HETEROGENEITY AND EQUILIBRIUM

by

Jason Matthew Shachat

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ECONOMICS

In partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 7

HETEROGENEITY AND EQUILIBRIUM

by

Jason Matthew Shachat

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ECONOMICS

In partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 7

UMI Number: 9738958

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by JASON MATTHEW SHACHAT

entitled HETEROGENEITY AND EQUILIBRIUM

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

MARK WALKER — Date: Dec 18, 1996

JAMES C. COX — Date: Dec. 18, 1996

R. MARK ISAAC — Date: Dec 18, 1996

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director MARK WALKER — Date: Dec 18 1996

STATEMENT OF AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: [illegible]

TABLE OF CONTENTS

LIST OF FIGURES....6
LIST OF TABLES....7
ABSTRACT....8
MIXED STRATEGY PLAY AND THE MINIMAX HYPOTHESIS....9
1 Introduction....9
2 Review of the Previous Literature....12
2.1 Experimental Research of Two-Outcome Constant-Sum Games....13
2.1.1 O'Neill's Experiment and Results....13
2.1.2 Brown and Rosenthal's Re-examination of the Data....14
2.1.3 Rapoport and Boebel....15
2.1.4 Mookherjee and Sopher....17
2.2 Earlier Experimental Attempts to Directly Examine Mixed Strategies....17
2.3 Summary....19
3 Experimental Design....20
3.1 The Three Treatments....21
3.2 Experimental Procedures....23
4 Data Analysis....23
4.1 Realized Actions....24
4.1.1 Aggregated Data of Realized Actions....25
4.1.2 Individual and Pair Level Data of Realized Actions....31
4.1.3 Contemporaneous Play within Pairs....32
4.1.4 Intertemporal Correlation of Actions....36
4.2 Observed Mixed Strategies....48
5 Concluding Remarks....57
HETEROGENEOUS RISK ATTITUDES AND EQUILIBRIUM IN OBSERVABLE GAMES....63
1 Introduction....63
2 The Gamble/Safe Game....66
3 Implications of the BNE of the Gamble/Safe Game....71
3.1 M&P Solution of the Gamble/Safe Game....73
4 A Purification of a Gamble/Safe Game....75
5 Theoretical Structure and Results....76
5.1 Existence of Equilibrium....78
5.1.1 Compactness and Convexity of the Strategy Space....81
5.1.2 Properties of the Expected Payoff Function....82
6 Pure Strategy Equilibrium....85
7 Concluding Remarks....86
8 Proofs of Propositions....87
AN EXPERIMENTAL INVESTIGATION OF THE INCENTIVES TO FORM AGRICULTURAL MARKETING POOLS....88
1 Introduction....88

TABLE OF CONTENTS - *Continued*

2 Experimental Design....................91
3 A Game Theoretic Model of the Pooling Experiments....................96
4 Discussion of Aggregated Data....................98
5 Introduction to the Analysis of the Individual Level Data....................100
6 Parametric Analysis of the Individual Level Data....................102
7 Nonparametric Analysis of the Individual Level Data....................107
8 Conclusion....................111
9 Game Theoretical Framework....................112
9.1 Homo/low Treatment....................114
9.2 Florida Treatment....................115
9.3 California Treatment....................116
9.4 Hetero/high Treatment....................118
10 The Random Effects Probit Model....................120
11 Nonparametric Analysis of the Experimental Data....................122
REFERENCES....................133

LIST OF FIGURES

FIGURE 1.1 Standard Deviation of the Joker Card versus the Average Standard Deviation of the Non-Joker Cards....59
FIGURE 1.2 Contour Plots of Expected Row Winning Percentage for Different Joker Mixed Strategy Profiles....60
FIGURE 1.3 Contour Plots of the Variance of the Number of Row Stage Wins for 60 Stage Games Under Different Joker Mixed Strategies....61
FIGURE 1.4 Average Sample Variance of Simulated Experiments as ε is Varied....62
FIGURE 2.1 Type Space of the Gamble/Safe Game....68
FIGURE 2.2 Feasible Equilibrium Threshold Types of the Gamble/Safe Game....71
FIGURE 2.3 Feasible Equilibrium Proportions of Gamble Play in the Gamble/Safe Game....73
FIGURE 2.4 Gamble/Safe Game LQRE Correspondences for Different δ....76
FIGURE 2.5 LQRE for the Gamble/Safe Game for Different δ Values....77
FIGURE 2.6 Sample Elements of the Type Space....80
FIGURE 3.1 Cost Schedule for High Cost Environments....127
FIGURE 3.2 Theoretical Implications....128
FIGURE 3.3 Pooling Frequencies CA versus Fla Treatment Cells....129
FIGURE 3.4 Pooling Frequencies for CA, Fla I, & Fla II....130
FIGURE 3.5 Pooling Frequencies for Low Cost Types....131
FIGURE 3.6 Pooling Frequencies for High Cost Types....132

LIST OF TABLES

TABLE 1.1 O'Neill's Game....13
TABLE 1.2 Rapoport and Boebel's Game....16
TABLE 1.3 Relative Frequencies of Card Choices in O'Neill's Experiment....26
TABLE 1.4 Relative Frequencies of Card Choices in Treatment 1....27
TABLE 1.5 Relative Frequencies of Card Choices in Treatment 2....28
TABLE 1.6 Relative Frequencies of Card Choices in Treatment 3....29
TABLE 1.7 Goodness of Fit Chi-square Tests....30
TABLE 1.8 Percentage of Row Wins in Different Experiments....30
TABLE 1.9 Chi-square Tests for the Difference Between Sample and Theoretical Variance of Number of Row Wins....31
TABLE 1.10 Relative Frequencies of Card Choices and Row Wins by Pairs....33
TABLE 1.11 Chi-Square Tests for Row and Column Independence....34
TABLE 1.12 Deviations of Row Wins from Expectations: Treatment 1....36
TABLE 1.13 Deviations of Row Wins from Expectations: O'Neill....37
TABLE 1.14 Deviations of Row Wins from Expectations: Treatment 2....38
TABLE 1.15 Deviations of Row Wins from Expectations: Treatment 3....39
TABLE 1.16 Independence and Contemporaneous Correlation Hypothesis Tests....43
TABLE 1.17 General Time Independence Hypothesis Tests....44
TABLE 1.18 Independence of Own Past Choices Hypothesis Tests....46
TABLE 1.19 Independence of Opponent's Past Choices Hypothesis Tests....47
TABLE 1.20 Mean Card Decks Submitted by Treatment and Role....49
TABLE 1.21 Average Card Decks Submitted by Group 1 Subjects....50
TABLE 1.22 Average Card Decks Submitted by Group 2 Subjects....51
TABLE 1.23 Average Card Decks Submitted by Group 3 Subjects....52
TABLE 2.1 Gamble/Safe Game....66
TABLE 2.2 Utility Version of the Gamble/Safe Game....67
TABLE 2.3 BNE Implied Frequencies of Actions in the Gamble/Safe Game....70
TABLE 2.4 Equilibrium Description for Purification Example....78
TABLE 3.1 Experimental Treatments and Parameters....94
TABLE 3.2 Random Effects Probit Estimations....104
TABLE 3.3 Likelihood Ratio Test Comparisons Within the Lee Framework....109

ABSTRACT

The research reported in this dissertation explores the observable effects that individual heterogeneity implies in strategic environments. The first chapter provides a focused experimental test of mixed strategy play in strictly competitive games. The experiment directly tests whether serial correlation results from subjects' inability to generate sequences of actions that appear to be time independent, or instead from the play of non-equilibrium strategies. This is achieved by allowing the subjects to generate actions via a simple randomizing device. It is found that serial correlation is not reduced and that subjects adopt a wide variety of non-equilibrium mixed strategies. This wide variety of mixtures potentially explains the seeming paradox of minimax winning proportions with a high variance of win rates across pairs of players.

In the second chapter a theoretical model is developed for simultaneous move games in which the observable outcomes are allocations of monetary payoffs or commodity bundles, not expected utility levels. It is assumed that the players' mappings from the uncertain money amounts or commodity bundle allocations to expected utility levels are heterogeneous and are private information. The third chapter applies this framework to investigate the incentives to form agricultural marketing pools.

Chapter 1: Mixed Strategy Play and the Minimax Hypothesis

1 Introduction

In recent years there has been a resurgence in experimental tests of the minimax hypothesis in games with unique mixed strategy solutions. This research has been largely motivated by Barry O'Neill's [31] experiment that incorporated a game with a unique mixed strategy minimax solution that is, in most instances, invariant to the way subjects transform dollar payoffs into expected utilities. O'Neill's game remedied a deficiency found in previous experimental research: the minimax solutions of experimental games had relied upon an assumption that subjects were risk neutral. Unlike previous research, the aggregate data from O'Neill's experiment were strikingly close to many of the implications of the minimax solution. However, a review of O'Neill's data by Brown and Rosenthal [9], hereafter B&R, shows that the data, especially when considered at the individual level, are *not* consistent with players adopting their minimax strategies. B&R's data analysis reveals that subjects' sequences of choices were correlated, that the distributions of action profiles did not adhere to the ones implied by the minimax equilibrium, and that there were significant differences in skill across players.

The degree to which these results are damaging evidence against the minimax hypothesis depends heavily upon the sources of subject's failure to adhere to minimax strategies. For example, it is not clear whether the serial correlation found in subjects' choices was the result of deficiencies in subjects' cognitive abilities to generate random sequences of actions that were independent and identically distributed (i.i.d.), or the result of repeated game strategies that deviated from the minimax solution. Also, if subjects were not adopting the minimax strategies in O'Neill's experiment, then it is difficult to identify what strategies were in fact adopted. The major difficulty in identifying the mixed strategies that subjects might have used is that the experimenter can only observe realizations from the mixed strategies' unobservable probability distributions. Moreover, since mixed strategies cannot be directly observed, it is difficult to ascertain why some subjects are more successful at playing the game than other subjects in the same role. Obviously, the different sources of the subjects' failure to adhere to the minimax equilibrium imply different degrees to which the minimax hypothesis is an appropriate positive behavioral model. Specifically, there is a significant difference between play that is *intended* to be equilibrium but fails, and play that is clearly *not intended* to be part of the minimax equilibrium.

This paper reports on a series of experiments designed to distinguish between the possible sources of the failure of subjects' behavior to adhere to the implications of the minimax hypothesis in O'Neill's experiment. This is achieved in large part by having the subjects play the same game as O'Neill, except that each subject selects his actions via a probability experiment which he controls. This new procedure is referred to as the "mixed strategy device". In other words, each subject explicitly chooses a "mixed strategy." This mixed strategy defines a probability distribution over the subject's action set from which a single realization is drawn. By introducing a simple randomizing device, any residual serial correlation in a subject's sequence of actions cannot be the result of the subject's inability to generate a sequence of identical and independently distributed (i.i.d.) random actions, but rather it must be from a deviation from minimax play. A second benefit of the mixed strategy device is that it can permit one to see the actual mixed strategies subjects play. This allows one to discriminate among alternative strategies that can generate data similar to that generated by the minimax strategies. Thus, the experiments presented here provide a definitive test of a unique mixed strategy minimax equilibrium, and also provide insights into the exact nature of subjects' deviations from their minimax strategies.

Before discussing the results of these experiments, let's consider how the issues have arisen from past research. Recall that O'Neill uses a game form in which the unique mixed strategy minimax solution was nearly impervious to the manner in which subjects map dollar payoffs to expected utilities. This was achieved through a two-outcome zero-sum game in which *any* strictly increasing expected utility function leads to the same behavior as a risk neutral one. O'Neill's experimental design was the first test of a mixed strategy minimax solution that did not rely upon restrictive assumptions on the curvature of subjects' expected utility functions. O'Neill found that several key statistics of his aggregate data were much closer to the implications of the minimax hypothesis than those of previous experiments. Thus, O'Neill's improved experimental design and new results constituted a significant breakthrough and contribution to understanding the empirical validity of mixed strategy play. O'Neill's study motivated several subsequent experimental studies and enlightened discussions.

The most penetrating of the discussions is B&R. B&R acknowledge that O'Neill's design is a dramatic improvement over previous designs, but question the nature of the analysis of the data. Specifically, B&R argue that the statistical analysis of these experimental data should be expanded to hypothesis tests on the individual level data, as the minimax hypothesis is a theory of *individual* behavior. Re-examining O'Neill's data, B&R find that many characteristics of the data, particularly at the individual level, are actually *inconsistent* with the implications of the minimax hypothesis. Some prominent discrepancies are the presence of contemporaneous and serial correlation found among players' actions, and an exceedingly high variance in the number of wins across pairs of opponents. B&R claim that the high variance in observed winning percentages is the result of the contemporaneous correlation. Although it is easy to see how contemporaneous correlation arises from players anticipating and best responding to opponents' serially correlated actions, the source of the serial correlation is not clear. It could be a result of intended equilibrium behavior that is flawed by subjects' cognitive inability to generate time independent random realizations, or it could result from the use of *non-equilibrium* repeated game strategies. B&R raise this as one of the key issues in the original O'Neill data (p. 1067):

"An objective for anyone studying these data is to judge whether the departures observed in the data appear more likely to have been caused by intentional departures from the equilibrium strategies, rather than by chance or unintentional departures."

The results of the experiments reported here provide some resolution of these issues. Specifically, in the analysis of the data, presented in Section 4, evidence is found that the serial correlation in the data results from subjects' pursuit of non-minimax strategies, thus confirming B&R's suspicions that the intended play of subjects is not their part of the minimax strategy profile. This conclusion is reached from the following results: (1) one cannot reject the hypothesis that the action choices in O'Neill's data and the realized actions in the experiments here come from the same population; (2) even when given a simple randomizing device, subjects' choices are serially and contemporaneously correlated, leading one to conclude that subject behavior is clearly *not* intended to be mixed strategy equilibrium play; (3) the data presented here also show that, in this game, subjects do adopt mixed strategies a significant proportion of the time when given the opportunity, but these mixed strategies are almost never the equilibrium ones; and (4) the wide variety of mixed strategies itself can explain some of the exceedingly high variance found in the winning percentages.

In the next section we review in more detail the related experimental work. Then the new experimental design is presented, in Section 3. Following this description. an analysis of the data is given, in Section 4. A few concluding remarks are provided. in Section 5.

2 Previous Experimental Literature

This section contains a further, more detailed review of O'Neill's experiment. This is followed by a discussion of three additional papers that are largely motivated by O'Neill experiments, including B&R's re-examination of O'Neill's data. Finally, a review is provided of two previous attempts to directly observe mixed strategies in a laboratory setting.

2.1 Two-Outcome Constant-Sum Game Experiments

2.1.1 O'Neill's Experiment and Results

O'Neill presents a game that overcomes previous difficulties testing the minimax hypothesis. Specifically, the minimax solution of the game is impervious to the way subjects assign utility values to experimental payoffs. In his experiment, 25 pairs of subjects played 105 repetitions of the game in Table 1.

		Column			
		Ace	2	3	Joker
Row	Ace	L	W	W	L
	2	W	L	W	L
	3	W	W	L	L
	Joker	L	L	L	W

Table 1: O'Neill's Game

The payoffs in the game matrix are for the Row player; W = 5 cents and L = -5 cents. Column's payoffs are the negative of Row's. The unique minimax equilibrium in O'Neill's game is the mixed strategy profile in which both players randomize over the set of cards (actions) $\{Ace, 2, 3, Joker\}$ with the probabilities $\{.2, .2, .2, .4\}$ respectively. If one normalizes the utility of 5 cents to equal one and the utility of -5 cents equal to zero, then the value of this game is .4, the probability that Row wins in equilibrium. Furthermore, one may easily verify that the only Nash equilibrium path of the repeated game is the one in which the players play the unique stage game equilibrium each period.

O'Neill conducted his experiment by having each pair of subjects use playing cards, sitting face-to-face. Each player was endowed with $2.50 in nickels. Each pair participated in 15 practice rounds without payments, followed by 105 rounds with payments.

O'Neill concludes that some of the features of the data are consistent with the minimax hypothesis. The Row players win 41% of the stage games, which is within a standard deviation of the theoretical prediction of 40% under the minimax hypothesis. Furthermore, the differences between the proportions of play of the Joker for all players, for the row players, and for the column players do not differ in a statistically significant manner from the predictions of these proportions. O'Neill does note that there are some features of the data that are not consistent with the implications of minimax play. The variance of the number of Jokers played across subjects is greater than that predicted by minimax. This is indicative of heterogeneous marginal probabilities of Joker play among the subjects.[1] A test on the number of runs of card choices also provides evidence that card choices are not independent across time. Finally, there also appears to be an upward *Ace* card bias. O'Neill conjectures that the *Ace* card possibly presents a focal point for subjects.

2.1.2 Brown and Rosenthal's Re-examination of the Data

B&R agree that O'Neill's design provides an excellent test of the minimax hypothesis. However, B&R also suggest that the testable implications of the minimax hypothesis extend beyond simple wins rates and binomial distributions over the individual actions. Moreover, minimax is a theory whose predictions are about individual behavior. Thus, the most important attribute of any data that is supportive of the minimax hypothesis is that it adhere closely to the model's predictions at the in-

[1] O'Neill presents a model of how individuals choose Joker card probabilities that could have generated this heterogeneous Bernouilli process.

dividual level. B&R re-examine O'Neill's data and find that the results are not as supportive of minimax play as they first appear to be. A summary of B&R's results is as follows:

1. The probability distributions over wins and action frequencies implied by the minimax hypothesis are frequently rejected by statistical tests of low power (making these rejections strong). The low power of these tests results from the presence of several "plausible" alternative models, such as the equiprobable model and the win weighted model (Rapoport and Boebel, 1992). These alternative models imply probability distributions that can generate empirical data similar to minimax behavior. The presence of these alternative models makes it difficult to get a sense of exactly how "close" play is to minimax.

2. There is contemporaneous correlation of action choices within pairs. This correlation significantly influences win proportions.

3. There exists a significant amount of serial correlation in the players' choices across time. However, it is impossible to determine, from the data, whether this is the result of players' inabilities to generate observations that appear to be i.i.d. or whether these correlated actions are deliberate.

4. The behavior of subjects does not appear to be converging to the predictions of the minimax hypothesis over time.

2.1.3 Rapoport and Boebel

Rapoport and Boebel [34], hereafter R&B, conduct an experimental test similar to O'Neill's. They adopt the game form in Table 2:

Payoffs are for the Row player; the Column player has opposite payoffs. This game is slightly more complicated than O'Neill's because each player has five possible actions instead of four. R&B's design contains a within-subject control, in which subjects play both as a Row and as a Column player. They find that this role-change treatment has no significant effect on observed play. Also, there is a between-

Row		Column O	Column I	Column F	Column L	Column C
	O	L	L	W	W	L
	I	L	W	L	W	L
	F	W	L	L	W	L
	L	W	W	W	L	L
	C	L	L	L	L	W

Table 2: Rapoport and Boebel's Game

subject control for the amount of the payoffs. In one set of experiments, W=\$10 and L=\$6, and in the second set of experiments, W=\$15 and L=-\$1. They find that the outcomes of the game are invariant to these payoff perturbations. The results of these experiments are very close to those of O'Neill's. Thus, these experiments extend the robustness and scope of the O'Neill results to a wider array of two person two-outcome constant-sum games, subject role assignments and experimental payment designs.

2.1.4 Mookherjee and Sopher

Matching pennies is a familiar two-outcome constant-sum game. In Mookherjee and Sopher's [28] paper on learning behavior in matching pennies, only one of their treatments involves knowledge of the opponent's payoff function and history of chosen actions. Hence, only this treatment is comparable to the results already discussed. With ten pairs playing forty repeated stage games, Mookherjee and Sopher find that the frequencies of outcomes for nine of the ten pairs are not statistically different from those predicted by the minimax hypothesis. Mookherjee and Sopher address the question of whether action choices are i.i.d. via a runs test. They reject i.i.d. play for five of the twenty subjects. However, when examining the effect this has upon winning percentages, there is little evidence that correlated play is being exploited by some of the subjects, unlike subjects in O'Neill's experiment.[2]

2.2 Experimental Attempts to Directly Examine Mixed Strategies

Ochs [30] and Bloomfield [8] conducted experiments on two person 2x2 games with unique Nash equilibria in strictly mixed strategies. Both experiments elicited mixed strategies by having pairs "play" some number K of stage game "rounds" within each period. At the start of each period, each subject is asked to choose a number between zero and K, which corresponds to the number of times that he wishes to play an element of his action space. Then the subjects are told that the computer will randomly order these actions for the K rounds of play. In Ochs' experiment the two strings of actions are played against each other element by element, and in Bloomfield's experiment the players' expected payoffs are calculated from the two "mixed" strategies and awarded to each subject.[3] These papers refer to this design

[2] See B&R pp. 1073-1074.

[3] In the instructions for Bloomfield's experiment the subjects are asked to select the number of times they want to play a particular action in the 50 rounds of play and then are told, "You do not have to worry about ordering your choices - the computer will do that for you (randomly)."

as the "mixed extension" of the original game.

The mixed extension at first appears to be a reasonable way to elicit mixed strategies, but upon closer inspection some fundamental problems appear in both applications. Consider Ochs' design. There are two potential reasons why this procedure may not be appropriate to use with O'Neill's game form. First, the procedure generates distributions of play that are different from that implied under mixed strategy play. One can see this by considering the game of matching pennies and its four-round mixed extension. The unique minimax solution in matching pennies is the mixed strategy profile in which both players play heads with probability one half, and in the four-round mixed extension of matching pennies the unique solution is for both players to play heads twice. Assume that players adopt their equilibrium strategies. Then in Ochs' design, a player's four element string of actions is determined by a random draw from a uniform distribution on the set of all strings with exactly two heads. This in turn generates a probability density function over the set of the number of possible wins for a player, $\{0, 1, 2, 3, 4\}$. For each respective number in the set, the probability of that number of wins is $\left\{\frac{1}{6}, 0, \frac{2}{3}, 0, \frac{1}{6}\right\}$. This is clearly different than the probability density function over the number of wins for a player generated by the subgame perfect equilibrium of the four period repeated game of matching pennies, which is $\left\{\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16}\right\}$. These different distributions over stage game wins may significantly affect the feedback that the subjects receive from mixed strategy play. This feedback may play a crucial role in whether subjects actually play mixed strategies. Thus, when conducting an empirical test of a mixed strategy equilibrium, this changing of the distribution over outcomes is not desirable.

A second potential problem of Ochs' design is that in two-outcome games it introduces the risk attitudes of subjects as an issue. Once again, consider the four period mixed extension of the matching pennies game. As was noted, there are no longer just two possible payoffs in the game, but five. This is also true for the four

Then in the paper, the reader is informed that period outcomes are determined as "the computer determines the expected frequencies of the four possible outcomes, rounding fractional outcomes to the nearest integer," and the expected earnings are awarded to the subjects.

period repeated game of matching pennies as well. However, if one assumes that the players' expected utility functions are additively separable across stage games, then the unique Nash equilibrium path does not change. This assumption is characteristic of this type of research and is reasonable, considering the independence of stage game outcomes in equilibrium[4]. However, in Ochs' design this assumption is dubious because the round games are not independent. Consider the following interpretation of the four period mixed extension of matching pennies. One can imagine a player's strategy as a choice of an urn's composition. Specifically, of the four balls to be placed in an urn, a player chooses how many will be heads. Then the two players' urns are matched and a ball is drawn from each. These two balls become the strategy profile of the first round game. The strategy profile for the second round is determined by now drawing a second ball from each player's urn, *without* replacing the first period ball. Hence, the play of the second round is dependent upon the play of the first round. This continues until all balls have been drawn. From the description, it is clear that the round games are not independent, and the assumption that expected utility is additive across round games is questionable. Ochs avoids this issue by utilizing a lottery payoff scheme that attempts to induce risk neutral preferences on the subjects.

The distribution over outcomes in equilibrium in Bloomfield's design also deviates from the distribution over outcomes in the corresponding mixed strategy equilibrium. Once again, consider the distribution of wins in the four period matching pennies game. In Bloomfield's design where the games are *never* played, but the expectation over outcomes is used as the payoff, the distribution is a spike at the level of payoffs a player receives for exactly two wins. This obviously distorts the feedback that subjects receive in the "mixed strategy" equilibrium. Thus, this design suffers from some of the same problems as Ochs'. Given the problems of the two versions of the mixed extension, it is desirable to design a new device that preserves both the independence and the distributions over outcomes implied by mixed strategies.

[4]This assumption on the form of subjects' expected utility functions is implicitly made in O'Neill and B&R.

2.3 Summary

Three major questions remain unanswered in the literature on two-outcome constant-sum games. First, does the correlation of subjects' action choices across time result from their use of non-equilibrium repeated game strategies or from the subjects' inabilities to generate i.i.d. sequences? Second, does the exceedingly large variance of win rates across subject pairs result solely from the contemporaneous correlation of subject actions? Finally, why do aggregate statistics conform more closely to minimax predictions than when they are disaggregated? Also, given the problems with the current methodologies of eliciting mixed strategies, can a better "mixed strategy" device be found?

In the next section we present an experimental design that provides some insight into these three unanswered questions. Moreover, the design incorporates a mixed strategy device that elicits mixed strategies without the weaknesses of the Ochs and Bloomfield designs. The design uses O'Neill's original two-outcome constant-sum game, but in an experimental setting with significant differences from O'Neill's.

3 Experimental Design

In this section, a new experimental design is described that allows one to address some of the questions left unanswered by the previous literature. Specifically, subjects play the O'Neill game, except that in some treatments subjects select probability distributions over the four actions, instead of selecting a single action. The mixed strategy device provides benefits to both the experimenter and the subjects. First, this device allows subjects to easily generate i.i.d. sequences of actions across stage games, or in other words, successfully execute intended mixed strategies. Second, the device also provides the researcher with a new view of how subjects may actually be playing the game. After describing and motivating the experiment's three treatments, I provide a brief discussion of the experimental procedures.

3.1 The Three Treatments

There are three between-subject treatments conducted in the experiments reported here. All three treatments utilize the O'Neill game, but have some common differences from O'Neill's design. First, all treatments are conducted on computer terminals on the University of Arizona's Economic Science Laboratory local area network. and are conducted with a matrix payoff framing. This differs from O'Neill's design. in which subjects played the game face-to-face with playing cards.[5] A second difference in my design is that the subject pairs engage in sixty stage games instead of 105.[6] A third difference involves renaming the elements of the action space to $\{Green, Red, Brown, Purple\}$, in order to avoid the previously observed Ace bias.[7] The final difference, common to all treatments, involves using larger stage game payoffs. The following payoff transformations are made: the Row player receives 90 cents for a win and zero for a loss, and the Column player receives 60 cents for a win and zero for a loss. In equilibrium, each players' expected stage game earnings are 36 cents.

In Treatment 1, I replicate O'Neill's experiment with the above changes in order to establish a benchmark for computerized experiments of this type. This treatment allows for identification of changes in behavior resulting solely from the move to the matrix representation of the game on a computer.

In Treatment 2, subjects are asked to give a probability distribution over the set of actions instead of picking an action. A single random draw is made from this distribution, and the realization becomes the subject's action. This mixed strategy device allows subjects to generate random play through a probability experiment that they control and conduct on the computer. Each subject has the option to

[5]I agree with O'Neill and B&R, that the card game framing presents the strategic nature of the game more clearly to subjects than the payoff table framing; however, the mixed strategy device can only be implemented with a computer.

[6]The subjects were informed during the instructions that there would be exactly 60 periods of play.

[7]In order to avoid confusion, actions will be referred to by the names used in O'Neill's experiment for the remaining exposition of the paper.

fill a "shoe" of 100 cards with any composition of actions (cards) he or she desires. Once the shoe is filled, the computer shuffles its cards and selects the top card as the action to be played. However, opponents are only shown the chosen action (card), not the mixed strategy (i.e., the composition of the shoe). This generation of a random outcome is in the spirit of how mixed strategies are motivated in the classical treatments of game theory, namely, players choose a probability distribution over the set of actions, and then draw a realization.[8]

The data generated in this treatment is informative in two ways. First, by eliciting actual mixed strategies that are not costly to the subject and do not reveal any new information to one's opponent, the data generated is useful for discriminating between play that is consistent with minimax or with alternative models. Second, one can determine whether the previously observed serial correlation in players' choices is deliberate or is a product of the difficulty in generating i.i.d. sequences.

Treatment 3 is the same as the second treatment except that now, after each play is completed, each subject is informed of the mixture his opponent chose as well as the action (card) his opponent played as a result of that mixture. This treatment permits one to gain insight into one potential reason why players who do not adopt minimax strategies are not exploited, driving their play towards the minimax strategy. The observable products of mixed strategies in previous environments have been sequences of realizations, and if a sequence is not sufficiently long it can be difficult for a player to detect when his opponent deviates from his minimax strategy [24]. Therefore, deviations from minimax strategies may not be fully exploited. Also, if a player deviates from his minimax strategy and is detected, it still may be the case that the exact nature of the deviation may not be discovered and thus not fully exploited. In Treatment 3, if a player submits his mixed strategy and it is not minimax, then his opponent will have higher quality information about the nature of

[8]A compiled demonstration program of the instructions for all three treatments is available upon request from the author. Also, one should note that while this treatment makes playing mixed strategies cognitivelly simple, playing a pure strategy is still less "expensive," as subjects can fill the shoe completly with one type of card via a single key stroke.

his deviation and can more fully exploit it. If this type of disciplining is fostered in this experimental environment, then it may lead to greater adherence of individual behavior to minimax strategies. Finally, if players are learning to play minimax (i.e. play is converging) in the previous environments, but learning is occurring too slowly to detect, then the stronger presence of disciplining in this environment could also increase this speed of convergence.

3.2 Experimental Procedures

All experiments were conducted during the summer of 1995 in the Economic Science Laboratory at the University of Arizona with undergraduate students recruited with telephone calls, or with fliers posted around campus, or by the experimenter visiting undergraduate classes. At the beginning of each session, the subjects were instructed not to talk and to take a seat in a carrel with a computer terminal. The carrels prevented subjects from seeing the other subjects' computer terminals. Then the subjects read the computerized instructions to themselves at their own pace. When all subjects had finished reading the instructions, subjects played sixty stage games against a fixed but unknown opponent. When the experiment was completed, each subject was paid his or her earnings privately.

The experimental sessions had between six and twelve participants each. All subjects in each session participated in the same treatment. Treatments 1 and 3 each had a total of fifteen pairs participate, and Treatment 2 had sixteen pairs participate. Thus, for each treatment a panel data set is generated in which each pair generates a sixty element time series.

4 Data Analysis

This section presents the results from two ways of analyzing the experimental data. First, I analyze data on realized actions in a fashion similar to the B&R analysis of the original O'Neill game data. Second, the mixed strategies submitted by subjects in Treatments 2 and 3 are analyzed. The two approaches are taken because of the possibility that subjects may choose to internally randomize their actions independently of the mixed strategy device. The main conclusions reached are:

1. One cannot reject the hypothesis that data from O'Neill's experiment and from the three treatments presented here come from the same population.

2. The contemporaneous correlation within pairs that affects win rates predominantly arises from the play of the Joker card. This contemporaneous correlation can contribute to a higher than expected variance of win rates across pairs.

3. Serial correlation is reduced when easily implemented randomization is introduced; however, some serial correlation does remain, and much of it derives from players conditioning their play on the past play of their opponents.

4. When mixed strategies are elicited from subjects, it appears that the majority of pure action choices are made with the Joker card.

5. When mixed strategies are selected, they are typically heterogeneous across subjects and are not minimax.

6. Heterogeneous mixtures do contribute to the high variance of win rates across pairs.

4.1 Realized Actions

The following data analysis of the realized actions compares the data from the experiments presented here to the original O'Neill data and to the minimax hypothesis. First, we examine how well the aggregate data adheres to the minimax implications under the assumption that each action is i.i.d. across all subjects. While this analysis provides an evaluation of how well the minimax hypothesis predicts the aggregate statistics of play across pairs, it does not adequately test the theory in the domain for which it was developed, namely the decisions of individuals engaged in strictly competitive games. Hence, a broad analysis is performed on the data at the individual pair level. The first subsection of analysis on the individual and pair level data focuses on conducting hypothesis tests suggested by the implications of minimax play under the assumption that each player's action choices are independent across time. The next subsection addresses the minimax equilibrium implication that players randomize independently of each other. Evidence is found that subjects' choices are contemporaneously correlated within pairs, and this results in win rates that statistically differ from those expected under the assumption of independent play. Finally, the minimax implication that subjects' choices should be independent across time is tested via a log-linear probability model of each pair's joint choices.

