50,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 60,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 70,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 80,	9/10th chance at 2
7/10th chance at 8, 3/10th chance at 5:	1/10th chance at 100,	9/10th chance at 2

Question 5

3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 13.60,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 15,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 16.60,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 18.60,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 21.20,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 25,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 30,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 36,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 44,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 60,	9/10th chance at 1
3/10th chance at 8, 7/10th chance at 2:	1/10th chance at 80,	9/10th chance at 1

Question 6

9/10th chance at 8, 1/10th chance at 6	7/10th chance at 10.80, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 11.20, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 11.60, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 12.00, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 12.40, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 13.00, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 13.60, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 14.40, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 15.40, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 16.60, 3/10th chance at 1
9/10th chance at 8, 1/10th chance at 6	7/10th chance at 18.00, 3/10th chance at 1

Question 7

1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 15.00, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 15.50, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 16.00, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 16.50, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 17.00, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 17.50, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 18.00, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 18.50, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 19.00, 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 19.50 3/10th chance at 8
1/10th chance at 50, 9/10th chance at 10:	7/10th chance at 20.00 3/10th chance at 8

Question 8

7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 30, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 40, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 50, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 60, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 70, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 80, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 100, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 120, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 140, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 160, 9/10th chance at 4
7/10th chance at 16, 3/10th chance at 10:	1/10th chance at 200, 9/10th chance at 4

Question 9

3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 3.40, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 3.75, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 4.15, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 4.65, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 5.30, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 6.25, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 7.50, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 9, 9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 11, 9/10th chance at .25

3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 15,	9/10th chance at .25
3/10th chance at 2, 7/10th chance at .50:	1/10th chance at 20,	9/10th chance at .25

Question 10

9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 2.70,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 2.80,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 2.90,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 3.00,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 3.10,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 3.25,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 3.40,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 3.60,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 3.85,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 4.15,	3/10th chance at .25
9/10th chance at 2, 1/10th chance at 1.50	7/10th chance at 4.50,	3/10th chance at .25

Question 11

1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 3.75,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 3.87,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.00,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.12,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.25,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.37,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.50,	3/10th chance at 2

1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.62,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.75,	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 4.87	3/10th chance at 2
1/10th chance at 12.50, 9/10th chance at 2.50:	7/10th chance at 5.00	3/10th chance at 2

Question 12

7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 7.50,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 10,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 12.50,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 15,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 17.50,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 20,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 25,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 30,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 35,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 40,	9/10th chance at 1
7/10th chance at 4, 3/10th chance at 2.50:	1/10th chance at 50,	9/10th chance at 1

Question 13

3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 7.50,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 8.25,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 9,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 9.75,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 10.50,	5/10th chance at 0

3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 11.25,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 12,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 12.75,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 13.50,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 14.25,	5/10th chance at 0
3/10th chance at 12, 7/10th chance at 3:	5/10th chance at 15,	5/10th chance at 0

Question 14

9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 7.50,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 8.25,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 9,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 9.75,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 10.50,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 11.25,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 12,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 12.75,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 13.50,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 14.25,	5/10th chance at 0
9/10th chance at 6, 1/10th chance at 4.50:	5/10th chance at 15,	5/10th chance at 0

Question 15

1/10th chance at 37.50, 9/10th chance at 7.50:	5/10th chance at 18.75,	5/10th chance at 0
1/10th chance at 37.50, 9/10th chance at 7.50:	5/10th chance at 19.50,	5/10th chance at 0
1/10th chance at 37.50, 9/10th chance at 7.50:	5/10th chance at 20.25	5/10th chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 21.00 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 21.75 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 22.50 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 23.25 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 24.00 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 25.50 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 27.00 $5/10^{th}$ chance at 0

$1/10^{th}$ chance at 37.50, $9/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 30.00 $5/10^{th}$ chance at

Question 16

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 18.75, $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 19.50, $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 20.25 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 21.00 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 21.75 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 22.50 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 23.25 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 24.00 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 25.50 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 27.00 $5/10^{th}$ chance at 0

$7/10^{th}$ chance at 12, $3/10^{th}$ chance at 7.50: $5/10^{th}$ chance at 30.00 $5/10^{th}$ chance at 0

Question 17

$3/10^{th}$ chance at 6, $7/10^{th}$ chance at 1.50: $5/10^{th}$ chance at 3.75, $5/10^{th}$ chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 4.12, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 4.50, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 4.87, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 5.25, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 5.62, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 6.00, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 6.37, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 6.75, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 7.12, 5/10th chance at 0

3/10th chance at 6, 7/10th chance at 1.50: 5/10th chance at 7.50, 5/10th chance at 0

Question 18

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 3.75, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 4.12, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 4.50, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 4.87, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 5.25, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 5.62, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 6.00, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 6.37, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 6.75, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 7.12, 5/10th chance at 0

