PRUDENCE, JUSTICE, BENEVOLENCE, & SEX:
Evidence from Similar Bargaining Games

John Van Huyck and Raymond Battalio

Printed October 28, 2001

Comments to john.vanhuyck@tamu.edu
Related research available at
http://erl.tamu.edu

Abstract: Most learning experiments involve repeated play of exactly the same situation and, hence, can not discriminate between learning to use a deductive principle and other forms of routine learning. In this paper, subjects confront a sequence of similar, but not identical, bargaining games all of which can be solved using the same deductive principles. Conventions based on these deductive principles emerge within 70 periods in 5 of 26 eight person cohorts. We found no economically significant differences between all male and all female cohorts.

Key Words: Bargaining, Equilibrium Selection, Learning, Evolutionary Games, Gender Differences.

JEL Classification: c72, c78, c92, d83.

Acknowledgments: Nick Rupp and Nick McKinney provided able research assistance. Jane Sell provided the recruiting protocol. Joanna McAfee and Carla Goar executed it. Charles Holt suggested the revised title. The original title was an allusion to Smith (1759). Gary Bolton, Catherine Eckel, John Kagel, Ernan Haruvy, and Dale Stahl made helpful comments. The National Science Foundation and the Texas Advanced Research Program provided financial support. Eric Battalio implemented the experimental design on the TAMU Economic Research Laboratory network. The first draft of this paper was written while Van Huyck was on a faculty development leave at the University of Pittsburgh.

I. INTRODUCTION

Learning requires recognizing the similarity between the current instance of a situation and past instances. Game theory is a useful way to identify strategic similarities between situations. When placed in a sequence of strategically similar games, do people recognize and exploit these similarities to learn to coordinate on a way to play?[1]

The experiment reported below uses a random sequence of scrambled payoff perturbed 2×2 bargaining games to discover if subjects learn to use a deductive bargaining concept to solve their coordination problem. All games have the following strategic properties: two efficient strict equilibria one of which is selected by a *Utilitarian* criterion (select the equilibrium with the largest payoff sum) and the other is selected by a *Rawlsian* criterion (select the equilibrium that maximizes the welfare of the worst off player); no other outcomes are contained in the set of efficient outcomes; security varies so that risk dominance selects the Utilitarian equilibrium half the time and the Rawlsian equilibrium half the time.

Subjects interact using an evolutionary protocol.[2] Repeated interaction allows subjects to learn to use a principle of justice to select an equilibrium if subjects fail to bring one into the laboratory that is sufficiently flexible to solve the posed asymmetric bargaining problems. The principle would have to be more sophisticated than equal division, since equal division is not a feasible outcome *ex post*.[3]

---

[1] Kreps (1990, p184) concludes his book by observing that game theory, "has brought a fairly flexible language to many issues, together with a collection of notions of 'similarity' that has allowed economists to move insights from one context to another and to probe the reach of those insights. But too often it, and in particular equilibrium analysis, gets taken too seriously at levels where its current behavioral assumptions are inappropriate." Rankin *et al.* (2000) reports an experiment in which efficiency, rather than risk dominance or security, emerges as a selection principle in similar stag hunt games. This contrasts with the widely documented coordination failure observed when exactly the same stage game is played repeatedly, see Battalio *et al.* (2001).

[2] Specifically, the experiment uses a one population pairwise random matching protocol.

[3] Equal division is a powerful behavioral regularity found in most situations with symmetric bargaining power. However, Van Huyck, *et al.* (1995) studied symmetric bargaining game forms using an evolutionary matching protocol and the same earnings matrix every period. Even though equal division was an efficient strict equilibrium, they observed unequal-division conventions emerging in communities of symmetrically endowed subjects. Their interpretation of these observations was that security undermined the salience of equal division and then routine learning drove the laboratory community to an unequal division convention.

Action labels and player roles are also scrambled to prevent players from using non-strategic details to solve their coordination problem. Hence, the only meaningful way to distinguish between actions is by using a deductive principle and it is always possible to label one action Utilitarian, the *U*-action, and the other action Rawlsian, the *R*-action. Do subjects behave as if they recognize the map between the current action labels and an abstract description of the game using *U*-action and *R*-action labels? If so, which role contingent convention is more likely to emerge: the *U*-convention or the *R*-convention?[4]

Of course, this design can not resolve a normative dispute between Harsanyi (1977) and Rawls (1971). Rather, it may demonstrate that subjects can learn to recognize either deductive principle in an environment that eliminates distracting and competing selection principles and coordinate their behavior on one of the two competing principles in small laboratory communities. Will a convention in which subjects use a deductive principle emerge?[5]

The widely documented fact that naive subjects use equal division to solve simple bargaining problems does not appear to reflect any deeper widely shared understanding of general principles of justice. Initial behavior was not *ex post* mutually consistent. However, conventions based on deductive principles emerge after 70 periods in 5 of 26 eight person cohorts. In four cohorts the U-convention emerged and in one cohort the R-convention emerged.

It is common to make generalizations based on the sex ratio of a group. Observing that behavior in dictator games varied systematically by gender classification, Andreoni and Vesterlund (2001, p307) conclude that "researchers would be wise to assure that their experimental findings are the result of economic incentives and not of varying sex compositions of their control and treatment groups." The experiment reported below varies the sex ratio. All male, all female, and mixed cohorts are used. Does the likelihood of observing a convention vary with the sex ratio? We found no economically significant differences between all male and all female cohorts.

---

[4] Van Huyck, Battalio, and Rankin (1997) found that subjects could coordinate on role contingent conventions in an evolutionary pure coordination game. See Crawford (1997) for a theoretical discussion and references to the theory literature.

[5] See Binmore (1998) for a naturalistic defense of Rawls.

II. ANALYTICAL FRAMEWORK

Figure 1 is a description of a 2×2 game. In order to generate a sequence of similar bargaining game forms, let all of the variables denote money drawn randomly from a joint distribution. The first requirement is that the values give a bargaining game satisfying the following conditions:

$$w > z > y > x \tag{1}$$

$$w + x > y + z \tag{2}$$

Let a,b,c,d be less than x. Specifically, w,x,y,z were drawn independently and uniformly from the interval [21,50] and a,b,c,d were drawn independently and uniformly from the interval [1,20], where units denote cents. The second requirement was a 5 cents salience rule-of-thumb, which amounts to $y - 5 \geq x$ and $w + x - 5 \geq y + z$. A bimatrix was filtered out if it did not satisfy conditions (1) and (2), the 5 cents salience condition, and a risk-dominance condition described below.