4.1.1 Aggregated Data of Realized Actions

As B&R point out, aggregated data from experimental games often conform well to the equilibrium prediction, despite the fact that the individual-level data do not. Not surprisingly, the aggregate data from the experiments here conform well to some of the equilibrium predictions.

If one assumes that action choices are i.i.d. across players and time, then an appropriate approach is to analyze the data via contingency tables. One can see the rough adherence to the minimax hypothesis by inspecting the contingency tables, found in Tables 3-6 (Table 3 is reproduced from B&R).[9]

First, a majority of the joint moves involving two non-Joker cards are less than one standard deviation away from the theoretical predicted frequencies. The frequencies of joint moves involving Joker cards also do not differ greatly from the theoretical predictions. Also, the marginal frequencies of actions for the Column and Row players are impressively close to the minimax predictions. However, a chi-square goodness-of-fit test of the experimental data to the joint probability distribution implied by the minimax solution clearly *rejects* the null hypothesis of minimax play for each of the experiments. The results of these tests are given in Table 7. Note that

[9] Several of the hypothesis tests and tables presented here include results from B&R. They are reproduced here for the reader's convenience.

		Column Players' Choices				Marginal Frequencies for Row Players
		Ace	2	3	Joker	
	Ace	.044	.043	.043	.091	.221
		(.040)	(.040)	(.040)	(.080)	(.200)
		.004	*.004*	*.004*	*.005*	*.008*
	2	.046	.038	.038	.092	.215
Row Players'		(.040)	(.040)	(.040)	(.080)	(.200)
Choice		*.004*	*.004*	*.004*	*.005*	*.008*
	3	.049	.032	.037	.085	.203
		(.040)	(.040)	(.040)	(.080)	(.200)
		.004	*.004*	*.004*	*.005*	*.008*
	Joker	.086	.065	.051	.158	.362
		(.080)	(.080)	(.080)	(.160)	(.400)
		.005	*.005*	*.005*	*.007*	*.010*
Marginal Frequencies		.226	.179	.169	.426	
for Column Players		(.200)	(.200)	(.200)	(.400)	
		.008	*.008*	*.008*	*.010*	

Numbers in Parenthesis represent the minimax predicted relative frequencies. Numbers in italics represent standard deviations for observed relative frequencies under the minimax hypothesis.

Table 3: Relative Frequencies of Card Choices in O'Neill's Experiment

		Column Players' Choices				Marginal Frequencies for Row Players
		Ace	2	3	Joker	
	Ace	.044	.043	.047	.068	.202
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	2	.047	.037	.048	.093	.224
Row Players' Choice		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	3	.050	.042	.031	.091	.214
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	Joker	.073	.061	.068	.157	.359
		(.080)	(.080)	(.080)	(.160)	(.400)
		.009	*.009*	*.009*	*.012*	*.016*
Marginal Frequencies for Column Players		.214	.183	.193	.409	
		(.200)	(.200)	(.200)	(.400)	
		.008	*.008*	*.008*	*.016*	

Numbers in Parenthesis represent the minimax predicted relative frequencies. Numbers in italics represent standard deviations for observed relative frequencies under the minimax hypothesis.

Table 4: Relative Frequencies of Card Choices in Treatment 1

Row Players' Choice		Column Players' Choices: Ace	2	3	Joker	Marginal Frequencies for Row Players
	Ace	.047	.058	.034	.072	.211
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	2	.047	.055	.045	.072	.219
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	3	.044	.042	.053	.067	.205
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	Joker	.080	.073	.065	.147	.365
		(.080)	(.080)	(.080)	(.160)	(.400)
		.009	*.009*	*.009*	*.012*	*.016*
Marginal Frequencies for Column Players		.218	.228	.197	.357	
		(.200)	(.200)	(.200)	(.400)	
		.008	*.008*	*.008*	*.016*	

Numbers in Parenthesis represent the minimax predicted relative frequencies. Numbers in italics represent standard deviations for observed relative frequencies under the minimax hypothesis.

Table 5: Relative Frequencies of Card Choices in Treatment 2

		Column Players' Choices				Marginal Frequencies for Row Players
		Ace	2	3	Joker	
	Ace	.061	.047	.037	.078	.222
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	2	.042	.043	.039	.100	.224
Row Players' Choice		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	3	.042	.047	.034	.087	.210
		(.040)	(.040)	(.040)	(.080)	(.200)
		.007	*.007*	*.007*	*.009*	*.013*
	Joker	.071	.074	.061	.137	.343
		(.080)	(.080)	(.080)	(.160)	(.400)
		.009	*.009*	*.009*	*.012*	*.016*
Marginal Frequencies for Column Players		.217	.211	.171	.401	
		(.200)	(.200)	(.200)	(.400)	
		.008	*.008*	*.008*	*.016*	

Numbers in Parenthesis represent the minimax predicted relative frequencies. Numbers in italics represent standard deviations for observed relative frequencies under the minimax hypothesis.

Table 6: Relative Frequencies of Card Choices in Treatment 3

all statistics are asymptotically distributed chi-square with 9 degrees of freedom.

Experiment	Chi-square Statistic	P-value
O'Neill	60.248	.000
Treatment 1	19.771	.019
Treatment 2	29.941	.000
Treatment 3	26.840	.001

H_0: Data is generated dy the joint probability distribution implied by minimax play.

Table 7: Percentage of Row Wins in Different Experiments

A second feature of the aggregate data that is strikingly close to the implications of minimax play is the win rates of the row players. Under the minimax solution the row player is expected to win 40% of the stage games. Win rates in the original O'Neill experiment, and in Treatments 2 and 3 are all within approximately one standard deviation of the prediction under the minimax hypothesis; Treatment 1 is within two standard deviations. These statistics are presented in Table 8.

Experiment	Row Win %	Predicted %	# of Obs.	Standard Deviation
O'Neill	.410	.400	2625	.010
Treatment 1	.433	.400	900	.016
Treatment 2	.417	.400	960	.016
Treatment 3	.390	.400	900	.016

Table 8: Percentage of Row Wins in Different Experiments

Another question one can ask of the data in the four contingency tables is whether they all came from the same population. This can be addressed by a hypothesis test comparing the differences in the cell proportions across treatments, where the null hypothesis is that the data from each treatment is generated by the same probability distribution. The appropriate chi-square statistic is calculated to be 57.433, has 45 degrees of freedom, and a p-value of .101. Hence, one cannot reject that the data all come from the same population unless a level of significance greater than 10% is chosen. This is surprising, given how different the framing and tasks are across the four experiments. Hence, one might suspect that the data generated in the new experiments will possess many of the same deviations from the implications of the minimax solution as did the O'Neill data.

4.1.2 Individual and Pair Level Data of Realized Actions

When B&R examined the O'Neill data at the pair level, they discovered significant evidence that the data were inconsistent with the minimax model. A similar approach is taken here to testing the minimax implications of the data at the individual and pair level. By examining the variance of the number of wins across pairs, and individual and pair frequencies of actions, one can clearly reject that play at the pair level is consistent with the minimax equilibrium.

If one assumes that play for all pairs coincides with the minimax equilibrium, then the probability that the row player wins a stage game is $p = .4$. Thus, each pair generates a realization of a binomial process and the variance of the number of row wins is $np(1-p)$, where n is the number of trials. Clearly, as seen in Table 9, the sample variance is much greater than the theoretical variance for all treatments, and an appropriate Chi-square test allows one to reject the hypothesis that the sample and theoretical variances are the same.

Experiment	Predicted Var.	Sample Var.	P-value	D.O.F.
O'Neill	25.2	41.5	.025	25
Treatment 1	14.4	42.67	.000	15
Treatment 2	14.4	27.4	.029	16
Treatment 3	14.4	24	.025	15

Table 9: Chi-square tests for the Difference Between Sample and Theoretical Variance of the Number of Row Wins

As B&R note, when conducting a hypothesis test at the pair level, the lower number of observations used in calculating the test statistic implies that the test will have lower power. Nevertheless, one can still reject the minimax hypothesis for many of the pairs. In Table 10, one can see the results of many tests conducted on the pair level data. For Treatments 1, 2, and 3, one rejects the joint probability distribution over action profiles implied by minimax play for 10, 9, and 11 pairs respectively. This is about two thirds of the pairs in each treatment. Furthermore, there are a significant number of rejections of the minimax-implied binomial probability for single cards. In fact, the rejection rate is much higher in Treatments 2 and 3 than in Treatment 1. This is surprising, in that Treatments 2 and 3 have subjects directly picking mixed strategies. There are many tests that can be done at this level of the data, as shown by B&R, that all lead to the rejection of the minimax hypothesis. Thus, without much hesitation, one can conclude that the presence of a device that allows subjects to easily randomize over their action sets does not bring the data of observed actions plausibly close to minimax play.

4.1.3 Contemporaneous Play within Pairs

Perhaps the most striking evidence B&R present against minimax play in O'Neill's experiment is that the observed action profiles are inconsistent with independent drawings from player specific stationary multinomial distributions. This is first demonstrated by comparing observed winning percentages to those one would expect to see if the players' choices had been made independently. B&R call the difference between these two frequencies the correlation effect. Statistical tests reveal that the correlation effect is significant. B&R claim that this correlation effect is the primary explanation of why the winning percentages diverge substantially from the prediction of minimax play and why the variance of Row wins across individual pairs is significantly large.

Interestingly, while the correlation effect is statistically significant, simple contingency table tests usually do *not* reject that Row and Column actions are independent. This paradox is resolved by demonstrating that all of the significant contemporaneous correlation in observed card play is constrained to the set of action profiles in which at least one subject plays his Joker card. More succinctly, while the proportion of adopted profiles in which there is at least one Joker card does not typically

Treatment 1

Pair	Winning % for Row Player	Row Player Action A	2	3	J	Column Player Action A	2	3	J	Notes*
1	.333	.267	.217	.233	.283	.300	.167	.217	.317	-
2	.567	.233	.250	.150	.367	.233	.100	.250	.417	c
3	.483	.150	.183	.350	.317	.233	.200	.167	.400	a,c
4	.350	.133	.167	.167	**.533**	.217	.167	.117	.500	c
5	.317	**.317**	.233	.117	.333	.117	.167	.267	.450	c
6	.417	.200	.267	.133	.400	.117	.217	.283	.383	-
7	.400	.183	.283	.167	.367	.233	.183	**.033**	**.550**	b,c
8	.350	.133	.200	.200	.467	.250	.267	.100	.383	-
9	.400	.183	.250	.233	.333	.167	.183	.383	**.267**	b,c
10	.550	**.350**	.150	.283	**.217**	.333	.167	.150	.350	a,c
11	.467	.167	.233	.217	.383	.200	.167	.267	.367	-
12	.417	.217	.183	.250	.350	.200	.300	.100	.400	-
13	.433	.183	.117	.250	.450	.200	.150	.217	.433	c
14	.450	.133	.333	.267	.267	.117	.117	.250	.517	a,c
15	.567	.183	.300	.200	.317	.300	.200	.100	.400	c

Treatment 2

Pair	Winning % for Row Player	Row Player Action A	2	3	J	Column Player Action A	2	3	J	Notes*
1	.467	.250	.283	.167	.300	.150	.267	.250	.333	c
2	.300	.117	.167	.283	.433	.233	.150	.133	.483	-
3	.667	**.317**	**.050**	**.333**	.300	.150	**.400**	.200	**.250**	a,b,c,d
4	.433	.167	.217	.167	.450	.283	.200	.100	.417	-
5	.317	.233	.217	.217	.333	.267	.100	.233	.400	-
6	.350	.200	.233	.117	.450	**.317**	**.350**	.183	**.150**	b,c
7	.433	**.083**	.267	.150	.500	.167	.100	.300	.433	c
8	.417	.283	.167	.183	.367	.233	**.317**	.150	.300	c
9	.317	.200	.133	.183	.483	.183	.233	.250	.333	-
10	.367	.300	**.450**	.217	**.033**	.200	.300	.183	.317	a,c
11	.333	.167	.200	.133	.500	.133	.167	.250	.450	-
12	.383	.167	.217	.167	.450	.217	.183	.183	.417	-
13	.400	.250	.217	.250	.283	.217	.250	.117	.417	c
14	.483	.117	.217	.183	.483	.183	.183	.150	.483	-
15	.650	.250	**.333**	**.317**	**.100**	**.333**	**.333**	**.317**	**.017**	a,b,c,d
16	.350	.283	.133	.217	.367	.217	.117	.150	.517	c,d

Treatment 3

Pair	Winning % for Row Player	Row Player Action A	2	3	J	Column Player Action A	2	3	J	Notes*
1	.417	.200	.183	.283	.333	.183	.283	.150	.383	-
2	.400	.200	.150	.300	.350	.217	.300	.133	.350	-
3	.500	.183	.233	.217	.367	.200	.233	.300	**.267**	c,d
4	.333	.233	.183	.133	.450	**.317**	.100	.283	.300	b,c
5	.233	**.317**	.250	.117	.317	.200	.217	.167	.417	a,c
6	.450	.133	.250	.200	.417	.267	.150	**.067**	.517	b,c
7	.317	.217	.217	.167	.400	.233	.183	.167	.417	-
8	.383	.233	.217	.200	.350	.250	.217	.167	.367	-
9	.433	.267	.150	**.317**	**.267**	.233	.267	.117	.383	a,c
10	.250	.167	.233	.283	.317	.233	.200	.133	.433	c,d
11	.333	**.317**	.250	.233	**.200**	**.367**	.200	**.067**	.367	a,b,c
12	.517	.200	**.350**	**.083**	.367	**.083**	.150	.200	.567	a,b,c
13	.383	.283	.267	.233	**.217**	.267	.183	.233	.317	a,c
14	.517	.267	.233	.200	.300	**.083**	.267	.250	.400	c
15	.383	.117	.200	.183	.500	.117	.217	.133	**.533**	c

*Bold face denotes rejection (at the .05 level) of minimax binomial model for the given card.
a Denotes rejection (at the .05 level) of minimax model for all cards chosen by row player.
b Denotes rejection (at the .05 level) of minimax model for all cards chosen by column player.
c Rejection (at the .05 level) of the joint distribution of player and column moves under the minimax solution.

Table 10: Relative Frequencies of Card Choices and Row Player Wins by Pairs

differ from what one would expect under independent play, once play is considered to be conditional upon at least one Joker card being played, the observed frequencies of action profiles do differ significantly from the expected frequencies. Thus, one can conclude that all significantly successful or unsuccessful attempts to guess an opponent's action involve the Joker card.

When examining contingency table data for contemporaneous correlation, a natural first step is to test for row and column independence. The appropriate chi-square tests do not allow one to reject the null hypothesis of Column and Row independence in any of the experiments at any standard levels of significance. These tests are reported in Table 11, in which all of the test statistics are distributed chi-square with nine degrees of freedom.

Experiment	Chi-square Statistic	p-value
O'Neill	13.510	.141
Treatment 1	9.661	.379
Treatment 2	13.469	.143
Treatment 3	7.272	.609

Table 11: Chi-square Tests for Row and Column Independence

If one looks at the data at the individual pair level, there is still little rejection of Row and Column action independence. Looking back at Table 10, one sees for Treatment 1, that there is not a single rejection of independence. For Treatments 2 and 3, there are two pairs each for which the null hypothesis of independence is rejected at the .05 level of significance. Clearly, when viewing the complete set of actions, there appears to be little significant covariance among pair action choices.

However, one should recognize that the tables have a natural partition: the cells can be partitioned into two types, those for which the Row player wins and those for which the Row player loses. This is the approach that B&R take towards addressing the contemporaneous correlation question. B&R decompose the deviations of the observed winning rates from the rates predicted by the minimax solution into two parts. The first part is the difference between expected Row winning percentage, assuming that the Row and Column play was independent, and the .4 rate predicted by the minimax solution. This is called the *mixture effect* by B&R. The second part is the difference between the observed Row winning percentage and the expected Row winning percentage, under the assumption that Row and Column play is independent. This is called the *correlation effect.* B&R conduct hypothesis tests for the significance of the two effects and find that the correlation effect is significant but the mixture effect is not. The B&R results are reproduced and the results for the new treatments are reported in Table 10.

Given that one cannot reject that the Row and Column actions are independent, and at the same time that the correlation effect is significant, an appropriate line of inquiry is to ask whether the correlation is constrained to a particular subset of the table cells. Specifically, given that the cells of the contingency tables can be partitioned by the payoffs, perhaps it is appropriate to further partition the cells into those that involve a play of at least one Joker card and those that do not. This may be a reasonable approach, given that the three non-Joker cards are strategically equivalent.

Tables 12 - 15 presents a decomposition of the correlation effect into a correlated play effect involving Joker play and a correlated play effect involving non-Joker play. The correlated play effect with non-Joker cards is the difference between the proportion of plays in which the Row player wins without either player using a Joker card and the actual proportion of such action profiles. A similar calculation gives the correlated play effect with a Joker card, except that it is calculated with action profiles that include at least one Joker. To test the significance of these two effects, one performs chi-square tests similar to the ones performed for the ordinary correlated effect. However, one conducts these tests on the conditional distributions of non-Joker and Joker play.[10] The results show that the Joker correlated play effect is significant in all four experiments and that the correlated non-Joker play is sig-

[10]Statistical tests show that for all experiments, the number of observed non-Joker action profiles do not statistically differ from those expected under independent play.

nificant in only Treatment 1. This result highlights the idea that subjects view the Joker card as strategically distinct from the other cards and that the appropriate way to model the data is to focus on whether or not a subject plays a Joker card.

Treatment 1

Pair	Observed Row Win %	Expected Row Win %	Mixture Effect	Correlated Play Effect	Correlated Play Effect (Non-Joker)	Correlated Play Effect (Joker)
1	.333	.413	.013	-.079	-.040	-.040
2	.567	.405	.005	.161	.098	.064
3	.483	.407	.007	.077	.053	.023
4	.350	.424	.024	-.074	-.024	-.050
5	.317	.410	.010	-.093	-.060	-.033
6	.417	.404	.004	.012	.016	-.003
7	.400	.386	-.014	.014	.032	-.018
8	.350	.401	.001	-.051	-.039	-.012
9	.400	.412	.012	-.012	-.023	.011
10	.550	.401	.001	.149	.075	.074
11	.467	.401	.001	.066	.023	.043
12	.417	.407	.007	.010	-.017	.027
13	.433	.398	-.002	.035	.047	-.012
14	.450	.371	-.029	.079	.067	.012
15	.567	.402	.002	.165	.125	.040
	Mean Absolute Value:		.009	.072	.049	.031
	Chi-Square Statistic:		.534	29.045	31.355	41.860
	Chi-Square P-value:		1.000	.016	.008	.000

Table 12: Deviations of Row Wins from Expectations: Treatment 1

4.1.4 Intertemporal Correlation of Actions

A key motivation of the new experimental design reported here is to discover the source of serial correlation found in O'Neill's data. This section uses log linear probability models to detect the presence of serial correlation of actions in O'Neill's experiment and in the three treatments presented here. The results of hypothesis tests conducted in this log-linear framework reveal that the presence of serial cor-

O'Neill's Experiment

Pair	Observed Row Win %	Expected Row Win %	Mixture Effect	Correlated Play Effect	Correlated Play Effect (Non-Joker)	Correlated Play Effect Effect (Joker)
1	.390	.418	.018	-.027	-.002	-.026
2	.295	.410	.010	-.115	-.044	-.071
3	.390	.437	.037	-.046	-.032	-.014
4	.419	.366	-.034	.054	.019	.035
5	.343	.406	.006	-.063	-.035	-.028
6	.419	.397	-.003	.022	.016	.006
7	.476	.408	.008	.069	.014	.055
8	.467	.414	.014	.053	.015	.038
9	.362	.413	.013	-.051	-.035	-.016
10	.390	.400	.000	-.010	.010	-.019
11	.390	.403	.003	-.013	.005	-.018
12	.543	.401	.001	.142	.112	.030
13	.410	.368	-.032	.042	.041	.001
14	.467	.427	.027	.040	.005	.035
15	.324	.395	-.005	-.071	-.045	-.027
16	.343	.396	-.004	-.053	-.020	-.034
17	.362	.400	.000	-.038	-.023	-.015
18	.486	.398	-.002	.087	.057	.031
19	.390	.386	-.014	.004	.006	-.002
20	.438	.402	.002	.037	-.008	.045
21	.476	.405	.005	.071	.030	.042
22	.400	.393	-.007	.007	.002	.005
23	.448	.402	.002	.046	.038	.008
24	.495	.407	.007	.088	.050	.038
25	.333	.377	-.023	-.043	-.008	-.035
	Mean Absolute Value:		.011	.052	.027	.027
	Chi-Square Statistic:		2.655	40.428	15.417	93.630
	Chi-Square P-value:		1.000	.026	.931	.000

Table 13: Deviations of Row Wins from Expectations: Oneill

Treatment 2

Pair	Observed Row Win %	Expected Row Win %	Mixture Effect	Correlated Play Effect	Correlated Play Effect (Non-Joker)	Correlated Play Effect Effect (Joker)
1	.467	.412	.012	.055	.038	.017
2	.300	.412	.012	-.112	-.069	-.043
3	.667	.466	.066	.201	.109	.092
4	.433	.401	.001	.032	.036	-.004
5	.317	.399	-.001	-.082	-.032	-.050
6	.350	.369	-.031	-.019	-.018	-.001
7	.433	.414	.014	.019	.002	.017
8	.417	.407	.007	.010	-.014	.023
9	.317	.392	-.008	-.075	-.064	-.011
10	.367	.436	.036	-.070	-.059	-.011
11	.333	.411	.011	-.078	-.069	-.008
12	.383	.402	.002	-.019	-.031	.013
13	.400	.399	-.001	.001	-.031	.032
14	.483	.412	.012	.071	.022	.050
15	.650	.592	.192	.058	.043	.015
16	.350	.386	-.014	-.036	-.030	-.006
	Mean Absolute Value:		.026	.059	.042	.024
	Chi-Square Statistic:		11.142	24.547	16.160	39.674
	Chi-Square P-value:		.801	.078	.442	.001

Table 14: Deviations of Row Wins from Expectations: Treatment 2

Treatment 3

Pair	Observed Row Win %	Expected Row Win %	Mixture Effect	Correlated Play Effect	Correlated Play Effect (Non-Joker)	Correlated Play Effect (Joker)
1	.417	.408	.008	.009	.003	0.006
2	.400	.417	.017	-.017	.006	-0.023
3	.500	.406	.006	.094	.025	0.069
4	.333	.390	-.010	-.057	-.022	-0.035
5	.233	.394	-.006	-.160	-.112	-0.049
6	.450	.411	.011	.039	.021	0.018
7	.317	.399	-.001	-.082	-.049	-0.033
8	.383	.401	.001	-.018	-.006	-0.012
9	.433	.415	.015	.018	.037	-0.019
10	.250	.401	.001	-.151	-.097	-0.054
11	.333	.398	-.002	-.065	-.042	-0.023
12	.517	.396	-.004	.120	.061	0.059
13	.383	.425	.025	-.042	-.056	0.015
14	.517	.406	.006	.111	.081	0.030
15	.383	.419	.019	-.035	-.035	0.000
Mean Absolute Value:			.009	.068	.044	.030
Chi-Square Statistic:			.473	29.730	17.768	29.288
Chi-Square P-value:			1.000	.013	.275	.015

Table 15: Deviations of Row Wins from Expectations: Treatment 3

relation is lower in the new experiments than in O'Neill's experiments. Thus, the computerization and mixed strategy execution procedures appear to reduce intertemporal correlation in subject play. However, there is still evidence of substantial serial correlation that is the product of players' choices depending upon their opponents' previous choices. This kind of dependency is inconsistent with intended minimax equilibrium play.

The models here only consider the probabilities that each player chooses Joker. This choice is made because of the small sample sizes and prior evidence that the major source of contemporaneous correlation occurs with the play of Joker cards. The initial step of the analysis is to define a joint probability distribution of Row's and Column's stage game play of Joker. The identification of this distribution will allow hypothesis tests to be conducted for a variety of potential dependencies in the data..

Define the following random variables,

$$y_{it} = \begin{cases} 1 & \text{if player } i \text{ plays Joker in period } t, \\ -1 & \text{if player } i \text{ does not play Joker in period } t, \end{cases}$$

where $i = r$ (row) or c (column). Now consider the following simple log linear probability model for the joint distribution of (y_{rt}, y_{ct}):

$$\ln \Pr(y_{rt}, y_{ct}) = \alpha + \alpha_1 y_{rt} + \alpha_2 y_{ct} + \alpha_3 y_{rt} y_{ct}. \tag{1}$$

Using the property that the sum of the probabilities of the four possible pairs of (y_{rt}, y_{ct}) equals one, one can solve for the constant α in equation (1):

$$\alpha = -\ln \sum_{supp(y_{rt})} \sum_{supp(y_{ct})} \exp(\alpha_1 y_{rt} + \alpha_2 y_{ct} + \alpha_3 y_{rt} y_{ct}).$$

Substituting for α, (1) can be represented as a multinomial logit model:

$$\Pr(y_{rt}, y_{ct}) = \frac{\exp(\alpha_1 y_{rt} + \alpha_2 y_{ct} + \alpha_3 y_{rt} y_{ct})}{\sum_{supp(y_{rt})} \sum_{supp(y_{ct})} \exp(\alpha_1 y_{rt} + \alpha_2 y_{ct} + \alpha_3 y_{rt} y_{ct})}. \tag{2}$$

The model given in (2) has three unknown parameters to describe the three free probabilities of the joint distribution. Extending this model to include each player's lagged actions is straightforward.

The extension of a log-linear probability model to be conditional upon exogenous variables is fairly simple, especially when the exogenous variables are dichotomous.[11] The interest in this section is to present a joint model of players' choices that depends upon the past choices of the pair. If one assumes that the probability of the pair (y_{rt}, y_{ct}) depends upon the last period choices, (y_{rt-1}, y_{ct-1}), then the log-linear probability model for (y_{rt}, y_{ct})can be expressed as:

$$\begin{aligned} \ln \Pr\left(y_{rt}, y_{ct} \mid y_{rt-1}, y_{ct-1}\right) &= \alpha + \alpha_1 y_{rt} + \alpha_2 y_{ct} + \alpha_3 y_{rt} y_{ct} + \\ &\quad \beta_1 y_{rt} y_{rt-1} + \beta_2 y_{rt} y_{ct-1} + \beta_3 y_{rt} y_{rt-1} y_{ct-1} + \\ &\quad \delta_1 y_{ct} y_{rt-1} + \delta_2 y_{ct} y_{rt-1} y_{ct-1} + \delta_3 y_{ct} y_{rt-1} y_{ct-1} + \\ &\quad \gamma_1 y_{rt} y_{ct} y_{rt-1} + \gamma_2 y_{rt} y_{ct} y_{ct-1} + \gamma_3 y_{rt} y_{ct} y_{rt-1} y_{ct-1}. \end{aligned} \tag{3}$$

This is a saturated model with twelve unknown parameters and free probabilities. Attempting to estimate all of the parameters in (3) is difficult when one has only sixty observations per pair. To overcome the short data problem, let's assume that the coefficients γ_1, γ_2, and $\gamma_3 = 0$. Then we can solve for α as a function of the remaining nine unknown parameters,

$$\alpha = -\ln \sum_{supp(y_{rt})} \sum_{supp(y_{ct})} \sum_{supp(y_{rt-1})} \sum_{supp(y_{ct-1})} \exp(\alpha_1 y_{rt} + \alpha_2 y_{ct} + \cdots + \delta_3 y_{ct} y_{rt-1} y_{ct-1}).$$

Now, by taking the exponential of (3) one obtains the following multinomial logit expression:

$$\Pr(y_{rt}, y_{ct} \mid y_{rt-1}, y_{ct-1}) = \frac{exp\left(\alpha_1 y_{rt} + \alpha_2 y_{ct} + \cdots + \delta_2 y_{ct} y_{ct-1} + \delta_3 y_{ct} y_{rt-1} y_{ct-1}\right)}{\sum_{supp(y_{rt})} \cdots \sum_{supp(y_{ct-1})} \exp(\alpha_1 y_{rt} + \cdots + \delta_3 y_{ct} y_{rt-1} y_{ct-1})} \tag{4}$$

Estimates of the unknown parameters in (4) can be obtained by maximum likelihood techniques.

The model of equation (4) can be used to test for various types of intertemporal and contemporaneous correlations of pair choices. Different types of stochastic

[11] For extended discussions and examples of this the reader is reffered to Lee [23], Nerlove and Press [29], and Goodman [13].

processes imply that different parameters are non-zero. Therefore, one can test for these processes by estimating restricted versions of (4) and conducting the appropriate likelihood ratio tests.[12] This task is completed for the three treatments reported here and for the original O'Neill experiments. But first, one should note that the larger number of stage games that O'Neill's subjects played leads to a higher power for the hypothesis tests. Therefore, the tests on O'Neill's data are conducted for both the first sixty periods (to allow proper comparison to the new experiments), and then for 105 periods (to allow for more power and comparison to B&R's results).

The first series of tests addresses the issue of time independence and contemporaneous correlation of Row and Column stage game play of Joker. The results of these tests are reported in Table 16. Test (1) is for the restriction that all of the unknown parameters in (4) are zero except α_1 and α_2. If both the Row and Column play of Joker are independent Bernoulli processes, then this restricted model is appropriate. In Treatments 1 and 3, the restricted model is rejected for four pairs in each treatment. In Treatment 2, six out of sixteen pairs reject the null hypothesis. In the original O'Neill experiment one can reject the hypothesis of joint independent play for eight pairs when only using the first sixty observations, and thirteen pairs when using all observations.[13] Hence, for all treatments, joint independent play is rejected for a significant number of pairs, even though it appears that the power of

[12] B&R take a similar approach in their investigation of the different types of correlations across subject choices. However, their approach is to estimate a logit equation for each subject's probability of playing Joker as a function of his previous two choices and of his opponent's *current* and previous two choices. Then they conduct hypothesis tests upon restricted versions of these equations. A slighlty different approach is taken here because of a potential simultaneity problem resulting from contemporaneous correlation, and the possibilty that the probabilies of the joint distribution will not sum to one. These potential problems are avoided here because I estimate the parameters of the *joint* distribution. Furthermore, this specification allows one to test for certain joint hypotheses that can not be tested for in the B&R framework. It should be also noted that only one-period lagged decisions are considerd here because of the relatively small sample sizes.

[13] This increase in the number of rejections is indicative that the hypothesis tests have greater power at 105 observations. This will be characteristic of almost all of the remaining hypothesis tests.

the hypothesis test is relatively low with only sixty observations.

Experiment	Test	Null Hypothesis	Pairs for which the Null Hypothesis is Rejected (α=.05)	Total
O'Neill (60 obs.)	(1)	α_3, β_1, β_2, β_3 δ_1, δ_2, δ_3 all=0	10,11,14,15,16, 17,21,25	8
	(2)	α_3, β_1, β_2, $\beta_3 = 0$	7,9,10,11,14,16,21,22,24,25	10
	(3)	α_3, δ_1, δ_2, δ_3 all=0	4,6,7,9,10,11,17,20,21,25	10
	(4)	α_3=0	9,10,11,21,25	5
O'Neill (105 obs.)	(1)	α_3, β_1, β_2, β_3 δ_1, δ_2, δ_3 all=0	2,7,8,9,10,11,12,14,16,17,21,24,25	13
	(2)	α_3, β_1, β_2, β_3 all=0	2,7,8,9,10,11,12,14,16,19,20,21,24,25	14
	(3)	α_3, δ_1, δ_2, δ_3 all=0	2,4,6,7,8,9,10,11,14,15,17,20,21,24,25	15
	(4)	α_3=0	2,8,9,11,21,24,25	7
Treatment 1	(1)	α_3, β_1, β_2, β_3 δ_1, δ_2, δ_3 all=0	3,8,10,13	4
	(2)	α_3, β_1, β_2, β_3 all=0	10,13	2
	(3)	α_3, δ_1, δ_2, δ_3 all=0	1,3,8,10	4
	(4)	α_3=0	10	1
Treatment 2	(1)	α_3, β_1, β_2, β_3 δ_1, δ_2, δ_3 all =0	1,2,4,9,11,13	6
	(2)	α_3, β_1, β_2, β_3 all=0	1,2,3,4,5,9,13	7
	(3)	α_3, δ_1, δ_2, δ_3 all=0	3,4,5,11,13,14,15	7
	(4)	α_3=0	3,5,9,1	4
Treatment 3	(1)	α_3, β_1, β_2, β_3 δ_1, δ_2, δ_3 all=0	7,8,11,13	4
	(2)	α_3, β_1, β_2, β_3 all=0	11,13	2
	(3)	α_3, δ_1, δ_2, δ_3 all=0	12,13	4
	(4)	α_3=0	13	1

Table 16: Independence and Contemporaneous Correlation Hypothesis Tests

Tests (2) and (3) are conducted with a null hypothesis that one of the players' Joker play is an independent Bernoulli process. For the Row player this corresponds to restricting all of the parameters associated with Row's current choice to be zero except for α_i. This is Test (2). Test (3) is a similar hypothesis test for Column. Again, rejections of the restricted models are more prevalent in Treatment 2 and in O'Neill's experiment than in Treatments 1 and 3. Also, one can always reject independent play for at least as many Column players as Row players in all experiments.

