2000-02R

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

COGNITION AND BEHAVIOR IN NORMAL-FORM GAMES: AN EXPERIMENTAL STUDY

BY

MIGUEL COSTA-GOMES

VINCENT P. CRAWFORD

AND

BRUNO BROSETA

DISCUSSION PAPER 2000-02R
JULY 2000

COGNITION AND BEHAVIOR IN NORMAL-FORM GAMES: AN EXPERIMENTAL STUDY[1]

Miguel Costa-Gomes (Harvard Business School and University of York)

Vincent P. Crawford (University of California, San Diego)

Bruno Broseta (University of Arizona and Organismo Publico Valenciano de Investigacion)

September 29, 1998; latest revision July 20, 2000

"Human experience, which is constantly contradicting theory, is the great test of truth."

—Dr. Johnson, quoted in James Boswell, *The Life of Samuel Johnson L.L.D.*

This paper reports experiments designed to study *strategic sophistication*, the extent to which behavior in games reflects attempts to predict others' decisions, taking their incentives into account. We studied subjects' initial responses to normal-form games with various patterns of iterated dominance and unique pure-strategy equilibria without dominance, using a computer interface that allowed them to search for hidden payoff information, while recording their searches. Monitoring subjects' information searches along with their decisions allows us to better understand how their decisions are determined, and subjects' deviations from the search patterns suggested by equilibrium analysis help to predict their deviations from equilibrium decisions.

Keywords: noncooperative games, experimental economics, strategic sophistication, cognition (*JEL* C72, C92, C51)

[1]We thank Andrew Chesher, Aaron Cicourel, John Conlisk, Graham Elliott, Daniel Friedman, David Grether, Frank Hahn, Eric Johnson, David Laibson, David Levine, José Machado, Amnon Rapoport, Stanley Reynolds, Alvin Roth, Larry Samuelson, Reinhard Selten, Jason Shachat, Joel Sobel, Dale Stahl, the editor, three referees, and especially Colin Camerer, Glenn Ellison, and Mark Machina for helpful advice; Mary Francis Luce for providing software; and Bill Janss and Dirk Tischer for research assistance. We are grateful for research support from the U.S. National Science Foundation (Costa-Gomes, Crawford, and Broseta), the University of California, San Diego, and the Russell Sage Foundation (Costa-Gomes and Crawford), the Alfred P. Sloan Foundation, the Banco de Portugal, Fundação Luso-Americana para o Desenvolvimento, and Fundação para a Ciência e Tecnologia (Costa-Gomes), the John Simon Guggenheim Memorial Foundation (Crawford), and the University of Arizona (Broseta). We also thank the Economic Science Laboratory at the University of Arizona, which made its facilities available for our experiments.

1. Introduction

Many unresolved questions about strategic behavior concern the extent to which it reflects players' analyses of their environment as a game, taking its structure and other players' incentives into account. This notion, which we call *strategic sophistication*, is the main difference between the behavioral assumptions of traditional noncooperative and cooperative game theory, which take it to be unlimited, and evolutionary game theory and adaptive learning models, which take it to be nonexistent or severely limited. That these leading theories rest on such different assumptions about sophistication highlights the need for more empirical work.[2]

Experiments have two important advantages in studying sophistication. They allow the control needed to test theories of behavior in games, which are highly sensitive to environmental details.[3] And, while sophistication and other aspects of cognition are normally studied indirectly, by inference from the models that best describe decisions, experiments make it possible to study sophistication more directly, by monitoring subjects' searches for hidden payoff information. This provides an additional lens through which to examine their strategic thinking.

This paper reports experiments designed to assess the sophistication of subjects' decisions and information searches in a series of 18 normal-form games with various patterns of iterated dominance and unique pure-strategy equilibria without dominance. To justify an analysis of subjects' behavior as their initial responses to each game, the design suppresses learning and repeated-game effects as much as possible. Our goals are to use subjects' information searches, in the light of the cognitive implications of alternative theories of behavior, to better understand how their decisions are determined; and to learn whether subjects' deviations from the search implications of equilibrium analysis help to predict their deviations from equilibrium decisions.

The paper makes two main contributions. It creates an experimental design to monitor subjects'

---

[2]The importance of sophistication is often downplayed because players can usually avoid the need to model others' decisions by observing their past decisions, and even unsophisticated learning models often converge to Nash equilibrium. Even so, sophistication is likely to exert important influences on convergence, limiting outcomes, and responses to changes in the environment through players' initial beliefs and the structures of their learning rules.

[3]There is a growing experimental literature that studies the principles that govern strategic behavior, surveyed in Kagel and Roth (1995) and Crawford (1997); see also Beard and Beil (1994), Brandts and Holt (1995), Cachon and Camerer (1996), Cooper, DeJong, Forsythe, and Ross (1990), Ho and Weigelt (1996), Ho, Camerer, and Weigelt (1998), Holt (1999), McKelvey and Palfrey (1992), Nagel (1995), Schotter, Weigelt, and Wilson (1994), Stahl and Wilson (1994, 1995), Van Huyck, Battalio, and Beil (1990), and Van Huyck, Wildenthal, and Battalio (2000).

searches for hidden payoffs in normal-form games, and it develops a unified theoretical and econometric framework for analyzing subjects' decisions and information searches.

Our theoretical and econometric framework is organized around a mixture model, in which each subject's behavior is determined, possibly with error, by one of nine decision rules or *types*, and each subject's type is drawn from a common prior distribution and remains constant over the 18 games he plays.[4] The possible types are required to be general principles of decision-making, applicable to a wide range of games. They are specified a priori, selected for their appropriateness as possible descriptions of behavior, theoretical interest, and separation of implications for decisions and information search. This structural approach to characterizing heterogeneous behavior in populations builds on the analyses of heterogeneous strategic behavior of Holt (1999) (which first appeared in 1990), Stahl and Wilson (1994, 1995) (which appeared about 1993; henceforth "S&W"), and Harless and Camerer (1995). Our model of decisions is closest to S&W's, who studied similar games, and some of our types are close relatives of theirs. Our error structure is closest to the models of heterogeneous individual decisions of Harless and Camerer (1994) and El-Gamal and Grether (1995) (henceforth "H&C" and "EG&G").

Four of our types are nonstrategic, in that they make no attempt to use their partners' incentives to predict their decisions: *Altruistic* seeks to maximize the sum of own and partner's payoffs over all decision combinations.[5] *Pessimistic* makes unrandomized "secure" or "maximin" decisions that maximize its minimum payoff over its partner's decisions. *Naïve* **(**S&W's *L1*, for *Level 1*) best responds to beliefs that assign equal probabilities to partner's decisions.[6] *Optimistic* makes "maximax" decisions that maximize its maximum payoff over its partner's decisions.

Five of our types are strategic: *L2* (a relative of S&W's *L2*) best responds to *Naïve*. *D1* (for *Dominance 1*) does one round of deleting decisions dominated by pure decisions and best responds to a uniform prior over its partner's remaining decisions. *D2* does two rounds of deleting decisions dominated by pure decisions and best responds to a uniform prior over its partner's remaining decisions.

---

[4]Some such structure is necessary for tractability, because in our 18 games there are more than 6 million possible individual decision histories and many more possible information search histories.

[5]*Altruistic* is nonstrategic, even though it takes its partner's payoffs into account, because it implicitly assumes that other subjects are also *Altruistic*, rather than trying to predict their decisions.

[6]Although we describe *Naïve* as nonstrategic, it might also reflect strategic decision-making with diffuse beliefs.

Equilibrium (a relative of S&W's *Naïve Nash*) makes equilibrium decisions (unique in our games). *Sophisticated* (S&W's (1995) *Perfect Foresight*, a relative of their *Rational Expectations*) best responds to the probability distribution of its partner's decision, operationalized by estimating the distribution, game by game, from the observed population frequencies in our experiment. All five strategic types exhibit some strategic sophistication. *Sophisticated* represents the ideal of a game theorist who also understands people, and so can predict how others will play in games with different structures, in which they may deviate from equilibrium. We include *Sophisticated* to learn whether any subjects have a prior understanding of other subjects' likely behavior that transcends simple, mechanical strategic decision rules.

Our approach to monitoring information search builds on the work of Camerer, Johnson, Rymon, and Sen (1993) and Johnson, Camerer, Sen, and Rymon (1996) (henceforth "C&J"). C&J presented two-person, three-period alternating-offers bargaining games to subjects in extensive form, using a computer interface called MouseLab.[7] The structure of the environment was publicly announced, except that the interface concealed the sizes of the "pies" subjects could share in the three periods, allowing subjects to look them up as often as desired, one at a time, and automatically recording their look-up sequences and decisions.[8] Our design adapts C&J's methods to monitor subjects' searches for hidden payoffs in two-person normal-form games, using MouseLab to present them as payoff tables in which subjects can look up their own and their partners' payoffs for each decision combination, as often as desired, one at a time.

Our analysis of information search also builds on C&J's work. With complete information their games have unique subgame-perfect equilibria, easily computed by backward induction. They argued that backward induction has a characteristic search pattern, in which subjects first look up the last-period pie size, then the second-last (perhaps re-checking the last), and so on, with most transitions

[7]MouseLab was developed to study individual decisions; see Payne, Bettman, and Johnson (1993, Appendix) and the user manual and software available at http://www-marketing.wharton.upenn.edu/~mouselab/. It can be viewed as an automated way of doing "eye-movement" studies like those used in experimental psychology. The first application to games other than C&J was Algaze [Croson] (1990), who briefly discussed the results of two trials using a normal-form MouseLab design similar to the one we later (independently) developed for this study.

[8]Subjects could not record the pie sizes and the frequency of repeated look-ups suggests they did not memorize them.

from later to earlier periods.[9] C&J observed systematic deviations from subgame-perfect equilibrium decisions, as in earlier studies. They added a cognitive dimension to the analysis by showing that subjects whose searches were closer to the backward-induction pattern tended to make and/or accept offers closer to the subgame-perfect equilibrium.

Our analysis, like C&J's, must model the relationship between information search and cognition. Specifying such a model is harder for normal-form games, which have a larger space of possible search patterns and a much richer set of plausible decision rules. Our specification is based on a procedural model of decision-making, in which a subject's type first determines his information search, possibly with error, and his type and search then jointly determine his decision, again with error. Each of our types is naturally associated with one or more algorithms that describe how to process payoff information into decisions. Using these algorithms as models of subjects' cognitive processes and invoking two conservative hypotheses about how cognition is related to search imposes enough structure on the space of possible look-up sequences to allow a tractable characterization of each type's search implications. This allows us to describe our subjects' noisy and heterogeneous information searches in a comprehensible way, without overfitting or excessively constraining the econometric analysis, and links subjects' searches to their decisions so that the analysis can identify relationships between them.

Our results can be summarized as follows. Subjects' decisions are too heterogeneous to be adequately described by any single decision rule, even allowing for errors. As in previous experiments, compliance with equilibrium decisions is high in games solvable by one or two rounds of iterated dominance, but much lower in games solvable by three rounds of iterated dominance or the circular logic of equilibrium without dominance.[10] In an econometric analysis of decisions alone, the types with the largest estimated frequencies are *L2*, *Naïve/Optimistic* (whose decisions are not separated in our games), and *D1*. The total frequency of strategic types is more than 70%, but most subjects' sophistication is better described by boundedly rational strategic types like *L2* or *D1* than by *Equilibrium* or *Sophisticated*, suggesting that few had a prior understanding of others' decisions that

[9]C&J supported this claim by showing that a control group, trained in backward induction but not in information search, and rewarded for identifying their subgame-perfect equilibrium decisions, came to exhibit such a pattern.
[10]See for example Beard and Beil (1994); Ho, Camerer, and Weigelt (1998); Ho and Weigelt (1996); McKelvey and Palfrey (1992); Nagel (1995); Schotter, Weigelt, and Wilson (1994); Selten (1998); S&W; Van Huyck, Wildenthal, and Battalio (2000); and other papers surveyed in Crawford (1997, Section 4).

transcends simple rules. The most frequent types all respect simple dominance and make equilibrium decisions in our simplest games, but switch to nonequilibrium decisions in some of our more complex games; this reconciles the sharp decline in equilibrium compliance in more complex games with the high frequency of strategic types.

Subjects' information searches are even more heterogeneous than their decisions. Our econometric analysis of decisions and search generally confirms the view of subjects' behavior suggested by our analysis of their decisions alone, with some significant differences. The most frequent estimated types are *Naïve* and *L2*, each accounting for nearly half the population. For games with a given strategic structure, our estimates of the type frequencies imply a simple, systematic relationship between subjects' deviations from the search implications of equilibrium analysis and their deviations from equilibrium decisions.

The shift toward *Naïve* when information search is taken into account, which comes mainly at the expense of *D1* and *Optimistic*, reflects the fact that *Naïve*'s search implications explain more of the variation in subjects' decisions and searches than *Optimistic*'s (which are too unrestrictive to be useful in our sample) or *D1*'s (which are more restrictive than *Naïve*'s, but too weakly correlated with subjects' behavior). *D1* does poorly relative to *L2*, even though their decisions are only weakly separated, because their search implications are strongly separated, and *L2*'s explain more of the variation in subjects' searches and decisions. The strong separation of *Naïve* from *Optimistic* and *L2* from *D1* via their search implications yields a significantly different interpretation of subjects' behavior than our analysis of decisions alone.

Overall, our econometric analysis suggests a strikingly simple view of subjects' behavior, with two of our nine types, *Naïve* and *L2*, comprising 65-90% of the population and a third, *D1*, 0-20%, in each case depending on one's confidence in our model of information search.