9/10th chance at 3, 1/10th chance at 2.25: 5/10th chance at 7.50, 5/10th chance at 0

Question 19

1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 9.37,	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 9.75,	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 10.12	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 10.50	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 10.87	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 11.25	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 11.62	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 12.00	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 12.75	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 13.50	5/10th chance at 0
1/10th chance at 18.75, 9/10th chance at 3.75:	5/10th chance at 15.00	5/10th chance at 0

Question 20

7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 9.37,	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 9.75,	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 10.12	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 10.50	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 10.87	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 11.25	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 11.62	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 12.00	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 12.75	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 13.50	5/10th chance at 0
7/10th chance at 6, 3/10th chance at 3.75:	5/10th chance at 15.00	5/10th chance at 0

Using whole numbers, please allocate 100 tickets, or chances, between the following selection of lotteries:

Allocation amongst lotteries 1:

______ $1/10^{th}$ chance of 50, $9/10^{th}$ chance of 5

______ $2/10^{th}$ chance of 55, $8/10^{th}$ chance of 0

______ $4/10^{th}$ chance of 15, $6/10^{th}$ chance of 5

______ $5/10^{th}$ chance of 20, $5/10^{th}$ chance of 0

______ $10/10^{th}$ chance for 8

Allocation amongst lotteries 2:

______ $1/10^{th}$ chance of 30, $9/10^{th}$ chance of 3

______ $2/10^{th}$ chance of 15, $8/10^{th}$ chance of 6

______ $4/10^{th}$ chance of 12, $6/10^{th}$ chance of 6

______ $5/10^{th}$ chance of 10, $5/10^{th}$ chance of 7

______ $10/10^{th}$ chance for 8

References

Angeletos, G.-M. and A. Pavan (2005). "The Social Value of Information and Coordination," mimeo.

Angeletos, G.-M. and A. Pavan (2007). "Efficient Use of Information and Social Value of Information." *Econometrica*, 75(4), 1103–1142

Banerjee, A.V. (1992). "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107, 797-817.

Bikhchandani, S. D. Hirschleifer and I. Welch (1992). "A Theory of Fads, Fashion, Custom and Cultural Change as Information Cascades." *Journal of Political Economy*, 100, 992-1026.

Birchby, J., G. Gigliotti, and B. Sopher (2007). **"**Consistency and Aggregation in Individual Decision Making under Uncertainty," working paper.

Blume, A., and A. Ortmann (2007). "The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria." *Journal of Economic Theory*, 132, 274-290.

Brandts, J., and D.J. Cooper (2007). "Its what you say, not what you pay: An experimental study of manager-employee relationships in overcoming coordination failure." *Journal of the European Economic Association*, 5, 1223-1268.

Camerer, C., Q. Nguyen, and T. Tanaka, (2009). "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *The American Economic Review,* forthcoming.

Carbone, E., and J. Hey (2000). "Which Error Story is Best?" *Journal of Risk and Uncertainty, 20*(2), 161–176.

Celen, B. S. Kariv and A. Schotter (2005). "An Experimental Test of Advice and Social Learning." Mimeo, New York University.

Charness, Gary, and M. Rabin (2002). "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817-869.

Chaudhuri, A., A. Schotter and B. Sopher (2009). "Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Games with Private, Almost Common and Common Knowledge of Advice," *Economic Journal*, 119(534), 91-122.

Chaudhuri, A., A. Schotter and B. Sopher (2006). "Advice and Learning in Individual Choice Under Uncertainty,"

Chaudhuri, A., S. Graziano and P. Maitra (2006). "Social Learning and Norms in a Public Goods Experiment with Intergenerational Advice," forthcoming, *Review of Economic Studies.*

Choi, S., D. Gale and S. Kariv (2004). "Behavioral Aspects of Learning in Social Networks: An Experimental Study." Mimeo, New York University.

Choi, S., D. Gale and S. Kariv (2005). "Learning in Networks: An Experimental Study." Mimeo, New York University.

Erat, S. and U. Gneezy (2012). "White Lies" *Management Science,* 58(4), 723-733.

Etchart-Vincent, N. (2004). "Is Probability Weighting Sensitive to the Magnitude of Consequences? An Experimental Investigation on Losses." *Journal of Risk and Uncertainty, 28*(3), 217–235.

Etchart-Vincent, N. (2009). "Probability Weighting and the 'Level' and 'Spacing' of Outcomes: An Experimental Study over Losses. *Journal of Risk and Uncertainty,* forthcoming.

Falk, Armin, E. Fehr and U. Fischbacher (2007). "Testing theories of fairness-Intentions matter." *Games and Economic Behavior*, 62, 287-303.

Fehr, H., T. Epper, A. Bruhin and R. Schubert, (2007). "Risk and Rationality: The Effect of Incidental Mood on Probability Weighting," Working Paper #0703, University of Zurich (Switzerland).