Test (4) is for contemporaneous correlation. The absence of contemporaneous correlation corresponds to setting the parameter $\alpha_3 = 0$. Given the contingency-table-test results for independence presented earlier, it is not surprising that for only a small number of pairs in each treatment can one reject the absence of contemporaneous correlation. This is also consistent with the earlier result that significant contemporaneous correlation is found only when one considers play that is *conditional* upon at least one player adopting a Joker card.

Experiment	Test	Null Hypothesis	Pairs for which the Null Hypothesis is Rejected (α=.05)	Total
O'Neill (60 obs.)	(5)	β_1, β_2, β_3, δ_1, δ_2, δ_3 all=0	1,4,6,7,9,10,11,14,15,16, 17,19,20,21,22,24,25	17
	(6)	β_1, β_2, β_3, all=0	9,10,11,14,16,17,21,22	8
	(7)	δ_1, δ_2, δ_3 all=0	4,6,7,9,10,14,17,19,20,21,25	11
O'Neill (105 obs.)	(5)	β_1, β_2, β_3, δ_1, δ_2, δ_3 all=0	2,4,6,7,8,9,10,11,12,14, 16,17,19,20,21,22,24,25	18
	(6)	β_1, β_2, β_3, all=0	7,8,10,11,12,14,16,17,20,21,22	11
	(7)	δ_1, δ_2, δ_3 all=0	2,4,6,7,8,9,10,14,15,17,18,20,21,23,25	15
Treatment 1	(5)	β_1, β_2, β_3, δ_1, δ_2, δ_3 all=0	3,8,10,13	4
	(6)	β_1, β_2, β_3, all=0	13	1
	(7)	δ_1, δ_2, δ_3 all=0	3,8,10	3
Treatment 2	(5)	β_1, β_2, β_3, δ_1, δ_2, δ_3 all=0	1,2,4,9,11,13	6
	(6)	β_1, β_2, β_3, all=0	2,4,9,13	4
	(7)	δ_1, δ_2, δ_3 all=0	4,10,11,13,14	5
Treatment 3	(5)	β_1, β_2, β_3, δ_1, δ_2, δ_3 all=0	8,11,13	3
	(6)	β_1, β_2, β_3, all=0	11	1
	(7)	δ_1, δ_2, δ_3 all=0	7,8,12,13	4

Table 17: General Time Independence Hypothesis Tests

Table 17 presents hypothesis tests for general dependence of players' actions on the actions made in the previous period. Test (5) considers a model in which neither Row nor Column play depends upon any of the data generated in the previous stage game. The corresponding log-linear probability model under this assumption is given in equation (2). For the three new experiments, the results of these tests are identical to the first tests presented in Table 16. One exception is for pair seven in Treatment 3, where the independent Bernoulli model is rejected but the time independent model is not. However, when examining the results of the hypothesis tests for the O'Neill experiment, one see that the new restricted model is rejected by seventeen pairs: the eight pairs that rejected the independent Bernoulli model and an additional nine pairs. It is a rather striking result that a significantly larger proportion of pairs in O'Neill's experiment exhibit this dependence upon past play than in the three other treatments- especially when one considers the detailed records of past play subjects have at their disposal on the computer in the new experiments as opposed to the hand-run protocol of O'Neill's experiment. When considering Column or Row play separately, as in Tests (6) and (7), one finds similar results of more serial correlation in O'Neill's experiment.

The next step of the hypothesis testing exercise is to examine specific types of serial correlation: namely, the dependence of a player's action upon his own previous action and on that of his opponent. Table 18 presents the results of hypothesis tests of the dependency of players' choices on their *own* previous choice. A specification that incorporates the independence of one's choice and his own previous choice can be made by eliminating the terms of equation (4) that capture the interaction of a choice and its lagged value. In other words, one can impose the constraints: $\beta_1 = \beta_3 = \delta_2 = \delta_3 = 0$. To test for this type of independence for the Row or Column player separately, one tests the restrictions that $\beta_1 = \beta_3 = 0$ and $\delta_2 = \delta_3 = 0$ respectively. The results of these tests ((7), (8), and (9)) show that few pairs in Treatments 1,2, and 3 reject these restrictions. However, in O'Neill's experiment, nine pairs (or twelve pairs, if all observations are used) reject the joint independence of players' choices and their own previous actions. When the tests are done for the two roles separately, one can reject the null hypothesis for ten column players and only three row players (there are twelve and seven rejections respectively when 105 observations are used). These results are qualitatively similar to those found by

B&R. This series of tests suggests that some of the original serial correlation that depended upon one's own past choices, found in O'Neill's data, has been removed by the simple change to the computerized environment. However, as B&R point out, one cannot conclude whether this type of serial correlation is the result of intended non-equilibrium play.

Experiment	Test	Null Hypothesis	Pairs for which the Null Hypothesis is Rejected (α=.05)	Total
O'Neill	(8)	β_1, β_3, δ_2, δ_3 all=0	6,7,9,11,14,17,20,21,22	9
(60 obs.)	(9)	β_1, β_3 all=0	9,11	3
	(10)	δ_2, δ_3 all=0	4,6,7,9,14,17,19,20,21,24	10
O'Neill	(8)	β_1, β_3, δ_2, δ_3 all=0	4,6,7,8,9,11,17,20,21,22,24,25	12
(105 obs.)	(9)	β_1, β_3 all=0	8,9,10,11,16,20,22	7
	(10)	δ_2, δ_3 all=0	4,6,7,8,9,14,17,20,21,23,24,25	12
Treatment 1	(8)	β_1, β_3, δ_2, δ_3 all=0	3	1
	(9)	β_1, β_3 all=0	none	0
	(10)	δ_2, δ_3 all=0	3	1
Treatment 2	(8)	β_1, β_3, δ_2, δ_3 all=0	4,11,13	3
	(9)	β_1, β_3 all=0	4,8,9,13	4
	(10)	δ_2, δ_3 all=0	1,4,10,11,15	5
Treatment 3	(8)	β_1, β_3, δ_2, δ_3 all=0	12,13	2
	(9)	β_1, β_3 all=0	none	0
	(10)	δ_2, δ_3 all=0	7,8,12,13	4

Table 18: Independence of Own Past Choices Hypothesis Tests

Serial correlation that is indicative of non-equilibrium play is the significant dependence of a player's choice on his opponent's lagged choice. The null hypothesis of the last series of hypothesis tests is that subjects' decisions do not posses this type of serial correlation. To construct the restricted model which is implied by this null hypothesis, restrict to the value zero those parameters of (4) which precede the product of a player's current choice with his opponent's lagged choice. The results of these tests are presented in Table 19. In the three new treatments, there is little rejection of joint independence, Test (11), of previous stage game opponents' choices. However, when one looks at the player roles separately, Test (12) for the Row player and Test (14) for the Column player, one third of the pairs in Treatment 1 have a player for which the null hypothesis is rejected. Likewise ten pairs in Treatment 2 and four pairs in Treatment 3 allow for the rejection of the null hypothesis for at least one player. In the original O'Neill experiment, similar results prevail. There are only four pairs that reject the hypothesis tested jointly, but 13 pairs that reject the model for at least one player.

Experiment	Test	Null Hypothesis	Pairs for which the Null Hypothesis is Rejected (α=.05)	Total
O'Neill	(11)	β_2, β_3, δ_1, δ_3 all=0	9,10,11,12	4
(60 obs.)	(12)	β_2, β_3 all=0	9,10,11,14,15,16,17,22	8
	(13)	δ_1, δ_3 all=0	18,19,21,25	5
O'Neill	(11)	β_2, β_3, δ_1, δ_3 all=0	2,4	2
(105 obs.)	(12)	β_2, β_3 all=0	4,7,8,10,11,12,14,16,17	9
	(13)	δ_1, δ_3 all=0	1,2,10,15,18,19,21,25	8
Treatment 1	(11)	β_2, β_3, δ_1, δ_3 all=0	10	1
	(12)	β_2, β_3 all=0	2,13	2
	(13)	δ_1, δ_3 all=0	3,8,10	3
Treatment 2	(11)	β_2, β_3, δ_1, δ_3 all=0	3,4	3
	(12)	β_2, β_3 all=0	1,2,4,8,9,11	6
	(13)	δ_1, δ_3 all=0	1,7,11,13,14,15	6
hline Treatment 3	(11)	β_2, β_3, δ_1, δ_3 all=0	3	1
	(12)	β_2, β_3 all=0	11	1
	(13)	δ_1, δ_3 all=0	7,8,13	3

Table 19: Independence of Opponent's Past Choices Hypothesis Tests

In summary, there are several points that arise in this series of hypothesis tests. First, a smaller proportion of pairs in Treatments 1 and 3 reject the restrictions than in O'Neill's experiment and Treatment 2. Second, when considering the results of the statistical tests, one should consider the evidence that the tests conducted with sixty observations appear to have relatively lower power than those conducted with 105 observations. Third, even with the lower power tests, independent play is rejected for significant proportions of pairs in all four experiments. Fourth, serial correlation that depends upon a player's own previous choice appears stronger in O'Neill's experiment than in any of the new experiments. Finally, serial correlation dependent upon opponent's previous actions does appear frequently when one considers it a pair at a time, and typically only one player in a pair exhibits this type of serial correlation. Thus, while serial correlation is not as strong in the three new treatments as in the original experiment, intertemporal and contemporaneous correlation that is inconsistent with independent play is found in a substantial number of pairs' realized actions.

4.2 Observed Mixed Strategies

In this section, attention is turned towards analyzing the mixed strategies that subjects selected in Treatments 2 and 3.[14] In these two treatments, the mixed strategy device is introduced to give the subjects the ability to generate independent and identically distributed sequences of actions. The mixed strategy device also produces an additional benefit: when a subject uses the device, he provides a new view of how he plays the game. An initial inspection of the data reveals that a large amount of play in these treatments does utilize the mixed strategy device. However, almost none of the play incorporates the minimax equilibrium proportions. In fact, only three subjects in Treatment 2, and none Treatment 3, play the minimax equilibrium strategy more than ten times, and none of these subjects plays the strategy more than twenty times. Moreover, out of the 3720 decisions made in the two treatments, only 103 decisions correspond to the mixed strategy equilibrium,i.e., only about 2.8% of the play is minimax. While the data here show that play is not consistent with minimax, it also suggests ways in which subjects' play deviates. This section proceeds to argue that behavior does not differ across the two treatments, but there do appear to be three distinct types of subject play in the two treatments. These differing styles of play suggest rationalizations of how play is correlated and why the mean of row win rates is close to forty percent but the variance of win rates is much higher than expected.

[14]In the appendix, one can find the complete set of data from these two treatments

		Card			
Treatment	Role	Ace	2	3	Joker
2	Row	22.72	20.80	21.78	34.39
		(26.79)	(26.34)	(26.76)	(40.34)
2	Column	24.29	22.86	20.67	33.87
		(24.93)	(21.12)	(19.00)	(36.06)
3	Row	22.68	21.07	22.43	33.82
		(21.65)	(17.94)	(20.46)	(32.27)
3	Column	20.93	20.88	17.83	40.36
		(21.02)	(19.23)	(17.72)	(36.05)

(Sample Standard Deviation for Card)

Table 20: Mean Card Decks Submitted by Treatment and Role

Earlier data analysis showed that the distributions of realized actions in Treatments 2 and 3 did not differ significantly. Now inspection of the actual subject choices does not show significant differences in subject behavior as well. Table 20 presents the average card deck selected by the Row and Column players in each treatment. Below each number is the standard deviation for the mean of the card. The compositions of the four card decks are fairly close to the minimax proportions; this is consistent with earlier results on realized actions. However, the standard deviations are very large for all cards, making it difficult to conclude that these decks are statistically different from one another. Perhaps a more informative way to view the data can be achieved by organizing it according to characteristics found at the individual level.

A natural way to partition the individual level data is by the tendency of players to choose strict mixed strategies. I assign subjects in an ad hoc way to one of three groups. Group 1 subjects are those who by the end of the experiment generally put 100 of the same card in their box each period, essentially only picking pure actions. The criterion used to select members of this group is that a subject plays a pure action in at least fourteen out of the last twenty periods. Group 3 subjects are those who almost exclusively play strictly mixed strategies. The criterion here is that a

			Card			
Treatment	Role	Pair	Ace	2	3	Joker
2	R	8	26.02 (11)	14.43 (5)	24.05 (10)	35.50 (19)
2	R	9	17.98 (10)	15.30 (7)	17.97 (9)	48.75 (29)
2	R	12	15.23 (8)	24.45 (13)	15.25 (8)	45.07 (27)
2	C	4	25.00 (12)	22.67 (10)	11.50 (1)	40.83 (16)
2	C	16	22.35 (13)	12.35 (7)	15.72 (9)	49.58 (29)
3	R	3	23.50 (4)	24.17 (4)	22.50 (4)	29.83 (11)
3	C	1	19.00 (5)	27.95 (8)	19.35 (6)	33.70 (12)
3	C	2	21.08 (3)	26.92 (1)	18.67 (2)	33.33 (19)
	Averages:		21.27 (8.3)	21.03 (6.9)	18.13 (6.1)	39.57 (20.3)

(Number of Periods in which submitted box was filled with only this card)

Table 21: Average Card Decks Submitted by Group 1 Subjects

subject plays a pure action nine times or fewer. Group 2 subjects are the remaining subjects who play a significant number of both pure and mixed strategies. Tables 20-22 present the partitioning of subjects into the three groups, each subject's average card deck choice, and the number of times he or she played each card as a pure action.

Although these classifications are ad hoc, I believe that they do reveal some striking characteristics in the data. Specifically, there is greater than expected variance for the number of Joker cards that subjects put in their boxes, indicating that subjects are speculating more on the expected payoff of the Joker cards than the other cards. This point is made even stronger when one looks at the percentage of pure action plays that are Joker in each of the three groups. Finally when subjects do play strictly mixed strategies, subjects play a wide variety of non-minimax strategies. However, it should be noted that several subjects almost exclusively play the focal equiprobable mixed strategy.

Treatment	Role	Pair	Card: Ace	2	3	Joker
2	R	2	13.03 (1)	19.42 (1)	25.55 (4)	42.00 (17)
2	R	3	33.50 (0)	2.33 (0)	33.50 (0)	30.67 (17)
2	R	4	16.57 (0)	21.62 (1)	16.28 (0)	45.53 (10)
2	R	6	21.83 (9)	19.00 (7)	14.83 (5)	44.33 (22)
2	R	7	17.58 (1)	24.25 (3)	16.83 (2)	41.33 (13)
2	R	10	27.82 (5)	41.43 (10)	28.33 (4)	2.42 (1)
2	R	11	19.72 (4)	19.50 (3)	14.65 (1)	46.13 (22)
2	C	1	17.47 (4)	27.28 (6)	22.85 (4)	32.40 (18)
2	C	5	20.02 (4)	15.68 (3)	23.35 (6)	40.95 (18)
2	C	11	18.55 (2)	15.55 (3)	23.05 (0)	42.85 (15)
2	C	14	16.82 (0)	19.62 (0)	19.58 (0)	43.98 (24)
3	R	1	21.47 (1)	22.02 (2)	25.73 (4)	30.78 (5)
3	R	4	23.63 (4)	19.03 (2)	13.90 (1)	43.43 (15)
3	R	5	28.67 (5)	18.82 (0)	20.85 (1)	31.67 (7)
3	R	9	26.18 (7)	15.57 (1)	27.92 (7)	30.33 (14)
3	R	11	32.98 (0)	21.33 (2)	24.37 (1)	21.32 (8)
3	R	15	18.22 (2)	19.32 (0)	16.42 (2)	46.05 (7)
3	C	4	25.08 (3)	16.42 (0)	21.33 (2)	37.17 (15)
3	C	6	19.62 (0)	16.30 (0)	14.23 (0)	49.85 (18)
3	C	8	21.35 (0)	21.42 (0)	19.98 (0)	37.25 (21)
3	C	10	20.00 (1)	21.65 (0)	17.60 (0)	40.75 (15)
3	C	11	34.35 (10)	16.10 (2)	11.72 (0)	37.83 (19)
3	C	12	11.83 (1)	15.00 (2)	16.33 (1)	56.83 (24)
3	C	14	18.00 (0)	19.92 (0)	19.42 (0)	42.67 (11)
	Averages:		21.85 (2.7)	19.52 (2.0)	20.36 (1.9)	38.27 (14.8)

(Number of Periods in which submitted box was filled with only this card)

Table 22: Average Card Decks Submitted by Group 2 Subjects

Treatment	Role	Pair	Card: Ace	2	3	Joker
2	R	1	26.22 (2)	21.88 (0)	22.65 (0)	29.25 (4)
2	R	5	21.87 (0)	21.97 (0)	21.82 (0)	34.35 (5)
2	R	13	19.53 (0)	18.63 (0)	25.65 (0)	36.18 (1)
2	R	14	17.67 (0)	17.33 (0)	17.43 (0)	47.57 (8)
2	R	15	29.28 (0)	30.52 (0)	31.37 (0)	8.83 (1)
2	R	16	20.77 (0)	19.98 (0)	19.60 (0)	39.65 (3)
2	C	2	19.82 (0)	19.38 (0)	18.30 (0)	42.50 (7)
2	C	3	24.47 (1)	25.22 (1)	25.17 (0)	25.15 (4)
2	C	6	32.35 (0)	30.28 (0)	23.68 (1)	13.68 (1)
2	C	7	20.77 (0)	21.60 (0)	20.77 (0)	36.87 (9)
2	C	8	25.75 (2)	26.58 (1)	24.08 (0)	23.58 (3)
2	C	9	24.75 (0)	24.83 (0)	24.58 (0)	25.83 (0)
2	C	10	19.12 (0)	18.52 (0)	15.17 (0)	47.20 (8)
2	C	12	16.25 (0)	17.17 (0)	16.92 (0)	49.67 (7)
2	C	13	24.07 (0)	24.60 (0)	10.25 (0)	41.08 (8)
2	C	15	33.82 (0)	33.87 (1)	30.82 (0)	1.50 (0)
3	R	2	24.00 (2)	15.95 (0)	23.62 (3)	36.43 (2)
3	R	6	9.67 (0)	21.53 (0)	23.50 (0)	45.30 (8)
3	R	7	19.13 (2)	23.60 (0)	20.35 (1)	36.92 (5)
3	R	8	21.18 (0)	21.85 (0)	19.88 (0)	37.08 (1)
3	R	10	26.92 (1)	19.10 (0)	24.50 (0)	29.48 (0)
3	R	12	16.22 (0)	23.27 (0)	23.12 (0)	37.40 (6)
3	R	13	24.67 (1)	26.65 (1)	25.47 (1)	23.22 (1)
3	R	14	23.75 (1)	23.92 (1)	24.33 (0)	28.00 (3)
3	C	3	28.48 (2)	21.88 (0)	20.27 (0)	29.37 (3)
3	C	5	20.03 (0)	19.98 (0)	20.00 (0)	39.98 (5)
3	C	7	22.32 (1)	19.15 (0)	16.98 (0)	41.55 (1)
3	C	9	18.37 (0)	27.18 (1)	15.90 (0)	38.55 (6)
3	C	13	24.58 (0)	24.58 (0)	25.42 (0)	25.42 (0)
3	C	15	9.83 (0)	18.72 (0)	10.25 (0)	61.20 (5)
	Averages:		22.19 (.50)	22.66 (.20)	21.39 (.20)	33.76 (3.83)

(Number of Periods in which submitted box was filled with only this card)

Table 23: Average Card Decks Submitted by Group 3 Subjects

Individual average card decks do not appear to differ significantly across the three groups, except for several Group 3 individuals who have lower Joker card averages. However, if one looks at the variance of the number of each card placed in the box, clearly one will see that the variance of the Joker card is much greater than one would expect. Figure 3 shows a scatter plot of the average of the three non-Joker cards' standard deviations and the Joker card's standard deviation for each subject. What should one expect to see in this relationship? If subjects played pure actions in the minimax proportions (i.e. they put 100 identical Joker cards in the box 24 times, and 12 times for each of the other cards), then the standard deviation of the Joker card would be 49.4, and any of the other three cards would be 40.34. The ratio of these two numbers is the slope of the reference line in Figure 3.[15] Not surprisingly, the data points for Group 1 subjects fall close to the line. However, for Group 1 and 2 subjects the data clearly lies above this line, with the Group 1 data more concentrated towards the origin. Thus, for these subjects the variance of the number of Joker cards placed in the box is almost across the board higher than one would expect. This is indicative that subjects in these groups are acting much more speculatively on the Joker card than on the other cards. The term "speculate" is intended to convey the idea that a subject forms a belief regarding his opponent's next action, and then for an action he evaluates wether it is a better to place 100 of that card in the box or zero of that card in the box.

(Insert Figure 3 here.)

One possible form of speculation is that if a particular card is a best response for a subject given his belief about his opponent's play, and if his confidence in this belief is strong enough, then he places 100 of that particular card in his box. Thus, if subjects speculate in this nature more for the Joker card than any of the other cards one should see that the percentage of pure actions that are Joker should be greater

[15]This is a conservative reference point, especially for subjects who play a large proportion of strictly mixed strategies. If a subject plays strictly mixed strategies the magnitude of these standard deviations would be smaller and the ratio of interest would be closer to 1:1.

than the proportion of all realized actions that are Joker. For Group 1 subjects, 48.9% of the pure actions were of the Joker card. For Group 2, 356 out of the 513 total pure actions were the Joker card. This is roughly 69.5%, and is certainly much greater than one would expect if all cards were speculated upon to the same degree. In fact, according to Table 17, almost one fourth of *all* subject choices were to play the Joker card as a pure action. Hence, for subjects in Group 2, there is clear evidence that subjects treated the Joker card as special and focused their attention on when it was appropriate to play this card. For Group 3 subjects, only a small percentage of their play involved using a pure action by definition. However, of the limited pure action play, approximately 81% was with the Joker card. The fact that one sees players who adopt both significant amounts of pure Joker play and significant amounts of strictly mixed strategies suggests that an appropriate behavioral model would have players mixing so as not to be exploited in some periods and then playing the Joker card in periods where they feel that it will exploit their opponents.

When a subject does play a mixed strategy it is clear that in most cases it is significantly different from the minimax strategy. In examining the data in the appendix, one notices that no single mixed strategy is predominantly played by the subjects. Furthermore, only a few people appear to lock onto a particular mixed strategy and play it exclusively. Of those who do, none of them are using the minimax strategy (although one subject does play the minimax strategy for the last four periods), but they instead typically adopt focal strategies such as the equiprobable one. In fact, when we examine the average selection for each Group 3 subjects, presented in Table 18, there appears to be a significant heterogeneity in subject choices. As B&R point out, subjects must jointly and significantly deviate from the minimax solution in order to substantially affect win rates. Hence, the subjects may not be costing themselves significantly with these deviations. However, this heterogeneity is suggestive of an alternative model of behavior that would rationalize a striking characteristic in the data, namely, that the average number of wins across pairs is close to minimax proportions, but the variance of wins is too great to be generated by the binomial model implied by the minimax hypothesis.

Consider the case where Column plays the Joker card with some probability x and plays the other cards with probability $\frac{1-x}{3}$, and assume that Row does the same except with probability y. I will refer to such strategies as Joker mixed strategies. Now consider the different Row win rates associated with different pairs of Joker mixed strategies, given in Figure 4. Clearly from the contour graph, the Row winning percentage is rather flat. Thus, a large variety of Joker mixed strategy profiles can generate winning percentages close to that under minimax play.

(Insert Figure 4 here.)

If we look at the corresponding variances of the number of stage game Row wins when 60 stage games are played, in Figure 5, we see that for any single mixed strategy profile the theoretical variance is always well below the sample variance calculated for any of the experiments. In fact the maximum theoretical variance is 15, and this is for strategy profiles that generate a 50% win rate. This is extremely small, considering that the smallest sample variance is 24 for Treatment 3. Thus, a model that assumes all subjects in each role use the same mixed strategy is clearly inconsistent with the data.

(Insert Figure 5 here.)

Let's loosen this restriction and assume that each player adopts a Joker mixed strategy, but that each player selects his stationary strategy by selecting an independent draw from the uniform distribution on the interval $[.4-\epsilon, .4+\epsilon]$. This draw becomes the probability with which he plays the Joker card. Since the expected strategy here is the minimax one, the expected win rate for the row play is 40%, independent of ϵ. But let us ask how the variance of the number of Row wins changes as ϵ grows.

A simulation is conducted to answer this question. I vary ϵ between .001 and .399 at .001 intervals. For each value of ϵ, 250 experiments are simulated. In each experiment 15 pairs play 60 stage games with O'Neill's payoff table. Each subject receives an independent draw from the uniform distribution on $[.4-\epsilon, .4+\epsilon]$. This determines a mixed strategy from which sixty random moves are generated. The strings of moves are matched and the total number of stage games won by the row player is calculated. Next the mean and the variance of the number of Row wins is calculated. Then the average of these statistics is taken across the 250 experiments. Figure 6 shows how the sample variance changes with ϵ. Clearly, the variance does increase with ϵ, but it does not significantly increase until after $\epsilon = .25$. While this story is appealing, it probably only explains some of the high variance of row win rates observed in O'Neill's experiment and the three experiments reported here. Consider that the sample variances from the experiments ranged from 24 to 43. If heterogeneous mixtures was the sole source of high variance in Row win rates, then the value of ϵ would have to be on the magnitude of at least .32, as is shown in Figure 6. In conclusion, the heterogeneity of subjects mixed strategies may lead to some increase in the variance of the Row win rates, but the heterogeneity does not appear to be great enough to account for all of the variance. However, this result, coupled with the presence of the correlation effect, may generate the high variances calculated from the data.

(Insert Figure 6 here.)

Thus, the new view of how people play in the O'Neill game gives us several results. First, play is not affected by whether or not one's opponent is shown his or her selected mixed strategy. Second, there is strong heterogeneity in the manner in which subjects use the mixed strategy device. Third, for almost all subjects there tends to be active speculation on the effectiveness of the Joker card. Fourth, the selected mixtures are heterogeneous, and this can contribute to the high variance of Row wins found in the data and allow still account for the average Row win rate to remain close to .4.

5 Concluding Remarks

This paper reports on a set of experiments designed to discriminate among the possible sources of the failure of the minimax hypothesis' predictions for the unique mixed strategy minimax equilibrium of the O'Neill game. First, the experimental design allows one to identify the causes of the serial correlation in subjects' action choices. Second, the design allows one to discriminate among a set of mixed strategies which generate distributions over actions and wins that are similar to the ones generated by the minimax strategy. This is accomplished by introducing a new methodology for eliciting mixed strategies, which overcomes some of the problems of previous experimental attempts to elicit mixed strategies.

The mixed strategy device allows subjects to select their actions by defining a probability distribution over the set of actions in a simple fashion. Thus, if serial correlation remains in subjects' actions, it must be the result of non-equilibrium repeated game strategies and not the product of subjects' inability to generate sequences of actions that are independently and identically distributed. The data reveal that when the device is available there is a smaller proportion of subjects whose play is serially correlated than in the original O'Neill experiment. However, a significant number of subjects' play is still found to be correlated with the past play of their opponents. This time dependency is indicative of intended behavior by subjects that is not minimax. Furthermore, I find within pairs that there is significant contemporaneous correlation of play that is conditional upon at least one player playing his strategically unique action (his Joker card). B&R conclude that this is a major source of the Row win rates' deviations from the binomial model implied by the minimax solution.

When inspecting the mixed strategies subjects select with the mixed strategy device, it is clear that subjects tend to focus on whether or not to play the joker card. Thus, the Joker is played as a pure action an inordinate amount of the time. However, when subjects did not play the Joker as a pure action, subjects tended to play a wide variety of mixed strategies that are not minimax. It is shown that this heterogenous adoption of mixed strategies can rationalize some of the characteristics of the aggregate data, namely, that the average number of Row wins across subjects is close to the minimax prediction, but that the variance of wins is too large to have been generated by the binomial model. Thus, heterogenous mixed strategy play amplifies the effects already noted from contemporaneous correlation on the distribution of Row winning percentages.

While the results presented do show that the minimax hypothesis does not adequately describe individual play in the O'Neill game, the results are also suggestive of ways to develop more appropriate theoretical models. It seems clear that a desirable theoretical model would incorporate the qualitative notion that subjects will play both mixed and pure strategies in the repeated game context. The data here also suggest that models of populations of players who adopt heterogenous strategies may rationalize some of the anomolous features of the data. In summary, while the minimax hypothesis fails to fully rationalize the behavior of individuals in these two-outcome strictly competitive games, there still does not exist a more robust alternative theory. Thus, it appears that the next step in this line of research should be the construction of theoretical models that will surpass the benchmark of empirical validity established by the minimax hypothesis.

Figure 1: Standard Deviation of the Joker Card versus the Average Standard Deviation of the Non-Joker Cards.