The rest of the paper is organized as follows. Section 2 describes our design. Section 3 discusses our theoretical and econometric framework. Section 4 reports preliminary statistical tests and subjects' patterns of compliance with equilibrium; estimates the types and error rates that best describe subjects' behavior, first using decisions alone, and then combining decisions and information search; and discusses aggregate patterns in search. Section 5 is the conclusion.

2. Experimental Design

This section describes our experimental design. First we discuss the overall structure, then the use of MouseLab to present games in normal form, and finally the games to be studied.

A. Overall structure

Our experiment consisted of two sessions of a Baseline treatment, B1 on 22 April 1997 and B2 on 21 July 1997, and one session each of two control treatments, OB (for "Open Boxes") on 24 July 97 and TS (for "Trained Subjects") on 22 July 97. We first describe the Baseline treatment, and then explain how the OB and TS treatments differed.[11]

To test theories of strategic behavior, the design must clearly identify the games to which subjects are responding. This is usually accomplished by having a "large" population of subjects repeatedly play a given stage game, randomly pairing them each period to suppress repeated-game effects. The learning such designs allow greatly reduces the noisiness of subjects' responses over time, but even unsophisticated learning tends to converge to equilibrium in the stage game, making it difficult to disentangle learning from sophistication. Our design studies sophistication in its purest form by eliciting subjects' initial responses to a series of 18 different games, with different partners and no feedback to suppress learning and repeated-game effects as much as possible.[12] Varying the games also helps to prevent subjects from developing preconceptions about their strategic structures, enhances our control of their information by making it impossible to remember current payoffs from previous plays, and more precisely identifies subjects' types.

The leading theories that address the issue of sophistication abstract from *contextual* factors, such as the social setting or how the game is framed, that might influence subjects' behavior. To focus sharply on such theories, our design suppresses contextual factors by using abstract decision labels and jointly randomized orders of decisions and games, framing all subjects as row players, and avoiding face-to-face or non-anonymous interaction.

Subjects were recruited from undergraduate and graduate students at the University of Arizona, with a completely different group for each session. To reduce noise, we sought subjects in courses that

[11]Appendix B reproduces the Baseline and TS instructions. The OB instructions are straightforward modifications of the Baseline instructions, available on request. These treatments were preceded by one session each of three pilot treatments, described in Appendix C. All appendices are available as pdf files at http://weber.ucsd.edu/~vcrawfor/.
[12]Designs that elicit initial responses have been used successfully by Beard and Beil (1994), Roth (1987), and S&W.

required quantitative backgrounds, but we disqualified all subjects who revealed that they had previously studied game theory or participated in game experiments.

In our Baseline treatment, after an instruction and screening process described below, the subjects were randomly divided into subgroups of Row and Column players, as nearly equal in size as possible. They were then randomly paired to play a common series of 18 two-person normal-form games. The order of games was the same for all subjects, randomized to avoid bias except that two larger games, added after the pilots, were placed last.[13] Subjects were given no feedback while they played the games. They could proceed independently at their own paces, but (unlike in S&W) they were not allowed to change their decisions once they were confirmed.[14]

To control subjects' preferences, they were paid according to their payoffs as follows. After the session each subject returned to the lab in private, and was shown the number of points he earned in each game, given his partners' decisions. He then drew a game number from a bag and was paid according to his payoff in that game, at the rate of $0.40 per point.[15] With game payoffs of 12-98 points, this made the average payment about $21; with an additional $5 for showing up and passing the test, average earnings were approximately $15 per hour.[16] Subjects never interacted directly, and their identities were kept confidential.[17]

The structure of the environment, except the payoffs, was made public knowledge by presenting instructions in handouts and on subjects' computer screens and announcing that all received the same instructions. The instructions avoided suggesting decision-making principles or decisions. During the

---

[13]Some of the pairings were repeated once in 18 games, in a game unknown to the subjects. The larger games were placed last to preserve comparability with the pilots, and because we feared (incorrectly) that they would confuse subjects. Row and Column subjects faced different orders of strategic structures, because the games had asymmetric structures. This asymmetry also avoids spurious correlation between their decisions, or between a subject's best response and his partners' most frequent decision, guarding against bias in favor of *Sophisticated*.

[14]The 18 games took subjects an average of 1-2 minutes each. Adding an hour for signing up, seating, instructions, and screening yielded sessions of 1½-2 hours, which we judged to be near the limit of subjects' attention spans.

[15]It is theoretically possible to control subjects' risk preferences using Roth and Malouf's (1979) *binary lottery* procedure, in which a subject's payoff determines his probability of winning a given monetary prize, as in Cooper et al. (1990) and S&W. We avoided this added complexity because risk preferences do not influence predictions based on iterated dominance or pure-strategy equilibrium, subjects often appear approximately risk-neutral for payoffs like ours, and results using direct payment are usually close to those using the binary lottery procedure.

[16]The analogous average earnings figures were $23 and $16 for OB subjects and $27 and $21 for TS subjects, who were paid an extra $4 for correctly answering the questions in an additional test, described in Appendix B.

[17]After all subjects had checked in, each picked an identification number from a basket. They were then told to seat themselves at the terminal in the lab with that number. After the session, a subject only needed to show his identification number to be paid. This made it clear (and they were told) that we would never know their identities.

session, subjects had unrestricted access to the payoffs via MouseLab. After reading the instructions, subjects were given an opportunity to ask questions and then required to pass a test of understanding. They were paid an additional $5 for showing up on time, and subjects who failed the test were dismissed.[18] The remaining subjects participated in four unpaid practice rounds before the main part of the session, in which they faced a balanced mix of strategic structures; at the end of the practice rounds they were told their partners' decision frequencies for two of the rounds.[19]

The OB treatment was identical to the Baseline treatment except that the games were presented via MouseLab with all payoffs continually visible, in "open boxes." Its purpose was to reveal whether subjects' responses are affected by the need to look up their payoffs.

The TS treatment was identical to the Baseline treatment (with "closed boxes") except that TS subjects were trained and rewarded differently. TS subjects were taught the relevant parts of noncooperative game theory via instructions on their computer screens, including dominance, iterated dominance, dominance-solvability, and pure-strategy equilibrium. (TS subjects, like Baseline subjects, received training in the mechanics of looking up their payoffs, but neither were trained in information search strategies.) TS subjects were rewarded for only correctly identifying their equilibrium decisions in the 18 games, independent of other subjects' responses.[20] The TS treatment was intended to provide a check on the extent to which deviations from equilibrium in the Baseline are due to subjects' cognitive limitations, and to reveal what *Equilibrium* Baseline subjects' information searches would be like, as a check on our model of information search.[21]

---

[18]The dismissal rates were 25%, 16%, and 53% for the Baseline, OB, and TS treatments respectively. We attribute the high failure rate among TS subjects to the difficulty of their task, comparable to learning 2-3 weeks of material in an undergraduate game theory course in about 45 minutes of programmed instruction.

[19]Section 4.A's statistical analysis suggests that the resulting variation in feedback from practice rounds across runs and treatments had a negligible effect on subjects' decisions in the actual games.

[20]For TS subjects the practice rounds were replaced by a test of understanding of dominance, iterated dominance, and equilibrium, with no feedback because other subjects' responses were irrelevant to their tasks. All TS subjects were made Row players because only 15 of 32 recruits passed the Understanding Test. This difference is inessential because TS subjects are paid for correct answers, not game payoffs, and the mix of strategic structures is similar for Row and Column players. So that TS subjects' searches would resemble those of Baseline *Equilibrium* subjects (who were not told that equilibrium was unique, or anything else about the structures of the games), TS subjects were encouraged to try to identify all equilibria by telling them that games can have multiple equilibria, and that to receive credit in a game they had to identify their equilibrium decision that gave them the highest payoff of any equilibrium.

[21]It would be interesting to run additional treatments, one with TS training but Baseline motivation, and one with no training but TS motivation, but we felt that the expected benefits of such treatments were lower than their costs.

B. Using MouseLab to present normal-form games

MouseLab was used to present games as follows. Each subject was framed as a row player and called "You" in all games, without regard to whether he was a Row or Column player as described here. A typical game appeared on a subject's screen as in Figure I, in which each player has two decisions, labeled # and * for the subject and & and @ for his partner. The subject's payoffs ("Your Points") are in the two left-most columns, and his partner's ("Her/His Points") are in the two right-most columns. The subject has opened the box that gives his payoff (42) when he plays # and his partner plays @.[22] The separation of the subject's and his partner's payoffs, emphasized by the legends at the bottom of the display, helps to distinguish them.[23]

In the Baseline and TS treatments, a subject could look up the payoffs as often as desired, one at a time, by using his mouse to move the cursor into its box and left-clicking. Before he could open another box or enter his decision, he had to close the box by right-clicking. Thus both opening and closing a box required conscious choice.[24] Subjects were not allowed to record the payoffs. A subject could enter and confirm his decision in the current game by moving the cursor into one of the boxes at the bottom of the display and left-clicking. A subject could move on to the next game only after confirming his decision; the cursor then returned to the top-center. MouseLab automatically records subjects' decisions, look-up sequences, and look-up durations. The OB treatment used MouseLab in exactly the same way as the Baseline, but with all payoffs continually visible, so subjects used the mouse only to enter and confirm their decisions.

Our display made subjects' information processing simpler than in CJ's games, by revealing payoffs directly rather than requiring subjects to deduce them from pie sizes, but it also made subjects'

---

[22]Different labels were always used for Row and Column subjects' decisions. The labeling and order of decisions were the same for all subjects (transposed for Column subjects), and were separately randomized for each game.

[23]By forcing subjects to look up their own and their partners' payoffs separately, our design makes it transparent that subjects are interacting with partners who have different goals and make their decisions independently. In S&W's design, all games are symmetric and the display makes no distinction between own and partner's payoffs.

[24]A box could be closed even after the cursor had been moved out of it. In the version of MouseLab C&J used, a box opens or closes, without clicking, when the cursor enters or leaves it. In preliminary trials with this version, subjects often rolled the mouse quickly across intervening boxes, which took longer than 0.017 seconds, the minimal duration MouseLab records. These "accidental" look-ups add a great deal of noise, and this version also yielded very large numbers of look-ups (100 or more in a 2x2 game), which decreases the discriminatory power of look-up patterns. Following C&J, the noise could be reduced by filtering out look-ups shorter than subjects' minimum perception time of approximately 0.18 seconds (Card, Moran, and Newell (1983)). However, the "click" version of MouseLab we used, which was supplied to us by Mary Francis Luce, solves both of these problems.

information searches more complex, with 8-16 payoffs that varied independently in three dimensions (up-down, left-right, and own-other's). We could have simplified subjects' searches by leaving some payoffs continually visible or publicly announcing a simple payoff structure (e.g. pure coordination or zero-sum), but this would have thrown away information about their cognitive processes, reducing our ability to discriminate among alternative theories.[25]

C. Games

Figure II displays the 18 games Baseline, OB, and TS subjects played as traditional payoff matrices, without decision labels, in an order that highlights their structural relationships.[26] Those relationships were disguised by small payoff shifts and a common, random ordering of all games but 9A and 9B: 3A, 6B, 2A, 8B, 8A, 5A, 4A, 7A, 4C, 7B, 4B, 3B, 2B, 6A, 5B, 4D, 9A, 9B. There are several pairs of *isomorphic* games, identical for Row and Column players except for transposition of player roles and small, uniform payoffs shifts: 3A, 3B and 2A, 2B; 4A, 4B and 4C, 4D; 5A, 5B and 6A, 6B; and 7A, 7B and 8A, 8B. Theories that abstract from context predict the same decisions in isomorphic games, and our design controls for all such effects but the order and labeling of decisions and small and non-salient payoff shifts. We find only insignificant differences in behavior across isomorphic games, which sometimes allows us to pool the data.

The games were chosen to separate the decisions of strategic and nonstrategic types as much as possible, given the need to provide adequate incentives and to vary strategic structures. Because the information search implications of strategic and nonstrategic types are sharply separated, this separation gives us the best chance to detect sophistication and facilitates our search for relationships between decisions and search.[27] We also sought to "stress-test" equilibrium predictions by

[25]Most theories of strategic decision-making implicitly require players to make sequences of binary payoff comparisons. As explained below, there is good reason to expect subjects to make most such comparisons via adjacent look-ups, but even so their look-up sequences necessarily include many adjacent look-ups that are not comparisons. One can imagine software that requires subjects to observe payoffs in pairs, which could be interpreted as binary comparisons; but such restrictions might not be desirable because deciding which payoff pairs to look at is a nontrivial task, which might distract subjects from thinking about the game. We address this issue below by deriving restrictions on information search that do not depend on precise identification of comparisons.

[26]Figure II orders subjects' decisions they way they saw them (transposed for Column subjects). Figure II also indicates types' predicted decisions by the types' initial letters along the margins of each payoff matrix. After the game number, it indicates the numbers of rounds of iterated pure-strategy dominance a (Row, Column) player needs to identify his own equilibrium decision, with 8 for non-dominance-solvable games.

[27]It also limits the separation that can be achieved between different strategic or nonstrategic types. *Naïve* and *Optimistic* decisions are not separated in any of our games, because this conflicted with our other goals. We omit

eliminating possible alternative rationales for equilibrium decisions; dominance, for example, always occurs with overlapping payoff ranges, so that subjects can reliably identify it only by looking up all payoffs involved. Finally, we avoided the artificial clarity of overly simple payoff structures and salient payoffs such as 0 and 100.