	U	*R*
U	*x* *w*	*b* *a*
R	*d* *c*	*z* *y*

Figure 1: Bargaining Game Forms ($w > z > y > x > a, b, c, d.$)

Under an abstraction assumption of own money motivated players, the games have two strict equilibria: *UU* and *RR*. Condition (1) implies that *RR* is the equilibrium that maximizes the welfare of the worst off player and will be denoted the *Rawlsian Equilibrium* or *R-equilibrium* for short. Condition (2) implies that *UU* is the equilibrium with the largest payoff sum and will be denoted the *Utilitarian Equilibrium* or *U-equilibrium.*

Two kinds of prudence will be important for understanding the experimental design: security and risk dominance. *Security* selects the action that maximizes own payoff in the worst outcome.[6] *Risk dominance* is equivalent to selecting the equilibrium with the larger basin of attraction under best response dynamics in 2×2 games.

Games with $a > c$ and $b > d$ will be denoted *R*-games. In *R*-games, the action implementing a player's best equilibrium is secure for both players.[7] Games with $c > a$ and $d > b$ will be denoted *U*-games. In *U*-games, the action implementing a player's worst strict equilibrium is secure for both players. If both players choose their secure action the outcome is neither mutually consistent nor efficient in *U* or *R* games. Nevertheless, Rapoport, Guyer, and Gordon's (1976) in their taxonomy of ordinal 2×2 games denote this the "natural outcome."[8]

Risk dominance uses more information than the ordinal ranking of the outcomes used to this point. It is possible for risk dominance to select either strict equilibrium in either game form. The final condition on the games was that the randomly generated games were filtered such that risk dominance selects the *U*-equilibrium in *U*-games and the *R*-equilibrium in *R*-games.

The similar bargaining games have a third equilibrium in mixed strategies, which varies with the random payoffs. Figure 2 plots the mixed equilibria for 35 *R*-games and figure 3 plots the mixed equilibria for 35 *U*-games. In the figures, p denotes the probability of the row player's *U* action and q denotes the probability of the column player's *U* action. The *U*-equilibrium is at (1,1) and the *R*-equilibrium is at (0,0) for all 35 *R*-games and all 35 *U*-games.

Notice that for *R*-games the mixed equilibrium (p^*,q^*) always lies to the northeast of the line defined by $p+q=1$, that is, $p^* + q^* > 1$, which means the best response basin of attraction for the *R*-equilibrium is larger than the basin of attraction for the *U*-equilibrium. This is the condition for risk-dominance to select the *R*-equilibrium and this is why the games are denoted *R*-games. The reverse is true for *U*-games.

A *convention* is usually defined to be a regularity in behavior amongst members of a community in a recurrent situation that is customary,

---

[6] In this paper, we use security to refer to the action that maximizes own payoff in the worst outcome and maximin to refer to the mixed strategy that maximizes own payoff in the worst outcome.

[7] Van Huyck, *et al.*'s (1995) stage game reduces to an *R*-game once the equal division action becomes extinct in the community.

[8] See their discussion of game #69, the *R*-game , and game #68, the *U*-game.

expected, and mutually consistent.[9] This definition combines facts about an observable situation (customary behavioral regularities) and properties of an abstract game derived from the observable situation (mutually consistent behavior). For our purposes the definition must be extended to recurring similar, but not identical, game forms.

Reinterpret p and q in figures 2 and 3 as the population frequency of the U-action when labeled row or column. It is always in a player's pecuniary interest to play the R-action for all population states (p,q) in the shaded region of figure 2 labeled R-convention for all 35 R-games. Similarly, it is always in a player's pecuniary interest to play the U-action for all population states (p,q) in the region labeled U-convention for all 35 R-games. Interpret the R and U convention of the U-games in figure 3 the same way. Denote behavioral states in the shaded regions labeled R or U convention to be observations of an R or U convention.

Probabilistic choice models given the same precision parameter predict different distortions away from the strict equilibria for the average R and U games. Figures 4 and 5 graph the logit response vector field of the average R and U bimatrix, see Fudenberg and Levine (1998). The logit response dynamic is as follows:

$$\dot{p} = \frac{e^{\lambda(1,-1).m.(q,1-q)}}{1 + e^{\lambda(1,-1).m.(q,1-q)}} - p$$

$$\dot{q} = \frac{e^{\lambda(p,1-p).M.(1,-1)}}{1 + e^{\lambda(p,1-p).M.(1,-1)}} - q$$

where m denotes the row player's matrix and M denotes the column player's matrix from the respective bimatrix representations R and U. The figures use a precision parameter of $\lambda=0.14$.

The red dots in the figure denote the fixed points of the logit response dynamic, which are logit equilibria, see McKelvey and Palfrey (1995). Notice that, when we account for the incentive players have to optimize, the equilibria of the average R-game are distorted in different ways than the average U-game. For example, given $\lambda= 0.14$, the U-equilibrium goes from (1,1) to (0.98,0.92) in the average U-game and from (1,1) to (0.96,0.67) in the average R-game.

---

[9] The definition is from Lewis (1969) who demonstrated that it is useful to model people using sounds, gestures, or symbols to communicate things to one another as a coordination game and he used game theory to explain what it means to say that language is a convention.

III. EXPERIMENTAL DESIGN

The existential description given in the previous section should not be confused with either the experimental description we use to explain the game to our subjects or with the subjective description they construct to think about the game. Appendix A provides the 70 bimatrix games used in the experiment. Units denote cents. Actions were labeled 1 and 2. Note that due to independent row and column action label scrambling the location of the U and R equilibria can be in any of the four cells of the bimatrix.

Twenty-six cohorts, each consisting of 8 subjects, played the sequence of bimatrix games reported in Appendix A. Each period a subject was randomly assigned a role, either row or column, and then randomly matched to play one of the other subjects in the cohort. Hence, a subject had a one in two chance of playing the same role and a one in seven chance of playing the same opponent next period. We have found that a one in seven chance of immediate repeated interaction is small enough for subjects to treat it as negligible, thus suppressing repeated game effects.