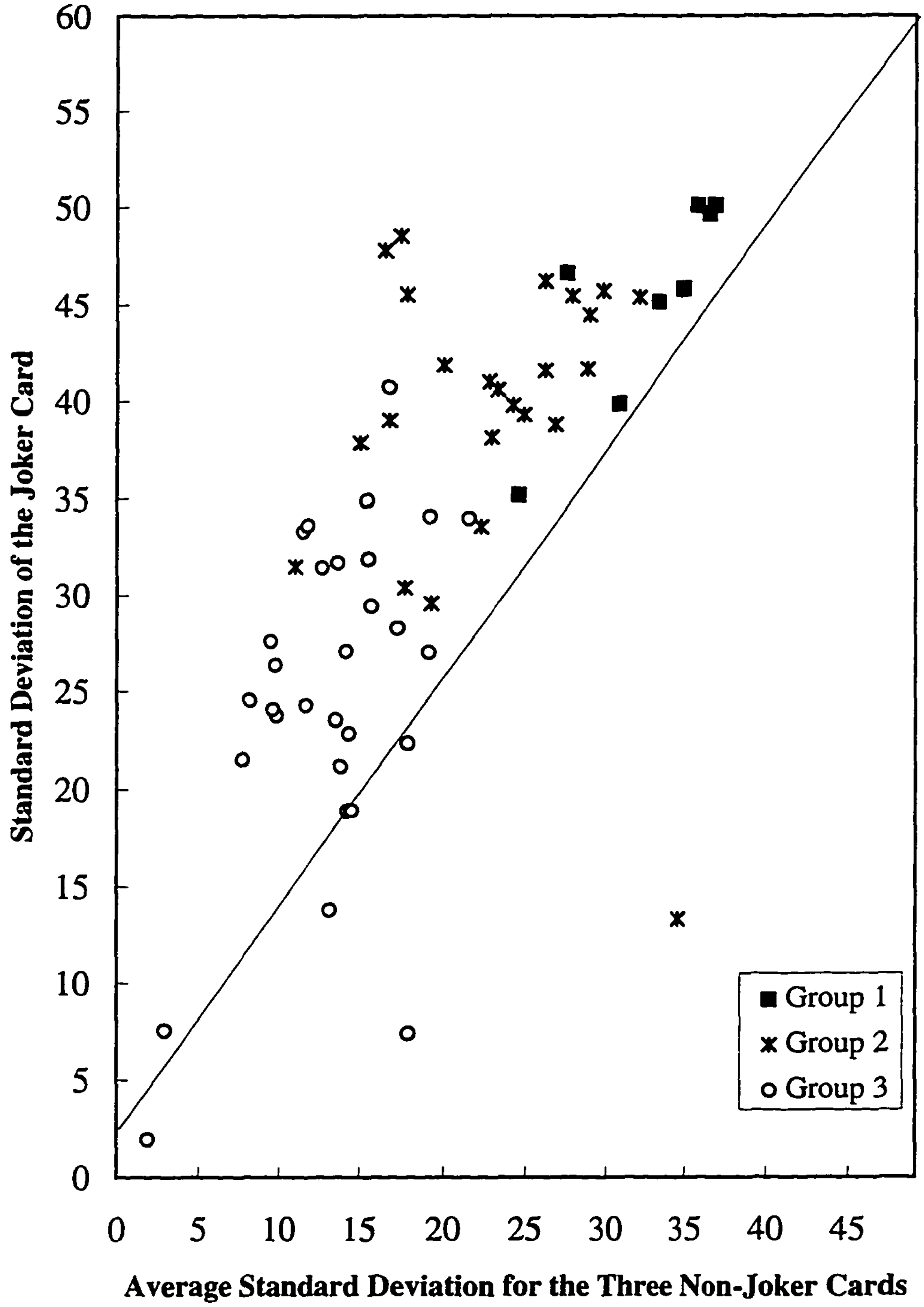


Figure 2: Contour Plots of Expected Row Winning Percentage for Different Joker Mixed Strategy Profiles

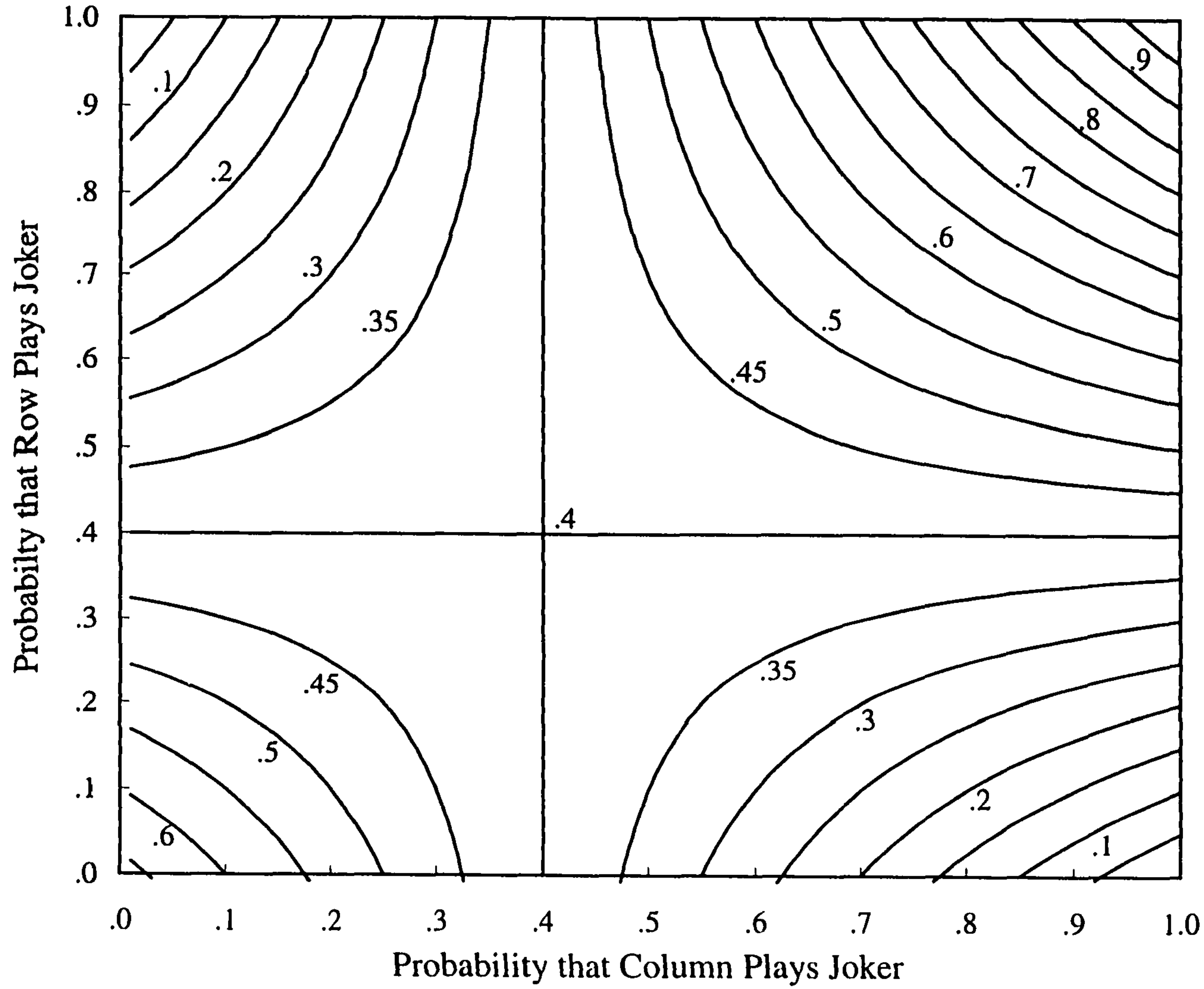


Figure 3: Contour Plots of the variance of the number of Row Stage Game Wins for 60 stage games under Joker mixed strategies

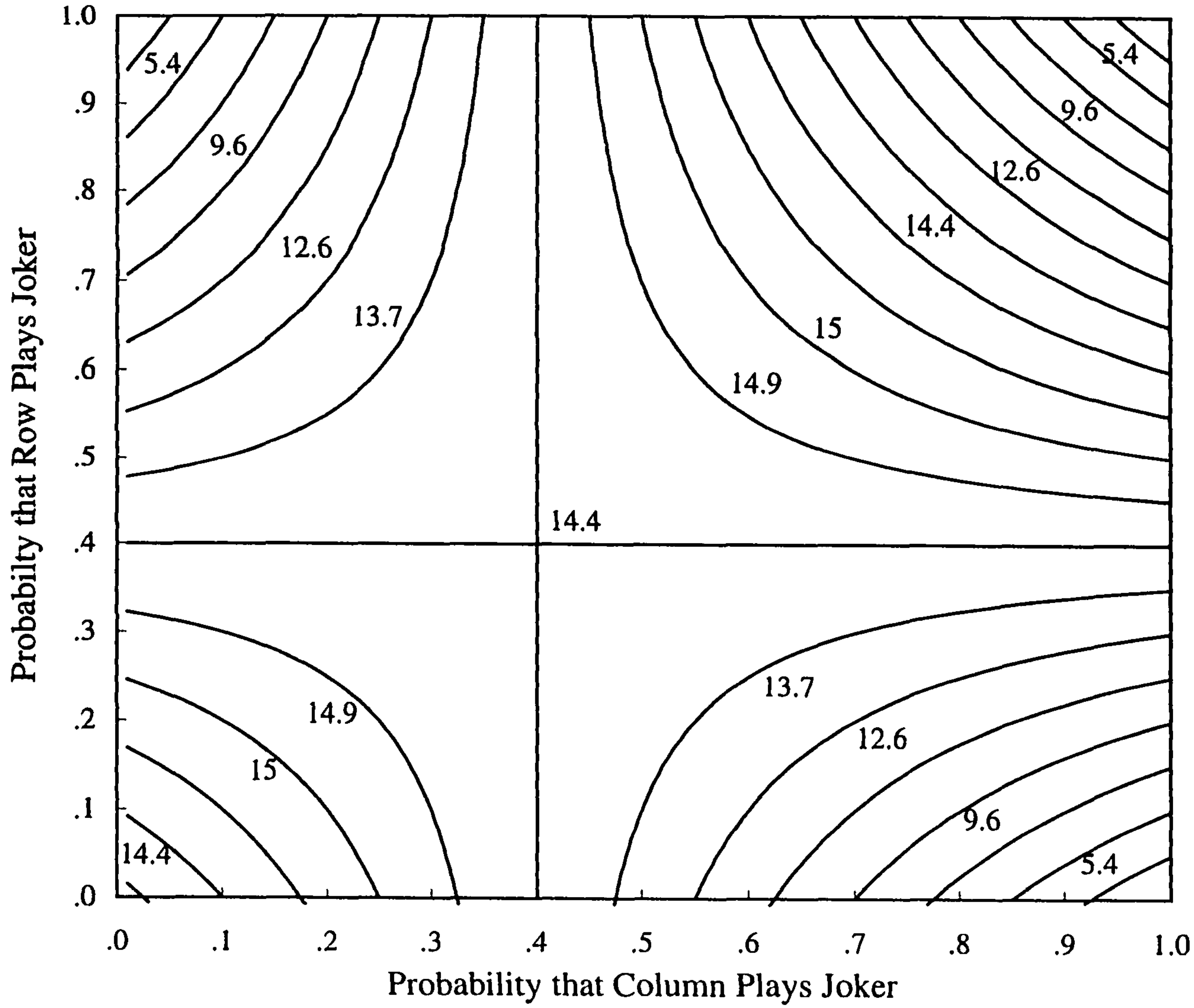
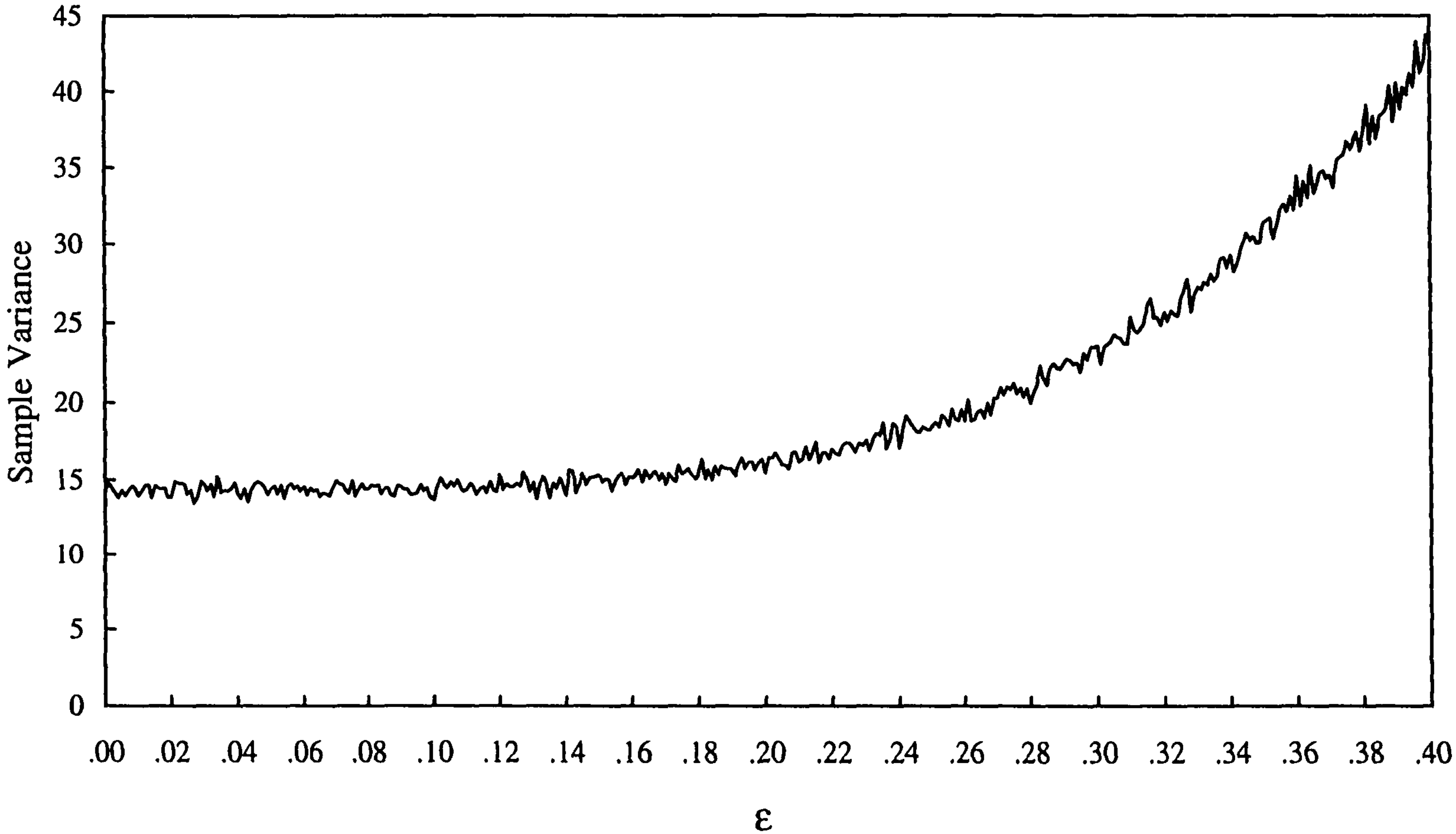


Figure 4: Average Sample Variance of Simulated Experiments as ε is Varied



Chapter 2: Heterogeneous Risk Attitudes and Equilibrium in Observable Games

1 Introduction

The analysis of experimental game data often leads to a conclusion that individual subjects do not play their respective parts of a Nash equilibrium. However, when the same individual data is aggregated, it is often consistent with Nash equilibrium. I do not know of any theoretical model with an equilibrium that simultaneously explains non-Nash strategy profiles at the individual level, but roughly Nash at the aggregate level. In this paper, a theoretical model is presented that could resolve this paradox.

In developing a theoretical model to rationalize this phenomenon, a general methodology is developed to analyze strategic situations in which preferences are not perfectly observable. This in turn provides a framework that applied researchers can use to derive statistical specifications for situations typically modeled as games.

At the heart of the analysis is a distinction between game theorists' analyses of games and the way games naturally occur. When Von Neumann and Morgenstern [40] constructed the framework for modern game theory, a crucial element was the identification of a mapping from observable outcomes to unobservable expected utility values. The existence of such a mapping allows an analyst to explore, both generally and tractably, the intricacies of strategic uncertainty and interaction found in games. However, when social scientists model naturally-occurring strategic situations, observable outcomes are not in terms of expected utility numbers but are instead *data*. These data may take the form of money, commodity bundles, or dollar profits, or even non-economic data such as a number of fatalities or reproduction rates.

The obvious difference between the study of abstract games and their naturally-occurring counterparts is the extent of knowledge which researchers and players possess regarding players' mappings from observable outcomes to expected utility values. An applied researcher usually lacks this knowledge and must either interpret observed outcomes as units of expected utility or accept that his model provides qualitative insights but has few testable implications (the model does not make clear predictions for the data generated from strategic situations.) I believe that neither approach adequately answers the question of how players actually play games.

The model in this paper begins with the assumption that outcomes from games are observable and that players' preferences over these uncertain outcomes may be heterogeneous, private information. For simplicity, all observable outcomes are assumed to be dollars. The analysis begins by considering the class of simultaneous move games for which the outcomes are allocations of dollars rather than expected utility levels. Attention is further restricted by only considering games in which outcomes are common knowledge but individuals' mappings from dollars to expected utility numbers are not commonly known. This class of games is referred to as *observable normal form games.* When an individual plays an observable game, his payoffs should be interpreted as arguments of an unknown von Neumann-Morgenstern (VM) expected utility function rather than as expected utility numbers.

To analyze this class of games, a new model of disturbed games is presented. It is assumed that players' unknown risk preferences are characterized by independent, identically distributed random variables. Each player is assumed to know his realized risk attitude when he plays, but to know only the distribution of other players' risk attitudes. The realized risk attitudes are the disturbances of the observable normal form game.

In characterizing the Bayes-Nash equilibrium of the game of incomplete information, two key ideas are developed. First, a player's action choice is seen to be determined not only by the inherent risk attitude of his type, but also by the location of his type relative to the support of the distribution of risk preferences. Second, the paradox previously described can be rationalized by near purifications of the disturbed game, i.e. where a large amount of probability mass is located around the risk neutral type.

The normal form games studied here are disturbed in a manner that differs from Harsanyi [15, 1973] and McKelvey and Palfrey [26, 1995, hereafter M&P]. In Harsanyi's model, every player receives a random variable with dimensionality equal to the number of possible pure strategy profiles. Each realized component of this random variable is added to the known and fixed utility of its corresponding strategy profile. M&P's model is similar except that each player receives a random variable of a dimension equal to the cardinality of his strategy set. Realized components of this random variable are added to the calculated expected utilities of the corresponding strategies, independent of other players' strategies. In the model presented here, the set of possible payoffs in the *observable* game is considered for each player. Then for each player, a random component is added to each element of this set while preserving the ordering of the fixed payoffs. Some of the effects of these different specifications will be highlighted in the examples presented in this paper.

The paper is organized in the following manner. In section two, a simple game is presented that highlights the effects of heterogeneous risk attitudes on play in games. Then a solution for the game is derived from the new framework. In section three, it is shown that the solution makes systematic and testable predictions, and the predictions are compared to those of the M&P solutions of the game. In section four, an example is presented that highlights how a small amount of heterogeneity in risk attitudes can generate data that are close to the complete information Nash equilibrium at the aggregate level, but are very different at the individual level. When this heterogeneity is collapsed upon the risk neutral type, the Nash equilibrium of this game arises as a purification similar to the purification presented in Harsanyi. The framework is formalized and analytical results are presented in section five. In section six, results on the existence of pure strategy equilibrium in games with private information are reviewed. Finally, in the concluding section, the results of this paper of placed are summarized and additional research topics are suggested.

2 The Gamble/Safe Game

Consider the two person 2×2 game in Table 1; in which δ is a real number between 0 and 1,and G stands for "Gamble" and S for "Safe".

		Column G	Column S
Row	G	$0, 1$	$1, \delta$
	S	$\delta, 0$	δ, δ

Table 1: Gamble/Safe Game

If one adopts the usual interpretation that the elements of the cells are expected utility values, this game has a unique Nash equilibrium in which each player plays a mixed strategy: Row plays R with probability δ and Column plays R with probability $1 - \delta$. When the table entries are allocations of VM utility, the matrix will be called the *utility version* of the game. Note that, although the example emphasizes a mixed strategy equilibrium, the results apply to pure strategy equilibrium as well.

When one takes the game matrix payoffs to be dollars, then it will be called the *observable version* of the game. In order to provide a complete and consistent analysis of the game one needs to provide a structure that adequately describes the manner in which players map dollar outcomes to their respective expected utilities. Furthermore, one needs to describe what each player knows about the other players' mappings. Also, one would like to specify what a non-participating observer knows about these mappings. The exposition here proceeds to define these structures and to illustrate an analysis of the observable version of the Gamble/Safe game.

If one assumes that a population of players of the observable version of the game has heterogeneous risk preferences, new considerations enter into the play of the game. If the population tends to be highly risk averse, for example, we see that this will tempt Row players to adopt their Gamble strategy, as the population on the whole places high value on safe strategies. Likewise, Column players will be tempted by their Gamble strategy when the population tends to be risk loving. This game underscores the idea that while the ***absolute*** intensity of individual risk preferences is important in game outcomes, the ***relative*** intensity of individual risk preferences is just as important.

All relevant risk attitudes for this game can be captured by the set of VM utility functions such that $u(0) = 0$, $u(1) = 1$, and the expected utility of δ takes on an intermediate value. Let us assume that a player's risk preferences (i.e., his expected utility of δ) are characterized by the random variable ω_i, with support equal to the unit interval and with an atomless distribution function F. Table 2 presents the utility version of the Gamble/Safe game.

		Column G	Column S
Row	G	$0, 1$	$1, \omega_C$
Row	S	$\omega_R, 0$	ω_R, ω_C

Table 2: Utility Version of the Gamble/Safe Game

Thus, the game is disturbed by allowing the expected utility of the Safe payoff to be stochastic. This random variable completely characterizes the set of types and beliefs necessary to define a game of in complete information. The type space for the game of incomplete information is the closed unit interval. Perhaps it is more informative to view the type space as it is presented graphically in Figure 2.

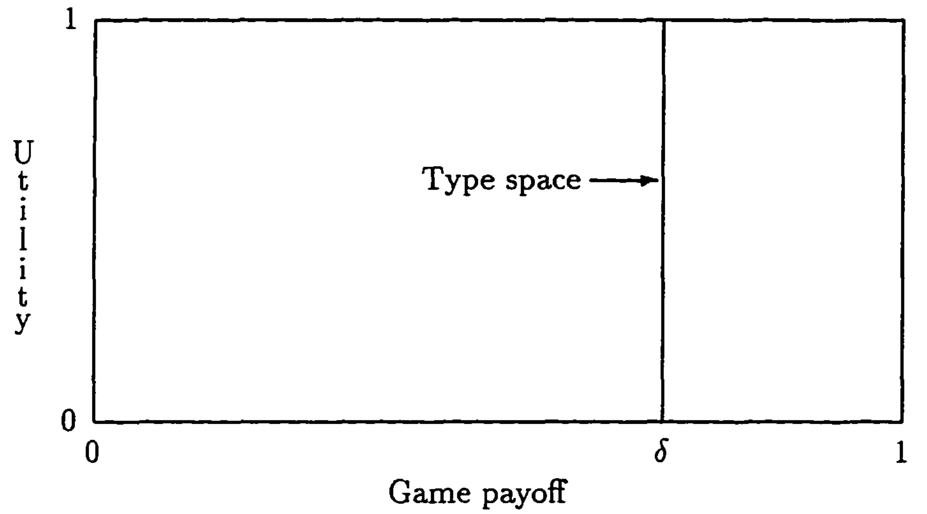


Figure 1: Type Space of the Gamble/Safe Game.

The x-axis values are levels of monetary payoff, and y-axis are corresponding levels of expected utility. For this decision problem, the relevant monetary outcomes are zero, δ, and one. Hence, any strictly increasing expected utility function is equivalent to a piece-wise linear function with a kink at the value of δ. The kink solely determines the relevant level of risk aversion. With the utility levels of zero and one fixed at zero and one respectively, the value of $u(\delta)$, i.e. the kink in the utility function, is constrained to lie between zero and one.

Now we need to define strategies for the disturbed version of the Gamble/Safe game. Let the set of behavioral strategies be the set of measurable functions

$$\varphi_i : [0,1] \longrightarrow [0,1], \quad i = R, C.$$

These functions have the following interpretation: given ω_i (a player's utility for the dollar amount δ), the behavioral strategy gives the probability that he chooses action R.

A natural solution for the observable game involves defining a Bayes-Nash equilibrium (hereafter BNE) for this game. To construct a BNE for this game, consider

Row's action choice given Column's behavioral strategy, φ_C. Row strictly prefers R to S when

$$\int_0^1 [1 - \varphi_C(\omega)]\, dF(\omega) > \omega_R. \tag{5}$$

The left hand side of (1) is the probability Column plays S, given φ_C. Now Row's type contingent best response correspondence, for which the the range is the set of mixed strategy probabilities of playing G, given φ_C and F, is

$$BR_R(\varphi_C, \omega_R) = \begin{cases} \{1\} & \text{if } \omega_R < \int_0^1 [1 - \varphi_C(\omega)]\, dF(\omega) \\ [0,1] & \text{if } \omega_R = \int_0^1 [1 - \varphi_C(\omega)]\, dF(\omega) \\ \{0\} & \text{if } \omega_R > \int_0^1 [1 - \varphi_C(\omega)]\, dF(\omega). \end{cases} \tag{6}$$

Thus, according to (2) there is a step function that is a best response to φ_C, in which Row plays G with probability one whenever his type, ω_R, is less than or equal to the probability Column plays S given φ_C, and Row plays S with probability one when his type, ω_R, is larger than the probability Column plays S given φ_C.

Column weakly prefers G to S, given φ_R, whenever

$$\int_0^1 \varphi_R(\omega)\, dF(\omega) \geq \omega_R. \tag{7}$$

The left hand side of (3) is the probability that Row plays the Gamble action, given his behavioral strategy. Thus Column will prefer to play G when his type is less than this probability. Column's best response correspondence given his type, ω_C, φ_{Row}, and F, is

$$BR_C(\varphi_R, \omega_C) = \begin{cases} \{1\} & \text{if } \omega_C < \int_0^1 \varphi_C(\omega)\, dF(\omega) \\ [0,1] & \text{if } \omega_C = \int_0^1 \varphi_C(\omega)\, dF(\omega) \\ \{0\} & \text{if } \omega_C > \int_0^1 \varphi_C(\omega)\, dF(\omega). \end{cases} \tag{8}$$

Hence, one can define a step function as a best response to φ_R in a similar fashion as was done for Row.

The unique BNE strategy profile for this game has Row and Column adopting step functions that are best responses to one another. The BNE is unique in the sense that any other equilibrium strategy profiles will have strategies that differ only for a subset of types with probability measure zero. One can easily verify that the following strategies are mutual best responses. others best response correspondence.

$$\varphi^*_{Row} = \begin{cases} 1 & \text{if } \omega_R \leq \omega^*_R \\ 0 & \text{if } \omega_R > \omega^*_R \end{cases} \text{ and } \varphi^*_C = \begin{cases} 1 & \text{if } \omega_C \leq \omega^*_C \\ 0 & \text{if } \omega_C > \omega^*_C \end{cases}, \text{ where} \tag{9}$$

$$\omega_R^* = 1 - F\left(\omega_C^*\right) \text{ and } \omega_C^* = F\left(\omega_R^*\right). \tag{10}$$

Any such pair of behavioral strategies will be referred to as the BNE of the Gamble/Safe game. Also, ω_R^* and ω_C^* will be called the threshold types.

Notice that one can reinterpret the utility version of the game, with the equilibrium probabilities over actions for each player, as presented in Table 3.

			Column	
			$F(\omega_C^*)$	$1 - F(\omega_C^*)$
			G	S
Row	$F(\omega_R^*)$	G	$0, 1$	$1, \omega_C$
	$1 - F(\omega_R^*)$	S	$\omega_R, 0$	ω_R, ω_C

Table 3: BNE Implied Frequencies of Actions in Gamble/Safe Game

If a large population plays this game in equilibrium, one would expect to observe players select actions with the frequencies equal to the probabilities that appear along the left and top border of Table 3. An observer unaware of the true underlying Bayesian game may interpret these frequencies as a mixed strategy profile for the undisturbed game. However, the players are in no sense mixing. The *behavior* generating these frequencies is not driven by an attempt to make one's opponent indifferent among a subset of his alternative actions, but rather by individual choices that are strict best responses, given a realization of the inherent underlying randomness of the game.

3 Implications of the BNE of the Gamble/Safe Game

This section presents some implications of the BNE of the Gamble/Safe Game. First, I present some restrictions on the values that the threshold types may take. Most of these implications on threshold values can only be tested when one assumes that F has a specific form. Second, restrictions are derived for the equilibrium values of $F(\omega_R^*)$ and $F(\omega_C^*)$. These restrictions on the values of F evaluated at the equilibrium threshold types, i.e. the equilibrium probabilities that each player chooses the Gamble action, are testable *without* assuming a parametric form for F.

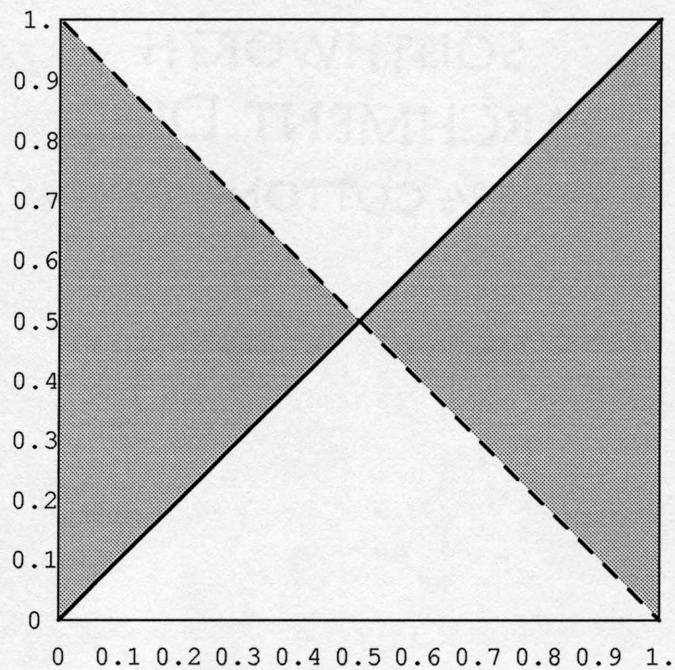


Figure 2: Feasible Equilibrium Threshold Types of the Gamble/Safe Game

A pair of equilibrium thresholds (ω_R^*, ω_C^*) is an element of the unit square. It is shown in the appendix that any equilibrium pair of thresholds must lie in the interior of one of the two shaded triangles or must be the point (.5, .5) of Figure 3. This implication is not testable without assuming a specific functional form for F. However, there is another implication presented in Figure 2 that is testable without a priori knowledge about F. Specifically, except in the case where the equilibrium pair is (.5, .5), there will be exactly two cases:

1. $\omega_C^* > \omega_R^*$. In this case, there is a continuum of types that would play Safe as a Row player but Gamble as a Column player. However, there are no types which would play Safe as a Column player but Gamble as a Row player.

2. $\omega_C^* < \omega_R^*$. In this case, there is a continuum of types that play Safe as a Column player but Gamble as a Row player. However, there are no types which would play Safe as a Row player but Gamble as a Column player.

We can use expression (6) to derive implications regarding the frequencies with which players from a population would adopt the Gamble action. Precisely, the admissible pairs of equilibrium frequencies, $[F(\omega_R^*), F(\omega_C^*)]$, are restricted in the unit square to lie in the interior of one of the two shaded triangles, or be the point (.5, .5), of Figure 3. The derivation of this is given in the appendix.

Figure 3 also presents a graphical representation of the Nash equilibria of the undisturbed versions of the game. As one lets δ vary within the interval $(0, 1)$, the Nash equilibrium follows the off diagonal (the dashed line) of the unit square. Thus, the Nash equilibrium always lies on the boundary of the admissible set of the $[F(\omega_R^*), F(\omega_C^*)]$. In the appendix, it is shown that $[F(\omega_R^*), F(\omega_C^*)]$ corresponds to the undisturbed Nash equilibrium in the case where $\delta = .5$ and the median of F is .5. Hence, frequencies from a population adopting the BNE of the Gamble/Safe game will generally *not* be the same as the Nash equilibrium frequencies.

One can also find several other testable implications of the BNE of the Gamble/Safe game in Figure 3. First, if over 50% of the Column players adopt the Gamble action, then 1) the proportion of Column Gamble play must be larger than that of the Row players, and 2) the sum of the two proportions must be larger than

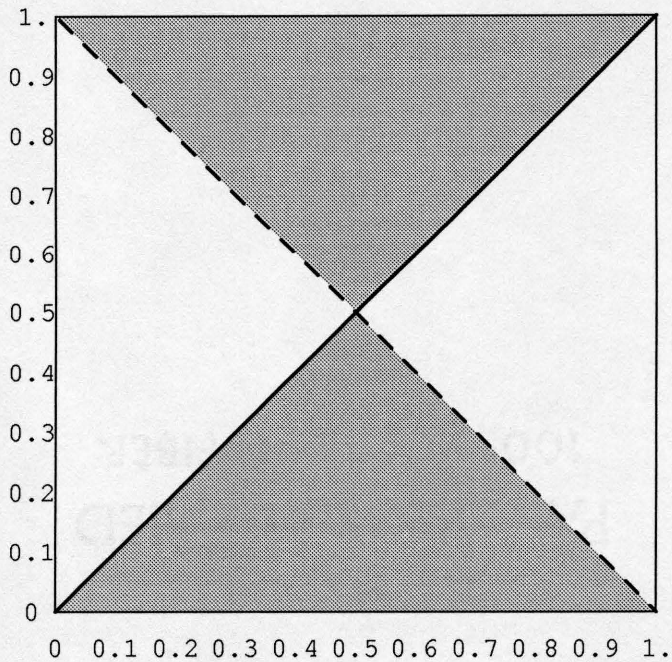


Figure 3: Feasible Equilibrium Proportions of Gamble Play in the Gamble/Safe Game

one. Secondly, if less than 50% of the Column play is Gamble, then 1) the proportion of Row play that is Gamble must be greater than that of Column play, and 2) the sum of these two proportions must be less than one.

3.1 M&P Solution of the Gamble/Safe Game

M&P also present a model of games that allows for simple but natural probabilistic action choice by players. The key to their framework is also a disturbed game coupled with a BNE. However, the nature of the M&P disturbances is different from the framework presented here. M&P assume that each player adds a random component whenever he evaluates the expected utility of an action. This component is assumed to be independent of the other player's action. In order to obtain implications for equilibrium in this framework, one must define a parametric form for these random vectors. M&P show that if each component of each vector is distributed independently Log-Weibull, then in equilibrium players adopt their actions according to multinomial logistic distribution functions.

M&P refer to this game solution as the Logistic Quantal Response equilibrium (LQRE). There are two main differences in the implications for play in the Gamble/Safe game between the framework presented here and the M&P framework. First, in the M&P framework, one must assume a specific parametric form for the random vectors. Second, if the random vectors are assumed, specifically, to be distributed independently Log-Weibull, as in M&P, the set of possible solutions lies within that of our framework.