3. Theoretical and Econometric Framework

This section discusses our theoretical and econometric framework, focusing on the two-person normal-form games of complete information in our design and assuming that subjects treat them as strategically independent. Recall that our model is a mixture model, in which each subject's behavior is determined, with error, by one of nine types, and each subject's type is drawn from a common prior distribution and remains constant over the 18 games he plays. We begin by discussing our specification of possible types, and then discuss their implications for decisions and information search. We use the following standard terminology. Players' decisions

are in *equilibrium* if each player's decision maximizes his expected payoff, given the other's decision. A player's decision *dominates* (respectively, *is dominated by*) another of his decisions if it yields a strictly higher (respectively, lower) payoff for any of the other player's decisions. A player's decision that dominates all of his other decisions is called a *dominant* decision. A decision is *iteratively undominated* if it survives iterated elimination of dominated decisions. A *round* of iterated dominance is defined as eliminating all dominated decisions for both players. A game is *dominance-solvable* (in k rounds) if each player has a unique iteratively undominated decision (that can be identified in k rounds of iterated dominance).[28] The iteratively undominated decisions in a dominance-solvable game are players' unique equilibrium decisions.

---

games with mixed-strategy or multiple equilibria, because they each raise issues interesting enough for a separate investigation. Our games with unique equilibria without pure-strategy dominance are dominance-solvable via mixed-strategy dominance, a necessary feature when one player has only two pure strategies. Although both kinds of dominance play the same role in epistemic justifications for equilibrium (Aumann and Brandenburger (1995)), we expected subjects to find mixed-strategy dominance less salient, and it seemed a small price to pay for simplicity.

[28]We sometimes distinguish between the numbers of rounds in which a game is dominance-solvable and the number that each player needs to identify his equilibrium decision; the former is of course the maximum of the latter two.

A. Types

In specifying the possible types, we started with representatives of general principles of decision-making that have played important roles in the literature and are relevant to our design. We then selected a subset a priori, which includes the six types used to analyze our data in Costa-Gomes, Crawford, and Broseta (1998) ("CGC&B"), *Altruistic*, *Pessimistic*, *Naïve* (S&W's *L1*), *Optimistic*, *Equilibrium*, and *Sophisticated*, which were specified before our experiment, plus three boundedly rational strategic types (*L2*, *D1*, and *D2*), which were added at the suggestion of reviewers to refine the interpretation of the apparent predominance of *Sophisticated* behavior in CGC&B's analysis.[29] This subset was chosen for appropriateness as possible descriptions of behavior, and to be large and diverse enough to describe our subjects' decisions and information searches without overly constraining the data analysis, yet small enough to avoid overfitting.[30]

If a subject comes to the experiment with a prior understanding of others' responses, it is possible for his decisions to be determined by a rule that depends on aspects of others' behavior that we, as analysts, must estimate. If not, his decisions must be determined by an algorithm for processing payoff information into decisions that can be implemented without input from others. To allow for the former possibility, we include one type, *Sophisticated*, to represent the ideal of a rational player who can predict how others will respond to games with different structures, and we require that all other types' decisions be identifiable without input from others. *Sophisticated* is defined, game by game, as best responding to the observed population frequencies of subjects' potential partners in our experiment, which we take as the best available estimates of the probability distributions from which subjects' responses are drawn.[31] This definition best represents the ideal we are interested in detecting, cleanly

[29]All nine types are defined in the Introduction. In defining *D1* and *D2* we disallow dominance by mixed strategies because discovering it involves algebra or mental averaging, which we judged too complex. "Dominance" means pure-strategy dominance unless otherwise noted. Our model implicitly allows a uniform random type like S&W's *L0*.

[30]One could construct an ad hoc type to mimic each subject's decision history exactly, but this would have little explanatory power. Moreover, because there are many alternative rationales for any given decision history, there would be no way to derive the cognitive implications of such types, which is an essential part of our analysis.

[31]To reduce sampling error, we base our definition on the pooled Baseline and OB frequencies, which differ only slightly; this makes the definition uniform across the Baseline and OB treatments, and differs from a definition using only the Baseline data in one of 18 games for each player role. Thus, for instance, a *Sophisticated* Row player's decision in a given game is his best reponse to the pooled Baseline and OB frequencies of Column subjects' decisions in that game. If there were an empirically reliable theory of players' responses to games like ours, we could use it to make *Sophisticated* a purely theoretical construct. But subjects' responses to games like ours vary in ways that cannot yet be predicted entirely by theory (Table II). This way to operationalize *Sophisticated* is implicit in a

addressing the issues raised by prior understandings without imposing structural restrictions on how others' decisions are determined.

The requirement that all other types' decisions can be identified without input from others rules out other types whose decisions depend on estimated parameters and versions of types that are defined taking others' decision noise into account, as in S&W's strategic types, which are defined as best responses to decisions with logit errors with estimated error rates.[32] Such types' decisions are determined simultaneously with others' decisions, and are unlikely to be descriptive of subjects' initial responses to games in the absence of a prior understanding. They also blur information search implications because the estimated parameters usually depend on all payoffs.

Accordingly, we replace S&W's *L2* and *Naïve Nash* by their noiseless analogs, our "*L2*" and "*Equilibrium*."[33] We also rule out S&W's (1995) *Worldly*, which best responds to an estimated population mixture of their *L1* and *Naïve Nash*; and their *Rational Expectations*, defined by plugging in an estimated population mixture of *L1*, *L2*, *Naïve Nash*, and *Worldly* decisions and finding an equilibrium in the reduced game among *Rational Expectations* players.[34]

B. Decisions

We begin with some observations that clarify the relationships among our types' predicted decisions and information searches in our games. The proofs are straightforward and are omitted.

Observation 1: No type but *Altruistic* ever makes a dominated decision in our games.

Observation 2: If it is common knowledge that all players are *Sophisticated*, then their decisions are common knowledge, and *Sophisticated* makes the same decisions as *Equilibrium* (Aumann and Brandenburger (1995), Theorem B). In games that yield sufficiently high frequencies of equilibrium decisions, *Sophisticated* makes the same decisions as *Equilibrium*.

---

method sometimes used to test for the rationality of subjects' responses, as in Roth and Murnighan (1982).

[32]It does allow types that depend on exogenously specified noise parameters, some of which might describe behavior better than our strategic types in games where deviations from equilibrium decisions by one player have extreme payoff consequences for the other. But it is not clear how one would specify such noise parameters a priori, and making them depend on the payoffs seems inappropriate for subjects who have not looked them up. As explained in Section 4.C, our error-rate analysis deliberately avoids fine-tuning the error structure in this way.

[33]We rule out our noiseless *L3*, which best responds to our noiseless *L2*, because in our games its decisions coincide with *Equilibrium*'s and it has similar information search implications. We also rule out higher-order *Lk* and *Dk* types as insufficiently separated from *Equilibrium* (Observations 4-5, Section 3.B).

[34]For similar reasons, we rule out types that are blends of other types (as in Stahl's (1999) "evidence-based" model). Such hybrids risk overfitting, particularly with the freedom allowed by estimating the distribution of types. Because

Observation 3: *Lk*, *k*=2,3,…, makes the same decisions as *Sophisticated* if the other players are all type *Lk-1*, but *Lk* and *Sophisticated* decisions can differ if the other players are a mixture of types *L0*, *L1*,…, and *Lk-1*.

Observation 4: *Lk*, *k*=1,2,3,…, never makes a decision that does not survive *k* rounds of pure- or mixed-strategy iterated dominance, and thus makes the same decisions as *Equilibrium* in any game that is solvable in *k* rounds of pure- or mixed-strategy iterated dominance.

Observation 5: *Dk*, *k*=1,2,…, never makes a decision that does not survive *k*+1 rounds of pure-strategy (but not mixed-strategy) iterated dominance, and thus makes the same decisions as *Equilibrium* in any game that is solvable in *k*+1 rounds of pure-strategy iterated dominance.

Observation 6: *Dk*'s, *k*=2,…, decisions can differ from *Dk*-1's only in games (dominance-solvable or not) with exactly *k*+1 rounds of iterated pure-strategy dominance, and only when the player's partner has a dominated decision in the $(k+1)^{st}$ round. In games with no pure-strategy dominance, *Dk*'s decisions are thus the same as *Naïve*'s.

Each of our types predicts a unique, pure decision for each player role in each game. Figure II summarizes the relationships among types' decisions, with types identified by initial letters (*D* = dominant decision, made by all types but *Altruistic*; *A* = *Altruistic*; *P* = *Pessimistic*; *N* = *Naïve*/*Optimistic*, not separated in any of our games; *D12* = *D1* and *D2*; *E* = *Equilibrium*; and *S* = *Sophisticated*). After the game number, Figure II also gives the numbers of rounds of pure-strategy iterated dominance the (Row, Column) player needs to identify his equilibrium decision.

Our strategic types' (*D1*, *D2*, *L2*, *E*, and *S*) decisions are usually separated from our non-strategic types' (*A*, *P*, and *N*) decisions when the player has no dominant decision. Although our strategic types' decisions cannot be (or for *S*, are not) separated from one another in our 12 games that are solvable in one or two rounds of dominance, they are well separated in the other 6 games. We now describe the patterns of separation, with *Equilibrium* as the reference point.

Sophisticated makes the same decisions as *Equilibrium* in games that yield high frequencies of equilibrium decisions (Observation 2). Given our subjects' decisions (Table II), separation occurs essentially in and only in our 3x2 games that are solvable by three rounds of pure-strategy dominance,

---

their information requirements are unions of those of their component types, they too blur search implications.

or have unique equilibria but no pure-strategy dominance: 5A, 5B, 8A, and 8B for Rows; 6B, 7A, and 7B (but not 6A) for Columns.

L2 makes the same decisions as *Equilibrium* in our games that are solvable by 2 rounds of pure- or mixed-strategy dominance (Observation 4): 2A, 2B, 3A, 3B, 4A, 4B, 4C, 4D, 9A and 9B for Rows and Columns; 6A, 6B, 7A, and 7B for Rows; and 5A, 5B, 8A, and 8B for Columns. *L2* and *Equilibrium* can and do make different decisions in our 3x2 games that are solvable by three rounds of pure-strategy dominance, or have unique equilibria but no pure-strategy dominance: 5A, 5B, 8A, and 8B for Rows and 6A, 6B, 7A, and 7B for Columns. Excluding 6A for Columns, these games are exactly those in which *Equilibrium* and *Sophisticated* decisions differ, and *L2* and *Sophisticated* make the same decisions in all but 6A for Columns.[35]

Lk, k=3,…, and *Equilibrium* make the same decisions in all of our games, because none of them are *not* solvable by three rounds of mixed-strategy iterated dominance (Observation 4).

D1 makes the same decisions as *L2* and *Equilibrium* in games that are solvable by two rounds of pure-strategy dominance (Observation 5): 2A, 2B, 3A, 3B, 4A, 4B, 4C, 4D, 9A, and 9B for Rows and Columns, plus 6A, 6B for Rows and 5A, 5B for Columns. *D1* could, but does not, differ from *L2* in our games that are solvable by three rounds: 5A, 5B for Rows and 6A, 6B for Columns. Because we disallow mixed-strategy dominance in defining *D1*, it makes the same decisions as *Naïve* in our games with unique pure-strategy equilibria but no pure-strategy dominance (Observation 6): 7A, 7B, 8A, and 8B for Rows and Columns. In those games *D1* and *L2* happen to make different decisions when the subject has two decisions (7A, 7B for Rows and 8A, 8B for Columns) but not when he has three (8A, 8B for Rows and 7A, 7B for Columns).[36]

D2's decisions can be separated from *D1*'s only in our games with exactly three rounds of iterated pure-strategy dominance, and only when the player's partner has a dominated decision in the third round (Observation 6): 5A, 5B for Rows and 6A, 6B for Columns. In those games *D2* makes the same decisions as *Equilibrium* and *L3* (Observation 5). Because we disallow mixed-strategy

[35]We take the separation of *L2* from *Equilibrium* and *Sophisticated* in 6A for Columns with a grain of salt because *L2*'s decision was within 4.55 points of being optimal for *Sophisticated*. At \$0.40 per point, this implies an expected-payoff difference of \$1.82, or just over \$0.10 ex ante, given the 1/18 probability that 6A determines payment.
[36]If we allowed mixed-strategy dominance, *D1* would make the same decisions as *L2* in all but our games that require three rounds of pure- or mixed-strategy dominance (5A, 5B, 8A, 8B for Rows and 6A, 6B, 7A, 7B for Columns);

dominance in defining *D2*, this equivalence need not extend to games that are solvable by three rounds of iterated mixed-strategy dominance: 8A, 8B for Rows and 7A, 7B for Columns. *D2* and *Equilibrium* make different decisions in those games, in which *D2* sides with *L2* and *Sophisticated*. *Dk*, $k=3,\ldots$, can be separated from *Dk*-1 only in games with exactly $k+1$ rounds of iterated pure-strategy dominance (Observation 6). Because we have no games with more than three rounds, *Dk*, $k=3,\ldots$, therefore makes the same decisions as *D2* in all our games.

We close this section by using the observed frequencies of Baseline and OB subjects' decisions to estimate the strength of their incentives to make their types' decisions. Each row of Table I gives the payoff over all 18 games of the average Row or Column Baseline or OB subject of a given type, measured according to its goal (maximax or maximin payoffs for *Optimistic* or *Pessimistic*, own plus other's payoff for *Altruistic*, and expected payoff for the other types), as a function of the type hypothesized to determine the subject's decisions (Table I's columns**)**. Each type's payoffs are expressed as percentages of the subject's payoff when that type determines its decisions, normalized to 100%. The right-most column gives the associated expected monetary value of 100% of each type's points. Table I shows that the incentives to make the decisions of one rather than another strategic type are weak, but the incentives to make the decisions of a given strategic type rather than those of a given nonstrategic type, or vice versa, are strong.