Up to this point we have assumed that own money maps into some cardinal measure of utility. There is evidence to indicate that subjects are influenced by not only their own money earnings but also the earnings of others. Andreoni and Miller (forthcoming) have developed a protocol that allows one to estimate individual subject preferences for own, m_s, and other, m_o, money earnings. Define players with $u(m_s,m_o) = m_s$ as *selfish*, with $u(m_s,m_o) = \min\{m_s, m_o\}$ as *egalitarian*, and $u(m_s,m_o) = m_s + m_o$ as *magnanimous*. Using their protocol Andreoni and Miller (forthcoming, table 3) classified 47 percent of their subjects as selfish, 30 percent as egalitarian, and 23 percent as magnanimous.

Andreoni and Vesterlund (2001) found statistically significant gender differences in this classification. They classified their females as 36 percent selfish, 53 percent egalitarian, 9 percent magnanimous, and 2 percent unclassified. They classified their males as 47 percent selfish, 25 percent egalitarian, and 27 percent magnanimous. This gender difference led Andreoni and Vesterlund to suggest "gender balanced" experimental conditions.

If it is common knowledge that both players in the bargaining game form depicted by figure 1 are egalitarian then RR, the equilibrium selected by selfish players using a Rawlsian selection criterion, is the unique payoff dominant equilibrium of the game played by egalitarian players. Similarly, if it is common knowledge that both players in the game are magnanimous then UU, the equilibrium selected by selfish players using a Utilitarian selection criterion, is the unique payoff dominant equilibrium of the game played by magnanimous players. Given the salience of payoff dominance as a selection principle, changing the composition of player types may influence the likelihood of the *U* or *R* convention emerging.

The experiment was designed to determine if gender is an effective instrument variable. We consider three gender treatments: 10 all female cohorts, 10 all male cohorts, and 6 mixed cohorts. The mixed cohorts were of two types: 3 used a call back method to ensure exactly 4 males and 4 females and 3 used our usual gender blind methods. Public knowledge of cohort gender composition was established only to the extent that participants could see everyone file into the laboratory and go to their stations.

The subjects made choices using a computer assisted graphical user interface available in the TAMU Economic Research Laboratory. The payoff matrix for each period was displayed on a monitor at the front of the room to ensure that the game form was common information. No preplay communication was allowed. After each repetition of the period game, subjects' earnings were calculated for that period. Subjects only observed the outcome of their own pairing. To further limit the usefulness of retrospective selection principles, there was no computer assisted record keeping of experienced payoff matrices or action combinations. Subjects did have access to pencil and paper and could have kept what ever records they felt would be useful. Almost all subjects chose not to keep any records whatsoever. (The instructions text file used by the graphical user interface is available at http://erl.tamu.edu.)

The subjects were undergraduate students at Texas A&M University in the 1997 summer and 1998 spring semesters. A total of 208 subjects participated in the twenty-six cohorts reported below. After the instructions, but before the experiment began, the subjects filled out a questionnaire to ensure that it was common information that everyone understood how to compute payoffs for themselves and their opponents. If any subject made a mistake on the questionnaire, the relevant part of the instructions were read again. The experiment takes about two hours to conduct. The average subject earns $24.66 in the *U* equilibrium, which was more than twice the minimum wage at the time.

IV. RESULTS

A. Initial Behavior

Overall, the R action is slightly more popular than the U action in period 1. It is played by 110 of the 208 subjects (53 percent). This almost uniform behavior suggests that neither the *U* nor the *R* convention was initially present in our subject pool.

Table 1 reports counts of initial behavior for single sex cohorts, that is, cohorts 1 to 20. The game form presented to subjects for period 1, a *U*-game, and for period 2, an *R*-game, are included for easy reference. Males play the R action slightly more frequently than females. This difference is not statistically significant. The probability that a result this extreme or more extreme could have happened by chance, or *p*- value, of Fisher's exact test for no gender differences in initial row role behavior in the first *U* and *R* game is 1.00 and 1.00 respectively, while in the column role it is 0.65 and 0.60 respectively.

While the *U* and *R* games have similar strict equilibrium points, subjects responded to the difference in security between the two games. The main difference in initial behavior between *U* and *R* games is column conformity to the *U*-convention. Subjects avoid the short end of the *U*-equilibrium when it is implemented using the least secure action as happens in the *R*-game's column role.

Only one quarter of the subjects choose the *U* action in the column role of the *R*-game in period 2. The row labeled Fisher's *p*-value in table 1 reports the Fisher's exact test of no difference in behavior between *U* and *R*-games for the 2×2 contingency table directly above the respective cells. Combining the all male and all female cohorts gives a *p*-value of 0.53 for the null hypothesis of no difference in row role behavior, but a *p*-value of 0.01 for the null hypothesis of no difference in column role behavior. Thus the two games give statistically significant differences in initial column behavior.[10]

B. Adaptive Behavior

Figure 6*a,b,c* plot three interesting examples of observed adaptive behavior. The plots divide the seventy periods into seven 10 period states.

---

[10] Given no gender differences in initial row role behavior a more powerful test of no row role differences uses all 26 cohorts, which gives a statistically significant difference in row role behavior at the 10 percent level. The *p*-value on column behavior falls to approximately 0.00.

Every 10 periods the 8 participants made five choices in a *U*-game and five choices in an *R*-game half in the row role and half in the column role. The *U*-action could have been chosen up to 20 times by participants in the row role, the horizontal axis, and up to 20 times by participants in the column role, the vertical axis. The open circle (square) counts the number of *U*-actions in *U*-games (in *R*-games) for a ten period state. The number within the open circle or square indexes the states with 1 denoting the first ten periods and 7 denoting the last ten periods.

Cohort 7 in figure 6 is an example of convergence to the *U*-convention in both the *U* and *R* games. Cohort 6 is an example of non-convergence. (Or perhaps it is an example of convergence to the mixed equilibrium, but notice that cohort 6's behavior in the *U*-games are still a long way from the mixed equilibrium.) Cohort 20 is an example of convergence to the *R*-convention, in fact the *R*-equilibrium*,* in both the *U* and *R* games. These are the best examples of the three cases in the data and none of them should be taken as the typical case. Appendix B available at http://erl.tamu.edu plots all 26 cohorts.