Let's consider the LQRE of the Gamble/Safe game. Let p be the probability that Row plays Gamble and q be the probability that Column plays Gamble. Row receives a realization of a two component random vector, $(\varepsilon_{R,G}, \varepsilon_{R,S})$, and Column receives a realization of a two component random vector, $(\varepsilon_{C,G}, \varepsilon_{C,S})$. The expected utility for Row's two actions, given any Column mixed strategy $(q, 1-q)$, is

$$U_R\,(Gamble) = 1 - q + \varepsilon_{R,G} \text{ and } U_R\,(Safe) = \delta + \varepsilon_{R,S}.$$

The expected utility for Column's two actions, given any Row mixed strategy, $(p, 1-p)$, is

$$U_C\,(Gamble) = p + \varepsilon_{C,G} \text{ and } U_C\,(Safe) = \delta + \varepsilon_{C,S}.$$

Now in the LQRE we assume that each component of each random vector is distributed independently Log-Weibull, with mean zero and variance $\frac{1}{\lambda}$. Each player plays the action which yields the highest expected payoff. From Luce (1950), it is known that the probability that each action is played follows the logistic distribution.

The key behavioral assumption in this framework is that the players play a BNE of the game of incomplete information. The probabilities that each player chooses the Gamble action in the BNE are the solutions to the following system of non-linear simultaneous equations.

$$p^* = \frac{e^{\lambda(1-q^*)}}{e^{\lambda(1-q^*)} + e^{\lambda\delta}} \text{ and } \tag{11}$$

$$q^* = 1 - \frac{e^{\lambda\delta}}{e^{\lambda p^*} + e^{\lambda\delta}}$$

A solution to expression (7) is a LQRE of the Gamble Safe game.

Figure 3.1 and Figure 3.1 present two different view of the families of LQRE as δ is fixed and λ varies. Figure 4 shows plots of the LQRE correspondences for different fixed δ and λ varying from zero to ∞. All of these curves originate from the center, the solution for which $\lambda = 0$, and connects to its corresponding Nash equilibrium,the case in which $\lambda = \infty$. Clearly all of these solutions are contained within the feasible regions that were found for this paper's framework. Figure 5 presents each of these solution correspondences as a function of λ. It is easily seen that when $\lambda = 0$, i.e. the variance is infinite, the LQRE is always the uniform distribution. As λ increases, the LQRE mixtures approach the Nash equilibrium mixtures. Interestingly, when $\delta = .5$, the LQRE of the Gamble/Safe game is always at the Nash equilibrium for all λ.

4 A Purification of a Gamble/Safe game

In this section, the Gamble/Safe example is used to demonstrate how the distribution over actions is affected as the heterogeneity of risk attitudes becomes more concentrated around risk neutrality. It turns out that the weak limit of the equilibrium distribtuions in such a gane is the Nash equilibrium of the utility version of the game. In the language of Harsanyi, this is a purification of the Nash equilibrium. The striking part of the example is how data can be generated in an equilibrium framework that is consistent with the originally presented paradox. Specifically, for a sequence of games in which the information structure converges to a spike on the risk-neutral-type the sequence of behaviors generated converges to the Nash equilibrium, but in every element of the sequence individual types do *not* play their part of the Nash equilibrium strategy profile.

Consider a version of the Gamble/Safe game in which $\delta = .25$. In the undisturbed game, the Nash equilibrium is where Row plays the Gamble action with probability

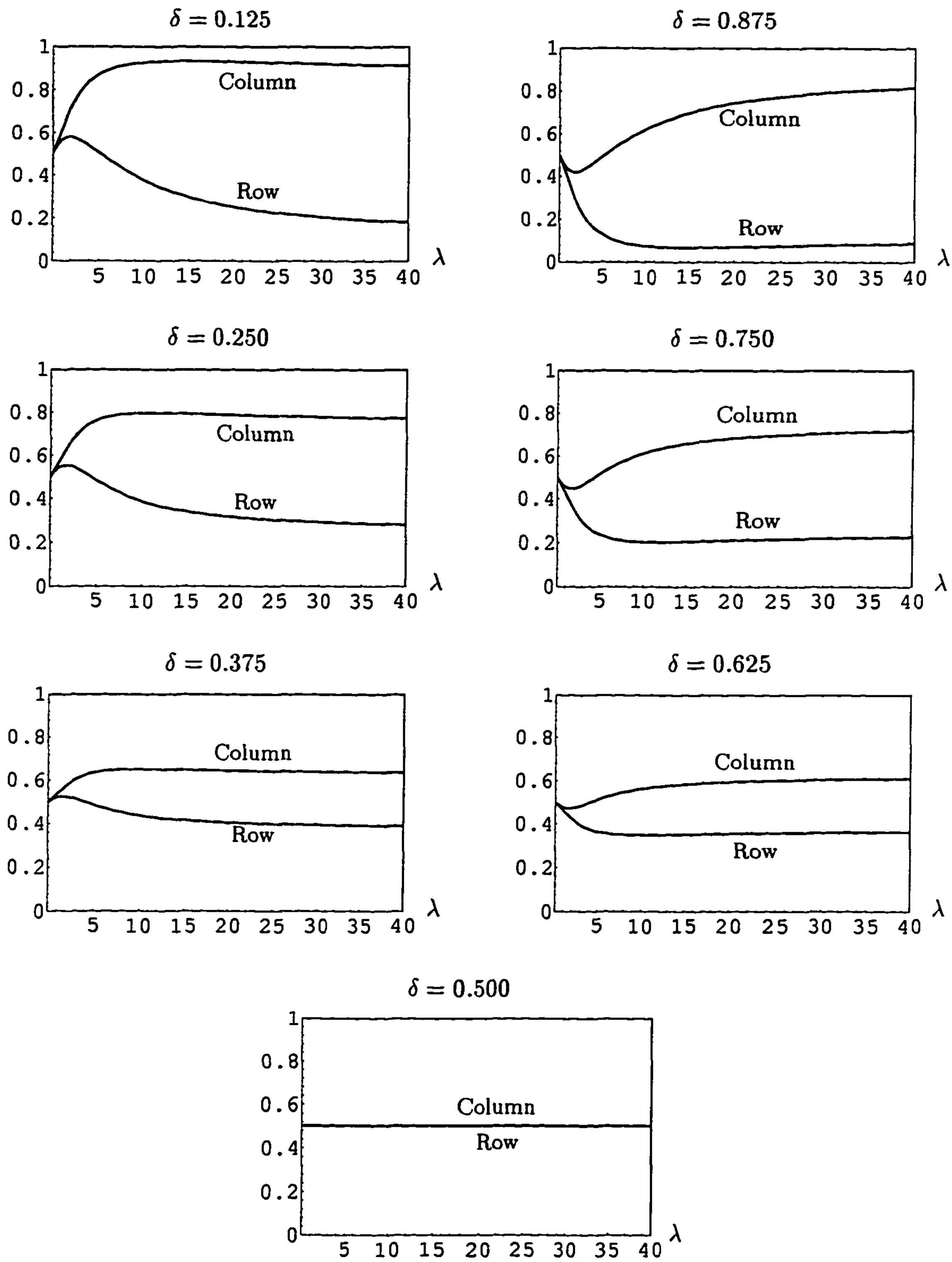


Figure 4: Gamble/Safe Game LQRE Correspondences for different δ

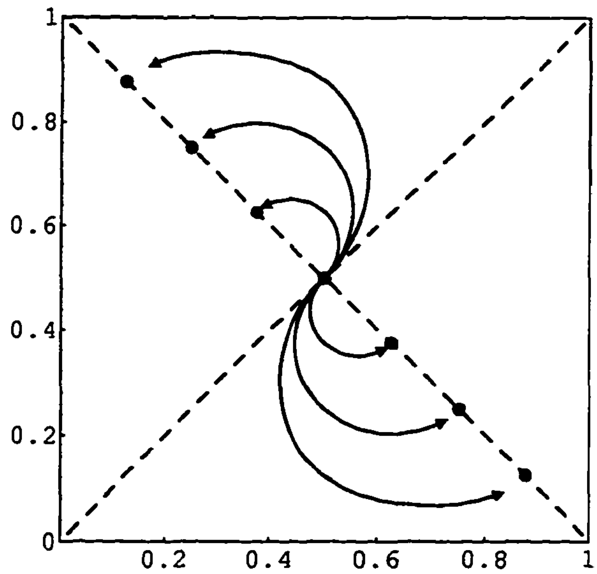


Figure 5: LQRE for the Gamble/Safe Game for different δ Values

$\frac{1}{4}$, and Column plays the Gamble action with probability $\frac{3}{4}$.

Now assume that the game is disturbed, and that ω_i, $i = R, C$, are independently and identically distributed uniformly on the interval $\left[\frac{1}{4} - \varepsilon, \frac{1}{4} + \varepsilon\right]$. Thus, $F(\omega_i) = \frac{\omega_i + \varepsilon - .25}{2\varepsilon}$. Utilizing (6), one can state the equilibrium threshold types as a function of ε:

$$\omega^*_{Row} = \frac{8\varepsilon^2 - 2\varepsilon + 1}{16\varepsilon^2 + 4} \text{ and } \omega^*_{Col} = \frac{8\varepsilon^2 + 2\varepsilon + 1}{16\varepsilon^2 + 4} \qquad (12)$$

Consider the case in which $\varepsilon = \frac{1}{4}$. Then $\omega^*_R = .2$ and $\omega^*_C = .4$. Thus, Row plays the Gamble strategy with 40% probability and Column play the Gamble strategy with 80% probability. These probabilities differ from those in the undisturbed game where Row and Column play their Gamble action with probabilities 25% and 75% respectively. Furthermore, there is a continuum of types, $(.2, .4]$, that play Safe as Row and Gamble as Column, but at the same time there are no types that adopt Gamble as Row and Safe as Column. If this was the true model generating the data one would likely reject the complete information hypothesis at the individual and at aggregate levels with a relatively small data set. However, if one allows ε to approach zero, one can see that the frequencies over actions do rapidly approach the complete

ε	ω_R^*	ω_C^*	$F(\omega_R^*)$	$F(\omega_C^*)$
.25	.2000	.4000	.4000	.8000
.24	.1993	.3943	.3943	.8007
.23	.1987	.3886	.3886	.8013
.22	.1984	.3827	.3827	.8016
.21	.1982	.3767	.3767	.8018
.20	.1983	.3707	.3707	.8017
.19	.1985	.3646	.3646	.8015
.18	.1990	.3584	.3584	.8010
.17	.1997	.3521	.3521	.8003
.16	.2007	.3458	.3458	.7993
.15	.2018	.3394	.3394	.7982
.14	.2033	.3331	.3331	.7967
.13	.2049	.3267	.3267	.7951
.12	.2069	.3203	.3203	.7931
.11	.2091	.3140	.3140	.7909
.10	.2115	.3077	.3077	.7885
.09	.2143	.3014	.3014	.7857
.08	.2172	.2952	.2952	.7828
.07	.2205	.2891	.2891	.7795
.06	.2240	.2831	.2831	.7760
.05	.2277	.2772	.2772	.7723
.04	.2317	.2715	.2715	.7683
.03	.2360	.2658	.2658	.7640
.02	.2404	.2604	.2604	.7596
.01	.2451	.2551	.2551	.7549
.008	.2461	.2541	.2541	.7539
.006	.2470	.2530	.2530	.7530
.004	.2480	.2520	.2520	.7520
.002	.2490	.2510	.2510	.7510
.001	.2495	.2505	.2505	.7505

Table 4: Equilibrium Description for Purification Example

information implications for the aggregate data (see Table 4). However, players of this game will almost always play a strict best response for their type.

5 Theoretical Structure and Results

In this section a general framework for analyzing observable games is described. Then a general existence of equilibrium result is presented for this framework. Finally, it is shown that for any sequences of games for which the distributions over subject types converge weakly to unit mass at risk neutrality, there is a subsequence of equilibrium distributions over actions that converge weakly to a Nash equilibrium of the utility version of the game. The possibility of multiple equilibrium in the utility version of the game leads one to prove this result only for subsequences.

Consider normal form games with the following structure:

1. There is a finite set of N players, $I = \{1, ..., N\}$, indexed by i.

2. Each player i has a set of possible actions, A_i, that is a compact metric space. Let $A = \times_{i=1}^{N} A_i$.

3. For any set S in a metric space, Let $\Im(S)$ denote the collection of S's Borel Sets.

4. Each player i has a continuous outcome function $\psi_i : A \to \Re$. Note that by this continuity assumption and the compactness of A, the range of ψ_i is compact (i.e.. closed and bounded in terval in $\Re$). Recall that one interpretation of ψ_i is that the possible outcomes of a game are allocations of money.

5. Every player receives a type that is a VM utility function over the compact range of the outcome function ψ_i. A type, denoted as u_i, is an element of a set of functions U_i. The elements of U_i have the following properties: for any $u_i \in U_i$, $u_i[\min \psi_i(A)] = 0$, $u_i[\max \psi_i(A)] = 1$, and u_i is a continuous strictly increasing function over the interval $[\min \psi_i(A), \max \psi_i(A)]$. The type set U_i for every i is a metric space, where the metric is the sup norm. One can easily verify that every element of U_i is a unique expected utility function, i.e. no element is an affine transformation of another element, and that every strictly increasing and continuous expected utility function over the range of outcomes has a representation in this set. The unique linear function in this space represents risk neutral preferences. Figure 6 presents a graphical depiction of some typical elements of this space.

6. Define the product space $U = \times_{i=1}^{N} U_i$.

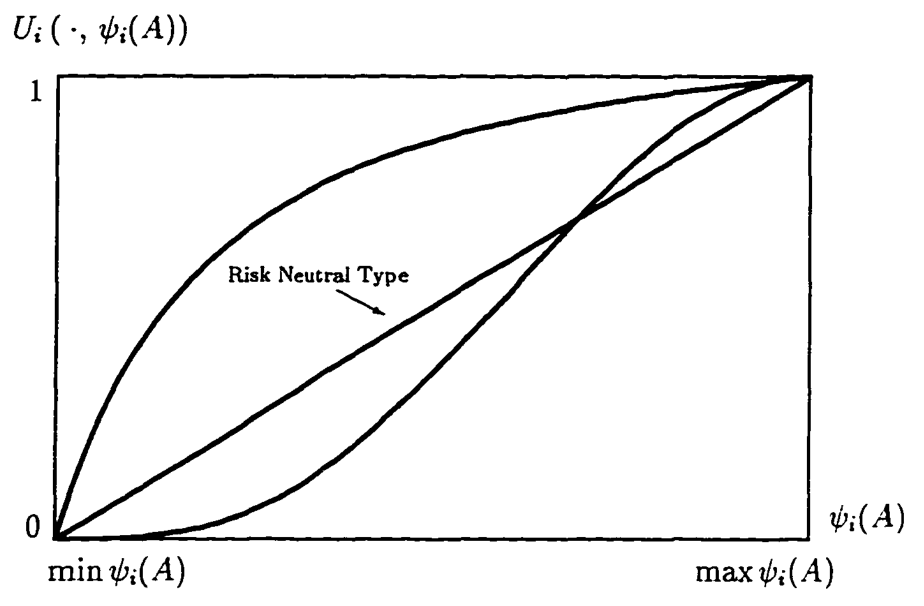


Figure 6: Sample Elements of the Type Space

7. For each potential type $u_i \in U_i$, the associated payoff function is a Borel measurable function $\pi_i : U \times A \rightarrow \Re$. We want a player's payoff function to depend only upon his own type and not those of the other players; thus, for every $u_i \in U_i$, any $a \in A$, and any $u_{-i} \in U_{-i} := \times_{j \in I \setminus \{i\}} U_i$, let $\pi_i(u, a) = u_i [\psi_i(a)]$. It is readily seen that π_i is continuous in each of its arguments.

8. For every U_i, there is a probability measure, η_i, on its Borel sets. Assume that for every $i = 1, ..., N$, η_i is independent of the other type space measures. Hence, $(U_i, \Im(U_i), \eta_i)$ is a probability measure space.

9. $\eta = \eta_1 \times \cdots \times \eta$ is called the information structure of the game.

All facets of the game are assumed to be common knowledge except the opponents' realized types.

In this game, it is assumed that players adopt behavioral strategies. Informally, a behavioral strategy is a function that tells a player what mixed strategy to play conditional upon his realized type. Formally, a behavioral strategy for a player i is a transition probability, b_i, with respect to $(U_i, \Im(U_i))$ and $(A_i, \Im(A_i))$. In other words, b_i is a function that maps $U_i \times \Im(A_i) \to [0, 1]$ such that,

1. $b_i(u_i; \cdot)$ is a probability measure on $(A_i, \Im(A_i))$ for almost every $u_i \in U_i$, and

2. $b_i(\cdot, \tau)$ is a Borel measurable function on U_i, for every $\tau \in \Im(A_i)$

Let B_i be the set of behavioral strategies for player i, and $B = \times_{i=1}^{N} B_i$. A behavioral strategy profile is $b \in B$.

Finally, for each player i his expected payoff to a behavioral strategy profile b is

$$E_{\pi_i}(b) = \int_U \int_{A_1} \cdots \int_{A_N} \pi_i(u, a) b_1(u_1, da_1) ... b_N(u_N, da_N) \eta(du).$$

Having fully defined the structure of the game, the rest of this section is spent demonstrating that an equilibrium always exists, and that the equilibrium correspondence is well behaved in sequences of games where the information structure weakly converges to a complete information game of risk neutral players.

5.1 Existence of Equilibrium

An equilibrium for the game is defined as a behavioral strategy profile b^*, in which for every player i,

$$E_{\pi_i}(b_i^*, b_{-i}^*) \geq E_{\pi_i}(b_i, b_{-i}^*), \text{ for every } b_i \in B_i.$$

To demonstrate the existence of such an equilibrium we will appeal to a common existence theorem and show that the game defined here satisfies all of the required conditions.

Theorem 1 (Glicksberg, 1952) *If for each $i = 1, \dots, N$, (i) the set of equivalence classes that partition B_i is a compact convex subset of a Hausdorff locally convex topological vector space L_i, and (ii) E_{π_i} is a continuous quasiconcave function in each of its variables, then there exists an equilibrium strategy profile b^*.*

The first task, carried out in section 5.1.1, will be to show that the set of *equivalence classes* of behavioral strategies is compact and convex. Then the continuity and quasiconcavity of the expected payoff functions will then be shown in section 5.1.2.

5.1.1 Compactness and Convexity of the Strategy Space

First, it is shown that the set of equivalence classes of behavioral strategies for a player is a compact convex subset of a Hausdorff locally convex topological vector space. If the type space is a complete separable metric space, then we can demonstrate compactness in the fashion of Milgrom and Weber [27, 1985]. In that paper, the set of strategies considered were joint distributions over the action-type product space. In that setting, the set of tight joint probability measures (i.e. the set of distributional strategies) on the type-action product space is compact in the classical topology of weak convergence [5]. However, the type spaces considered in this paper's model are not complete.[16] Hence, to demonstrate the compactness of B_i a different approach, Balder [4, 1988], is taken.

What follows is a description of a set for which the set of all transition probability measures with respect to $(U_i, \Im(U_i))$ and $(A_i, \Im(A_i))$ is a subset. The set of uniformly finite transition measures with respect to $(U_i, \Im(U_i))$ and $(A_i, \Im(A_i))$, denoted L_i, is the set of all functions, $\sigma : U_i \times \Im(A_i) \to \Re$ such that

1. $\sigma(u_i; \cdot)$ is a signed bounded measure on $(A_i, \Im(A_i))$ for every $u_i \in U_i$
2. $\sup_{u_i \in U_i} |\sigma(u_i; \cdot)|(A_i) < +\infty$. (the sup of the total variation of σ)
3. $\sigma(\cdot; \tau)$ is a $\Im(U_i)$-measurable function for every $\tau \in \Im(A_i)$.

Clearly, B_i is a subset of L_i.

[16]To see that U_i is not complete, let the range of the outcome function be $[0, 1]$ and consider a sequence of functions, $u_i^n(x) = (1/n)2x$ if $x \leq 1/2$ and $u_i^n(x) = (2-1/n)x+(1/n-1)$ if $1/2 < x \leq 1$. This Cauchy sequence converges to $u_i(x) = 1/2$ if $x \leq 1/2$ and $u_i(x) = 2x-1$ if $1/2 < x \leq 1$, which is not in the set U_i.

We next proceed to formally define sets of equivalence classes within L_i and B_i. Consider the elements of L_i such that $|\sigma(u_i;\cdot)|(A_i) = 0$ for almost every $u_i \in U_i$; denote this set $\aleph$. Two elements, x and $y \in L_i$ are equivalent if and only if $x - y \in \aleph$. The set of equivalence classes generated in this manner is the *quotient space* generated by $\aleph$, and is normally denoted $L_i/\aleph$. For notational simplicity we will continue to use the notation β_i, σ and L_i, but one should now interpret this notation as a typical element of the equivalence of σ, a typical element of the equivalence class of b_i and the quotient space $L_i \setminus \aleph$.

It is a well known fact that L_i is a Hausdorff convex topological vector space, for eaxample see Choquet [11, 1969, 15.5]. It should also be pointed out that L_i is a normed linear space with the essential sup norm with respect to η_i. The essential sup norm with respect to η_i is defined as

$$\|\sigma\|_\infty = \inf\left\{\mathrm{M} : \eta_i\left\{u_i : \sup_{u_i \in U_i} |\sigma(u_i;\cdot)|(A_i) > \mathrm{M}\right\} = 0\right\}. \tag{13}$$

In the construction of the topology of weak convergence of probability measures. a key component is the set of test functions used in defining convergence in measure. For a probability measure space $(X, \Im(X))$ and the set of bounded signed measures $M(X)$, the usual set of test functions considered is the set of continuous bounded real valued functions on X, $\mathcal{C}(X)$.[17] We want to define convergence in measure and equip L_i with a topology in a similar manner as in this case.

The following is a construction of test functions for L_i. Consider the set of continuous real valued functions on A_i (Note that each element is bounded in the sup norm. Denote this set $\mathcal{C}(A_i)$. Next, consider the set of all functions g such that

1. $g : U_i \longrightarrow \mathcal{C}(A_i)$

2. g is $\Im(U_i)$ measurable

3. $\int_{U_i} |g(u_i)| d\eta_i < \infty$

[17]The set of functions, $M(X)$, is the dual space of $\mathcal{C}(X)$.

Denote the set of all g by $\mathcal{G}_{A_i}^{U_i}$. If we let $\|g\|_1 = \int_{U_i} \|g(u_i)\|_\infty d\eta_i$, then $\mathcal{G}_{A_i}^{U_i}$ is a normed linear space.[18]

The set $\mathcal{G}_{A_i}^{U_i}$ is not only useful in defining the convergence of sequences in L_i, but also useful in constructing a topology on L_i. This is accomplished by recognizing that the set L_i is the set of all continuous bounded linear functionals on $\mathcal{G}_{A_i}^{U_i}$, under the inner product $\psi : L_i \times \mathcal{G}_{A_i}^{U_i} \longrightarrow \Re$, where

$$\psi(\sigma, g) = \int_{U_i} \int_{A_i} g(u_i)\sigma(u_i, da_i)d\eta_i. \tag{14}$$

Under this condition, L_i is called the *dual* of $\mathcal{G}_{A_i}^{U_i}$. [19]

Lemma 1 *L_i is the dual of $\mathcal{G}_{A_i}^{U_i}$.*

Proof of this lemma can be found in various texts, for example see [21, 2, 36], and is basically an extension of the Riesz Representation Theorem.

The *weak** topology on L_i, is defined as the coarsest topology such that all of the functionals $\psi(\sigma, g)$ are continuous. It is in this topology that we will show that B_i is compact. First, by the Banach-Alaoglu theorem [36] the unit ball in L_i is compact. Next,we need to show that B_i is a closed subset of the unit ball.

A natural notion of convergence falls out of the construction of the weak* topology on L_i. Namely, the sequence $\{\hat{\sigma}_n\}$ *converges weak** to $\hat{\sigma}$ if for every $g \in \mathcal{G}_{A_i}^{U_i}$, $\psi(\hat{\sigma}_n, g) \to \psi(\hat{\sigma}, g)$. Now we argue that B_i is a closed set in the weak* topology. First, assume that $\psi(\sigma_n, g) \to \psi(v, g)$, for every $g \in \mathcal{G}_{A_i}^{U_i}$, for $v \in L_i$, $v \notin B_i$, and $\forall\, n$, $\sigma_n \in B_i$. We proceed by the considering the following element of $\mathcal{G}_{A_i}^{U_i}$:

$$1_\tau = \begin{cases} h(a_i) = 1 & \forall a_i \in A_i \quad \text{if } u_i \in \tau \subseteq U_i \\ h(a_i) = 0 & \forall a_i \in A_i \quad \text{if } u_i \notin \tau \subseteq U_i \end{cases}$$

Direct calculation reveals that $\psi(\sigma_n, 1_\tau) = \eta_i(\tau)$ for all n and $\tau \in \Im(U_i)$. Thus it must be true that $\psi(v, 1_\tau) = \eta_i(\tau)$. Recalling that $v \notin B_i$ there is some $\overline{\tau} \in \Im(U_i)$ such that

$$\psi(v, 1_{\overline{\tau}}) \neq \eta_i(\overline{\tau}).$$

[18]The norm denoted $\|.\|_\infty$ is the sup norm.

[19]It is also the case that $M(X)$ is the set of all continuous bounded linear functionals on $\mathcal{C}(X)$, and thus is the dual of $\mathcal{C}(X)$.

This is a contradiction. Hence, B_i is a closed set in the weak* topology.

Finally, by (8), B_i is a subset of the unit ball. Also, the convexity of B_i is straight forward to demonstrate. These arguments demonstrate that B_i is a compact subset of a Hausdorff locally convex topological vector space.

5.1.2 Properties of the Expected Payoff Function

In order to complete the exercise of demonstrating existence of equilibrium all that is necessary is to show that E_{π_i} is continuous and quasiconcave. The first thing needed is an extension that allows us to extend previous results to the product spaces A and U. Consider the product mapping $(\sigma_1, \sigma_2) :\rightarrow (\sigma_1 \times \sigma_2)$ where $(\sigma_1 \times \sigma_2)(u_1, u_2; \Im(\alpha_1 \times \alpha_2) = \sigma_1(\alpha_1)\sigma_2(\alpha_2)$, for $\alpha_1 \in \Im(A_1)$ and $\alpha_2 \in \Im(A_2)$.

Lemma 2 (Balder, 2.5) *The product mapping $(\sigma_1, \sigma_2) :\rightarrow (\sigma_1 \times \sigma_2)$ is continuous relative to the weak* topologies.*

Coupling this lemma with the fact that for every i, $\pi_i \in \mathcal{G}_A^U$, the continuity of E_{π_i} is guaranteed. Also, by the properties of the the integral, E_{π_i} is a linear function and thus quasiconcave. This shows that all of the conditions in the Glicksberg existence theorem are satisfied.

6 Pure Strategy Equilibrium

An appealing interpretation of a behavioral strategy is that it identifies an action (possibly mixed) a player chooses conditional upon his realized type. Various empirical studies [38, 35] have raised serious concerns as to whether mixed strategy equilibrium are accurate descriptions of how individuals behave. Hence, one would like to know under what settings the framework in this paper yields pure strategy equilibrium. A summary of the literature on this topic is now presented.

In Radner and Rosenthal [33, 1982], study games with finite player sets, finite action sets,and in which each player receives a realization of a random variable. It is demonstrated that if each player's random variable has an atomless distributions and if the random variables form a mutually independent set, then a pure strategy equilibrium exists. Furthermore, a stronger result is also shown: for any mixed strategy equilibrium, there exists a pure strategy equilibrium which yields the same level of payoffs to each player and also generates the same distribution over actions.[20] It should be noted that players' payoffs may or may not depend upon the realized values of the random variables. In Kahn and Sun [1, 1995], these results are generalized to countable action spaces.

Without countability of the action space, one can find still find a pure strategy profile that guarantees each player a payoff arbitrary close to that received in the mixed strategy equilibrium. The main results on this topic can be found in Aumann et al.[3, 1983], and Milgrom and Weber.

7 Concluding Remarks

This paper suggests several possible paths to further develop the implications of heterogeneous risk attitudes in observable games. First, one can establish the class of normal form game in which every Nash equilibrium can be approached as the limits of sequences of games in which the distribution over risk attitudes converges to risk neutrality. This would extend Harsanyi's purification results to games with countably infinite action spaces. Second, one can experimentally explore the game used in the examples to investigate the empirical validity of the approach presented. Also, one can attempt to address the issue of whether individuals carry the same risk type with them through sequences of strategic situations, or develop a stochastic process of types that affects the equilibrium of the game. Finally, one can develop econometric specifications based upon the model of disturbed games for cross sectional and panel data sets.

[20]This is what Radner and Rosenthal, as well as Milgrom and Weber, refer to as "purification", however I want to reserve that term for the concept defined in Harsanyi, namely, the approach of a sequence of equilibria of a disturbed game to that of the undisturbed game as the size of the disturbances goes to zero.

8 Proofs of Propositions

The derivations of the testable implications of the Bayes Nash equibilirum of the Gamble/Safe game are now presented.

Proposition 1 *If* $\omega_R^* > \omega_C^*$, *then*

1. $F(\omega_C^*) < .5$.

2. $\omega_R^* > .5$.

3. $F(\omega_R^*) + F(\omega_C^*) < 1$.

4. $\omega_R^* + \omega_C^* > 1$.

5. $Median\ of F > .5$.

Proof 1. From equation (6), $1 - F(\omega_C^*) > F(\omega_R^*)$. This implies $F(\omega_C^*) < .5$.
Proof 2. Result 1. implies that $1 - F(\omega_C^*) > .5$. By substituting from (6), $\omega_R^* > .5$.
Proof 3 and 4. The results are immediately calculate from equation (6).
Proof 5. One can deduce that $\omega_C^* < .5$ from 4., and combining this with 1., 2. and (6) one gets that $F(\omega_R^*) < .5$. Hence, the median of F is strictly less than one half.

Proposition 2 *If* $\omega_R^* < \omega_C^*$, *then*

1. $F(\omega_C^*) > .5$.

2. $\omega_R^* < .5$.

3. $F(\omega_R^*) + F(\omega_C^*) > 1$.

4. $\omega_R^* + \omega_C^* < 1$.

5. $Median\ of F < .5$.

Proofs: The proofs are the same as in proposition 1 except with the inequalities reversed.

Chapter 3: An Experimental Investigation of the Incentives to Form Agricultural Marketing Pools

1 Introduction

A *marketing pool* is an agreement between producers wherein participants exchange their output for shares of the cooperative's profit. These cooperatives are prevalent in agricultural industries. For example, in 1989 U.S. agricultural marketing pools earned over fifty-three billion dollars in revenues [37] and accounted for twenty-six percent of total U.S. farm marketings [22]. The fact that pooling has been (and continues to be) a significant empirical phenomenon implies that it offers participants advantages over independent activity. These advantages may include economies in marketing, inventory control, and price stabilization for individual producers. Furthermore, once firms form pools, it is conceivable that they will attempt other types of cooperative behavior. For example, a pool of producers may be able to act as a monoposony in input markets or as a monopoly in output markets.

With potentially large returns from pooling, one may ask why producer cooperatives are not more common in markets in which individual firms have little market power.[21] This paper argues that the incentive to pool depends on individual characteristics of firms. When these individual characteristics interact with market characteristics, pools may become more or less likely. This paper reports results from experimental and theoretical research that addresses a historical case study [18]. This study attempts to determine when a marketing cooperative is incentive compatible for firms by identifying key individual and industry characteristics that determine the success of attempts to form large pools.

[21]In this paper we attempt to answer this question from the perspective of individual firm incentives. Besides incentive compatibility, antitrust laws and the restrictions these present affect the number of marketing pools formed. For example, contemporary antitrust cases that have centered on the issues of cartelization and marketing pools include U.S. vs. Appalachian Coal (228 U.S. 344(1933)) and U.S. vs. Socony-Vacuum Oil Co. (310 U.S. 150 (1940)).

Hoffman and Libecap (hereafter HL) explore the pooling issue by examining the United States orange industry from the late 1920's through the early 1940's. At that time, there were two major orange growing regions, Florida and California. In California, 90pools, while at the same time only limited pools formed in Florida. HL hypothesizes that the reasons for the asymmetric pooling have to do with differences in the respective market structures. Namely, the driving forces behind pooling in California and the lack of pooling in Florida resulted from differences in risk and cost structures in the two regions. Regardless of the region, "risk" for the orange growers was in terms of the random price received for output. By pooling, firms reduce risk by reducing the variance of the market price. We will call this motive the *insurance* incentive, and it was different for the two regions. The second industry characteristic that differed across the two regions was economies of scale in marketing and shipping. Since individual growers were small, it is likely that no single seller could realize these economies. We will refer to this motive as the *cost* incentive. Further elaboration of these two differences in the industry structure and why an experiment is necessary to test the pooling hypothesis will be provided below.