C. Information search

Recall that our analysis of decisions and information search takes a procedural view of decision-making, in which, in each game, a subject's type determines his information search, with error, and his type and search then determine his decision, again with error. The link between decisions and information search depends on how cognition influences search, which is difficult to model because there is little theory to guide a specification, the space of possible look-up sequences is enormous, and our subjects' sequences are noisy and highly heterogeneous.[37] Many aspects of subjects' searches might be related to their decisions (Section 4.C suggests several possibilities), but these circumstances make it

---

and *D2* would make the same decisions as *Equilibrium* and *L3* in all our games.

[37]To our knowledge these issues have been considered only in passing, by C&J and Algaze [Croson] (1990). Related issues have been discussed in the computational complexity literature, particularly for iterated dominance, by Knuth, Papadimitriou, and Tsitsiklis (1988) and Gilboa, Kalai, and Zemel (1993), among others; but their analyses focus on

intractable to take the full richness of our search data into account. They also make it difficult to specify a parametric model of search with confidence, and suggest that any such model is more than usually likely to introduce distortions.

We therefore take a more conservative approach, which promises to be more robust. We introduce two simple hypotheses about how cognition is related to information search, suggested by the search behavior in C&J's control treatment and our TS treatment, which impose minimal restrictions to avoid arbitrarily imputing inconsistency to subjects whose cognitive processes we cannot observe. We then use the algorithms for processing payoff information into decisions associated with our types to derive their search implications under our hypotheses, showing that they are strongly separated across types. In Section 4.D's econometric analysis, we use subjects' compliance with types' search implications, along with their decisions, to estimate their types.

Identifying a type's decision requires a set, or sometimes one of several alternative sets, of operations on payoffs. For instance, depending on the structure, *Equilibrium* decisions can only be identified by checking for dominance or iterated dominance among own and/or other's decisions, checking directly for pure-strategy equilibria, or combining those methods, all of which involve only pairwise ordinal payoff comparisons.[38] With minor exceptions, identifying other types' decisions also requires only pairwise ordinal payoff comparisons.[39] We call payoffs and operations on groups of payoffs *look-ups* and *comparisons* (abusing terminology because some operations are more complex) and we call the look-ups and comparisons in some minimal set required to identify a type's decision in a game that type's *relevant* look-ups and comparisons for the game.[40] We presume that a subject's look-ups in a game are determined by his type's relevant look-ups and comparisons, and his search and memorization costs. Our hypotheses are:

---

identifying ways to compute the equilibrium of a game in a number of operations that is polynomial in its size, which yields algorithms that seem to us much too subtle to be descriptive of subjects' cognitive processes.

[38]Our designs avoid ties, multiple equilibria, and games for which plausible predictions involve mixed strategies.

[39]This is true of most but not all (e.g. risk-dominance) notions in normal-form noncooperative game theory. The exceptions are that *Sophisticated* and *Naive* may compute expected payoffs via left-right look-ups in own payoffs, which in some games requires averaging three or four payoffs, and *Altruistic* must add own and other's payoffs and compare the totals across decision combinations.

[40]Sometimes there are alternative minimal sets of look-ups and comparisons, but the feasible sets are usually nested, so that the minimal sets are unique. When there is more than one minimal set, we allow a subject to use any one of them, requiring only that it be the same for look-ups and comparisons.

Occurrence: For a given type in a given game, each look-up in some minimal set needed to identify the type's decision appears at least once in the subject's look-up sequence.

Adjacency: For a given type in a given game, Occurrence is satisfied and each comparison in some minimal set needed to identify the type's decision is represented by an adjacent look-up pair (or group, in the exceptional cases) at least once in the subject's look-up sequence.[41]

For a given type and game, Adjacency implies Occurrence by construction. Occurrence is uncontroversial because a subject who does not make all the look-ups a type requires cannot identify its decision with certainty.[42] However, it has limited discriminatory power because it is likely to be satisfied by chance even for moderately long look-up sequences. Our TS subjects satisfied *Equilibrium* Occurrence in 98% of all game-subject pairs: 97% of those in which they made the equilibrium decision and 100% of the few pairs in which they did not.

Adjacency has greater discriminatory power because it is less likely to be satisfied by chance, but it is more controversial because a subject who violates it may still recall enough non-adjacent payoffs to identify his type's decision. In the extreme case where subjects scan and memorize a game's payoffs at the start, the order of their look-ups might be completely unrelated to their cognitive processes, making Adjacency useless. But Adjacency is useful when repeated look-ups are less costly than memory, so that subjects perform comparisons one at a time, acquiring the information for each comparison by adjacent look-ups, storing the results in the simplest form that suffices for the rest of the analysis, and otherwise relying on repeated look-ups rather than memory.[43] The results of C&J's control treatment and our TS treatment suggest that our Baseline subjects complied with Adjacency most of the time.[44] TS subjects looked up most payoffs repeatedly, and 50% of their adjacent pairs corresponded

[41]Our original proposal discussed an Efficiency hypothesis, which combined Adjacency with the requirement that the look-up policy minimize the expected total number of look-ups. This hypothesis was suggested by C&J's control subjects, whose look-ups were usually in the last-period-first order that minimized the total number needed to identify their subgame-perfect equilibrium offers. Although Efficiency implies potentially useful restrictions, we omit it here because its implications for our types are subtle, and seem unlikely to be satisfied often enough to be useful.
[42]A clever subject might be able to identify it with high probability, knowing that all payoffs are between 0 and 100.
[43]Adjacency respects our inability to use MouseLab to distinguish look-up pairs that are adjacent by coincidence from those that are associated with comparisons. We interpret "simplicity" as follows: The ordinal ranking of a pair of payoffs is simpler than the numerical payoffs, and a dominance relationship between decisions or the fact that a decision combination is an equilibrium are simpler than the corresponding sets of payoff comparisons.
[44]Identifying subgame-perfect equilibrium decisions in C&J's alternating-offers bargaining games requires pairwise ordinal payoff comparisons involving simple functions of the pie sizes, and is similar to *Equilibrium*'s task in our games. C&J's control subjects usually looked up the third-period pie size first, then the second-period pie size,

to comparisons relevant for *Equilibrium*, close to the maximum given that all but the first and last look-ups belong to two adjacent pairs. TS subjects satisfied *Equilibrium* Adjacency in 89% of game-subject pairs: 94% of those in which they made equilibrium decisions and 47% of those in which they did not.

We now characterize the implications of Occurrence and Adjacency for each type and kind of game in our design, discussing all games from the point of view of the Row player.

Altruistic only needs to compare the totals of own and other's payoffs for each possible decision combination. *Altruistic* Occurrence therefore requires all own and other's look-ups, and *Altruistic* Adjacency requires comparisons (in the sense of adjacent pairs) of own and other's payoffs for each decision combination.

Naïve, *Optimistic*, and *Pessimistic* decisions all depend only on the player's own payoffs. *Naïve* needs to compare expected payoffs for its decisions for a uniform prior over other's decisions. This can be done by running expected-payoff totals, up-down column by column or left-right row by row; we allow either method, but rule out mixtures. *Naïve* Occurrence thus requires all own payoffs. *Naïve* Adjacency normally requires a complete set of either up-down or left-right comparisons of own payoffs for all other's decisions.[45]

Optimistic or *Pessimistic* only need to identify the maximax or unrandomized maximin decision. The maximax decision can be identified by scanning all own payoffs in any order, keeping a record of the highest found so far. The maximin decision can be identified only by left-right comparisons. *Optimistic* and *Pessimistic* Occurrence thus require all own (and only own) payoffs, with two exceptions: If *Optimistic*'s look-ups include all own payoffs for all but one own decision and a higher payoff for the remaining decision, *Optimistic* Occurrence requires no more look-ups for the latter decision; and if *Pessimistic*'s look-ups include all own payoffs for one own decision and an own payoff for another decision that is lower than the minimum payoff for the former decision, *Pessimistic*

---

sometimes returning to the third, and then the first-period pie size, with most transitions from later to earlier periods. Our TS treatment was much less conducive to memorization than C&J's, with 8-16 payoffs versus three pie sizes.

[45] *Naïve* may be able to avoid some comparisons by identifying dominance among own decisions; this can be done by up-down comparisons that have the same requirements as comparing expected payoffs. *Naïve* may also be able to avoid some comparisons by eliminating decisions that its comparisons show to have lower expected payoffs than another decision. In game 3A, for instance, *Naïve* Occurrence requires either the comparisons (75, 42) and (48, 89) or the comparisons (75, 48) and (42, 89). In game 5A *Naïve* Occurrence requires either the left-right comparisons (53, 24), (79, 42), and (28, 71) or one of two alternative sets of up-down comparisons: (53, 79), (24, 42), (79, 28),

Occurrence requires no more look-ups for the latter decision. *Optimistic* Adjacency is vacuous, and *Pessimistic* Adjacency requires a set of left-right comparisons sufficient to identify the maximin decision.[46]

Equilibrium can identify its predicted decision by checking directly for pure-strategy equilibrium or, in dominance-solvable games, by checking for iterated dominance. Checking directly for equilibrium can be done for each possible decision combination separately or by *best-response dynamics*, which rules out some combinations using the fact that only best responses can be part of an equilibrium. Checking each combination requires an up-down comparison of own payoffs and a left-right comparison of other's payoffs. In games that are dominance-solvable, *Equilibrium* can also use iterated dominance or a combination of iterated dominance and equilibrium-checking. For our games, the minimal set or sets of look-ups or comparisons have a simple characterization that depends only on whether the game is dominance-solvable.

In our dominance-solvable games, there is only one way to perform iterated dominance, and the sets of look-ups and comparisons it requires are always contained in the sets required for checking directly for equilibrium (by either method). In such games *Equilibrium* Occurrence (Adjacency) requires all look-ups (comparisons) needed to identify the game's iterated dominance relationships between own or other's decisions, excluding those that can be eliminated using dominance relationships identified elsewhere in the iteration.[47]

In our non-dominance-solvable games, there is never any pure-strategy dominance, and the sets of look-ups or comparisons required to check for equilibrium via best-response dynamics are always contained in the sets required to check each possible decision combination separately. In such games *Equilibrium* Occurrence (Adjacency) thus requires the look-ups (comparisons) for an up-down

---

and (42, 71) or (53, 28), (24, 71), (79, 28), and (42, 71).

[46]In game 3A *Optimistic* Occurrence requires only the look-ups 75, 42, and 89, because 89 > max {75,42}; *Pessimistic* Occurrence requires only the look-ups 48, 89, and 42, because 42 < min {48,89}; and *Pessimistic* Adjacency requires only the comparison (48,89), again because 42 < min {48,89}.

[47]In Game 3A *Equilibrium* Occurrence requires all other's look-ups to identify that the other player has a dominant decision, plus own look-ups 75 and 48 to identify its best response. Similarly, *Equilibrium* Adjacency requires the comparisons (51, 27), (80, 68), and (75, 48).

comparison of own payoffs and a left-right comparison of other's payoffs for each possible decision combination, except those that can be eliminated as never best responses.[48]

For *Sophisticated*, the characterization of the minimal set(s) of look-ups or comparisons for our games depends only on whether the player has a dominant decision. If so, *Sophisticated* Occurrence (Adjacency) requires only the look-ups (comparisons) needed to identify its dominant decision, namely those for all up-down comparisons of that decision's own payoffs with each other own decision's own payoffs.

If *Sophisticated* does not have a dominant decision (whether or not the game is dominance-solvable), it needs to form beliefs and compare the expected payoffs of its decisions. Although *we* approximate *Sophisticated* beliefs by the observed decision frequencies in our experiment, *Sophisticated* must base them on prior knowledge of others' typical responses to strategically similar games. We assume that forming *Sophisticated* beliefs requires identifying all of the game's dominance and iterated dominance relationships and its set of equilibria, because it is clear from the literature or Table II that subjects' responses normally depend on them. We assume this makes it unnecessary for *Sophisticated* to compare expected payoffs for any of its decisions that are dominated by pure decisions. Like *Naïve*, *Sophisticated* can compare the expected payoffs of its undominated decisions via running totals, either up-down column by column or left-right row by row. We allow either method, but rule out mixtures. Thus, when *Sophisticated* does not have a dominant decision, *Sophisticated* Occurrence requires all own and other's look-ups. *Sophisticated* Adjacency requires all the comparisons *Equilibrium* Adjacency requires, plus any additional ones needed to identify all dominance relationships among own and other's decisions, plus a complete set of the comparisons associated with either all up-down comparisons of own expected payoffs for undominated decisions, or all such left-right comparisons.[49]

---

[48] In game 7A *Equilibrium* Occurrence requires all but the look-ups 18 and 96, to identify own best responses to Left and Right and other's best responses to Bottom and Top. Similarly, *Equilibrium* Adjacency requires comparisons (89, 63), (89, 30), (63, 24), (87, 65), (76, 37), and (37, 32). Note that for non-dominance-solvable games, *Equilibrium* might not need to identify the dominance relationships among its own decisions.

[49] In Game 3A *Sophisticated* Adjacency requires the comparisons (51, 27), (80, 68), and (75, 48) to identify the game's iterated dominance relationships, plus (42, 89) to complete the comparison of own decisions' expected payoffs. In Game 7A *Sophisticated* Adjacency requires all the comparisons *Equilibrium* Adjacency requires, (89, 63), (89, 30), (63, 24), (87, 65), (76, 37), and (37, 32), plus (30, 63) and (18, 96) to complete the identification of dominance relationships among own and other's decisions.