Figure 4 and 5 illustrate the logit response vector field for the average *U* and *R* game respectively. In order to compare behavior with these figures, we construct what we call an empirical vector field illustrating how behavior changed from the first five instances of the *U* or *R* game to the last five instances of the game. Figure 7 graphs the empirical vector field for all 26 cohorts. Each cohort contributes one green (grey) vector for *U*-games and one blue (black) vector for *R*-games. Vectors represent the change from the first to the last ten period state. Cohort 20 is the easiest to find since it is the only cohort to converge to the *R*-equilibrium, compare figure 6*c* and 7.

The most striking feature of figure 7 is the initial separation of behavior between the *R* and *U* games. Again, this reflects the reluctance of column players to select the *U*-equilibrium when it is implemented with their least secure action. In some cohorts, subjects get over this and become sufficiently confident in the *U*-equilibrium to conform to the *U*-convention. In fact, the closest state to perfect conformity with the *U*-equilibrium occurs in the *R*-games, see cohort 7 in figure 6*a*.

Most vectors point towards the *U*-convention. Some cohorts even begin within the *R*-convention and escape to the *U*-convention. Amongst *R*-games a good number seem to be converging to a mixed equilibrium. This may be a medium run phenomenon. There are also numerous examples of violations of the Logistic Response Vector Field.

C. Emerged Conventions

Table 2 classifies terminal states by emerged convention. Four cohorts (7,11,17,21) converge to the U convention for both U and R games. One cohort (20) converges to the R convention for both U and R games. Neither convention emerges for either game in three cohorts (4, 6, 26). The remaining 18 cohorts are mixed cases. None of the cohorts were even close to approaching the most interesting convention in which the *U* - convention is used for *U*-games and the *R*-convention is used for *R*-games, that is, a convention based on the deductive principle of risk dominance.

A *U*-convention for the *U*-game emerged in 15 of 26 cohorts, while the *R*-convention emerged once. An *R*-convention for the *R*-game emerged in 8 of 26 cohorts, while the *U*-convention emerged in 4 cohorts. Hence, the *U*-convention emerges twice as often as the *R*-convention when restricting attention to conformity with risk dominance. Moreover, within a cohort the emergence of the *U*-convention spreads from the *U*-game to the *R*-game more frequently than the *R*-convention spreads from the *R*-game to the *U*-game.

Table 3 sorts emerged conventions by the cohort's gender composition. While the *U*-convention emerges least frequently in the *U*-games amongst all female cohorts and the *R*-convention emerges least frequently in the *R*-games amongst the mixed cohorts, these differences are not statistically significant. Fisher's exact test of the null hypothesis of no difference by gender gives p- values of 0.37 for *U*-games and 1.00 for *R*-games. Including the mixed cohorts reduces the p- value to 0.24 and 0.39 respectively. The emerged convention is not related to the gender composition of the cohorts.[11]

Table 4 examines earnings differences by gender composition. Results under three theoretical solution concepts--maximin, *R*-equilibrium, and *U*-equilibrium–are also included for comparison.[12] The difference in mean earnings per capita between all female and all male cohorts is 3 cents. Mixed cohorts earned about 30 cents more on average and experienced a

---

[11] Ignoring which convention emerges does not give statistically significant results. A convention emerges in thirteen all male and nine all female cohort/games out of 20 possible trials respectively. The Fisher's exact test of no gender difference in the frequency of emerged conventions gives a p- value of 0.34.

[12] The maximin solution concept is similar to security except that it assigns a mixed (rather than pure) strategy to each player, which maximizes their earnings in the worst outcome. It provides a lower bound on a rational player's earnings. Cohort 14 is within 7 cents of this lower bound.

lower standard deviation. Subjects achieved a little less than 70 percent of the feasible earnings under the *U*-equilibrium, which costs the average subject more than $7.63. The standard deviation in earnings also exceeds the value under the *U*-equilibrium suggesting that actual behavior was not only inefficient but also unnecessarily unequal.

Rank the single sex cohorts by mean per capita earnings. The rank sum for all male cohorts is 106 and for all female cohorts is 104, giving a *p*-value of 0.97 for the null hypothesis of no gender difference using the Wilcoxon 2 sample rank sum test. If there are gender differences, they don't influence average efficiency.

V. CONCLUSIONS

Initially players are particularly reluctant to conform to the Utilitarian convention when doing so gives them a small share and, given their role, requires them to use their least secure action. This economically and statistically significant difference in behavior is predicted by probabilistic choice models and remains an important characteristic of the data through out the convergence process.

Conventions based on the strategic similarities between the games do sometimes emerge. In 5 of 26 cohorts the same convention emerged for all of the similar games within the 70 periods of the evolutionary game. Given convergence in both types of games, the Utilitarian convention emerged 4 times as often as the Rawlsian convention. We never observed the most interesting case, which would have been a Rawlsian convention in games in which risk dominance selects the Rawlsian equilibrium and the Utilitarian convention in games in which risk dominance selects the Utilitarian equilibrium, that is, a convention based on the deductive principle of risk dominance.

Our experiment reveals no evidence of economically significant gender differences in bargaining behavior. The difference in mean earnings between all male and all female cohorts was 3 cents and the difference in standard deviation was 2 cents.

REFERENCES

James Andreoni and John H. Miller, "Giving According to GARP: An Experimental Study of Rationality and Altruism," forthcoming in *Econometrica*, .

James Andreoni and Lise Vesterlund, "Which is the fair sex? Gender differences in altruism,"*Quarterly Journal of Economics*, 116(1), February 2001, 293-312.

Simon P. Anderson, Jacob K. Goeree, and Charles A. Holt, "Minimum-effort Coordination Games: an equilibrium analysis of bounded rationality," forthcoming in *Games and Economic Behavior.*

Raymond Battalio, Larry Samuelson, and John Van Huyck, "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games, *Econometrica*, 63(9), May 2001, 749-764.

Ken Binmore, *Game Theory and the Social Contract: Just Playing*, (Cambridge, MA: The MIT Press, 1998).

Vincent Crawford, "Theory and Experiment in the Analysis of Strategic Interaction," 206-242 in David Kreps and Ken Wallis, editors, *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. I*, Econometric Society Monographs No. 27, (New York, NY: Cambridge University Press, 1997).

Lawrence E. Fouraker and Sidney Siegel, *Bargaining Behavior*, (New York, NY: McGraw-Hill Book Company, Inc., 1963).

Drew Fudenberg and David K. Levine, *Theory of Learning in Games*, MIT Press, 1998.