Differences in the insurance incentive for Florida and California growers arose from the heterogeneity of subjective price expectations. In California, there were only two varieties of oranges harvested, homogeneous growing conditions, and the fruit could be left on trees for a considerable length of time. These factors led to a homogenous product for California growers and expected seasonal mean prices that were constant across growers (HL, p.208). In Florida, five varieties of oranges were grown, growing conditions varied across the region and the fruit could not be stored on the tree because of warm nights. Furthermore, with different sub-seasons for the different varieties, Florida oranges were a predictably heterogeneous product throughout a given growing season. This led to predictable differences in mean prices across sub-seasons in Florida.[22] Thus, orange growers in Florida had heterogeneous

[22]Hoffman and Libecap p.209 demonstrate that the mean price of Florida oranges followed a predictable U-shaped pattern in any given season. Thus growers whose fruit matured early or late in the season possessed a higher expected price than a grower whose fruit matured in the middle

price expectations, in contrast to California growers.

Economies of scale in shipping and marketing may have led to cost differences between the two regions. The fact that California had fewer varieties of oranges made storing and processing produce simple tasks. In Florida, these same tasks required a more complicated technology because a large cooperative would handle a larger variety of oranges. Therefore, marginal marketing cost for a large California marketing cooperative may have been decreasing for some levels of output, while marginal cost for a large Florida cooperative was likely increasing for comparable levels of output.

The markets for transporting oranges from the two regions to the major markets of the East and Midwest were regulated. However, California growers may have had potential economies in bargaining and lobbying for lower transportation rates. The only way for California growers to get their output to many major markets was to transport the fruit by rail. Thus, if large cooperatives had not existed, growers probably would have had a difficult time in negotiating with the two main transcontinental railroads (Union Pacific and Southern Pacific) and in forming successful lobbies with the I.C.C., which regulated the interstate fare structure. On the other hand, Florida growers could ship their fruit via railroad, trucks, and ships. This larger class of alternatives made the shipping market for Florida oranges more competitive and large lobby groups may have not been necessary to obtain competitive shipping rates. Consequently, differing returns to organized lobbying efforts may have resulted from a difference in the number of shipping alternatives available to the two regions.

The preceding analysis considered here suggests compelling reasons for large, season long-pools in California, but fewer large, season long-pools in Florida. More generally, the hypothesis examined here presents a powerful framework for providing insight into the pooling behavior within an industry. In this instance, HL believe they have isolated the few essential explanatory variables behind a phenomenon (namely, differing price expectations and potential economies in marketing), but they lack suf-

---

of the season.

ficient data to test their conjectures. Still, one can test their theory through the implementation of a laboratory experiment incorporating these explanatory variables. Such an experiment can provide a strong test for the pooling theory if it requires subjects to make repetitive decisions, as in the naturally occurring economy. This experimental data would then allow one to test for the dynamic consistency of the economic model's predictions, which one could not validate with cross sectional field or experimental data. Hence, an economics experiment could possibly qualitatively validate HL's conjecture about the historical phenomenon of interest.

This paper develops and verifies a theoretical hypothesis from a historical case study using the tools of Bayes-Nash game theory, experimental economics, [23]and discrete choice panel data econometrics. Due to the applied nature of the pooling problem, we design and conduct an experiment that strives to possess as much parallelism[24] as possible. Second, we extend the existing theory to allow for heterogeneous risk attitudes over uncertain profits by using a game theoretic model with incomplete information. Then we develop, estimate, and conduct hypothesis tests on statistical models of the experimental data to evaluate the validity of our theory. The discussion that follows is motivated by the quest to discover the motives for pooling and to separate these motives from the properties of the data that are generated by the limitations of the experimental environment.

2 Experimental Design

The experiments reported in this paper [25] are designed to capture the effects of cost and insurance incentives on firms' decisions to join marketing pools that do not

[23]For an extensive survey of the industrial organization experimental economics literature, we refer the reader to Holt's [20] survey article.

[24]According to Smith [39], parallelism in a valid economics experiment implies that the experimentaldesign has captured the essential features of a naturally occurring institution being modeled.

[25]All experiments were conducted at the Economic Science Laboratory (ESL) at the University of Arizona. All subjects were undergraduates enrolled at the University of Arizona during the summer or fall semestersof 1993. For a complete set of all instructions, please contact the authors.

engage in price or quantity fixing. In all experiments, subjects are asked to make a repeated simple binary choice. For ten decision making periods, ten individuals are each given a single unit of a fictitious good which they must sell.[26] They must choose whether to sell their units with other participants' units in a pool or to sell their units by themselves. Prior to making this decision the subjects can meet in a neutral place for up to five minutes and discuss their decisions, after which they must return to a private decision making area. An individual's decision is always private information and the only information the experimenter reveals is the number of individuals who decide to sell their units in the pool. Once this decision is made, the cost of selling the unit and the revenue received (and hence the profit) are determined for each subject. At the end of the experiment, experimental dollar profits are converted to US dollars at a ratio of five to one.

Four variations of this basic design are utilized. First, a set of experiments designed to mimic the California scenario is conducted. These are collectively referred to as the California treatment. The price a subject receives for a unit is determined by his or her choice. If a subject chooses to sell a unit in the pool, and at least one other person chooses the same, then the subject receives a fixed price of 7.5 experimental dollars. Otherwise, the price received is a random variable that places equal probability on the set of prices, $\{2.50, 7.50\}$.[27] Thus, every participant in this environment has the same random price, i.e. there are homogeneous price expectations. The cost a subject incurs for producing a unit depends upon the size of the pool with which he sells his unit. These cost are given in Figure 1. Note that any unit sold by itself is considered to be a pool of size one. As in the case of the California growers, this cost is decreasing in the size of the pool. If these economies of scale are present, then we will call the environment a high cost environment.

[26]By giving each subject a single unit we are hard wiring the choice that each producer produces at capacity and that the pool does not attempt to exercise monopolistic power by quantity fixing.

[27]A draw from a bingo cage, with twenty balls numbered one through twenty, determines price realizations. If a number between one and ten is drawn the high price is realized. Otherwise a low price is realized.

A second variation conducted is an environment whose purpose is to mimic the conditions present in the Florida orange market. In this environment there are two types of experimental agents. These agent types differ in the support of the random price received for a unit. Five individuals receive a price of either 12.5 or 2.5 experimental dollars with equal probability. These individuals are said to have a low price support. The other five individuals receive a price of either 12.5 or 7 experimental dollars with equal probability, and these individuals have a high price support.[28] We refer to these differing random prices as a heterogeneous price expectation environment. If any subject chooses to sell a unit in the pool, and at least one other subject chooses to do the same, then a fixed price of 7.5 experimental dollars is received for the unit instead of the random price. The cost of producing a unit is .5 experimental dollars for each subject, irrespective of how the unit is sold. This absence of economies of scale is referred to as a low cost environment.

Two more variations of the basic experimental design are conducted in order to obtain information on how the interaction of the cost and insurance incentives may affect pooling behavior. One set of experiments (or a treatment) has heterogeneous price expectations as in the Florida treatment but also a high cost environment. A second set of experiments utilizes the homogeneous price expectation environment of the California treatment with the low cost environment of the Florida treatment.

Table 1 summarizes the experimental design used in this research. T is the number of periods per experiment. N is the number of subjects per experiment. R is the number of replications conducted for that treatment. PS is the price support for subjects' units. CS is the cost of producing a unit.

From now on we will refer to the treatment cells in the table below by name. The upper left cell is the Homo/low treatment, the upper right cell is the California

[28]These more advantageous price supports were awarded by the relative performance of subjects on a simultaneously administered computerized trivia quiz. The quiz was conducted after the instructions were read but before the ten decision making periods started. Hoffman and Spitzer [19] first used this technique to reinforce the sense of property rights for more advantageous positions or endowments in experiments.

	Low Cost	High Cost
Homogeneous Price Expectations	$T = 10$, $N = 10$, and $R = 3$ PS={2.50, 12.50} for all subjects CS=.5	$T = 10$, $N = 10$, and $R = 3$ PS={2.50, 12.50} for all subjects CS=see Figure 1
Heterogeneous Price Expectations	$T = 10$, $N = 10$, and $R = 5$ PS={2.50, 12.50} for 5 subjects and {7, 12.50} for 5 subjects CS=.5	$T = 10$, $N = 10$, and $R = 5$ PS={2.50, 12.50} for 5 subjects and {7, 12.50} for 5 subjects CS=see Figure 1

Table 1: Eperimental Treatments and Parameters

treatment, the lower left cell is the Florida treatment and the lower right cell is the Hetero/high treatment. Furthermore, we make the distinction between individuals with different price supports in the Florida and Hetero/high treatments. If a subject has a "low" price support (i.e. their "bad" price is 2.50), then he or she is appropriately said to participate in either a Florida I or Hetero/high I treatment. On the other hand if a subject has a "high" price support (i.e., their "bad" price is 7), then he or she is appropriately said to participate in either a Florida II or Hetero/high II treatment.

This experimental design generates six observationally distinct experimental agent types. In this paper we will frequently refer to treatments via an index e. The set of experimental agent types will be indexed as $e = 1, \ldots, 6$ such that: $e = 1$ when subjects participate in the Homo/low treatment, $e = 2$ for the California treatment, $e = 3$ for the Florida I treatment, $e = 4$ for the Florida II treatment, $e = 5$ for the Hetero/high I treatment, and $e = 6$ for the Hetero/high II treatment.

In summary, the experimental design is powerful because it allows us to evaluate more than just a direct comparison of pooling behavior in California and in Florida industry-like conditions. The design also allows for a full comparison of the effects on pooling of the insurance incentive with no cost incentives and no adverse selection. Specifically, we isolate the insurance incentive in the Homo/low treatment and develop a baseline of pooling behavior solely driven by producer risk attitude. Additionally, we can test the effect of economies of scale on pooling behavior in the presence of adverse selection via the Hetero/high treatments.

In developing the experimental design we hope to obtain as much parallelism as possible, but there are almost always constraints in designing economic experiments that lead to hard choices. In this particular experiment, the most binding constraint is the relatively small number of individuals we can engage in a decision making period relative to the number of growers who actually had to make pooling decisions in California and Florida. In recreating the insurance benefit, experimental subjects receive a random price for their output by selling their output individually, or they receive a certain price by selling their output in a pool with other subjects. Collapsing the variance of the pool's price whenever at least two individuals participate in the pool is a reasonable characteristic of the experimental design for the following reason. If in the real world, n individuals receive prices that are identically and independently distributed and they form a pool, the variance of the pool's price is $1/n$ times the original variance. With a large population of orange growers this variance approaches zero relatively quickly. On the other hand, when attempting to recreate the cost incentive in the high cost environments we allow cost to be a nondegenerate function of the pool size. We feel this an appropriate characteristic of the design because it is not reasonable to have shipping costs drop to marginal cost too rapidly as individuals join the pool because this would imply there exists extremely high marginal gains to bargaining power. Hence, if these strong economies of scale were present we would not expect the industry to be characterized by price taking producers.

3 A Game Theoretic Model of The Pooling Experiments

In what follows, the structure of the observed experimental environment and conjectures on unobservable individual characteristics are analyzed in a Bayes-Nash equilibrium framework. Individuals' attitudes toward uncertainty are central to the economic question at hand. When recruiting subjects from a population to participate in an experiment, it is reasonable to expect that the subjects have different preferences over uncertain outcomes. Likewise, risk attitudes may differ within a population of orange growers. Unfortunately, controlling the risk attitudes of experimental agents is a difficult task for the experimentalist[29], and there is no way to a priori observe attitudes.[30] Individual risk attitudes are the consequence of one's life experiences, and when sampling subjects from a population the experimenter also samples risk attitudes. Our choice of the expected utility function used in the model below is driven by this conjecture. The assumed expected utility function is a form that exhibits constant absolute risk aversion. We use this particular function because it can characterize an individual's risk attitude by a single parameter.[31] We will treat this parameter as a random variable, part of an individual's player type (along with the observable experimental type), and a key source of incomplete information. An

[29]For an example of experimental economic research that attempts to control subject risk attitude see Berg, Dichaut, et al. (1986).

[30]In our strive for parallelism, controlling subject risk attitudes via lottery payoffs is not desirable in this set of experiments.

[31]There are other classes of two-parameter expected utility functions that have the desired monotonicity of risk attitude in the non-payoff parameter, e.g., the class of power functions. However. since we do not vary the payoffs within the experiment we can not empirically distinguish between these different classes.

individual's expected utility over uncertain incomes is defined as

$$U(\Pi_i, \mu_i) = \begin{cases} 1 - e^{\Pi_i \mu_i} & \text{if } \mu_i > 0 \\ \Pi_i & \text{if } \mu_i = 0 \\ e^{\Pi_i \mu_i} - 1 & \text{if } \mu_i, 0 \end{cases}$$

The part of individual's type profile that characterizes his or her risk attitude is μ_i, (also this is his absolute risk aversion coefficient) which is a random variable that is independently and identically distributed across individuals. One can easily show that the degree of risk aversion is monotonic in μ_i and risk neutrality occurs when $\mu_i = 0$.

The formal game theoretic model, the derivation of equilibrium and the derivation of the testable implications are fully presented in Appendix A. The results are intuitive and will be presented here in a less rigorous manner.

An individual grower's attitude toward risk is key to the empirical problem at hand. Obviously, an orange grower's choice to exchange random revenue for certain revenue by joining a pool is influenced by his or her risk attitude. Additionally, if there is uncertainty about the risk attitudes of other growers, and some of the benefits of belonging to a pool depend strongly upon membership size (as in the economies of scale arguments), then a grower's attitude towards this uncertainty will influence his or her marketing choices.

Our experimental environments and assumptions about the heterogeneity of subject risk attitudes fully define Bayes-Nash games. Pure strategies in these games are functions that map risk attitudes, μ_i, to choices of joining the pool or not joining the pool. The symmetric Bayes-Nash solutions for each of the experimental environments consist of step functions that are best responses to one another. These step functions are of the following form: if a player's risk parameter is below a certain level, which we call a *threshold*, then he or she will choose not to pool. Likewise, he or she will join the pool if his or her μ_i is greater than or equal to the risk threshold. A threshold for experimental agent type e will be denoted μ_e^*. The predictions of this model are summarized in Figure 2 and by the following expressions.

- $\mu_2^* < \mu_1^* = \mu_3^* < \mu_6^* < \mu_4^*$ and

- $\mu_2^* < \mu_1^* = \mu_3^* < \mu_6^* < \mu_4^*$

These two statements have the following testable implications. The degree of risk aversion required to make joining the pool the best response in equilibrium is the greatest for individuals who are Florida II experimental agents. The Hetero/high II agents have then next greatest equilibrium thresholds. Florida I and Homo/low agents have identical equilibrium threshold values, which are lower than the above mentioned agents. Finally, all four of the above mentioned agent types have higher thresholds than both California and Hetero/high I experimental agents. With respect to the proportion of subjects choosing to join the pool, the ranks of these frequencies will have the opposite ordering of the thresholds. In terms of orange growers, these predictions are consistent with the observations made in HL in regard to pooling rates across the two growing regions, namely that California growers are more likely to pool than Florida growers.

An additional result of the theoretical framework and experimental design is that treatments incorporating low costs environments are weakly dominance solvable, but the high cost environments are only solvable by the more demanding solution concept of Bayes-Nash equilibrium. Low cost environments are weakly dominance solvable because of the lack of variance in the pool's price and the fact that the size of the group that a subject sells his or her unit in does not affect the cost of the unit. Thus, in low cost environments a subject's belief about the other player's type does not influence the subject's expected utility for each action except in the case were he or she believes that none of the other subjects will join the pool with probability one. In this case, the expected utility of the two possible actions will be the same and thus will not remove an action from the best response correspondence.

4 Discussion of Aggregated Data

In this section of the paper, we present a discussion of the data in terms of the aggregated proportions of subjects choosing to pool. In the previous section, we presented a model that makes specific predictions over the ranking of pooling participation rates for different experimental agent types. Next we present graphs that provide visual evidence that these predictions hold true.

The phenomenon reported in HL of higher pooling frequencies amongst California growers is "replicated" in our experiment. The participation rates in the California treatment cell are much higher than those in the Florida treatment cell. This is easily seen by examining Figure 3. This result is certainly consistent with HL; however, we want to add a caveat to the conclusion. The parallelism in this experiment is based upon the qualitative nature of the induced incentives but not on the parametric form of these induced incentives. Still, the results are certainly consistent with the naturally occurring data in a qualitative sense.

Despite this lack of parametric parallelism, the consistency of the ordering of pooling rates to those predicted in the model may provide a justification for using the model presented as a basis for making qualitative predictions about this particular naturally occurring phenomenon. This adherence to the theoretical predictions is seen by examining figures 4, 5, and 6. Figure 4, presents a slightly disaggregated view of Figure 3. Clearly, Florida II subjects pool very little, and Florida I agents pool less than California subjects (except in period six). The pooling participation rates for low cost environment agents are presented in Figure 5. Although Florida II and Homo/low agent types should have the same pooling rates, Homo/low agents participate at a slightly lower rate and the time trend appears to show less variance. The pooling rates for the Hetero/high II agents are clearly the second lowest amongst agent types, California and Hetero/high I pooling frequencies are higher than all other types, and California agents pool at a slightly higher rate than Hetero/high I individuals according to the pooling rates presented in Figure 6. Almost all of the results are consistent with the predictions of our model.

5 Introduction to the Analysis of the Individual Level Data

The analysis of the pooling problem in the game theoretic section and in HL is static; however, the experimental design of this study, and the real phenomenon it attempts to explain, both involve individuals repeatedly facing the same choice over time. Hence, if our theoretical framework is to be an adequate modeling tool, then individuals must possess the same preferences across decision making periods. One way in which individual preferences may fail to have the desired dynamic consistency is if one's past participation (or lack of participation) endogenously affects the current likelihood that one chooses to join an agricultural pool. This phenomenon is commonly referred to as state dependency in the econometrics literature. Since the data generated by our experimental design is dynamic at the individual level. we can conduct statistical inference that scrutinizes individual behavior for dynamic inconsistencies in individual preferences. Thus, our study provides a tougher test of the pertinent economic models than a study that utilizes purely cross sectional data because possible dynamic inconsistencies can not be identified with cross sectional data.[32]

Confronting our models with this test is inevitable in an experimental setting. Conducting experiments that generate cross sectional data by having subjects make exactly one decision may be problematic in our experimental environments. In the mental experiment conducted, from which the theory was derived, a world was assumed where subjects completely understood the game structure, and knew their own types and the prior distributions of all types. Therefore, the mental experiment only required a single decision making period. Unfortunately, in the laboratory, the subjects often are initially unfamiliar with the new decision making environment and institution in which they are asked to participate. Thus, subjects may engage in trial and error type decisions in order to familiarize themselves with the new setting.

[32]For a full discussion of the ability to identify dynamic characteristics in discrete choice panel data sets, we refer the reader to Heckman [17].

Other types of disequilibrium initial conditions that lead to further complications of the data may exist. For example, if a subject is uncertain of the environment, he or she may also be unsure of his or her own preferences in a new environment. Also, if a player is unsure of his or her own type (i.e., his preferences), then it is doubtful that he or she initially knows the common prior over types. Under these circumstances, it is hoped that by repetition of the same task these problems will dissipate and the static theory will emerge as a good predictor of behavior.

The above statements may justify the use of multiple time periods in an experiment, but it does not give us (the experimenter) the authority to ignore the potential dynamic processes that may be entrenched in the data. This is especially true when viewing the data from a dynamic perspective provides a stronger test of the theory. However, one must be careful to separate the dynamic effects that are general to all pooling scenarios from those that are only characteristics of the experimental environment. There is a subtlety in the work presented here: the theory allows for basic heterogeneity in subject preferences, but does not allow for subject preferences to be affected deterministically by previous decisions. The presence of state dependency would be problematic for the applicability of our models. Yet we believe that the inconsistencies of subject choices that result from the changing of disequilibrium initial conditions over time are not as problematic as it would seem; it is only an artifact of the experimental environment that can be analyzed and isolated.

Using these guidelines, it is important that we correctly model and test the experimental data for these dynamic effects. With respect to the state dependency issue, if an econometrician does not account for heterogeneity in the data, this will result in an upward bias of the effects of state dependency and an erroneous rejection of the theory. Therefore, most of the statistical inference conducted in this paper concentrates on this issue, by attempting to distinguish between state dependency and heterogeneity in the data.

The next two sections of this paper explore the data at the individual level because the implications and predictions of our model are about individual behavior. First a parametric random effects probit model of individual behavior that controls for the heterogeneity of subjects is estimated and a test for state dependency is conducted. Next, we conduct a nonparametric evaluation of the data in order to test for different possible stochastic processes in the discrete panel data set that can be indicative of the presence of state dependency or heterogeneity. It turns out that the nonparametric approach justifies the parametric approach taken and may prove to be a useful tool in diagnosing the nature of other experimentally generated data in the future.

6 Parametric Analysis of The Individual Level Data

In this section of the paper, we present the results of the estimation and hypothesis testing conducted upon the data. According to the theoretical model, the unconditional probability that an individual joins a marketing pool is the probability that an individual "draws" a risk aversion parameter greater than the equilibrium threshold of the environment in which he participates. Hence, the empirical validity of our model with respect to the original orange grower's problem relies upon the assumption that pooling participation rates rely solely upon the distribution of risk attitudes in the set of growers and is independent of such factors as state dependency, wealth effects, past realizations of prices, and past participation rates of the pool (if the environment is in equilibrium). Therefore, we develop specifications that incorporate and test for the significance of these factors. Surprisingly, we find that none of these factors is significant, except in some environments in which the previous price realization significantly affects the probability of choosing to pool. The following analysis leads us to the conclusion that, on average, our framework does provide adequate qualitative predictions of participation in experimental marketing pools.

Each of the homogeneous price expectation cells was conducted three times, the Florida treatment cell was conducted five times, and the Hetero/high treatment cell was conducted seven times. This resulted in a size of thirty individuals (N) for the homogeneous price expectation environments, twenty-five individualsfor both California agent types, and thirty-five individuals for both Hetero/high agent types.

Also, each individual has ten observations (T) of decisions to join the pool or not join the pool. In addition to the observationon the pooling decisions, we also have observations on the following variables: accumulated earnings, price realizations and the size of the pool. The resulting data set is a panel of NxT observations.

The purpose of our estimation procedures is to estimate the risk aversion thresholds for the six distinct experimental agent types and possibly identify the effects other variables have on the probability a given subject joins the pool in a given period. The models estimated are random effects probit regressions. At this point we direct the interested reader to Appendix B for the derivations of the likelihood functions, the empirical formulation and the discussion of estimation procedures of the identified index functions.

We report the estimated random effects probit models for each of the experimental agent types in Table 2. The dependent variable in each model is the probability that an individual i chooses to pool in period t. The following notation is used in the tables: Constant is the reduced form of the risk aversion threshold parameter, y_{it} is an indicator variable ($y_{it} = 1$ if subject i chooses to join the pool in period t and $y_{it} = 0$ if subject i chooses not to join the pool in period t), HP Lag is a dummy for a high price realization in period t, and Rho is a transformation of the interclass correlation coefficient which represents the correlation of decisions across periods. The likelihood ratio test conducted for each parsimoniously selected model is against an alternative general model that is fully specified in Appendix C. Results of these procedures are summarized in Table 2.

The first striking result in Table 2 is that the estimated thresholds on the risk parameters (which are the constants in the regressions) for experimental agent types are ordered exactly as the theory predicted. Specifically, California and Hetero/high I have the two lowest thresholds, Homo/low and Florida I have thresholds not significantly different from zero, Hetero/high II has the second highest estimated threshold, and Florida II has the highest estimated threshold. This is evidence in favor of the testable implications of the model presented in this paper. Specifically, we were successful in predicting the ordered rankings of required levels of individual risk aversion

Homo/low Treatment

Parameter	Estimate	Std. Err.	t-stat.	P-value
Constant	-.131	.165	-.799	.212
HP Lag	-.408	.182	-2.243	.012
Rho	.717	.107	6.70	0

Log Likelihood Value	-149.5
L.R. Test Statistic	.734
Degrees of Freedom	5
P-value	.981

California Treatment

Parameter	Estimate	Std. Err.	t-stat.	P-value
Constant	-.742	.276	-2.69	.004
y_{it-2}	-.150	.238	-.63	.264
Rho	.519	.135	3.86	0

Log Likelihood Value	-134.92
L.R. Test Statistic	2.69
Degrees of Freedom	5
P-value	.747

Florida I Treatment

Parameter	Estimate	Std. Err.	t-stat.	P-value
Constant	-.170	.238	-.712	.238
y_{it-2}	.621	.261	2.38	.009
HP Lag	-.825	.213	-3.874	0
Rho	.514	.107	4.827	0

Log Likelihood Value	-121.4
L.R. Test Statistic	1.436
Degrees of Freedom	4
P-value	.838

Hetero/high I Treatment

Parameter	Estimate	Std. Err.	t-stat.	P-value
Constant	-.596	.145	-4.098	0
HP Lag	.449	.164	2.74	.003
Rho	.386	.079	4.885	0

Log Likelihood Value	-179.2
L.R. Test Statistic	.836
Degrees of Freedom	5
P-value	.975

Florida II Treatment

Parameter	Estimate	Std. Err.	t-stat.	P-value
Constant	2.307	.207	11.156	0
Rho	1.935	.1152	16.793	0

Log Likelihood Value	-68.58
L.R. Test Statistic	9.328
Degrees of Freedom	6
P-value	.156

Hetero/high II Treatment

Parameter	Estimate	Std. Err.	t-stat.	P-value
Constant	1.268	.176	7.206	0
HP Lag	-.452	.193	-2.349	.001
Rho	.351	.107	3.289	0

Log Likelihood Value	-128.95
L.R. Test Statistic	1.154
Degrees of Freedom	5
P-value	.949

Table 2: Random Effects Probit Estimations

for successful pooling across different environments.

Our theory is also supported by the noticeable absence of state dependence in the estimated models, as shown by the general ability to exclude or the lack of significance of lagged indicators of past decisions in the regression models. This is indicative that choices to join a pool are not structurally changing across decision periods and that our static models are appropriate predictors of behavior. In fact, there are only two instances where second order state dependence can not be rejected as part of the specification: this is in the California and Florida I environments. Furthermore, the y_{it-2} variable is significant only in Florida I model. In this case it is positive and, hence, it is positively related to the probability that an individual will join a marketing pool.

A third, slightly negative result is that, except for the California and Florida II treatments, we could not reject the inclusion of the variable HP Lag in our specifications. Moreover, the HP Lag coefficient is significant at the five percent level of significance and negative in the Homo/low, Florida I and Hetero/high II environments. This indicates that individuals may believe that the realizations of prices are not independent and in fact are negatively correlated. However, contrary to this fact is the significant positive coefficient of the HP Lag variable in the Hetero/high environment. The positive sign of this coefficient is indicative of beliefs that the probability of a high price realization is a positively serially correlated event.

One of the most interesting results of this analysis is the values that Rho takes for different experimental agents. The Larger Rho is, the more stable individual choices are across time periods. In other words, a large value of Rho is indicative that a typical subject tends to either be in the pool or not in the pool across time and does not switch states frequently. The largest value of Rho is for Florida II agents and this is not surprising since severe risk aversion is required for one to optimally participate in the pool in this experimental environment. Obviously, the lack of pooling amongst these subjects is not remarkable. Homo/low agents have the next highest Rho. This not surprising because this environment is weakly dominant solvable and subjects do not have to rely upon the participation of others to provide the full benefits of the pool.

The random effects probit regressions also reveal evidence that observed heterogeneity that does not affect the strategic uncertainty of pooling decision may reduce the stability of individual choices across decision periods. Specifically, a Florida I subject, who participates in a weakly dominance solvable environment, has an estimated Rho smaller than that estimated for a Homo/low subject. The only difference between a Florida I subject and a Homo/low subject is the observed and induced heterogeneity associated with price expectations between a Florida I subject and the other subjects within the Florida treatment. Apparently, this observed heterogeneity may "destabilize" subject choices across time, relative to an environment where participants are observationally homogeneous as in the Homo/low treatment. In fact. the Rho estimated from the Florida I regression is approximately the same magnitude as the Rho estimated for the California treatment, where unobserved heterogeneity introduces strategic uncertainty. Thus, remarkably, both unobserved and observed heterogeneity appear to empirically affect the stability of individual choices.

Meanwhile, the two lowest estimated values of Rho are for both Hetero/high agents. This demonstrates how the presence of both observable and unobservable heterogeneity may seriously affect the stability of individual behavior.

The results of this section demonstrate that the predictions of HL are qualitatively accurate, but a deeper understanding of the cooperative participation is achieved by the theoretical extension to a world with unobserved heterogeneity. We gain the insight that partial participation in a marketing cooperative is individually rational and can be a persistent phenomenon when producers have heterogeneous preferences over uncertain profits.

7 Nonparametric Analysis of the Individual Level Data

The random effects probit specifications estimated earlier presented a composite test of state dependency and an i.i.d. normal error structure. If the i.i.d. normality assumption of the error terms is incorrect then we may have incorrectly concluded that there is no state dependency in the experimental data. Therefore, we choose to scrutinize the data for state dependence and heterogeneity in a framework that does not rely upon any of the assumptions made in the analysis of the previous section. The framework adopted here is a nonparametric approach for discrete panel data, suggested by Lee [23]. This framework allows us to distinguish among several stochastic processes that may be the data generating process. Our theoretical model makes a prediction that only one of these processes should generate the data. In particular, the micro-model predicts that a heterogeneous Bernoulli model should characterize that data, and this is inconsistent with state dependency.

Lee argues that one can gain insight into the data generating process of a discrete choice panel data set by examining the sequences of choices from a finite data set for run patterns. Lee's empirical model and procedure involve using a log-linear probability model, much like an ANOVA, as a general specification. Next, Lee identifies four potential discrete time stochastic processes that could generate the random sequences of occupied states observed: the homogeneous Bernoulli, the heterogeneous Bernoulli, the Markov, and the mixed Markov models. Furthermore, each of the four processes is indicative of whether heterogeneity and/or state dependency is present in the data. State is inconsistent with either Bernoulli process.

Lee demonstrates that each of these processes is nested within the log-linear probability model by placing appropriate constraints upon the order terms. One can estimate these specifications by a multinomial logit procedure and then conduct classical hypothesis tests on the restrictions. In this section, we employ the likelihood ratio test to test specific stochastic processes against the general model. Next, we will present a general discussion of the four stochastic processes that may be driving the sequences of choices made by experimental subjects and then discuss the results of an extensive series of hypothesis tests conducted on the experimental data. In Appendix C, we present a detailed discussion of the framework employed and the development of the econometric specification used. We find that no model is clearly more appropriate than the heterogeneous Bernoulli, and this provides some support for our choice of parametric specification utilized previously.

The four competing stochastic processes are the homogeneous Bernoulli, the heterogeneous Bernoulli, the first order Markov, and the mixed first order Markov. Each can easily be described by the transitional probabilities they impose on the experimental data set.

The homogeneous Bernoulli requires that the probability a given subject chooses to join the pool in a given decision period must be the same for all individuals and all time periods, or in other words the marginal probabilities of decisions are constant across individuals and time. If we do not reject this model, there is evidence in favor neither of heterogeneity nor of state dependence in the data.

The heterogeneous Bernoulli model implies that individuals have differing marginal probabilities, but a given individual's marginal probability must be constant across time. This process is indicative of heterogeneity but no state dependency in the data.

The first order Markov process requires that conditional probabilities over choices in a given period be completely determined by the choices of the previous time period, and these are completely described by a Markov transition matrix. This process requires that every individual's behavior follows the same transition matrix while the mixed Markov model allows for individuals to follow different transition matrices. The Markov model is consistent with data characterized by only state dependency, while the mixed Markov model is evidence of both heterogeneity and state dependency.

When estimating these models with our data, we are limited by the small number of individuals relative to the number of time periods. We discuss this issue in the Appendix, and circumvent it by the estimation of two models for each experimental treatment; one for periods two through six and another for periods seven through ten. We report the results of the hypothesis test in Table 3.