Because *L2* is defined as best response to *Naïve*, we base its information search implications on identifying other's *Naïve* decision and *L2*'s best response to it. We require this even when *L2* has a dominant decision, and we do not allow the use of Observation 4 to simplify the search in other dominance-solvable games. *L2* Occurrence thus requires all other's look-ups, plus all own look-ups for other's *Naïve* decision. *L2* Adjacency requires a complete set of either up-down or left-right comparisons of other's payoffs for all other's decisions, plus all up-down comparisons of own payoffs for other's *Naïve* decision.[50]

In games solvable by k+1 rounds of iterated dominance, Occurrence and Adjacency for *Dk*, k = 1,2, are the same as for *Equilibrium*, because iterated dominance requires smaller sets of look-ups and comparisons than equilibrium-checking or best-response dynamics. In games with no pure-strategy dominance, Occurrence (Adjacency) for *Dk*, k = 1,2, requires the look-ups (comparisons) needed to check for other's dominance plus the look-ups (comparisons) for *Naïve*, namely all own look-ups (a complete set of either up-down comparisons of own decisions' payoffs, two at a time, or of left-right comparisons of own decisions' expected payoffs).[51]

4. Analysis of Decisions and Information Search

This section analyzes subjects' decisions and information searches. Section 4 of CGC&B provides more detail, and our decision data are in Appendix D. Section 4.A reports preliminary statistical tests. Section 4.B reports aggregate compliance with equilibrium. Section 4.C presents a maximum likelihood error-rate analysis of Baseline and OB subjects' decisions, estimating their types and error rates. Section 4.D generalizes this analysis to use Baseline subjects' compliance with our types' information search implications, along with their decisions, to estimate their types and error rates. Section 4.E discusses aggregate patterns in subjects' information searches.

[50]In Game 3A *L2* Occurrence requires the look-ups 51, 27, 80, 68, 75, and 48. *L2* Adjacency requires either the set of comparisons (51, 27) and (80, 68) or the set of comparisons (51, 80) and (27, 68) to identify other's *Naïve* decision, plus the comparison (75, 48) to identify *L2*'s best response to it.

[51]In Game 7A *D1* and *D2* Occurrence require all own and other's look-ups; and *D1* and *D2* Adjacency require the comparisons (89, 63), (89, 30), (32, 37), and (76, 37) to check for other's dominance, plus either the set of comparisons (87, 65), (18, 96), and (63, 24) or the set (87, 18, 63), (65, 96, 24) to identify *L2*'s *Naïve* decision.

A. Preliminary statistical tests

In this section we test for aggregate differences in subjects' decisions across the two runs of the Baseline treatment, B1 and B2; across the Baseline, OB, and TS treatments; and across player roles in isomorphic games. These tests confirm simplifying restrictions suggested by theory and answer questions that are helpful in evaluating our methods. Because the tests compare categorical data from independent samples with no presumption about how they differ, we use Fisher's exact probability test, conducting the tests separately for each game, pooling the data for all subjects in each player role, and for some purposes pooling the data for subjects with isomorphic player roles in different games.[52] Details can be found in CGC&B, Section 4.A.[53]

The tests reveal no differences in subjects' decisions in the B1 and B2 runs that are significant at the 5% level except in game 4C for Column subjects, well within the limits of chance for 36 comparisons. Accordingly, from now on we pool the data from the Baseline runs. The tests also reveal no differences between Baseline and OB subjects' decisions that are significant at the 5% level except in game 6A for Column subjects, again well within the limits of chance. We therefore pool Baseline and OB data when necessary to obtain adequate sample sizes. As expected, there are noticeable differences between Baseline and TS subjects' decisions in 16/18 games, which are significant at any reasonable level in 4/6 games where the subject had three decisions and at the 5% level in 9 games in total. There are no differences between Row and Column subjects' decisions in isomorphic games that are significant at the 5% level except in games 4B and 4D in the Baseline and 9A and 9B in OB, about what would be expected by chance. We therefore pool the data across isomorphic games when necessary to obtain adequate sample sizes. Because these tests include several pairs of isomorphic games that were widely separated in the sequence (5B and 6B, by 12 games; 5A and 6A, by 7; and 7B and 8B, by 5), and we did not control for decision order and labeling across isomorphic games, they provide some assurance that learning and decision labeling and order had little effect on subjects' decisions.

[52]These tests have low power because of our small sample sizes. Conducting tests separately for each game is fully justified only if subjects' decisions are statistically independent across games, which is unlikely because some games are related. However, the correct test without independence (comparing the distributions of entire decision histories) is impractical, and testing game by game allows us to gauge the differences across games and treatments.

[53]In the TS treatment, we exclude the 3 out of 15 TS subjects who revealed by their comments or exit questionnaires that they did not try to identify equilibria. CGC&B gives the results for the full TS sample, which are similar.

B. Aggregate compliance with dominance, iterated dominance, and equilibrium

We now examine subjects' decisions in the aggregate for compliance with dominance, iterated dominance, and equilibrium in the different kinds of games we study.

Table II reports subjects' rates of equilibrium compliance in the B, OB, and TS treatments, pooling the data from isomorphic games, with population fractions in parentheses. The games are grouped by the complexity of the strategic reasoning they require, measured by the number of rounds of iterated pure-strategy dominance needed to identify the subject's equilibrium decision. Baseline and OB subjects' compliance rates are similar across games of similar complexity; and holding complexity constant, the number of own or other's decisions has little effect. Compliance with equilibrium is quite high for initial responses to abstractly framed games, in most cases well above random. As in previous experiments, compliance is highest in games that can be solved by one or two rounds of iterated dominance, and subjects played dominant decisions with frequencies near 90%.[54] But compliance falls steadily as complexity increases, dropping below random in our 3x2 games that are dominance-solvable in three rounds or our 3x2 games with unique equilibria but no pure-strategy dominance.[55]

These results are consistent with subjects' initial responses to games in other experiments, where subjects typically comply with 1-3 rounds of iterated dominance. However, S&W found much higher equilibrium compliance for symmetric 3x3 games solvable by three rounds of iterated pure-strategy dominance or with unique pure-strategy equilibria but no (pure- or mixed-strategy) dominance (68% and 57%, respectively) than we found for 3x2 games of comparable complexity (11-22% and 18-28% in the Baseline and OB, respectively).[56] This difference may stem from S&W's use of symmetric player roles and payoff displays and round-number payoffs, or from our attempt to separate strategic from nonstrategic decision rules as sharply as possible.

---

[54]Interestingly, in 3x2 games with unique pure-strategy equilibria and dominance only via mixed strategies, subjects played dominated decisions with frequencies (10% Baseline, 4% OB) similar to those for pure-strategy dominance.

[55]In most cases compliance is slightly higher in OB than in the Baseline. Although this is unlikely to be due entirely to chance, the difference is too small to be significant in our samples.

[56]The papers cited in the introduction give more evidence from dominance-solvable games. Our results for games with unique equilibria but no pure-strategy dominance are consistent with evidence from other settings summarized by Selten (1998, Section 5), which tends to favor decision rules that employ step-by-step reasoning (such as iterated dominance) over what Selten calls "circular concepts" (such as our *Equilibrium* type in non-dominance-solvable games, our *Sophisticated* type, and, as explained in Section 3.A, all of S&W's strategic types).

TS subjects identified their dominant decisions with frequencies well above 90%. In striking contrast to Baseline and OB subjects, their equilibrium compliance rates fell only slightly in more complex games, averaging well over 90% even in games in which Baseline and OB compliance fell below random. This suggests that Baseline and OB subjects' low compliance in complex games was unlikely to be due to the difficulty of looking up payoffs via MouseLab or cognitive limitations. This leaves several possible explanations for the difference: TS subjects' training in identifying equilibria or their higher dismissal rate (see footnote 18); bounded rationality, in the form of decision rules that do not fully analyze others' incentives; a widespread prior understanding of others' decisions like that reflected in our *Sophisticated* type, coupled with a failure of common knowledge that most subjects are *Sophisticated*; or a combination of these. We now turn to a more detailed econometric investigation of the latter possibilities.

C. Econometric analysis of decisions

In this section we conduct a maximum likelihood error-rate analysis of Baseline and OB subjects' decisions. Recall that our econometric model is a mixture model in which each subject's type is drawn from a common prior distribution over nine types and remains constant for all 18 games.[57] Combining evidence from different patterns of deviation from types' decisions requires an error structure, which we specify as neutrally as possible, in the spirit of H&C's and EG&G's error-rate analyses. We combine *Naïve* and *Optimistic* in this section because their decisions are not separated in our games.[58] We include both *Sophisticated* and *L2*, even though their decisions are separated only in one game for Column subjects, because we pool the data for Row and Column subjects. With these exceptions, any two of our types make different decisions in at least 2/18 games in each player role, and most types are separated much more than that.[59]

Let $i = 1,..., N$ index the subjects in a treatment, let $k = 1,\ldots,K$ index our types, and let $c = 2$, 3, or 4 be the number of a subject's possible decisions in a given game. We assume that a type-k subject normally makes type k's decision, but in each game he makes an error with probability $e_k \in [0, 1]$, type k's *error rate*, in which case he makes each of his c decisions with probability $1/c$. For a type-

[57]We are grateful to Glenn Ellison for suggesting this approach.
[58]*Naïve* and *Optimistic* are distinguished by their information search implications in Section 4.D.
[59]Some overlap is inevitable because all types seek higher own payoffs and all but *Altruistic* pick dominant decisions.

k subject, the probability of type *k*'s decision is then 1 - (*c* - 1)e_k/*c* and the probability of any single non-type *k* decision is e_k/*c*.[60] We assume errors are independently and identically distributed **(**"i.i.d."**)** across games and subjects. Our error structure resembles EG&G's and H&C's, but it permits *c* = 2 to vary in the sample and it allows type-dependent error rates, which is important because the cognitive difficulty of identifying decisions may vary with type.

The likelihood function can be constructed as follows. Let T^c denote the total number of games in which subjects have *c* decisions; in our designs $T^2 = 11$, $T^3 = 6$, and $T^4 = 1$ for both Row and Column subjects (Figure II). Let x_k^{ic} denote the number of subject *i*'s decisions that equal type *k*'s in games in which he has *c* decisions, with $x_k^i \equiv (x_k^{i2}, x_k^{i3}, x_k^{i4})$, $x^i \equiv (x_1^i, ..., x_K^i)$, and $x \equiv (x^1, ..., x^N)$. Let $p \equiv (p_1, ..., p_K)$, where $\sum_{k=1}^{K} p_k = 1$, denote subjects' common prior type probabilities; and let $e \equiv (e_1, ..., e_K)$ denote the types' error rates. Given that a game has one type-*k* decision and *c*-1 non-type-*k* decisions, the probability of observing a particular sample with x_k^i type-*k* decisions when subject *i* is type *k* can be written:

$$L_k^i(e_k \mid x_k^i) = \prod_{c=2,3,4} [1-(c-1)e_k/c]^{x_k^{ic}} [e_k/c]^{T^c - x_k^{ic}}. \tag{4.1}$$

Weighting the right-hand side of (4.1) by p_k, summing over *k*, taking logarithms, and summing over *i* yields the log-likelihood function for the entire sample:

$$\ln L(p,e \mid x) = \sum_{i=1}^{N} \ln \left[\sum_{k=1}^{K} p_k \prod_{c=2,3,4} [1-(c-1)e_k/c]^{x_k^{ic}} [e_k/c]^{T^c - x_k^{ic}} \right]. \tag{4.2}$$

With eight types the model has 15 independent parameters: seven independent type probabilities p_k and eight type-dependent error rates e_k. The influence of x_k^{ic} on the estimated p_k is proportional to ln[(1-(*c*-1)e_k/*c*)/(e_k/*c*)] = 0, which is increasing in *c* for $e_k < 1$ and decreasing in e_k,

[60]Our specification constrains the probability of type *k*'s decision to be at least 1/*c*, but this is never binding. EG&G "take off for guessing" in this way when *c* = 2, in effect writing the probability of type *k*'s decision as $1 - e_k/2$, while H&C write it as $1 - e_k$. When *c* is constant the difference is only notational, but in our model it is substantive. In CGC&B we considered a specification that nests H&C's and EG&G's, in which the probability of each decision conditional on an error is d/*c* for some d ? [0, 1]. d is weakly identified, and we could reject neither d = 0 nor d = 1. We set d = 1 here (and there) for simplicity.

approaching 0 as $e_k \to 1$. Thus, in estimating the p_k, type-k decisions are taken as evidence of type k only to the extent that the estimated e_k suggests that they were made intentionally rather than in error, and are accordingly discounted for higher values of e_k and smaller values of c.[61]

Under our assumptions maximum likelihood yields consistent parameter estimates, which summarize the model's implications for subjects' prior type probabilities and the extent to which the types explain the variation in their decisions. We computed parameter estimates separately for the Baseline and OB treatments. Table III's left-hand columns give the estimated type probabilities and type-dependent error rates.[62] In OB the most frequent types are *L2*, *D1*, *Naïve/Optimistic*, and *Equilibrium*. In the Baseline the most frequent type is *L2*, followed by *Naïve/Optimistic* and *D1*. In each case the total frequency of strategic types is greater than two-thirds, but most subjects' sophistication is better described by boundedly rational types like *L2* or *D1* than *Equilibrium* or *Sophisticated*.[63] Except for *D1* in the Baseline, whose error rate is 70%, the error rates for types with positive estimated probabilities range from 16-29%—comparable to those EG&G and H&C (1994) found for decisions under uncertainty, and reasonably low for initial responses to games. Because the estimated probability that a subject of type k makes type k's decision is $1 - (c\text{-}1)e_k/c$, these error rates imply that OB and non-*D1* Baseline subjects made their types' decisions with probabilities ranging from 0.86 to 0.92, 0.81 to 0.89, or 0.79 to 0.88 for $c = 2$, 3, or 4; for *D1* Baseline subjects the analogous probabilities are 0.65, 0.53, or 0.47.