John C. Harsany, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations,* (New York, NY: Cambridge University Press, 1977)

David M. Kreps, *Game Theory and Economic Modeling*, (Oxford, UK: Clarendon Press, 1990).

Richard D. McKelvey and Thomas R. Palfrey, "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1) July 1995, 6-38.

Frederick W. Rankin, John B. Van Huyck, and Raymond C. Battalio, "Strategic Similarity And Emergent Conventions: Evidence from Scrambled Payoff Perturbed Stag Hunt Games," *Games and Economic Behavior*, 32(2) August 2000, 315-337

Anatol Rapoport, Melvin J. Guyer, and David G. Gordon, *The 2×2 Game* (Ann Arbor, MI: The University of Michigan Press, 1976).

John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University

Press, 1971).

Adam Smith, *The Theory of Moral Sentiments*, (London, UK: A. Millar, 1759).

John B. Van Huyck, Raymond C. Battalio, and Frederick W. Rankin, "On the Origin of Convention: Evidence from Coordination Games," *The Economic Journal* 107(442) May 1997, 576-97.

J. Van Huyck, R. Battalio, S. Mathur, P. Van Huyck, and A. Ortmann, "On the Origin of Convention: Evidence from Symmetric Bargaining Games," *International Journal of Game Theory* 24(2) 1995, 187-212.

Appendix A: The 70 Similar Bargaining Games (bimatrix form)

```
#matrix  1 U
 46  26    9  10
 11  19   31  35
#matrix  2 R
 18  20   50  26
 32  33    8   5
#matrix  3 U
  6   3   44  21
 26  34   15  13
#matrix  4 R
 44  23   18  16
  1   3   30  32
#matrix  5 U
 29  38   14  18
  3   7   50  23
#matrix  6 R
  8   4   27  34
 50  21   15  17
#matrix  7 U
 49  21    2   1
 20  19   26  30
#matrix  8 R
 47  22   13  20
  3   7   28  31
#matrix  9 U
 15  18   28  30
 47  23   10   3
#matrix 10 R
 30  37   10   1
 13  14   49  25
#matrix 11 U
 10   3   50  27
 32  39   14  18
#matrix 12 R
 17  18   47  27
 32  35    4   8
#matrix 13 U
 43  24    9   4
 14  15   29  31
#matrix 14 R
 50  27   19  19
  9   5   34  35
#matrix 15 U
  8   3   50  22
 28  34   16  17
#matrix 16 R
 11  14   49  22
 28  38    2   8
#matrix 17 U
  2   4   46  24
 29  36   11  16
#matrix 18 R
 28  32    3   3
 17  19   47  23
#matrix 19 U
 28  32   12  15
  7   5   48  23
#matrix 20 R
  9   4   32  33
 45  25   16  19
#matrix 21 U
 20  20   30  35
 49  25    5   8
#matrix 22 R
  4   6   31  33
 44  25   15  20
#matrix 23 U
 26  35   12  14
  3   6   45  21
#matrix 24 R
 49  26   14  11
 10   2   32  35
#matrix 25 U
 27  33   16  14
  5   3   47  21
#matrix 26 R
 46  23   15  19
  3   1   29  33
#matrix 27 U
  8   9   50  24
 30  36   12  15
#matrix 28 R
 18  19   45  21
 27  30    4   1
#matrix 29 U
 18  19   29  30
 46  22    4   5
#matrix 30 R
 50  23   16  20
  4   4   32  35
#matrix 31 U
 19  16   33  35
 50  26    5   5
#matrix 32 R
 28  31    7   6
 15  15   41  23
#matrix 33 U
 29  31   16  20
  3   3   47  23
#matrix 34 R
 50  27   11  20
  8   5   33  34
#matrix 35 U
  4   5   49  23
 31  36   15  16
#matrix 36 R
 26  29   10   3
 12  13   42  21
#matrix 37 U
 31  38   13  13
  1   1   49  25
#matrix 38 R
 27  30   10   1
 12  18   46  22
#matrix 39 U
 13  20   30  32
 47  21    9   9
#matrix 40 R
 44  23   13  18
  9   6   29  33
#matrix 41 U
 14  19   26  28
 43  21    3   2
#matrix 42 R
 32  37    3   2
 14  13   50  24
#matrix 43 U
 28  33   13  20
  9   9   44  23
#matrix 44 R
 29  31    3   6
 17  17   42  24
#matrix 45 U
 41  24    8   2
 15  15   29  30
#matrix 46 R
 12  13   48  25
 30  35    8   2
#matrix 47 U
 48  21    4   3
 11  14   28  29
#matrix 48 R
  8   8   31  33
 45  24   12  16
#matrix 49 U
 30  33   16  18
  8   8   47  23
#matrix 50 R
 10   8   28  29
 42  21   13  14
#matrix 51 U
  3   1   45  21
 27  28   11  18
#matrix 52 R
  3   9   27  31
 44  22   11  17
#matrix 53 U
 10   8   49  22
 32  34   17  15
#matrix 54 R
  3   8   27  28
 45  22   18  18
#matrix 55 U
 26  37   15  12
  8   4   50  21
#matrix 56 R
 47  27   11  20
  3   5   33  34
#matrix 57 U
 44  22    1   2
 12  11   29  31
#matrix 58 R
 10   2   29  32
 46  22   14  19
#matrix 59 U
 32  36   17  18
  8   5   49  27
#matrix 60 R
 50  28   16  18
  6   6   34  35
#matrix 61 U
 10  10   50  26
 32  34   15  12
#matrix 62 R
  2   5   27  33
 50  21   12  14
#matrix 63 U
  1  10   45  25
 31  32   18  20
#matrix 64 R
 45  24   17  19
  2   3   29  33
#matrix 65 U
 10   1   48  28
 33  35   11  12
#matrix 66 R
 15  16   49  21
 27  35    5   5
#matrix 67 U
 11  12   33  35
 47  28   10   5
#matrix 68 R
  4   2   31  32
 48  26   16  20
#matrix 69 U
  3   1   47  22
 28  34   16  14
#matrix 70 R
 28  37    6   2
 17  15   49  23
```