	Homogeneous Bernoulli 14 D.O.F.		Heterogeneous Bernoulli 11 D.O.F.		Markov 8 D.O.F.		Mixed Markov 2 D.O.F.	
Player Type	T1	T2	T1	T2	T1	T2	T1	T2
Homo/low	**25.73**	**35.82**	18.71	**22.58**	14.38	6.61	0	0
	(.028)	(.001)	(.066)	(.020)	(.072)	(.580)	(1.000)	(1.000)
California	**35.71**	**38.07**	**20.22**	**20.14**	**27.68**	**23.53**	0	0
	(.001)	(.001)	(.042)	(.043)	(.001)	(.003)	(1.000)	(1.000)
Florida I	**31.71**	20.66	**36.69**	18.3	**27.77**	14.62	0	**7.088**
	(.004)	(.111)	0	(.075)	(.001)	(.067)	(1.000)	(.029)
Florida II	9.676	4.499	0	1.387	0	(.001)	0	0
	(.886)	(.993)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Hetero/high I	20.29	10.44	12.19	16.9	**17.1**	**24.91**	0	**8.947**
	(.121)	(.729)	(.350)	(.111)	(.028)	(.002)	(1.000)	(.011)
Hetero/high II	22.76	**35.05**	**21.08**	**23.51**	**17.21**	9.501		0
	(.064)	(.001)	(.033)	(.015)	(.028)	(.302)	(1.000)	(1.000)

T1 = periods 3, 4, 5, and 6, and T2 = periods 7, 8, 9, and 10

Chi-square test statistics and (P-value)

Bold face indicates a rejection at the 5% level of significance

Table 3: Likelihood Ratio Test Comparisons Within the Lee Framework

Table 3 reports the likelihood ratio test statistics and the corresponding p-values from the series of hypothesis tests described above. We conduct all hypothesis tests at a five percent level of significance. The first result to notice is that the mixed Markov model is only rejected in a few instances. This result implies that the underlying process that generates individual choice probabilities is complex and may require the estimation of a specification that is also complex. However, our theoretical model indicates that this process should be relatively simple.

These frequent failures to reject may be due to two reasons. First, the test statistic has only two degrees of freedom, and secondly the restrictions placed on the general model by the mixed Markov model are frequently satisfied trivially because there are no observations for which the constraints are applicable. The next noticeable result is the inability to reject any of the models for the Florida II individuals. We believe this result is due to the small variance in the dependent variable, which is an artifact of the severe risk aversion required for an individual to ever join a pool in this environment.

The rest of the specification test results provide some problematic conclusions. First, notice that all but two environments reject the first order Markov model in periods two through six (the two exceptions are the Homo/low subjects and the Florida II individuals). However, in the last four decision making periods only two out the six environments reject the first order Markov model; these are the Hetero/high I and California environments. This indicates that state dependence has an increasing presence over the course of an experiment. This result is the opposite of what one might expect to happen as subjects become more familiar with the experimental environment.

The results of the hypothesis tests for the heterogeneous Bernoulli model also provide inconclusive results. The model is rejected for three of the six environments for both decision periods two through six, and the last four decision periods. We fail to reject the model in both sets of periods only for the Florida II and Hetero/high I environments. Similarly, we reject the heterogeneous Bernoulli model for both sets of periods for the California and Hetero/high II environments. The Homo/low treatment fails to reject this particular model for the first set of periods, but does not reject the heterogeneous Bernoulli model for the last four time periods. Finally, the hypothesis test for the Florida I environment has the opposite result, we reject for the first set of periods and fail to reject for the second set of periods.

With respect to the hypothesis test for the homogeneous Bernoulli model, we reject the model in both sets of periods for both the California and the Homo/low environments. On the other hand, we fail to reject this model for either sets of periods in the Florida II and Hetero/high I environments. In the Hetero/high II environment, we fail to reject the model for the first set of periods but do reject the model for the second set of periods. We observe the opposite result for the Florida I environment: we reject the model for periods three through six and fail to reject the model for the last four periods.

We believe that the series of hypothesis tests fails to reject our previous conclusion that the our data generating process is the heterogeneous Bernoulli model. However, we can not ignore dynamic complexities in our data not yet identified. At this time, we are comfortable concluding that the theory and the nonparametric statistical results indicate that our random effects probit model is the appropriate regression framework to choose. Nevertheless, the nonparametric results merit a few additional comments.

Clearly there is an inability to reject all stochastic processes except for the heterogeneous Bernoulli, which is an implication of our model. Yet, this is not surprising since this analysis does not allow for choices to depend upon observable variables, which are allowed for in the random effects probit models. Thus, some of the deviations from the implications of our theory found in the analysis of this section may be controlled for in the models of the previous section. However, the serial correlation observed in some the models is troubling since we tended to reject the presence of state dependency in the previous section. This serial correlation may be indicative of a serial correlated error term in the random effects probit models.

8 Conclusion

We use experimental economics to explain why California orange growers formed season-long pools, while Florida growers did not.[33] HL argue that this asymmetric pooling behavior was the result of different insurance and cost incentives. We construct an experimental environment that controls for the insurance incentive of the reduced variance in revenue offered by the pool and the economies of scale in shipping rates that constitute the cost incentive. We also extend HL's model to allow for heterogeneous risk attitudes amongst producers. This allows for individually rational behavior that is consistent with persistent partial participation in marketing cooperatives. After examination of the experimental data, we find that the original

[33]However, this is not the first research to examine an economic history topic with an economic experiment. For an earlier example see Binger, Hoffman and Libecap [6].

insights provided by HL were qualitatively correct, and our model reasonably predicts the partial pooling behavior observed and tracks the qualitative differences in participation rates across treatments.

Our results imply the following four points. First, the experimental results replicate the qualitative aggregate predictions of HL. Second, the game theoretic model of individual behavior we present makes strong predictions about partial pooling, and these predictions are consistent with results obtained from using discrete choice panel data techniques. Third, the nonparametric analysis of the data provides evidence (although not definitive) that the static models examined in this paper are fairly robust in the dynamic experimental environment and that the random effects probit is appropriate to analyze our experimental data. Finally, by applying experimental data and a Bayes-Nash equilibrium model to the analysis of a historical event, we uncover unanticipated complications in the individual decision making process which determine the persistence and composition of a marketing cooperative.

The results of this paper indicate several avenues for future research. For example, it is worthwhile to ask how heterogeneous producer attitudes towards uncertain profits affect cooperative and cartel behavior, especially when the cooperative or cartel engages in market quantity restrictions. Furthermore, will partial cooperative or cartel participation be sustainable and persistent in this framework? Also, are other historical phenomena (for which there may be scant data) amenable to investigation by experimental methods? Finally, in terms of experimental data analysis, can we improve the econometric techniques commonly used in order to control and evaluate dynamics imbedded in the data?

9 Game Theoretical Framework

The following is the common Bayesian game structure for all four of the experimental treatment cells:

- A set I of 10 players, whose number is indexed with the lower case i. Specifically, $I = 1, 2, ..., 10$. Assume that each player knows his or her own index.

- Each player i has a type pair $(\mu_i,\, p_i)$.
- The support of the random variable μ_i is the real numbers. The jointrandom variables μ_i, $i = 1, \ldots, 10$, are independently identically distributed with probability density function f, and cumulative distribution function F. Each individual is revealed the realization his μ_i before the game but only knows the probability distribution function F from which the other nine μ_i's are drawn.
- The second part of the type is p_i, which is an element of the set $\bar{P}$. Each $p_i \in \bar{P}$ is a random variable which places equal probability on the elements of the set $\bar{P} = \{\underline{p_i}, \bar{p}\}$. Player i only learns of the realization of his random variable once all players have committed to actions. The assignment of p_i is given by the function $G : I \mapsto \bar{P}$, which is common knowledge.
- For every player i, the set of feasible actions is $A_i = \{0, 1\}$. The actions are interpreted as follows: $a_i = 0$ is a choice of player i to not participate in the pool and $a_i = 1$ is a choice of player i to participate in the pool
- Player i's pure strategy space, S_i, is the set of measurable functions that map elements of the type space,$\Re \times \bar{P}$, to elements of A_i.
- The set of pure strategy profiles is $S = \prod_{i=1}^{10} S_i$.
- The outcome of the game for each player i is given by the function Π in which

$$\Pi(S, p_i) = \begin{cases} P^C - C\left(\sum_{j=1}^{10} s_j(\mu_j)\right) & \text{if } s_i(\mu_i) = 1 \text{ and } \sum_{j=1}^{10} s_j(\mu_j) \geq 2 \\ p_i - C(1) & \text{otherwise} \end{cases}$$

  In the above expression C is always non-increasing, convex and positive. Also P^C is a constant.

- Each player i has a Von Neuman Morgenstern utility function $u_i : S \times \Re \times \bar{P} \mapsto \Re$ in which

$$u_i(s,\, \mu_i,\, p_i) \equiv U\left(\Pi(s, p_i),\, \mu_i\right) = \begin{cases} 1 - e^{-\mu_i \Pi} & \text{if } \mu_i > 0 \\ \Pi & \text{if } \mu_i = 0 \\ e^{-\mu_i \Pi} - 1 & \text{if } \mu_i < 0 \end{cases}$$

Note that this utility function exhibits constant absolute risk aversion, where player i's coefficient of absolute risk aversion is $\mu_i$for all changes in income.

Now we have the five-tuple that completely describes the Bayesian game, Γ, for which $\Gamma = \{I, (\mu_i, p_i), F, S, U\}$. Next we derive game theoretic solutions for each of the experimental treatments.

9.1 Homo/low Treatment

In this treatment, for every i, $G(i) = p$, for which $supp(p) = \{2.50, 12.50\}$, $C(\cdot) = .50$, and $P^C = E(p)$ or 7.5. This information is public knowledge. Now we will show that this symmetric game has a unique[34] weakly dominant strategy s', for which

$$s'(\mu_i, p) = \begin{cases} 1 & \text{if } \mu_i \geq 0 \\ 0 & \text{if } \mu_i < 0 \end{cases} \tag{15}$$

Consider for a player i an arbitrary deleted strategy profile, $s_{-j} = \prod_{j=1, j\neq i}^{1} 0$, and any distribution function F. Furthermore, consider the set of possible sizes of the pool excluding player i, $K_{-i} = \{1, \ldots, 9\}$. Let H be the probability distribution function on the set as determined by F and s_{-j}. For any player i, choosing to join the pool is a best response if

$$U\left(\Pi\left((1, s_i), p\right), \mu_i\right) \geq U\left(\Pi\left((0, s_i), p\right), \mu_i\right).$$

Written more explicitly,

$$H(0)^9.5[U(12, \mu_i) + U(2, \mu_i)] + [1 - H(0)^9]U(7, \mu_i) \geq .5[U(12, \mu_i) + U(2, \mu_i)] \text{ or}$$

$$U(7, \mu_i) \geq [U(12, \mu_i) + U(2, \mu_i)] \tag{16}$$

Expression (2) holds with equality only when $\mu_i = 0$ or $H(0)^9$. Otherwise, (2) holds strictly when $\mu_i > 0$ and strictly fails when $\mu_i < 0$. These properties hold true independent of the deleted strategy profile and F. Hence, s' is a weakly dominant

[34]Unique in the sense that there is not another weakly dominant strategy that differs on a set of positive measure

strategy for all players. Furthermore, the homo/low environment is weakly dominance solvable by the strategy profile $S' = (s', \ldots, s')$ since s' is a strict best response for some generic s_{-j} and priors.

9.2 Florida Treatment

In this treatment for $i = 1, \ldots, 5$, $G(i) = p$, and, for $i = 6, \ldots, 10$, $G(i) = p'$, for which $supp(p') = \{7, 12.50\}$. The first five players are Florida I agents and the second five players are Florida II agents. To illustrate what is known by a player about the p_i's consider the following example: a Florida I agent would know that he and four other people have "low" price supports and five other people have "high" price supports. Also, identical to the last treatment cell, $C(\cdot) = .50$, and $P^C = 7.5$, or $P^C = E(p)$.

This game has a weakly dominant strategy, s'' and is uniquely weakly dominance solvable. The unique weakly dominant strategy, s'' has the following type contingent form: $s''(\mu_i, p)$ =(1) and

$$s''(\mu_i, p') = \begin{cases} 1 & \text{if } \mu_i \geq \mu_4^* \\ 0 & \text{if } \mu_i < \mu_4^* \end{cases} \tag{17}$$

By arguments identical to the ones in the previous subsection, one can see that (1) is a weakly dominant strategy for Florida I individuals. Hence, one would also predict that $\mu_1^* = \mu_3^*$.

One can show that (3) is a weakly dominant strategy for Florida II players by related arguments. Choosing to pool is a best response for player i if

$$U\left(\Pi\left((1, s_i), p'\right), \mu_i\right) \geq U\left(\Pi\left((0, s_i), p'\right), \mu_i\right). \tag{18}$$

Equivalently,

$$U(7, \mu_i) \geq .5[U(12, \mu_i) + U(6.5, \mu_i)]. \tag{19}$$

If $H(0)^9 = 1$, then expression (5) holds with equality for all values of μ_i and either actions is a best response. Assume that $H(0)^9 \neq 1$.

The utility of the degenarate lottery on the left hand side of (5) is larger than the utility of the lottery on the right hand side if the certainty equivalent is also larger. A certainty equvalient can be expressed as the sum of the expected value and the risk premium. Hence, we can rewrite (5) as

$$7 \geq 9.25 - \rho(.5, \{6.5, 12\}, \mu_i), \tag{20}$$

where ρ is the risk premium of the lottery determined by the probability distribution of the lottery, its support, and μ_i. By the theorem of Pratt [32], ρ is strictly increasing in μ_i. Hence, the difference between the right and left hand side of (6) is strictly decreasing in μ_i. Therefore, a best response for any player i to any distribution over types and deleted strategy profile is the step function with the step at μ_4^*, the type with which (6) holds with equality. These arguments show that s'' is a weakly dominant strategy for the Florida treatments in which $\mu_4^* \approx 1.385$. Also, for some beliefs and deleted strategy profiles this strategy is a strict best response, this environmet is has a unique weakly dominance solution.

9.3 California Treatment

In this treatment, for every i, $G(i) = p$, $C(z) \approx .5(6.01 - z^{.7})$, and $P^C = 7.5$. The exact cost schedule is given in Figure 1. This information is public knowledge.

In deriving a Bayes-Nash equilibrium of this environment, we first establish that for any player a best response to any deleted strategy profile is a step function of the form previously introduced. For player i, joining the pool is a best response for if

$$U\left(\Pi\left((1, s_i), p\right), \mu_i\right) \geq U\left(\Pi\left((0, s_i), p\right), \mu_i\right).$$

This expression written more explicitly is

$$H(0)^9.5[U(10,\mu_i)+U(0,\mu_i)]+\sum_{j=1}^{9} H(j)U(7.5-C(j+1),\mu_i) \geq .5[U(10,\mu_i)+U(0,\mu_i)]$$

Rearranging this expression

$$\sum_{j=1}^{9} H(j)U(7.5 - C(j+1), \mu_i) \geq [1 - H(0)^9].5[U(10,\mu_i) + U(0,\mu_i)]. \tag{21}$$

Recognizing that $\sum_{j=1}^{9} H(j) = 1 - H(0)^9$, we substitute in (7) and obtain

$$\sum_{j=1}^{9} H(j)U(7.5 - C(j+1), \mu_i) \geq \sum_{j=1}^{9} H(j).5[U(10, \mu_i) + U(0, \mu_i)]. \qquad (22)$$

If the difference between the right and left hand sides of (8) is strictly decreasing in μ_i, then a best response is of the form we set out to find. This easily seen by recognizing that for any i

$$U[7.5 - C(j+1), \mu_i] - .5[U(10, \mu_i) + U(0, \mu_i)]$$

is in fact strictly decreasing in μ_i, by the arguments of the last section.

Now we will show that this game has a symmetric Bayes-Nash equilibrium with strategy profile such that $\forall i \in I$,

$$\tilde{s}(\mu_i, p') = \begin{cases} 1 & \text{if } \mu_i \geq \mu_2^* \\ 0 & \text{if } \mu_i < \mu_2^* \end{cases} \qquad (23)$$

(9) The value of μ_2^* in (9) is the point at which an individual is indifferent between choosing to be in the pool or not, assuming that all other players are playing the strategy in (9). This value is found by solving the following equation for μ_2^*;

$$[F(\mu_2^*)]^9.5U(10, \mu_2^*) + \sum_{j=1}^{9} \binom{9}{j} [1 - F(\mu_2^*)]^j [F(\mu_2^*)]^{9-j} U[7.5 - C(j+1), \mu_2^*]$$
$$= .5U(10, \mu_2^*). \qquad (24)$$

The left hand side of equation (10) gives the expected utility of money from a choice join the pool. The first term on the left hand side of (10) is the probability that no one else joins the pool times the expected utility of the outcome function. The second term is the probability of exactly one other person choosing to pool times the expected utility of the corresponding outcome function, sum with similar higher order participation. The right hand side of (10) is the expected utility of choosing not to pool. Solving this equation for μ_2^* requires a specification for the cumulative distribution function F.

However, we can see that threshold μ_2^* is smaller than the threshold μ_1^*; striclty so if we assume that $F(\mu_2^* + \omega)$ for some $\omega > 0$ to be more explicitly described shortly. Slightly alter the current game as follows; if two or more people choose to pool, individuals choosing to pool receive a price of 7.50 and pay the cost, C(2)=2.19. Notice that the expected utility of choosing to pool is not as large as in the original game because one loses the additional benefit received when strictly more than two choose to pool. Now, a player i would choose to pool if

$$[F(\mu_2^* + \omega)]^9 .5U(10, \mu_i) + \left(1 - [F(\mu_i)]^9\right) U(5.31, \mu_i) = .5[U(10, \mu_i) + U(0, \mu_i)]$$

$$\text{or } U(5.31, \mu_i) \geq .5[U(10, \mu_i) + U(0, \mu_i)] \tag{25}$$

In the above statements, let $\mu_2^* + \omega$ be the type that would be indifferent between pooling and not pooling in this altered game. Equality in expression (11) implies that the utility function is strictly convex and that $\mu_2^* + \omega < 0$. It is clear from the structure of the new game that $\omega > 0$, since we reduced the benefit recieved from joining the pool and opting out of the gamble. Thus, μ_2^* must be strictly negative. From this argument we can conclude that the threshold value in this treatment is indeed lower than for the Homo/low types.

9.4 Hetero/high Treatment

In this treatment for $i = 1, \ldots, 5$, $G(i) = p$, for $i = 6, \ldots, 10$, $G(i) = p'$, and $P^C = 7.5$. For the rest of this section the first five players will be referred to as group 1 and the second five players will be referred to as group 2.

Best responses are once again step functions of the typical nature. Establishing this result for this treatment involves argument that are identical to those in the California treatment. A symmetric Bayes- Nash equilibrium in this treatment will consist of μ_i-type contingent step function strategies that are mutual best responses. Such a strategy is

$$\hat{s}(\mu_i, p) = \begin{cases} 1 & \text{if } \mu_i \geq \mu_5^* \\ 0 & \text{if } \mu_i < \mu_5^* \end{cases} \text{ and } \hat{s}(\mu_i, p') = \begin{cases} 1 & \text{if } \mu_i \geq \mu_6^* \\ 0 & \text{if } \mu_i < \mu_6^* \end{cases} \tag{26}$$

The values of μ_5^* and μ_6^* in (12)are the types which are indifferent between actions conditional upon the p type. These values are found by solving the following two simultaneous equations

$$[F(\mu_5^*)]^4[F(\mu_6^*)]^5(.5U(10,\mu_5^*) - U(7.5 - C(1),\mu_5^*))+$$
$$\sum_{k=0}^{5} \begin{pmatrix} 5 \\ k \end{pmatrix} \sum_{j=0}^{4} \begin{pmatrix} 4 \\ j \end{pmatrix} [1 - F(\mu_6^*)]^k [F(\mu_6^*)]^{5-k} [1 - F(\mu_5^*)]^j [F(\mu_5^*)]^{4-j} \tag{27}$$
$$U[7.5 - C(j+k+1),\mu_5^*] = .5U(10,\mu_5^*).$$

$$[F(\mu_5^*)]^5[F(\mu_6^*)]^4(.5U(10,\mu_6^*) - U(7.5 - C(1),\mu_6^*))+$$
$$\sum_{k=0}^{4} \begin{pmatrix} 4 \\ k \end{pmatrix} \sum_{j=0}^{5} \begin{pmatrix} 5 \\ j \end{pmatrix} [1 - F(\mu_6^*)]^k [F(\mu_6^*)]^{4-k} [1 - F(\mu_5^*)]^j [F(\mu_5^*)]^{5-j} \tag{28}$$
$$U[7.5 - C(j+k+1),\mu_6^*] = .5U(10,\mu_6^*).$$

The left hand side of (13) gives the expected utility of money from a choice to join the pool given one is a member of group 1, and the right hand side is the expected utility of money from a choice not to pool given one is a member of group one. Similar interpretations can be given to equation (14). Another result that can be derived for this treatment is that any Florida II individual choosing to pool would also choose to pool if he was a Hetero/high II, i.e. $\mu_6^* < \mu_4^*$, if we assume $F(\mu_6^* + \alpha)$. Consider the slightly altered game where if two or more choose to pool, they receive the pooling price of 7.50 and pay the cost of 2.19. In this game, the expected utility of choosing to pool is not as large as in the original game because we loose the potential additional cost benefit of several people joining in the pool. Now one would choose to pool if

$$[F(\mu_6^* + \alpha)]^9 .5U(10,\mu_i) + \left(1 - [F(\mu_6^* + \alpha)]^9\right) U(5.31,\mu_i) > .5U(10,\mu_i) \text{ or}$$

$$U(5.31,\mu_i) > .5U(10,\mu_i) \tag{29}$$

In expression (15) let $\mu_6^* + \alpha$ be the value of the type for which a player would be indifferent between pooling and not pooling in the slightly altered game. Obviously, $alpha > 0$. The above expression holds true when $\mu_6^* + \alpha \approx .844$. The final testable implication is that $\mu_5^* < \mu_1^*$. This can be derived by the identical argument used to obtain $\mu_2^* < \mu_1^*$.

10 The Random Effects Probit Model

We next use the theory discussed in the previous appendix to derive an empirical formulation. Let y_{it} be individuali's action choice in period t, for which $y_{it} = 1$ is a choice to join the pool and $y_{it} = 0$ is a choice not to join the pool. Now we can define the probability that an individual i will choose to be in the pool for period t conditional upon experimental agent type, e, as follows

$$\Pr(y_{it} = 1|e, u_i) = \Pr(\mu_i + \eta_{it} \geq \mu^* X_{it}). \tag{30}$$

Where X_{it} is a vector of dummy variables for experimental agent types and μ^* is a vector of experimental agents' risk aversion parameter thresholds. Assume that f is normal with mean $\bar{\mu}$ and varinace $\sigma^2 = 1$.[35] We also assume that for each i, μ_i is independent. Furthermore, for every t and i, η_{it} is an identical and independently distributed normal error term with mean zero and variance σ_η^2 , and is also independent of each μ_i. Also, X_{it} and μ^* will sometimes be expanded to include other explanatory variables and their coefficients. Now we can rewrite (16) as

$$\Pr(y_{it} = 1|e, u_i) = \Pr\left(\frac{\eta_{it}}{\sigma_\eta} \geq \frac{\mu^*}{\sigma_\eta} X_{it} \frac{\mu_i}{\sigma_\eta}\right) \tag{31}$$

Defining the interclass correlation coefficient $\rho = \sigma_\mu^2/\sigma_\mu^2 + \sigma_\eta^2$, one can easily show that $\sigma_\mu/\sigma_\eta = (\rho/1-\rho)^{\frac{1}{2}}$. Finally, let $\hat{\mu}$be the 6x1 vector with typical element $\hat{\mu}_e = \mu_e^* - \bar{\mu}/\sigma_\eta$. Expression (17) is equivalent to

$$\Pr(y_{it} = 1|X_{it}, \mu_i) = \Pr\left(\frac{\eta_{it}}{\sigma_\eta} \geq \hat{\mu}' X_{it} - \mu_i \left(\frac{\rho}{1-\rho}\right)^{\frac{1}{2}}\right) \tag{32}$$

Conditioning upon X_it and μ_i, the likelihood of a sequence of responses, denoted $y(i)$, for an individual i for the ten experimental periods is expressed as follows

$$\Pr(y(i)|X_it, \mu_i) = \prod_{t=1}^{10} \Phi\left[\left(\hat{u}' X_{it} - \mu_i \left(\frac{\rho}{1-\rho}\right)^{\frac{1}{2}}\right)(1-2y_{it})\right], \tag{33}$$

where $\Phi()$ is culmulative distribution function for the standard normal.

[35]This is not a restrictive assumption but a normalization.

Recognizing that X_{it} is a fixed constant for all periods as is $\mu_i\,(\rho/1-rho)^{\frac{1}{2}}$, there is an identification problem. Therefore, in order to obtain estimates of the treatment thresholds, we would like to maximize the likelihood in (19) unconditional upon μ_i, or the expected likelihood,

$$\Pr(y(i)|X_it,\mu_i) = \int_{-\infty}^{\infty}\prod_{t=1}^{10}\Phi\left[\left(\hat{u}'X_{it}-\mu_i\left(\frac{\rho}{1-\rho}\right)^{\frac{1}{2}}\right)(1-2y_{it})\right]f(u)d\mu. \qquad (34)$$

We can now write the joint log-likelihood function to be maximized for all individuals who are experimental agent type e as

$$L = \sum_{i=1}^{N}\ln\int_{-\infty}^{\infty}\prod_{t=1}^{10}\Phi\left[\left(\hat{u}'X_{it}-\mu_i\left(\frac{\rho}{1-\rho}\right)^{\frac{1}{2}}\right)(1-2y_{it})\right]f(u)d\mu. \qquad (35)$$

Expression (21) is a random effects probit model. Efficient estimation of these models is discussed in [16],[10] and [25]. The small sample properties of this model are discussed in [14].

The main attraction of the random effects probit is its ability to estimate the intercept terms while controlling for heterogeneity in subjects. However, due to the possibility of state dependency, we will want to include lagged dependent variable into this specification, as well as other variables that can control for specific heterogeneity in the subjects.

One plausible interpretation of the error term η_{it}, is that subjects are uncertain of their own exact risk attitudes, and this could lead to normally distributed error terms. A second possibility is that the subjects' preferences are not the same in each of the ten decision periods and are changing stochastically.

If preferences over uncertain incomes change deterministically there are several sources of changing preferences such as wealth effects, changes in subjective probabilities, changes in the prior over the player type distribution and state dependency.

In order to test for these for these different possible explanatory variables of subjects' risk attitudes, we estimate the following specification for each experimental agent type:

$$L = \sum_{i=1}^{N}\ln\int_{-\infty}^{\infty}\prod_{t=1}^{10}\Phi\left[h(\cdot)(1-2y_{it})\right]f(u)d\mu, \text{ where}$$

$$h(\cdot) = \hat{\mu}_e + \gamma_1 y_{it-1} + \gamma_2 y_{it-2} + \gamma_3 \text{laggp}_{it-1} + \gamma_4 \text{hp}_{it-1} + \gamma_5 \text{mhp}_{it-1} + \gamma_6 \text{ae}_{it-1} - \mu_i \rho \quad (36)$$

with the following definitions: laggp is the number of individuals in the group besides i, hp is a dummy variable for a high price realization, mhp is a dummy variable for a high price realization and i chose to be in the pool, and ae is i's accumulated earnings in the experiment in U.S. dollars.

This large specification was estimated for each experimental agent type, following which the most parsimonious model that could not be rejected by a likelihood ratio test was estimated for each experimental type.

11 Nonparametric Analysis of the Experimental Data

From this point forward we will use the following set up and notation. Let $Y = (Y_1, \ldots, Y_T)$ be a finite sequence of identical random variables with each having the support $M = \{1, -1\}$, where $Y_t = 1$ represents a choice to join the pool in period t and $Y_t = 1$ is the choice to not join the pool in period t. Also, let y_t denote the realization of Y_t. We will refer to the choices as states. Thus, in this paper we will only consider the case where the set of possible states has only two elements; however, Lee presents a framework that is general to any finite number of states. For any finite s equence of responses, let $P^{1,2,\ldots,T}(y_1, y_2, \ldots, y_T)$ denote the probability of the event $Y_1 = y_1, Y_2 = y_2, \ldots, Y_T = y_T$. Furthermore, let $P^t(y_t)$ denote the marginal probability of $Y_t = y_t$. Now we can discuss the four distinct data generating processes.

The independent Bernoulli process is the most basic stochastic process, as is it is stationary, each realization is independent, and all sequences come from an identical joint probability distribution function. Lee identifies two properties that characterize the independent Bernoulli process:

Definition 1 *Independence property (I)*

$$P^{1,2,\ldots,T}(y_1, y_2, \ldots, y_T) = P^1(y_1)P^2(y_2)\cdots, P^T(y_T)$$

$$\text{for } y_t = -1, 1 \text{ and for all } t = 1, 2 \ldots, T$$

Definition 2 *Marginal Homogeneity (MH)*

$$P^t(k) = P^s(k) \; k = -1, 1 \text{ and for all } t, s = 1, 2, \ldots, T$$

From these properties it is apparent that the marginal probabilities for the occupancy of state y_t is the same for all periods an can be denoted $P(y_t)$ with out a time subscript.

The second stochastic process to be considered is the heterogeneous Bernoulli. The heterogeneity is the result of the marginal probabilities of the independent Bernoulli model randomly distributed across the population by a non-degenerate probability distribution function $v(p)$. This process will no longer satisfy the independence property (I). However, a result of De Finetti (1975) shows that this process satisfies a property called exchangeability. According to Lee a sequence of realized responses, $(y_1, y_2, \ldots, y_T)$, are exchangeable if they satisfy

Definition 3 *Exchangeability (E)*

$$P^{1,2,\ldots,T}(y_1, y_2, \ldots, y_T) = P^{1,2,\ldots,T}\left(y_{c(1)}, y_{c(2)}, \ldots, y_{c(T)}\right)$$

where c is any permutation of the set $1, 2, \ldots, T$.

The most simple stochastic process with state dependence is the first order Markov chain. A first order Markov chain is characterized by the property that the conditional probability distribution of Y_t given the information $Y_{t-1}, Y_{t-2}, \ldots$ is equal to the probability of Y_t conditional upon only Y_{t-1}. Given, that we only have two states in the model here, we can completely describe the first order process by a 2x2 transition matrix, which all individuals in the population follow. Lee characterizes this process by the property:

Definition 4 *Markov Property (M)*

$$P^{1,2,\ldots,s}(y_1, y_2, \ldots, y_{s-1}, y_s) = P^{1,2,\ldots,s}(1, 1, \ldots, y_{s-1}, y_s) = P^{s-1,s}(y_{s-1}, y_s)$$

If the transitional probability matrix of the first order model is not identical for each member of the population but instead are randomly distributed amongst the population by a non-degenerate probability distribution, then Lee refers to this as the mixed Markov process by Lee. This process will not satisfy the Markov property but will satisfy a weaker property called partially exchangeability. Let $\theta = (\theta_1, \theta_2, \ldots, \theta_T)$ and $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$ be finite string of realized states, then θ and ξ are said to be similar, denoted $\theta \sim \xi$, if and only if they start in the same state, change states the same number of times, and they end in the same state as they started. The condition of partial exchangeability can be state as

Definition 5 *Partial Exchangability (PE)*

$$P^{1,2,\ldots,T}(\theta_1, \theta_2, \ldots, \theta_T) = P^{1,2,\ldots,T}(\xi_1, \xi_2, \ldots, \xi_T), \text{ whenever } \theta \sim \xi$$

This concludes the definitions and descriptions of the important properties of the four competing stochastic processes. The nonparametric log linear probability model in which the four processes are nested is presented next.