The most surprising aspects of our Baseline and OB results so far are the large frequency of subjects whose decisions suggest some sophistication (72-80% in Table III), and the fact that equilibrium compliance is still only 11-52% in our most complex games (Table II). Section 3.B's

[61]This discounting makes maximum likelihood estimates of the p_k differ from those that maximize the expected number of correctly predicted decisions. It plays a central role in Section 4.D's analysis, where it allows us to combine evidence from decisions and information search under weak assumptions about how type determines search.

[62]Here and below, the complexity of the estimation made it impractical to compute standard errors.

[63]Our estimates sharply separate *L2* from *Sophisticated*, even though their decisions are separated only for Column subjects in one game, because *L2* decisions predominate in that game, and we estimate one set of type-dependent error rates for all subjects, which give *L2* most of the credit for *L2*'s and *Sophisticated*'s common decisions in the rest of the sample. *L2* prevails over *D1* for similar reasons, although they predict subjects' decisions roughly equally well in the few games in which they are separated. CGC&B's econometric model avoids this "contagion" via error rates by estimating a separate type for each subject based only on his own decisions, and yields less unequal estimates of the p_k. Definitively distinguishing *L2* and *D1* from *Sophisticated* will require further experiments.

analysis suggests a simple explanation for this gap. In our simplest games, most Baseline and OB subjects make decisions that coincide with *Equilibrium*'s and *Naïve*'s; but in our more complex games in which *Equilibrium* and *Naïve* decisions are separated, their decisions tend to coincide with *Naïve*'s. *Naïve/Optimistic*, *L2*, and *D1* all display this kind of variation, which allows them to describe how many subjects' decisions vary across simple and complex games.[64]

Our estimates can also be used to characterize the model's implications for individual subjects' types, which allows us to assess the precision with which subjects' decisions identify their types and facilitates our analysis of cognition and information search at the individual level. To do this, we specify an uninformative prior over the entire parameter vector and compute a Bayesian posterior conditional on each subject's decision history, as in EG&G and S&W. Tables VI and VII in Appendix E summarize Baseline and OB subjects' posterior type probabilities, conditional on their decision histories. 38 of 45 Baseline subjects have one type with posterior probability at least 0.90: 20 *L2*, 9 *Naïve/Optimistic*, 5 *D1*, 2 *Equilibrium*, and 2 *Altruistic*. 20 of 27 OB subjects have one type with posterior probability at least 0.90: 7 *L2*, 5 *Naïve/Optimistic*, 4 *D1*, and 4 *Equilibrium*. Thus, given our specification, most subjects can be assigned to one type.

It is interesting to compare these results with S&W's (1994, 1995) classifications of their subjects by type. S&W (1995) found no evidence of *Rational Expectations* or *Perfect Foresight* (our *Sophisticated*) subjects. Excluding those two types, they found 38 of 48 subjects for which one type has posterior probability at least 0.90: 17 *Worldly*, 9 *L1* (our *Naive*), 6 *L0* (random), 5 *Naïve Nash* (equivalent except for others' decision noise to our *Equilibrium*), and 1 *L2* (equivalent except for others' decision noise to our *L2*). Haruvy, Stahl, and Wilson (1999) modified S&W's (1995) design in an attempt to identify *Optimistic* and *Pessimistic* subjects and found no *Pessimistic* subjects but a few *Optimistic* subjects, most of whom would have been identified as *Worldly* in an analysis like that of S&W (1995). S&W (1994) conducted a similar analysis of data generated by a design closely related to that of S&W (1995), allowing types *L0*, *L1*, *L2*, and *Naïve Nash* (but not *Worldly*, *Optimistic*, or

[64]These non-hybrid, boundedly rational types mimic the effects of a hybrid type that switches from *Equilibrium* to *Naïve* in complex games, which suggests that also allowing hybrid types would lead to overfitting. *D1*'s decisions coincide with *Naïve*'s in all the games in *L2*'s do, plus our 2x3 games with unique equilibria but no pure-strategy dominance. *D1* and *L2* respect two rounds of dominance, pure-strategy for *D1* and pure- or mixed-strategy for *L2*. *L2*'s advantage over *D1* turns on this difference; their decisions are the same in our games solvable by three rounds.

Pessimistic). They found 35 of 40 subjects for which one type has posterior probability at least 0.90: 18 *L2*, 9 *Naïve Nash*, and 8 *L1*.

S&W's type estimates for decisions are generally similar to ours. The main difference is that S&W (1995) identify many subjects as *Worldly* and few as *L2* (by their definition, in which subjects respond to others' decision noise), while we exclude S&W's *Worldly* type a priori and identify many subjects as *L2* (by our noiseless definition) and some as *D1* (also noiseless).[65] This difference might stem from our assumption of uniform (rather than logit) decision errors, from our use of asymmetric games with separate payoff tables for Row and Column each players (rather than symmetric games displayed without distinguishing own and other's payoffs), or from our a priori exclusion of types like *Worldly* that depend on estimated parameters and/or others' decision noise, and thus implicitly assume that subjects have prior understandings of others' initial responses to games. S&W's (1994) analysis, in which excluding *Worldly* makes the estimated type mix much closer to ours, suggests that the last difference is the most important.[66]

D. Econometric analysis of decisions and information search

This section generalizes our econometric model of decisions to analyze Baseline subjects' decisions and information searches.[67] We now distinguish *Naïve* and *Optimistic*, so there are nine types. We consider two alternative models, to allow readers with different degrees of confidence in our assumptions about information search to draw appropriate conclusions from our estimates.

The first model, which we call our model of decisions with compliance-conditional error rates, allows Section 4.C's type-specific error rates to depend on compliance with types' Occurrence and Adjacency as described below. This model makes weak assumptions about how type determines search, at the cost of using the information in subjects' searches in a limited way, allowing a type's decisions to count less as evidence for that type when they come with the "wrong" look-ups for the type, but otherwise not taking into account how well the observed searches are explained by the type's Occurrence and Adjacency. The second model, which we call our model of decisions and information

[65] *D1*, which S&W (1994, 1995) did not consider, is weakly separated from *Naïve Nash* in S&W's (1994) design.
[66] In Section 3.A we argued that the idea of "worldliness" is better represented by a type like *Sophisticated*, which cleanly addresses the issue of prior understandings without imposing structural restrictions on how others' decisions are determined or raising delicate specification issues; and that types' decisions should otherwise be identifiable entirely without input from other subjects. We view this as an important theoretical constraint.

search, generalizes Section 4.C's model to obtain a joint error-rate model of decisions and search, with the latter again represented by compliance with Occurrence and Adjacency. This model makes stronger distributional assumptions, but uses both decisions and search to estimate subjects' types, treating them as symmetrically as possible.

For tractability, and to minimize the need for structural restrictions, we categorize compliance with Occurrence and Adjacency discretely, as follows. For each subject, type, and game, we compute the percentages of the type's look-ups required by Occurrence, and the type's comparisons required by Adjacency, that appear at least once in the subject's look-up sequence. We then sort these percentages into four categories: B_H, 100% compliance with Occurrence and 67-100% compliance with Adjacency; B_M, 100% compliance with Occurrence and 34-66% compliance with Adjacency; B_L, 100% compliance with Occurrence and 0-33% compliance with Adjacency; and ~A, anything less than 100% compliance with Occurrence.[68] Thus we use a coarser grid for Occurrence than for Adjacency, in effect assuming that compliance with Adjacency is meaningless without 100% compliance with Occurrence. This is a reasonable simplification because a subject who violates a type's Occurrence cannot identify its decision with certainty, but memory may allow a subject who violates a type's Adjacency to identify its decision. For *Optimistic*, whose Adjacency is vacuous, there are only two distinct categories, A and ~A. We treat *Optimistic* category A as the union of vacuous categories B_H, B_M, and B_L and smooth compliance across them as explained below, to prevent our model from arbitrarily favoring *Optimistic* because it has fewer categories. We index a subject's compliance with a given type's Occurrence and Adjacency in a game by j, where $j = A$ or 0 for *Optimistic* categories A or ~A and $j = H, M, L$, or 0 for other types' categories B_H, B_M, B_L, or ~A. Compliance j for type k will be called *type-k compliance j*, or just *compliance j* when the type is clear. Sums and products over j are taken over H, M, L, and 0, with obvious adjustments for *Optimistic*.

Recall that we assume that in each game, a subject's type determines his information search with error, and his type and information search then determine his decision, again with error. Accordingly,

---

[67]The information search data from the Baseline and TS treatments are available on request.

[68]100% compliance with a type's Adjacency implies 100% compliance with its Occurrence, but less than 100% compliance with Adjacency implies no simple restrictions on compliance with Occurrence. Although we describe compliance as a continuous percentage, it too is discrete. Adjacency, for instance, requires from 0-8 comparisons for different games and types. This makes the precise locations of the boundaries between our categories unimportant.

our model of decisions and information search allows a subject's deviations from his type's decision and search in a given game to be correlated, but it assumes that, given c and k, the deviations are i.i.d. across games and subjects. We describe the joint probability distribution of decisions and search by specifying *un*conditional compliance probabilities and compliance-conditional decision error rates for each type. Let z_{kj} denote the probability that a subject of type k has type-k compliance j in a game, where $\sum_j z_{kj} = 1$ for all k. A type-k subject normally makes type k's decision, but in each game, given type-k compliance j, he makes an error with probability e_{kj} ? [0, 1], type k's *error rate with compliance j*, in which case he makes each of his c decisions with probability 1/c. For a type-k subject with compliance j, the conditional probability of type k's decision is then $1-(c-1)e_{kj}/c$ and the conditional probability of each non-type k decision is e_{kj}/c. For a type-k subject, the unconditional probability of type k's decision *and* type-k compliance j is $z_{kj}[1-(c-1)e_{kj}]/c$ and the unconditional probability of any single non-type k decision *and* type-k compliance j is $z_{kj}e_{kj}/c$. This error structure implies a joint probability distribution over decisions and searches that is fully general, with three exceptions: It constrains how the probabilities of deviations from types' decisions and searches vary with c, it assumes that subjects' searches influence their decisions only through compliance with Occurrence and Adjacency as we categorize them, and it smoothes compliance across *Optimistic*'s vacuous categories in a particular way. Despite the asymmetries in our notation, these are the only asymmetries in the model's treatment of decisions and search.

Let $z_k \equiv (z_{kH}, z_{kM}, z_{kL}, z_{k0})$, or (z_{kA}, z_{k0}) for *Optimistic*; and let $z \equiv [z_{kj}]$. Let $e_k \equiv (e_{kH}, e_{kM}, e_{kL}, e_{k0})$, or (e_{kA}, e_{k0}) for *Optimistic*; and let $e \equiv [e_{kj}]$. Generalizing Section 4.C's notation, let T_{kj}^{ic} denote the number of games in which subject i has c decisions and type-k compliance j, so that $\sum_j T_{kj}^{ic} = T^c$ for all i, k, and c. We equalize the likelihood for *Optimistic* with that of a hypothetical type whose compliance is randomly distributed across the missing categories by reallocating the T_{kj}^{ic} for *Optimistic* category A evenly across categories B_H, B_M, and B_L, so that all types get equal credit for actual success in predicting subjects' search compliance.

Let x_{kj}^{ic} denote the number of games in which subject *i* has *c* decisions, type-*k* compliance *j*, and makes type *k*'s decision, so $\sum_j x_{kj}^{ic} = x_k^{ic}$ for all *i*, *k*, and *c*. Let $\overline{T}_k^{\,i} \equiv [T_{kj}^{ic}]$, $\overline{T}^{\,i} \equiv (\overline{T}_1^{\,i},...,\overline{T}_K^{\,i})$, and $\overline{T} \equiv (\overline{T}^{\,1},...,\overline{T}^{\,N})$; and let $\overline{x}_k^i \equiv [x_{kj}^{ic}]$, $\overline{x}^i \equiv (\overline{x}_1^i,...,\overline{x}_K^i)$, and $\overline{x} \equiv (\overline{x}^1,...,\overline{x}^N)$. The probability of observing a particular sample with T_{kj}^{ic} and x_{kj}^{ic} when subject *i* is type *k* can be written:

$$L_k^i(\boldsymbol{e}_k, \boldsymbol{z}_k \mid \overline{x}_k^i, \overline{T}_k^i) = \prod_j \prod_{c=2,3,4} \boldsymbol{z}_{kj}^{T_{kj}^{ic}} [1-(c-1)\boldsymbol{e}_{kj}/c]^{x_{kj}^{ic}} [\boldsymbol{e}_{kj}/c]^{T_{kj}^{ic}-x_{kj}^{ic}}. \tag{4.3}$$

Weighting the right-hand side of (4.3) by p_k, summing over *k*, taking logarithms, and summing over *i* yields the log-likelihood function for the entire sample:

$$\ln L(p, \boldsymbol{e}, \boldsymbol{z} \mid \overline{x}, \overline{T}) = \sum_{i=1}^{N} \ln \left[\sum_{k=1}^{K} p_k \prod_j \prod_{c=2,3,4} \boldsymbol{z}_{kj}^{T_{kj}^{ic}} [1-(c-1)\boldsymbol{e}_{kj}/c]^{x_{kj}^{ic}} [\boldsymbol{e}_{kj}/c]^{T_{kj}^{ic}-x_{kj}^{ic}} \right]. \tag{4.4}$$

The log-likelihood function (4.2) for our model of decisions alone can be obtained from (4.4) by removing the conditioning of $\boldsymbol{z}_{kj}$ and $\boldsymbol{e}_{kj}$ on *j*, collecting terms, and summing T_{kj}^{ic} and x_{kj}^{ic} over *j*. The log-likelihood function for our model of decisions with compliance-conditional error rates can be obtained from (4.4) by removing the terms in $\boldsymbol{z}_{kj}$ to obtain:

$$\ln L(p, \boldsymbol{e} \mid \overline{x}, \overline{T}) = \sum_{i=1}^{N} \ln \left[\sum_{k=1}^{K} p_k \prod_j \prod_{c=2,3,4} [1-(c-1)\boldsymbol{e}_{kj}/c]^{x_{kj}^{ic}} [\boldsymbol{e}_{kj}/c]^{T_{kj}^{ic}-x_{kj}^{ic}} \right] \tag{4.5}$$

The model of (4.4) has 67 independent parameters: 9 type probabilities p_k less one adding-up restriction; 34 compliance-conditional error rates $\boldsymbol{e}_{kj}$ (4 compliance categories for each of 9 types, less 2 for *Optimistic*); and 25 unconditional compliance probabilities $\boldsymbol{z}_{kj}$ (34 compliance categories less one $\sum_j \boldsymbol{z}_{kj} = 1$ restriction for each of 9 types). Similarly, the model of (4.5) has 42 independent parameters. In both models, decisions influence the estimates as in (4.2)'s model, with a new twist because the error rates are compliance-contingent. As before, the influence of x_{kj}^{ic} on the estimated p_k decreases with $\boldsymbol{e}_{kj}$, approaching 0 as $\boldsymbol{e}_{kj} \to 1$, so the estimated p_k discounts decisions for the probability that they were made in error. But now, to the extent that the estimated $\boldsymbol{e}_{kj}$ decrease with *j* as theory

suggests, the estimated p_k discounts type-k decisions as evidence for type k more when they occur with the wrong look-ups for type k.[69] Except for sampling error the estimated e_{kj} will be independent of aspects of compliance that do not affect decisions, so the estimated p_k will ignore such aspects. This feature of our model allows us to estimate which aspects of types' Occurrence and Adjacency help to predict subjects' decisions, and how, while simultaneously using that information to estimate subjects' type frequencies.