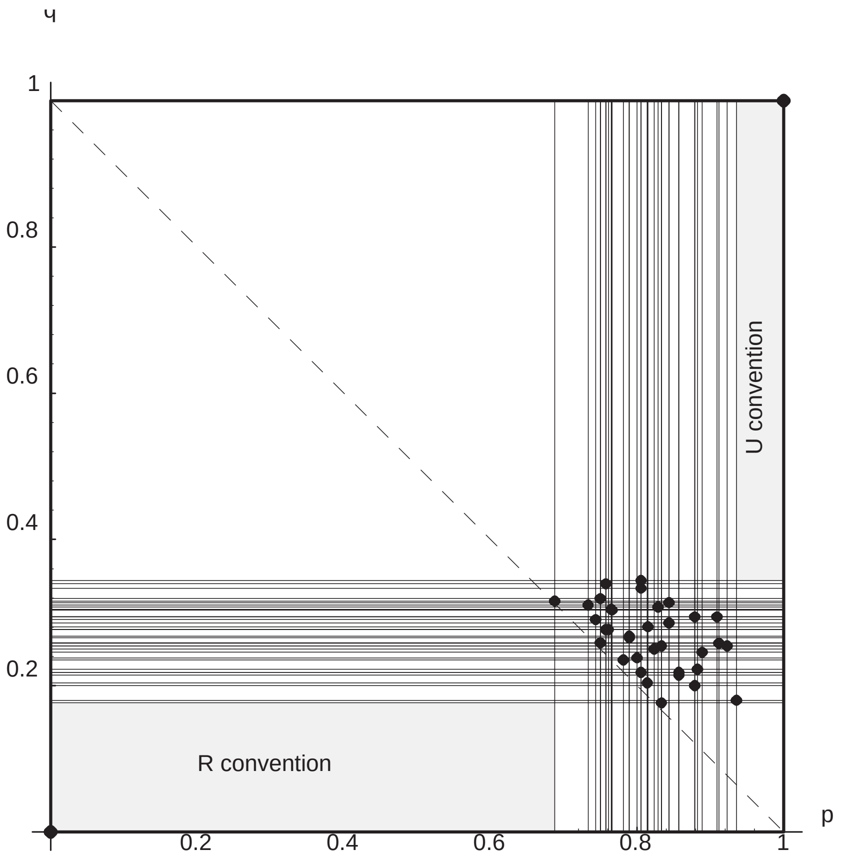

Figure 2: Mixed Equilibrium for 35 *R*-games.

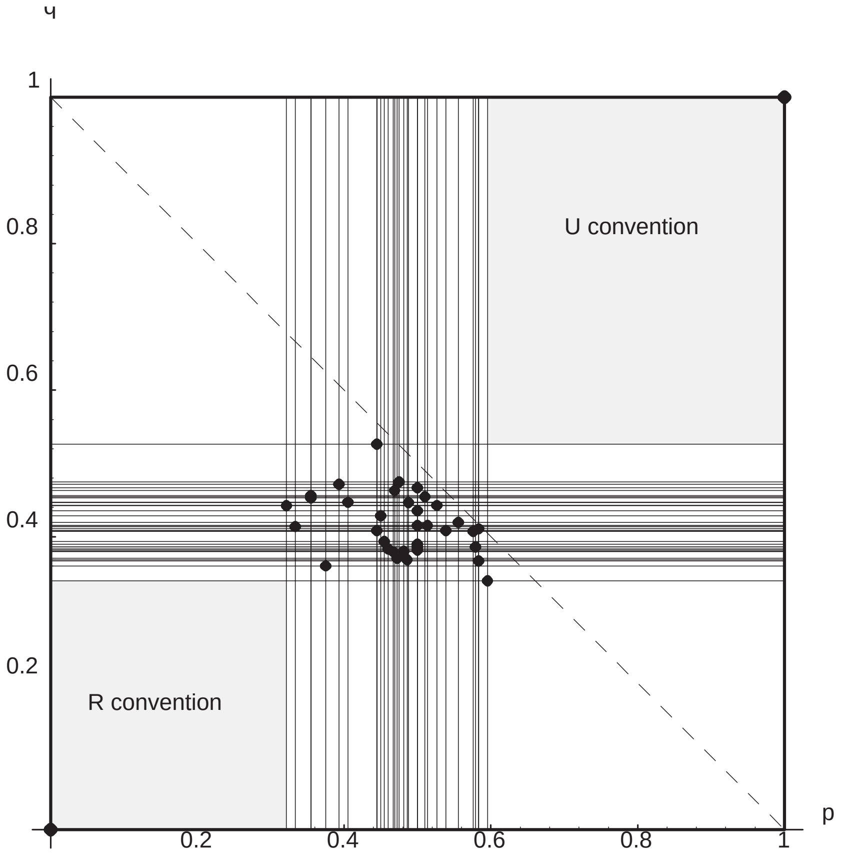

Figure 3: Mixed Equilibrium for 35 *U*-games.

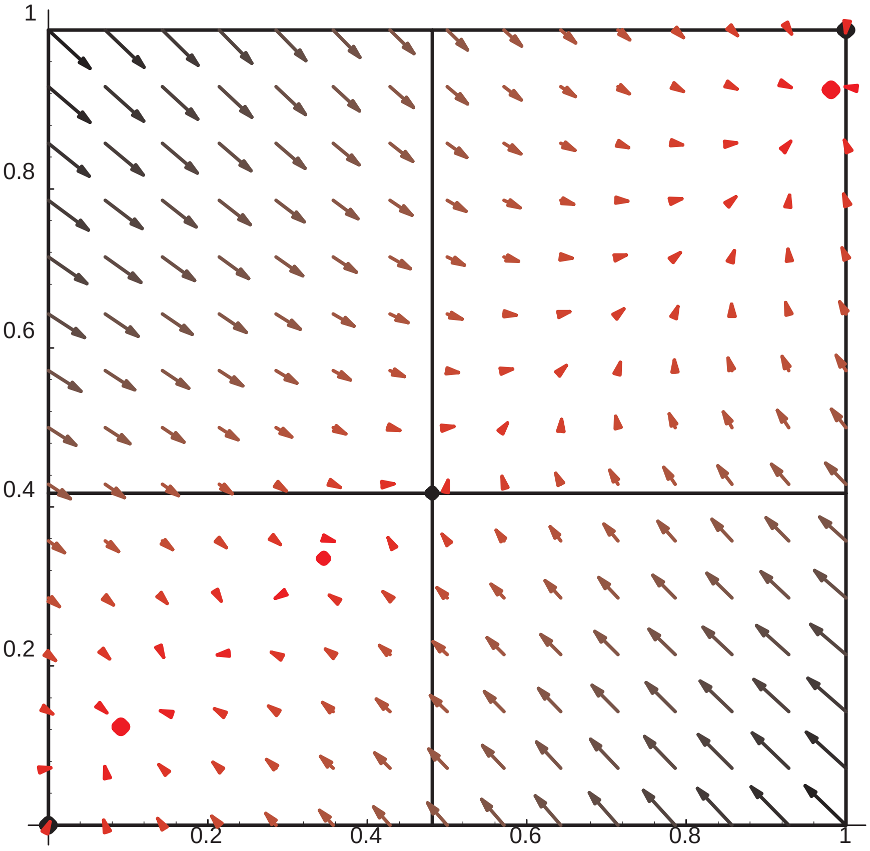

Figure 4:Logit Response Vector Field for average *U* matrix. Black dots denote Nash equilibria. Red dots denote Logit equilibria when λ = 0.14.