Gale, D. and S. Kariv (2003). "Bayesian Learning in Social Networks." *Games and Economic Behavior*. 45, 329-346.

Goyal, S. (2005). "Learning in Networks" in Group *Formation in Economics: Networks, Clubs and Coalitions*. G. Demange and M. Wooders, eds., Cambridge University Press.

Granovetter, M. (1978). "Threshold Models of Collective Action." *American Journal of Sociology*. 83(6), 1420-1443.

Granovetter, M. (1985). "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology*. 91(3), 481-510.

Hey, J. (2005). **"**Why We Should Not be Silent about Noise." *Experimental Economics, 8*(4), 325–345.

Hey, J. and C. Orme (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data." *Econometrica, 62*(6), 1291–1326.

Jackson, M. and A. Watts. (2002). "The Evolution of Social and Economic Networks," Journal of Economic Theory, 106, 265-295.

Jackson, M. (2005). “A Survey of Network Formation Models: Stability and Efficiency,” in *Group Formation in Economics: Networks, Clubs and Coalitions*. G. Demange and M. Wooders, eds., Cambridge University Press.

Jacobsen, S. and Petrie, R. (2009). “Learning From Mistakes: What do Inconsistent Choices over Risk Tell Us?” *Journal of Risk and Uncertainty, 38*, 143–158.

Karni, Ed, T. Salmon, and B. Sopher (2008). “Individual Sense of Fairness: an Experimental Study,” *Experimental Economics*, Springer, 11(2), 174-189.

Kirilenko, A. (2005) “Network Knowledge and Convergence to Uniformity,” mimeo.

Kosfeld, M. (2004). “Economic Networks in the Laboratory: A Survey.” *Review of Network Economics*, 3, 20-41.

Loomes, G. (2005). “Modeling the Stochastic Component of Behaviour in Experiments: Some Issues for the Interpretation of Data.” *Experimental Economics, 8*(4), 301–323.

Loomes, G., P. Moffatt and R. Sugden, (2002). “A Microeconometric Test of Alternative Stochastic Theories of Risky Choice. *Journal of Risk and Uncertainty, 24*(2), 103–130.

Mesoudi, A., A. Whiten, and K. Laland (2006). “Towards a Unified Science of Cultural Evolution.” *Behavioral and Brain Sciences*, forthcoming.

Morris, S. and H. S. Shin (2002). “Social Value of Public Information.” *The American Economic Review*, 92(5), 1521-1534.

Nyarko, Y., A. Schotter and B. Sopher (2006). “On the Informational Content of Advice: A Theoretical and Experimental Study,” forthcoming, *Economic Theory*.

Prelec, D. (1998). “The Probability Weighting Function,” *Econometrica*, *66*(3), pp. 497-527.

Quiggin, J. (1982). “A Theory of Anticipated Utility,” *Journal of Economic Behavior and Organization, 3*, 323–343.

Rabin, Matthew (1993). “Incorporating Fairness into Game Theory and Economics,” *American Economic Review*, 83(5), 1281-1302.

Schotter, A. and B. Sopher (2003). “Social Learning and Coordination Conventions in Intergenerational Games: An Experimental Study,” (with Andrew Schotter), *Journal of Political Economy,* 111(3), 498-529.

Schotter A. and B. Sopher (2006). “Advice and Behavior in Intergenerational Ultimatum Games: An Experimental Approach,” forthcoming, *Games and Economic Behavior.*

Sopher, B. and G. Gigliotti (1993). **“**Intransitive Cycles: Rational Choice or Random Error? An Answer Based on Estimation of Error Rates with Experimental Data.” *Theory and Decision*. 35, 311–336.

Sopher, B. and J. Narramore (2000). “Stochastic Choice and Consistency in Decision Making under Risk: An Experimental Study.” *Theory and Decision*. 48, 323–350.

Smith, L. and P. Sorensen (2000). “Pathological Outcomes of Observational Learning.” *Econometrica*, 68(2), 371-398.

Tversky, A. and D, Kahneman, (1992). “Advances in Prospect Theory: Cumulative Representation of Uncertainty.” *Journal of Risk and Uncertainty, 5*, 297–323.

Van Huyck, J. B., R. C. Battalio and R. O. Beil, (1990). “Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure.” *American Economic Review*, 80, 234-248.

Van Huyck, J. B., R. C. Battalio and R. O. Beil, (1991). "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games." *Quarterly Journal of Economics*, 106, 885-911.

Weber, R. and C. Camerer (2003). "Cultural Conflict and Merger Failure: An Experimental Approach," *Management Science*, 49(4), 400-415.

Young, H.P. (1993). "The Evolution of Conventions," *Econometrica*, 61, 57-84.

Young, H.P. (1998). *Individual Strategy and Social Structure.* Princeton University Press, Princeton, NJ.