In our model of decisions and information search, search also has a direct influence on the estimated type frequencies, in that (4.4)'s log-likelihood favors types k for which the T_{kj}^{ic}, and thus the estimated z_{kj}, are more concentrated on particular levels of type-k compliance j, because such types' search implications explain more of the variation in subjects' look-up patterns.[70] Such types are favored without regard to whether the levels of j on which the T_{kj}^{ic} are concentrated are high or low: Theory suggests that they should be concentrated on *high* levels, but we do not impose such restrictions. Instead we use the unrestricted estimates of z_{kj} as a diagnostic, placing more confidence in a type's estimated frequency when the restrictions are satisfied.

To see how these effects work more concretely, consider *Naïve* and *Optimistic*, which in our games make the same decisions and have almost the same Occurrence. Even so, (4.4)'s and (4.5)'s models can distinguish them because *Naïve* Adjacency is restrictive while *Optimistic* Adjacency is vacuous. Although it may seem that this must favor *Optimistic* in the estimates, there are two effects at work, both of which favor *Naïve* in our sample. To the extent that subjects' *Naïve/Optimistic* decisions are more concentrated on high levels of compliance for *Naïve* than for *Optimistic*, the e_{kj} terms in (4.4) or (4.5) favor *Naïve* because its Adjacency is more useful in predicting subjects' decisions. And

[69]We stress that our analysis allows rather than assumes this, in that which look-ups are wrong is determined by unrestricted estimates of the e_{kj}. This flexibility is important, given how little is known about information search.

[70]Our model of decisions and information search can be criticized because it assigns the same meaning to a given degree of dispersion of compliance across categories for each type (while allowing different *levels* of compliance for different types), even though some types' information search requirements are harder to satisfy. This aspect of our error-rate analysis is a considered response to the difficulty of modeling search, but it bears emphasis that it rests on untestable assumptions about the link between cognition and search that are stronger than those required for our model of decisions with compliance-conditional error rates. Also, in replacing CGC&B's five compliance categories by the four used here, we found that estimates for our model of decisions and information search (but not our model of decisions with compliance-conditional error rates) are somewhat sensitive to the number of categories. We presume that this sensitivity would be eliminated with more categories, but this is computationally impractical.

to the extent that subjects' searches satisfy *Naïve* Adjacency more than randomly, the z_{kj} terms in (4.4) also favor *Naïve*. Similar but more complex considerations distinguish *L2* and *D1* in both models, even though both respect two rounds of iterated pure-strategy dominance and make very similar decisions in our games.

Table III's two right-most columns report maximum likelihood estimates of p_k, z_{kj}, and e_{kj} for the Baseline treatment, using (4.5)'s model of decisions with compliance-contingent error rates and (4.4)'s model of decisions and information search. Under our assumptions both models yield consistent estimates, but the latter's are more efficient because (4.4)'s model uses more information. For (4.5)'s model we also report estimates of z_{kj} conditional on the p_k estimated from decisions alone, to indicate the compliance frequencies on which the e_{kj} estimates are based.

The estimates for our model of decisions with compliance-contingent error rates are very close to those for (4.2)'s model of decisions alone, with nearly the same ordering of type frequencies. The main difference is the strong separation of *Naïve* and *Optimistic*. A likelihood ratio test cannot reject the hypothesis that $e_{kj} \equiv e_k$ for all j and k (p-value 0.20) or, a fortiori, that e_{kj} is weakly increasing in j for all k.[71] These results weakly support the implication of our theory that subjects with higher compliance tend to make their types' decisions more frequently, suggesting that there are systematic relationships between subjects' deviations from search patterns associated with equilibrium analysis and their deviations from equilibrium decisions.

Naive and *L2* have high compliance, error rates that decrease with higher compliance (with a low-frequency exception for *L2*), and low error rates when compliance is high, all as suggested by theory. For those types the implied frequencies of noncompliance are generally lower than the corresponding error rates, which supports the interpretation that subjects made their estimated types' decisions intentionally, except for errors. By contrast, *D1* has fairly high compliance and high error rates that decrease with compliance (with a low-frequency exception); and *Altruistic* and *Equilibrium* have low compliance and error rates that decrease with compliance (with an exception for *Equilibrium*).

[71]This test has low power because several types have zero estimated frequencies, so the associated error rates are not identified. If we had imposed those zero restrictions a priori, the p-value would have been 0.015.

These results suggest that the estimated frequencies of *Naive* and *L2* are reliable, but they give somewhat less reason for confidence in that of *D1* and little reason for confidence in those of *Altruistic* or *Equilibrium*.

The estimates for our model of decisions and information search generally confirm the view of subjects' behavior suggested by the other two models, with some significant differences. *Naïve* and *L2* now have the largest estimated frequencies, each around 45%, and *D1* has disappeared. The shift toward *Naïve*, which comes mainly at the expense of *Optimistic* and *D1*, reflects the fact that *Naïve* compliance explains more of the variation in subjects' searches and decisions than *Optimistic*'s, which is too unrestrictive to be useful in our sample, or *D1*'s, which is more restrictive than *Naïve*'s, but less correlated with subjects' behavior. Again, a likelihood ratio test cannot reject the hypothesis that $e_{kj} \equiv e_k$ for all j and k (p-value 0.99), or that e_{kj} is weakly increasing in j for all k. These results also weakly support the theory's implication that subjects with higher compliance make their types' decisions more frequently.

Altruistic, *Naïve*, and *Optimistic* now have high compliance and estimated error rates that decrease with compliance (with two low-frequency exceptions for *Naive*). For these types compliance is generally high enough to support the interpretation that subjects made their estimated types' decisions intentionally, except for errors. By contrast, *L2* has high compliance but estimated error rates that increase with higher compliance as often as they decrease (although with low frequencies), and *Pessimistic* and *Sophisticated* have very high error rates that decrease with compliance. These results suggest that the frequencies of *Altruistic*, *Naïve*, and *Optimistic* estimated using our model of decisions and information search are reliable, but they give less reason for confidence in the estimated frequency of *L2* and very little reason for confidence in those of *Pessimistic* and *Sophisticated*. With these qualifications, our analysis suggests that there are large frequencies of *Naïve* and *L2* subjects, totalling from 67-89% of the population, and a frequency of *D1* subjects from 0-20%, in each case depending on one's confidence in our assumptions about information search. In any case, there are at most traces of our other six types.

The estimates can again be used to compute Bayesian posteriors for individual subjects' types by specifying an uninformative prior and conditioning on their decision and search histories. The second

entries in the cells of Table VI in Appendix E summarize the implications of our model of decisions and information search for subjects' posterior type probabilities. 43 of 45 Baseline subjects have a posterior probability for one type of at least 0.90: 19 *L2*, 19 *Naïve*, 2 *Pessimistic*, 1 *Altruistic*, 1 *Optimistic*, and 1 *Sophisticated*.[72] Thus, observing search allows us to assign more subjects to a single type with confidence (43) than our analysis of decisions alone (38), even though explaining subjects' decisions and searches simultaneously is a harder task. Focusing on modal type probabilities, observing search allows us to identify 11 subjects previously estimated to be *Naïve*/*Optimistic* as *Naïve* (10) or *Optimistic* (1). It changes another 16 subjects' modal posterior types: 6 from *L2*, 2 from *D1*, 1 from *Equilibrium*, and 1 from *Altruistic* to *Naïve*; 2 from *D1* and 1 from *Equilibrium* to *L2*; 2 from *D1* to *Pessimistic*; and 1 from *D1* to *Sophisticated*.[73] Finally, it sharpens the identification of 11 *L2* subjects, clouds the identification of 3 *L2* subjects, and leaves 4 posteriors (3 *L2* and 1 *Altruistic*) unchanged at 1.000.

E. Aggregate patterns in subjects' information searches

In this section we describe the aggregate patterns in subjects' information searches, using 13 simple measures. We hope that this approach, which imposes less structure. will convey more of the information in our search data, and thereby indicate the possibilities for further analysis.

The 13 measures are: the average total numbers of look-ups per game in own and other's payoffs; the average numbers of consecutive look-ups in own and other's payoffs, or *string lengths*; the average look-up durations in own and other's payoffs, in seconds, or *gaze times*; the frequencies with which own payoffs are inspected first, and inspected last; the frequencies of look-up transitions from own to own payoffs, and from other's to other's payoffs; the frequencies of up-down transitions in own payoffs and left-right transitions in other's payoffs, conditional on remaining in own or other's payoffs, respectively; and the frequency of *altruistic* transitions, from a given decision combination in own payoffs to the same one in other's payoffs, or vice versa.

Appendix A explains how to derive our types' implications for the measures under simple assumptions about the relationship between cognition and search. The top half of Table IV gives our types' theoretical implications for each measure, derived game by game and then averaged over games

[72]The remaining two subjects have modal posterior probabilities of 0.74, one on *L2* and one on *Naïve*.

[73]We find the comparatively high frequency of changed type estimates unsurprising, given that a subject's decisions and information searches over 18 games each constitute only one observation for the purpose of estimating his type.

because they do not vary much across games, with the implications of random look-ups as a benchmark.[74] Table IV shows that the implications differ systematically across three groups of types: our other-regarding but nonstrategic *Altruistic* type; our other nonstrategic types, *Pessimistic*, *Naïve*, and *Optimistic*; and our strategic types, *L2*, *D1*, *D2*, *Equilibrium*, and *Sophisticated*. These differences are large enough to have a chance to show up in aggregate data.

The bottom half of Table IV summarizes the measures for Baseline and TS subjects, also averaged over games, first with all subjects in each treatment pooled and then with Baseline subjects sorted by their most likely types estimated from decisions alone (Table III).[75] Several interesting patterns are apparent even at this level of aggregation. TS subjects have more own and other's look-ups, longer string lengths, and shorter gaze times than Baseline subjects, all of which suggest that TS subjects perform more systematic analyses. TS subjects also have many more own up-down and other's left-right transitions, which are characteristic (under Adjacency) of algorithms for identifying *Equilibrium* or *Sophisticated* decisions, and more generally of strategic thinking in the normal form. These differences suggest that the methods by which theorists analyze normal-form games do not come naturally to subjects without training in game theory. Both TS and Baseline subjects have more other's left-right than own up-down transitions, which suggests that our display generates some bias in favor of left-right transitions.

The search measures for Baseline subjects sorted by most likely type are often several times higher than Table IV's theoretical bounds, but the two vary roughly in proportion across types. *Altruistic* subjects have more own and other's look-ups, shorter string lengths, fewer own-to-own and other's-to-other's transitions, and, unsurprisingly, more altruistic own-to-other's transitions than other subjects. Every type but *Altruistic* has more own than other's look-ups, with the largest difference for *Naïve/Optimistic* (*Equilibrium* is a close second). Every type has (slightly) longer own than other's gaze times. Own payoff first exceeds 58% for every type, and 70% for all but *L2*; and own payoff last exceeds 54% for all types but *Altruistic*.

[74]Random look-ups are defined as independently and uniformly distributed given their total number, which is set equal to the observed total for each game-subject pair and then treated as exogenous.

[75]The measures were computed by first computing an average for each subject over all 18 games (unweighted by size of game) and then averaging over subjects (unweighted by length of look-up sequence).

Table V shows how the search measures compare to our types' implications for TS and Baseline subjects, and for Baseline subjects sorted by their most likely types as estimated from decisions alone. The top part of Table V gives TS and Baseline subjects' aggregate rates of compliance with our types' Occurrence (100% minus the percentage for $j = 0$) and Adjacency (the percentage for $j = 1$) restrictions. Compliance rates are calculated and categorized for each subject, type, and game and then aggregated across games and subjects. TS and Baseline subjects differ sharply in compliance with *Equilibrium*, *D1*, and *D2* Occurrence and Adjacency; with *Sophisticated* and *Pessimistic* Adjacency; and to a lesser extent with *L2* Occurrence; but comparatively little in compliance with the other, nonstrategic types' Occurrence and Adjacency.