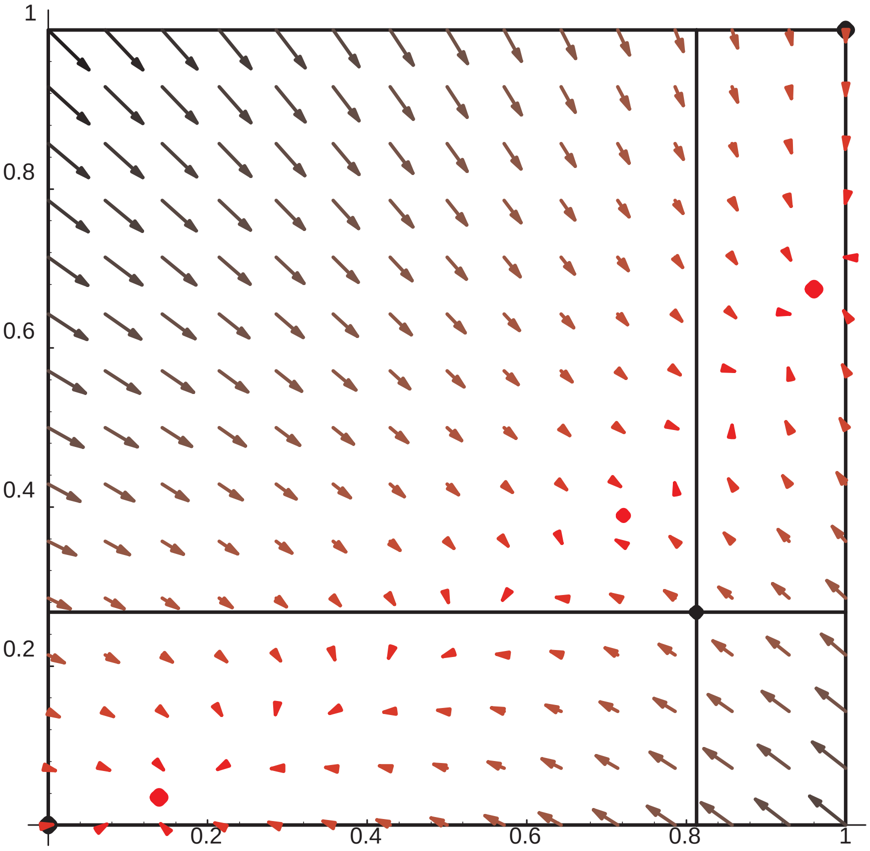

Figure 5: Logit Response Vector Field for average *R* matrix. λ=0.14.

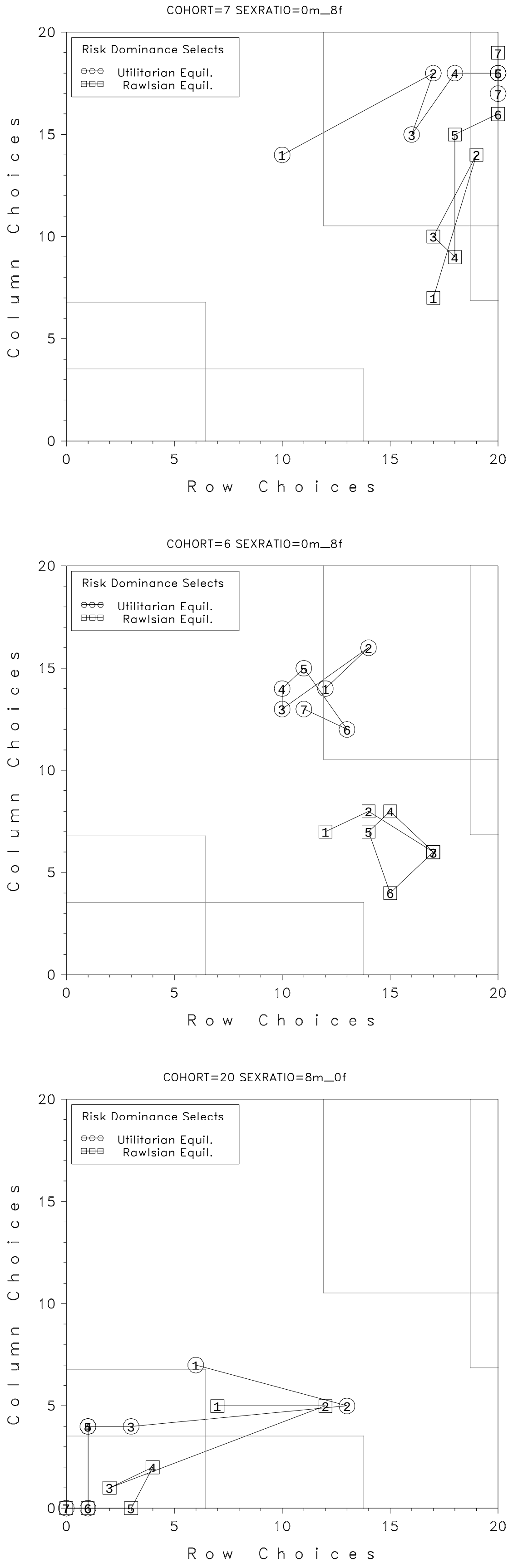

Figure 6: Three examples of observed adaptive behavior.