All four of the above stochastic processes can be expressed as a special case of a model where the natural logarithms of the probabilities of sequences are expressed as linear functions of unknown parameters. The most general saturated log linear probability model for a finite length T of responses from a two element space state. 1,-1, according to Bishop et al. [7] is

$$\ln P^{1,2,\ldots,T}(y_1, y_2, \ldots, y_T) = \alpha + \sum_{r=1}^{T} \sum_{t_1=1}^{T-r+1} \sum_{t_2>t_1}^{T-r+2} \cdots \sum_{t_r>t_{r-1}}^{T} \alpha_{t_1 t_2 \cdots t_r} y_{t_1} y_{t_2} \cdots y_{t_r}, \quad (37)$$

where the coefficients satisfy the following constraint:

$$\sum_{y_{t_1} \in M} \cdots \sum_{y_{t_T} \in M} P^{1,2,\ldots,T}(y_1, y_2, \ldots, y_T) = 1$$

Solving for α we obtain.

$$\alpha = -\ln \sum_{y_{t_1} \in M} \cdots \sum_{y_{t_T} \in M} \exp\left\{ \sum_{r=1}^{T} \sum_{t_1=1}^{T-r+1} \sum_{t_2 > t_1}^{T-r+2} \cdots \sum_{t_r > t_{r-1}}^{T} y_{t_1} y_{t_2} \cdots y_{t_r} \right\}. \quad (38)$$

Substituting (24) into (22) we obtain a multinomial logit,

$$P^{1,2,\ldots,T}(y) = \frac{\exp\left\{\sum_{r=1}^{T} \sum_{t_1=1}^{T-r+1} \sum_{t_2>t_1}^{T-r+2} \cdots \sum_{t_r>t_{r-1}}^{T} \alpha_{t_1 t_2 \cdots t_r} y_{t_1} y_{t_2} \cdots y_{t_r}\right\}}{\sum_{y_{t_1} \in M} \cdots \sum_{y_{t_T} \in M} \exp\left\{\sum_{r=1}^{T} \sum_{t_1=1}^{T-r+1} \sum_{t_2>t_1}^{T-r+2} \cdots \sum_{t_r>t_{r-1}}^{T} y_{t_1} y_{t_2} \cdots y_{t_r}\right\}} \quad (39)$$

In order to get a better feel for this formulation, consider the following two period specification:

$$\ln P^{1,2}(y_1, Y-2) = \alpha + \alpha_1 y_1 + \alpha_2 y_2 + \alpha_{12} y_1 y_2,$$

for which the constraint $\sum_{y_{t_1} \in M} \sum_{y_{t_2} \in M} P^{1,2}(y_1, y_2) = 1$ implies that

$$\alpha = -\ln \sum_{y_{t_1} \in M} \sum_{y_{t_2} \in M} \exp\{\alpha_1 y_1 + \alpha_2 y_2 + \alpha_{12} y_1 y_2\}.$$

Substitution yields the following multinomial logit specification:

$$P^{1,2}(y_1, y_2) = \frac{exp\{\alpha_1 y_1 + \alpha_2 y_2 + \alpha_{12} y_1 y_2\}}{\sum_{y_{t_1} \in M} \sum_{y_{t_2} \in M} \exp\{\alpha_1 y_1 + \alpha_2 y_2 + \alpha_{12} y_1 y_2\}}$$

In the above framework the coefficients α_1 and α are referred to as the first order terms and $\alpha_{1,2}$ is the second order term. By placing linear restrictions upon these and higher order terms we can create nested specifications that correspond to the four stochastic processes presented earlier. The following describes these specifications, and one can find proofs of the following statements in the appendix of Lee's paper.

The homogeneous Bernoulli model is characterized by the independence property (I) and the marginal homogeneity property (MH). The independence property requires that all second and higher order terms equal zero. Furthermore, the marginal homogeneity property requires that all first order terms be equal.

The heterogeneous Bernoulli model is characterized by the exchangeability property (E). This property requires all terms of the same order to be equal. For example, consider the case where $T = 3$. Then (E) implies that $\alpha_1 = \alpha_2 = \alpha_3$ and $\alpha_{12} = \alpha_{13} = \alpha_{23}$.

The first order Markov model must satisfy the (M) Markov property. In turn, this implies all terms in the log linear probability model are zero except the first order terms and second order terms that have successive time period subscripts.

Finally, the mixed model requires that the probabilities of similar strings are equal. In other words, the process satisfies partial exchangeability (PE). The linear restrictions here are best treated on a case by case basis, by recognizing the equivalent strings and then setting up the linear constraints accordingly.

To conduct the hypothesis test for a specification, one first estimates the unrestricted and the restricted models by maximum likelihood. Then one can conduct any of the following asymptotically equivalent specification tests; the Wald, Likelihood Ratio, and Lagrange Multiplier test.

Usually, when attempting to use the framework above to investigate the heterogeneity and/or state dependency of the experimental data set for each experimental agent type, there is a problem with a large T and a relatively small N. The procedure above relies upon a fixed T and an increasing N for its asymptotic results. For our data, T=10 and the log linear probability model has 210-1 free probabilities, or possible sequences, and the same number of coefficients to be estimated. Given that there are only 30 observed sequences per model, we have a serious shortfall of observations if we want to estimate the full model. One of Lee's suggestions for this scenario is to look at several successive time lengths. Since a T of at least four is required to test the mixed model, we choose to estimate and test models for period sets 3,4,5,6 and 7,8,9,10 for each experimental agent type. Then a likelihood ratio test is conducted to compare the restricted models to the unrestricted models.

Figure 1: Cost Schedule for High Cost Environments

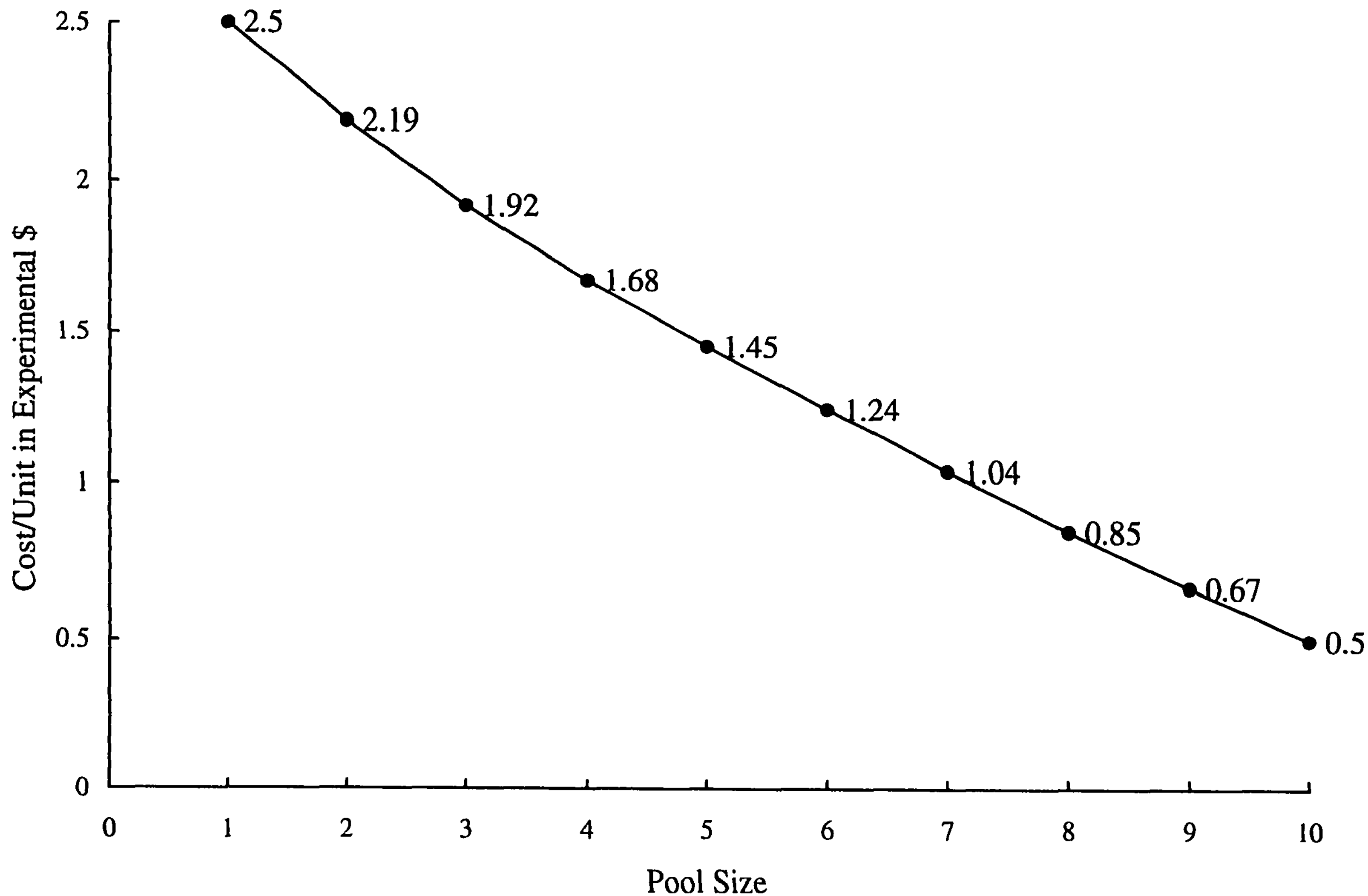
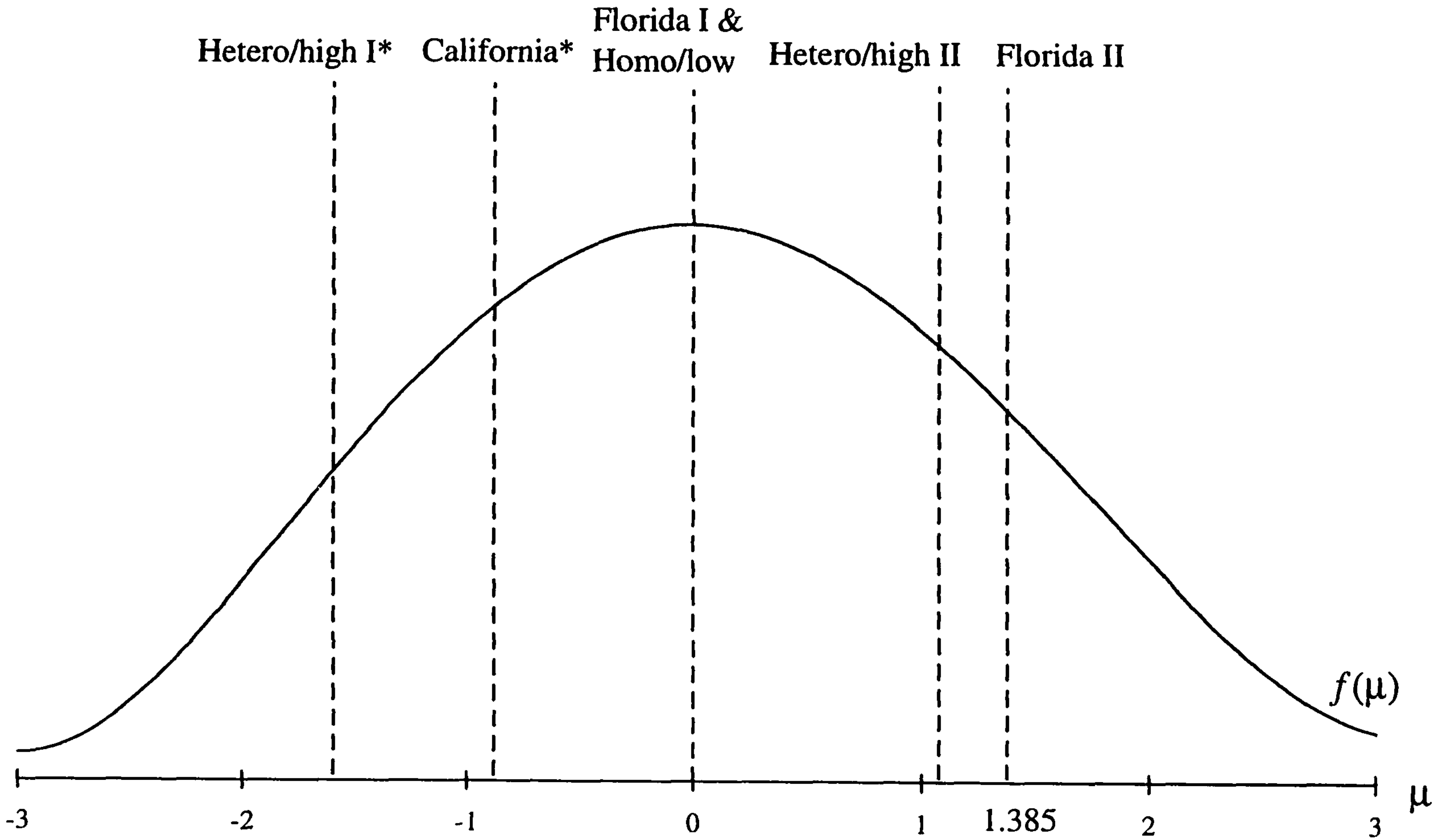


Figure 2: Theoretical Implications



*Note: The order of these two may be switched

Figure 3: Pooling Frequencies CA versus Fla Treatment Cells

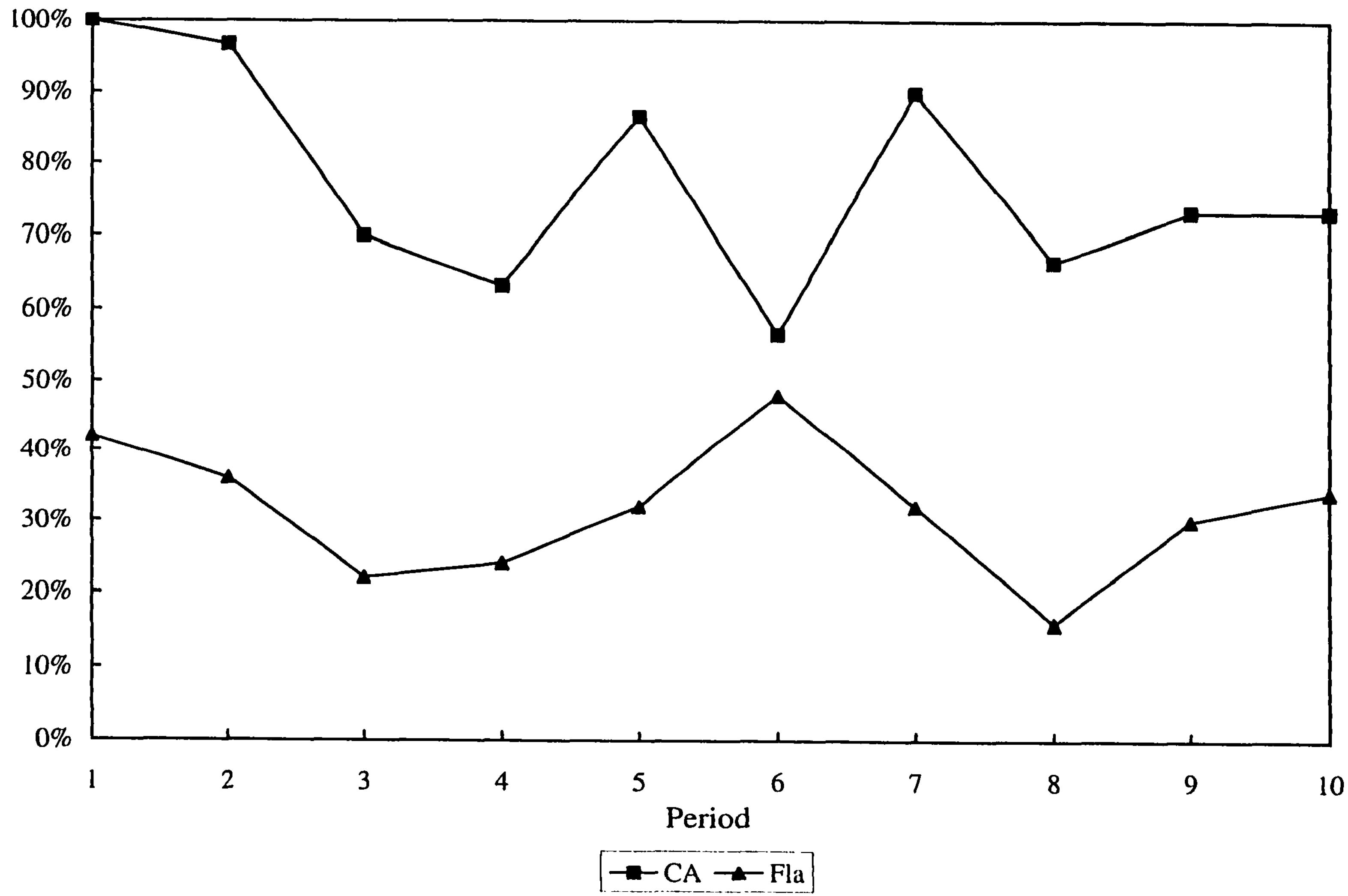


Figure 4: Pooling Frequencies for CA, Fla I, & Fla II

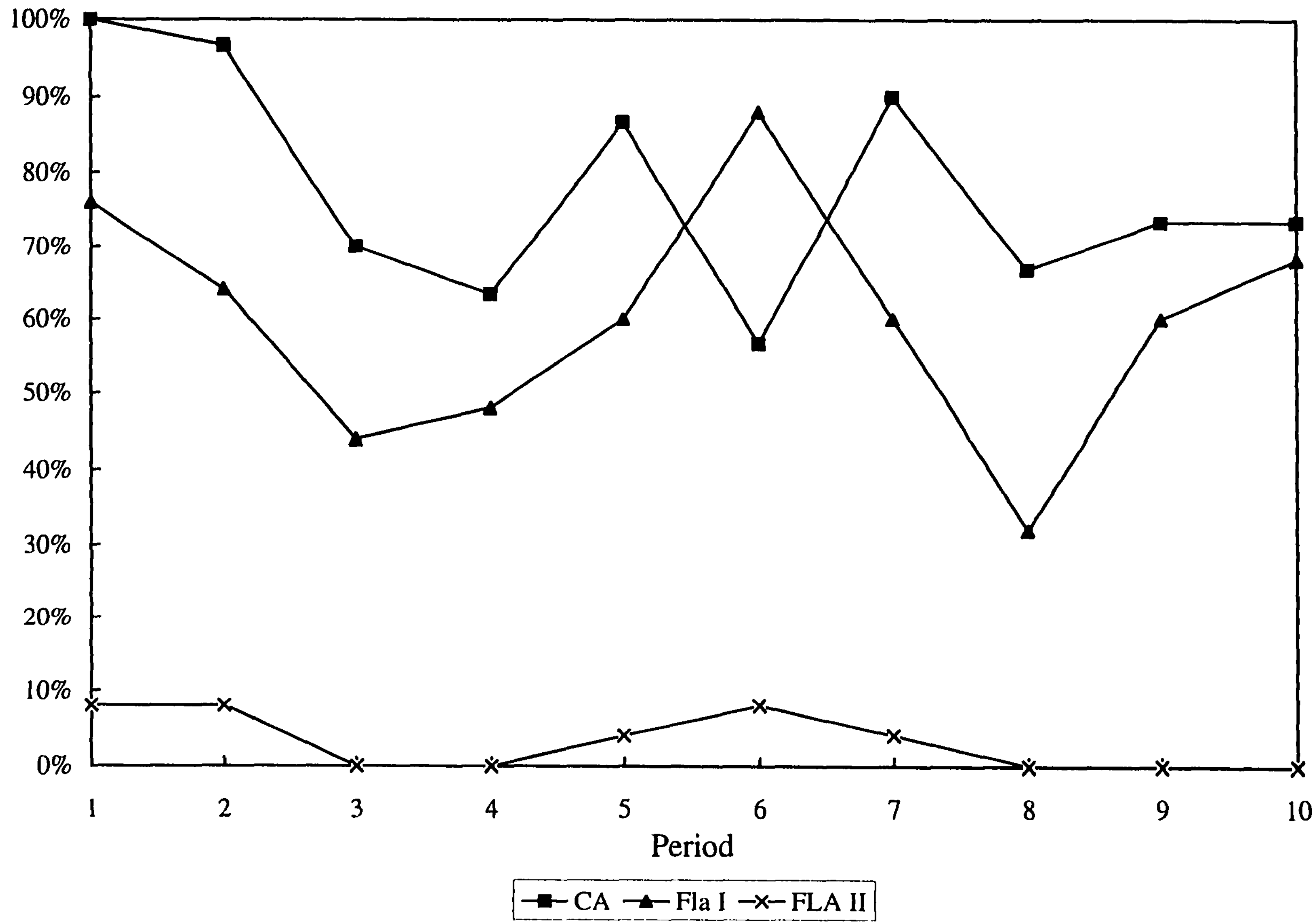


Figure 5: Pooling Frequencies for Low Cost Types

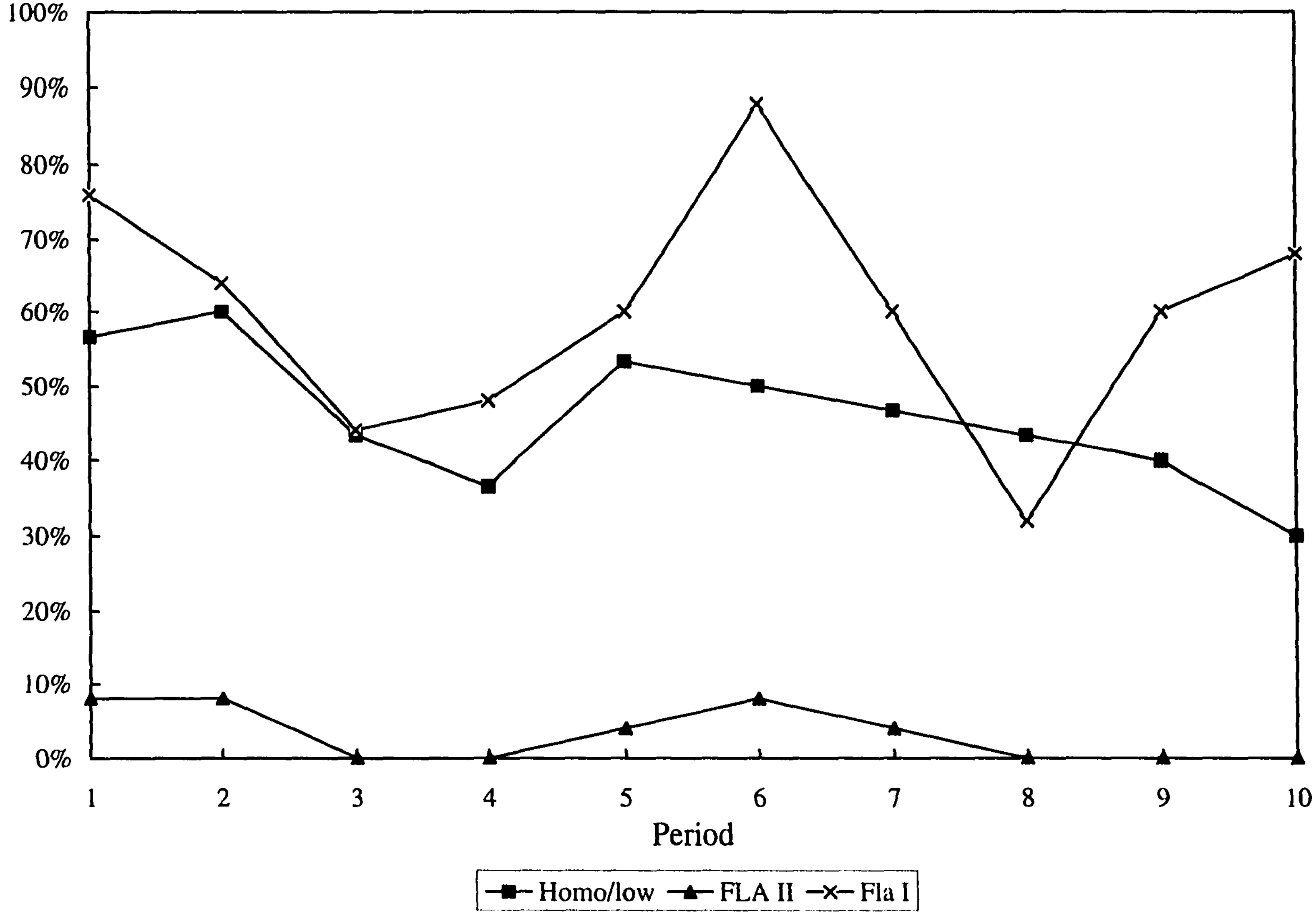
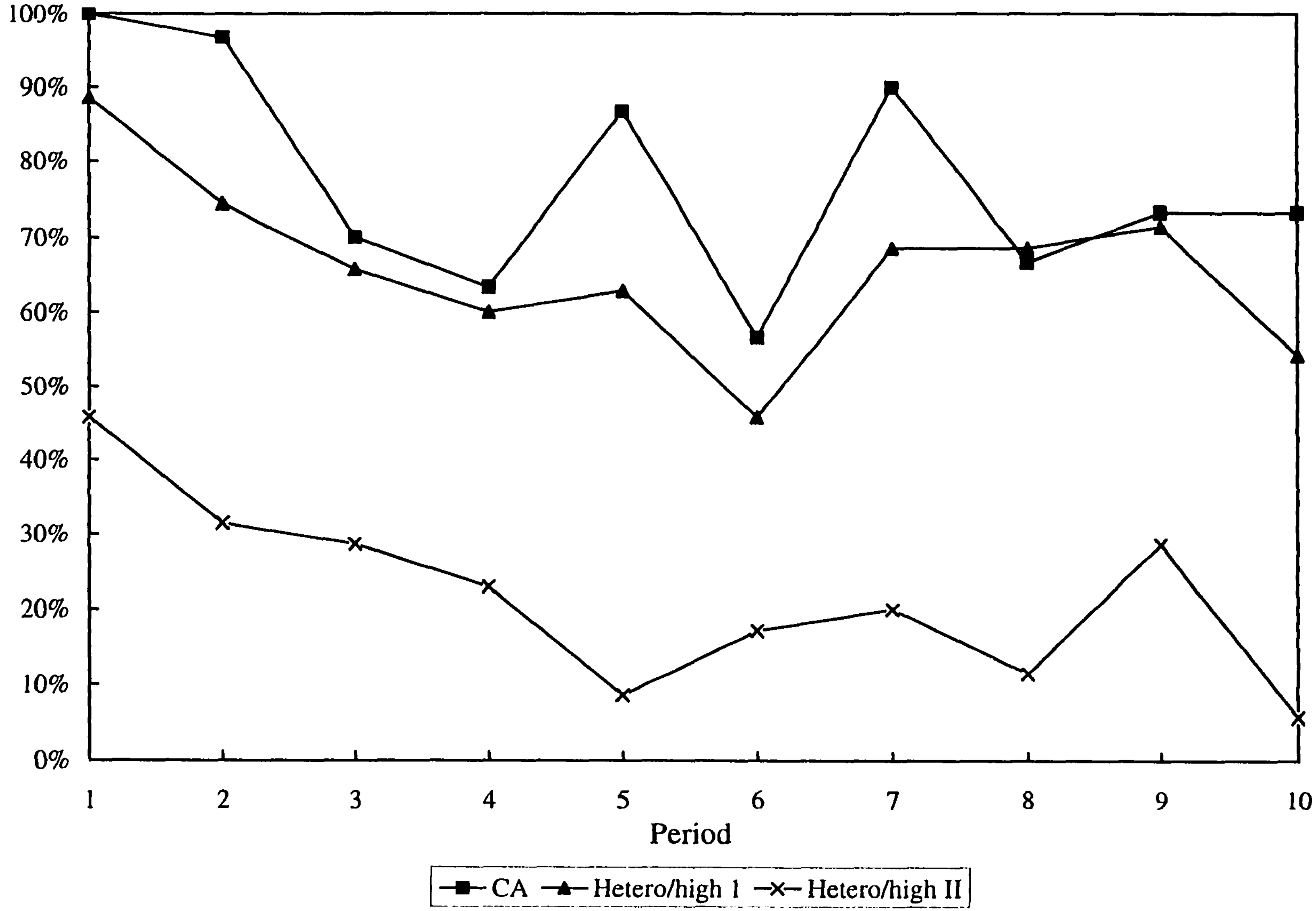


Figure 6: Pooling Frequencies for High Cost Types



References

[1] Mohammed Ali Khan and Yeneng Sun. Pure strategies in games with private information. *Journal of Mathematical Economics*, 24:633–653, 1995.

[2] Robert B. Ash. *Real Analysis and Probability.* Academic Press, 1972.

[3] Robert J. Aumann, Y. Katznelson, Roy Radner, Robert W. Rosenthal, and B Weiss. Approximate purification of mixed strategies. *Mathematics of Operation Research*, 8:327–341, 1983.

[4] Eric J. Balder. Generalized equilibrium results for games with incomplete information. *Mathematics of Operation Research*, 13:265–276, 1988.

[5] Patrick Billingsly. *Convergence of Probability Measures.* John Wiley and Sons. Inc., 1968.

[6] Brian Binger, Elizabeth Hoffman, and Gary Libecap. Experimental methods to advance historical investigation: An examination of cartel compliance by large and small firms. University of Arizona Discussion Paper, 1988.

[7] Y.M. Bishop, S.E. Fienberg, and P.W. Holland. *Discrete Multivariate Analysis: Theory and Practice.* MIT Press, Cambridge, 1975.

[8] Robert Bloomfield. Learning a mixed strategy equilibrium in the laboratory. *Journal of Economic Behavior and Organization*, 25:411–36, 1994.

[9] James Brown and Robert Rosenthal. Testing the minimax hypothesis: A reexamination of o'neill's game experiment. *Econometrica*, 58:1065–1081, 1990.

[10] John S. Butler and Robert Moffit. A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica*, 50(3):761–764, 1982.

[11] G. Choquet. *Lectures on Analysis.* Benjamin Reading, 1969.

[12] I. Glicksberg. A further generalization of kakutani's fixed point theorem with applications to nash equilibrium points. *Proceedings of the National Academy of Sciences, U.S.A.*, 38:170–172, 1952.

[13] Leo A. Goodman. A modified multiple regression approach to the analysis of dichotomous variables. *American Sociological Review*, 37:28–46, 1972.

[14] David K. Guilkey and John L. Murphy. Estimation and testing in the random effects probit model. *Journal of Econometrics*, 59:301–317, 1993.

[15] John C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed strategy equilibrium points. *International Journal of Game Theory*, 2:1–23, 1973.

[16] James J. Heckman. Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence. *Annales de L'INSEE*, 30/31:227–269, 1978.

[17] James J. Heckman. Statistical models for discrete panel data. In Charles F. Manski and Daniel McFadden, editors, *Structural Analysis for Discrete Data with Econometric Applications*, pages 114–178. MIT Press, Cambridge, 1981.

[18] Elizabeth Hoffman and Gary Libecap. Political bargaining and new deal agricultural policies: Citrus marketing orders in the 1930's. In Claudia Goldin and Gary Libecap, editors, *The Regulated Economy: A Historical Approach to Political Economy*, pages 189–221. University of Chicago Press, 1994.

[19] Elizabeth Hoffman and Mark L. Spitzer. Entitlements, rights, and fairness: Some experimental evidence of subjects' concepts of distributive justice. *The Journal of Legal Studies*, 42(2):258–298, 1985.

[20] Charles Holt. Industrial organization: A survey of laboratory research. In John Kagel and Al Roth, editors, *Handbook of Experimental Economics*, pages 114–178. MIT Press, Cambridge, 1995.

[21] A. Ionescu-Tulcea and C. Ionescu-Tulcea. *Topics in the Theory of Lifting.* Springer-Verlag, 1969.

[22] C. Kraenzie. Co-ops' share of grains, milk, cotton up in 1989. *Farmer Cooperatives*, 58(2):4–7, 1991.

[23] Lung-Fei Lee. Nonparametric testing of discrete panel data models. *Journal of Econometrics*, 34:147–177, 1987.

[24] L. Lopes. Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8:626–636, 1982.

[25] G.S. Maddala. Limited dependent variable models using panel data. *Journal of Human Resources*, 22(3):307–336, 1987.

[26] Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibrium for normal form games. *Games and Economic Behavior*, 10:6–38, 1995.

[27] Paul R. Milgrom and Robert J. Weber. Distributional strategies for games with incomplete information. *Mathematics of Operation Research*, 10:619–632, 1985.

[28] Dilip Mookherjee and Barry Sopher. Learning behavior in an experimental matching pennies game. *Games and Economic Behavior*, 7:62–91, 1994.

[29] M. Nerlove and Press S. J. Multivariate log-linear probability models for the analysis of qualitative data. Discussion Paper no. 1 (Center for Statistics and Probability, Northwestern University), 1976.

[30] Jack Ochs. Games with unique mixed strategy equilibria: An experimental study. *Games and Economic Behavior*, 10:202–217, 1994.

[31] Barry O'Neill. Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences, U.S.A.*, 84:2106–2109, 1987.

[32] John Pratt. Risk aversion in the small and the large. *Econometrica*, 32:122–136, 1964.

[33] Roy Radner and Robert W. Rosenthal. Private information and pure-strategy equilibria. *Mathematics of Operation Research*, 7:401–409, 1982.

[34] Amnon Rapoport and R. Boebel. Mixed strategies in strictly competitive games: A further test of the minimax hypothesis. *Games and Economic Behavior*, 4:261–283, 1992.

[35] Amnon Rapoport and David V. Budescu. Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology, General*, 121:352–363, 1992.

[36] Micheal Reed and Barry Simon. *Methods of Modern Mathematical Physics, I: Functional Analysis*. Academic Press, 1980.

[37] R. Richardson and et. al. Business volume sets $77 billion record, but farmer cooperative income declines. *Farmer Cooperatives*, 58(8):10–11, 1991.

[38] Jason Shachat. Mixed strategy play and the minimax hypothesis. University of Arizona, 1995.

[39] Vernon L. Smith. Microeconomic systems as an experimental science. *American Economic Review*, 72:923–955, 1982.

[40] John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, third edition, 1944.