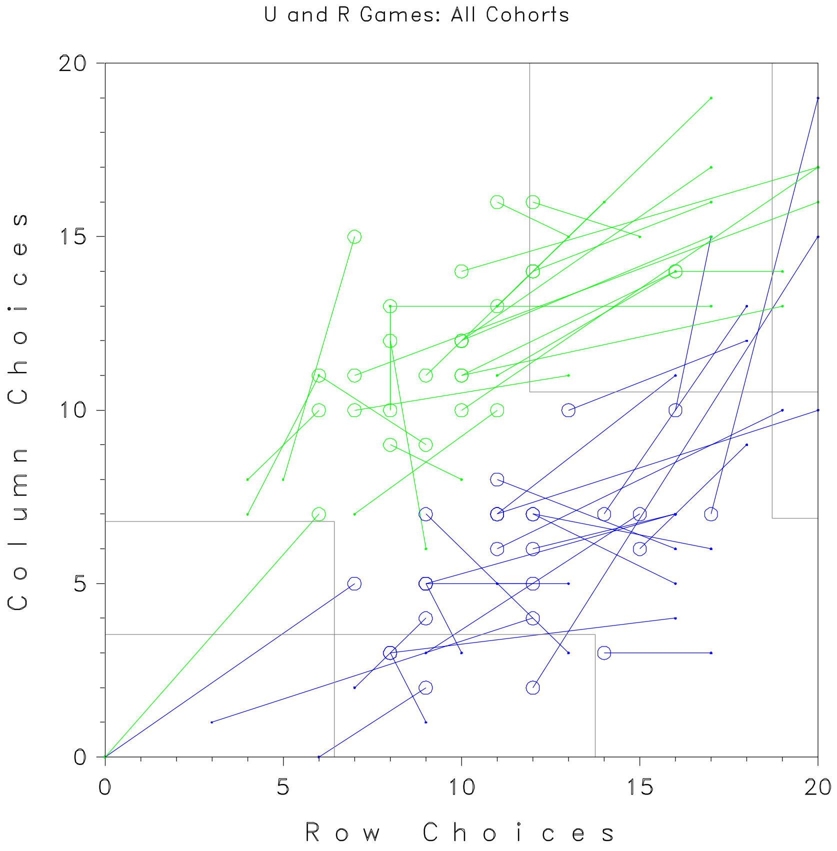

Figure 7: Empirical vector field for all 26 cohorts: Vectors are from first to last five period state: Green (grey) is for *U*-games and blue (black) is for *R*-games. Specifically, green open circles denote counts from odd periods in the interval 1 to 10 and green dots denote counts from odd periods in the interval 61 to 70. Blue (black) open circles denote counts from even periods in the interval 1 to 10 and blue dots denote counts from even periods in the interval 61 to 70. The axes run from 0 to 20 and count the frequency of the *U*-action by the four players assigned to a role in the current period, either row or column, for five periods.

Table 1: Gender Differences in Initial Behavior: Single Sex Cohorts Only

	U-action	*R-action*
U-action	26 46	10 9
R-action	19 11	35 31

Period 1 Game Form: a *U*-game

	R-action	*U-action*
U-action	20 18	26 50
R-action	33 32	5 8

Period 2 Game Form: an *R*-game

	Row Role						Column Role						Total	
	Female		Male		Total		Female		Male		Total			
Game	U	R	U	R	U	R	U	R	U	R	U	R	U	R
U-action	20	22	19	22	39	44	19	11	16	8	35	19	74	63
R-action	20	18	21	18	41	36	21	29	24	32	45	61	86	97
Fisher's p-value	0.82		0.65		0.53		0.10		0.09		0.01		0.26	
Total	40		40		80		40		40		80		160	

Table 2: Classification of Terminal State into Emerged Conventions

Cohort	Sex Ratio	*U* games Terminal State	*U* games Emerged Convention	*R* games Terminal State	*R* games Emerged Convention
1	all female	{6,11}	none	{13,3}	*R* convention
2	all female	{9,6}	none	{3,1}	*R* convention
3	all female	{17,16}	*U* convention	{16,7}	none
4	all female	{11,11}	none	{11,5}	none
5	all female	{5,8}	none	{7,2}	*R* convention
6	all female	{11,13}	none	{17,6}	none
7	all female	{20,17}	*U* convention	{20,19}	*U* convention
8	all female	{13,11}	*U* convention	{16,6}	none
9	all female	{7,7}	none	{9,3}	*R* convention
10	all female	{15,15}	*U* convention	{16,7}	none
11	all male	{20,17}	*U* convention	{20,15}	*U* convention
12	all male	{17,13}	*U* convention	{18,13}	none
13	all male	{16,14}	*U* convention	{13,5}	none
14	all male	{14,16}	*U* convention	{17,3}	none
15	all male	{4,8}	none	{6,0}	*R* convention
16	all male	{20,16}	*U* convention	{17,15}	none
17	all male	{19,13}	*U* convention	{19,10}	*U* convention
18	all male	{4,7}	none	{9,1}	*R* convention
19	all male	{8,13}	none	{10,3}	*R* convention
20	all male	{0,0}	*R* convention	{0,0}	*R* convention
21	4 each	{17,15}	*U* convention	{20,10}	*U* convention
22	4 each	{19,14}	*U* convention	{18,12}	none
23	4 each	{17,17}	*U* convention	{16,11}	none
24	6m 2f	{17,19}	*U* convention	{18,9}	none
25	8m 0f	{13,15}	*U* convention	{16,5}	none
26	5m 3f	{10,8}	none	{16,4}	none

Table 3: Emerged Conventions by Cohort Gender Composition.

	U games			R games		
	U convention	R convention	None	U convention	R convention	None
All female	4	0	6	1	4	5
All male	6	1	3	2	4	4
Mixed	5	0	1	1	0	5
	15	1	10	4	8	14

Fisher's Exact Test prob value 0.24　　Fisher's Exact Test prob. value 0.39

Table 4: Single Sex Cohort Ranks of Mean Earnings Per Capita

Gender Composition	Mean Earnings	Percent of *U*-equil.	Earnings Range	Standard Deviation
Female	$16.71	0.68	$9.24	$1.67
Male	$16.74	0.68	$7.36	$1.69
Mixed	$17.03	0.69	$7.65	$1.44
Theory				
Maximin	$15.11	0.61	$0.90	$0.16
R-Equilibrium	$21.95	0.89	$1.34	$0.19
U-Equilibrium	$24.66	1.00	$6.75	$1.02