The bottom part of Table V gives Baseline subjects' aggregate rates of compliance with subjects sorted by most likely type as in Table IV. These results suggest that Occurrence by itself doesn't discriminate very well, mainly because most subjects usually comply with most types' Occurrence. By contrast, Adjacency (which includes Occurrence by definition), particularly category B_H (100% compliance with Occurrence and 67-100% compliance with Adjacency), discriminates well even at this aggregate level. In general, subjects whose most likely type is k have higher rates of compliance in type k's category j = H than other subjects, with minor exceptions for *Naïve/Optimistic* and larger exceptions for *D1*. The contrast between these results and the sharp separation of TS from Baseline subjects in the top part of Table V suggests that there are important differences between TS subjects and "naturally occurring" *Equilibrium* Baseline subjects (Baseline subjects whose most likely type is *Equilibrium*).

5. Conclusion

This paper reports the results of experiments designed to study strategic sophistication, in which subjects play a series of 18 two-person normal-form games with varying structures, using a computer interface called MouseLab that records their searches for hidden payoff information along with their decisions. Our results show that it is feasible to study subjects' cognitive processes by monitoring their information searches along with their decisions in normal-form games, and that the richness of the search possibilities in normal-form games makes such an analysis a powerful complement to C&J's analysis of information search in extensive-form games. Our analysis shows that it is possible to give a coherent,

unified account of subjects' searches and decisions by incorporating the cognitive implications of decision rules into a simple model based on a procedural view of decision-making and conservative assumptions about the relationship between cognition and search. The model gives a clearer view of how subjects' decisions are determined, and shows that their deviations from the search implications of equilibrium analysis help to predict their deviations from equilibrium decisions.

More generally, our analysis suggests that strategic behavior can be better understood by searching for sets of simple, generalizable decision rules that describe players' decisions in a variety of environments, using their cognitive requirements to derive their implications for information search, and constructing a unified explanation of both aspects of players' behavior. We hope that the theory and the tools for measurement and data analysis discussed here will be useful in expanding the view of economic agents from economic decision-makers, to economic and informational decision-makers, and eventually to economic and cognitive decision-makers.

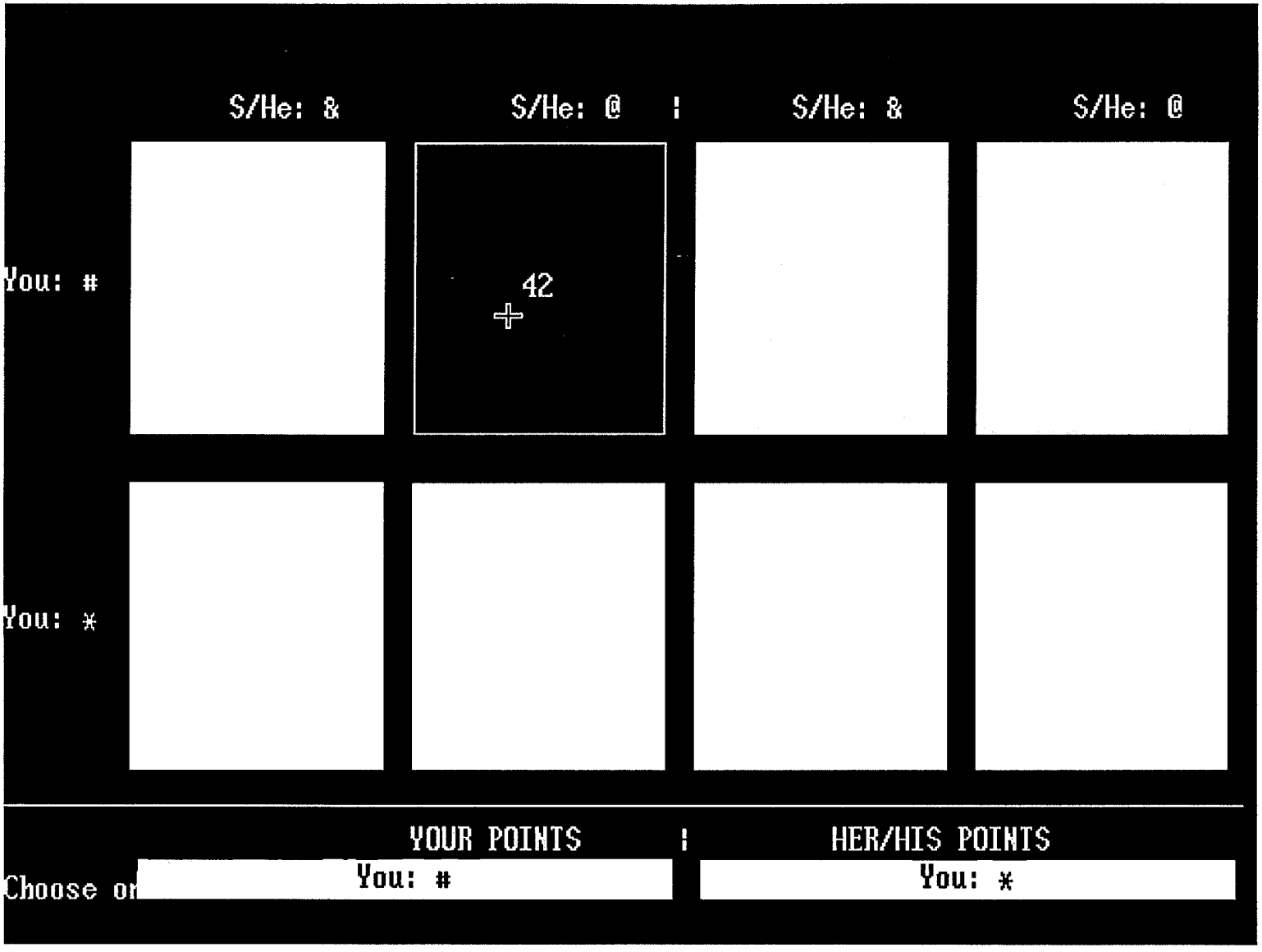

Figure I. MouseLab screen shot

Figure II. Games

2A (1,2)	*A,P,N*	*D12,L2,E,S*
A	72,93	31,46
D	84,52	55,79

2B (1,2)	*A,P,N*	*D12,L2,E,S*
D	94,23	38,57
A	45,89	14,18

3A (2,1)	*D*	*A*
D12,L2,E,S	75,51	42,27
A,P,N	48,80	89,68

3B (2,1)	*D*	*A*
A,P,N	21,92	87,43
D12,L2,E,S	55,36	16,12

4A (2,1)		*D*	*A*
A,P,N	59,58	46,83	85,61
D12,L2,E,S	38,29	70,52	37,23

4B (2,1)	*A*	*D*
D12,L2,E,S	31,32	68,46
P	72,43	47,61
A,N	91,65	43,84

4C (1,2)	*D12,L2,E,S*	*A,P,N*
	28,37	57,58
A	22,36	60,84
D	51,69	82,45

4D (1,2)	*D12,L2,E,S*	*P*	*A,N*
D	42,64	57,43	80,39
A	28,27	39,68	61,87

5A (3,2)	*A,P,N*	*D12,L2,E,S*
A	53,86	24,19
P,N,D1,L2,S	79,57	42,73
D2,E	28,23	71,50

5B (3,2)	*A,P,N*	*D12,L2,E,S*
A	76,93	25,12
D2,E	43,40	74,62
P,N,D1,L2,S	94,16	59,37

6A (2,3)	*A*	*D2,E,S*	*P,N,D1,L2*
D12,L2,E,S	21,26	52,73	75,44
A,P,N	88,55	25,30	59,81

6B (2,3)	*D2,E*	*A*	*P,N,D1,L2,S*
A,P,N	42,45	95,78	18,96
D12,L2,E,S	64,76	14,27	39,61

7A (8 ,8)	*N,D12,L2,S*	*A,P*	*E*
L2,E,S	87,32	18,37	63,76
A,P,N,D12	65,89	96,63	24,30

7B (8 ,8)	*N,D12,L2,S*	*A,P*	*E*
A,P,N,D12	67,91	95,64	31,35
L2,E,S	89,49	23,53	56,78

8A (8 ,8)	*L2,E,S*	*A,P,N,D12*
E	72,59	26,20
A,P	33,14	59,92
N,D12,L2,S	28,83	85,61

8B (8 ,8)	*L2,E,S*	*A,P,N,D12*
A,P	46,16	57,88
E	71,49	28,24
N,D12,L2,S	42,82	84,60

9A (1,2)	*D12,L2,E,S*	*A,P,N*
	22,14	57,55
	30,42	28,37
A	15,60	61,88
D	45,66	82,31

9B (2,1)			*A*	*D*
A,P,N	56,58	38,29	89,62	32,86
D12,L2,E,S	15,23	43,31	61,16	67,46

Table I. Types' Payoff Incentives

Decisions Type	*A*	*P*	*N*	*O*	*L2*	*D1*	*D2*	*E*	*S*	Expected $ Value
A	100%	122%	125%	125%	145%	141%	140%	142%	146%	17.11
P	82%	100%	102%	102%	119%	115%	114%	116%	119%	20.93
N	80%	98%	100%	100%	116%	113%	112%	113%	117%	21.38
O	80%	98%	100%	100%	116%	113%	112%	113%	117%	21.38
L2	69%	84%	86%	86%	100%	97%	96%	98%	100%	24.87
D1	71%	87%	89%	89%	103%	100%	99%	101%	103%	24.13
D2	71%	87%	89%	89%	104%	101%	100%	101%	104%	23.95
E	71%	86%	88%	88%	103%	99%	99%	100%	103%	24.19
S	69%	84%	86%	86%	100%	97%	96%	97%	100%	24.93

Table II. Percentages of Decisions that Comply with Equilibrium by Type of Game

Type of Game (rounds of dominance for player to identify own equilibrium decision)	Baseline	OB	B+OB	TS
2x2 with dominant decision (1)	85.6%	92.6%	88.2%	100.0%
(2A, 2B for Rows; 3A, 3B for Cols.)	(77/90)	(50/54)	(127/144)	(24/24)
2x3 with dominant decision (1)	82.2%	100.0%	88.9%	100.0%
(4D for Rows; 4B for Cols.)	(37/45)	(27/27)	(64/72)	(12/12)
3x2 with dominant decision (1)	86.7%	92.6%	88.9%	100.0%
(4C for Rows; 4A for Cols.)	(39/45)	(25/27)	(64/72)	(12/12)
4x2 with dominant decision (1)	88.9%	96.3%	91.7%	100.0%
(9A for Rows; 9B for Cols.)	(40/45)	(26/27)	(66/72)	(12/12)
2x2, partner has dominant decision (2)	61.1%	79.6%	68.1%	95.8%
(3A, 3B for Rows; 2A, 2B for Cols.)	(55/90)	(43/54)	(98/144)	(23/24)
2x3, partner has dominant decision (2)	62.2%	63.0%	62.5%	100.0%
(4A for Rows; 4C for Cols.)	(28/45)	(17/27)	(45/72)	(12/12)
3x2, partner has dominant decision (2)	60.0%	55.6%	58.3%	83.3%
(4B for Rows; 4D for Cols.)	(27/45)	(15/27)	(42/72)	(10/12)
2x4, partner has dominant decision (2)	73.3%	70.4%	72.2%	100.0%
(9B for Rows; 9A for Cols.)	(33/45)	(19/27)	(52/72)	(12/12)
2x3 with 2 rounds of dominance (2)	62.2%	68.5%	64.6%	100.0%
(6A, 6B for Rows; 5A, 5B for Cols.)	(56/90)	(37/54)	(93/144)	(24/24)
3x2 with 3 rounds of dominance (3)	11.1%	22.2%	15.3%	87.5%
(5A, 5B for Rows; 6A, 6B for Cols.)	(10/90)	(12/54)	(22/144)	(21/24)
2x3, unique equilibrium, no dominance	50.0%	51.9%	50.7%	91.7%
(7A, 7B for Rows; 8A, 8B for Cols.)	(45/90)	(28/54)	(73/144)	(22/24)
3x2, unique equilibrium, no dominance	17.8%	27.8%	21.5%	91.7%
(8A, 8B for Rows; 7A, 7B for Cols.)	(16/90)	(15/54)	(31/144)	(22/24)

Table III. Parameter Estimates for OB and Baseline Subjects (— vacuous)

	Decisions Alone (*Naïve* and *Optimistic* Parameters Constrained Equal)		Decisions with Compliance-conditional Error Rates	Decisions and Information Search
Treatment (log-likelihood)	**OB (-246.44)**	**B (-446.39)**	**B (-433.23)**	**B (-852.02)**
Type				
Altruistic p_k	0.000	0.044	0.089	0.022
z_{kj}, *j=H,M,L,0*	—	—	(0.04,0.02,0.36,0.57)	0.89,0.00,0.00,0.11
e_k or e_{kj}, *j=H,M,L,0*	—	0.253	0.26,0.63,0.79,0.82	0.00,—,—,0.66
Pessimistic p_k	0.000	0.000	0.000	0.045
z_{kj}, *j=H,M,L,0*	—	—	—	0.47,0.00,0.00,0.53
e_k or e_{kj}, *j=H,M,L,0*	—	—	—	0.60,—,—,1.00
Naïve p_k	0.199	0.240	0.227	0.448
z_{kj}, *j=H,M,L,0*	—	—	(0.97,0.02,0.01,0.01)	0.95,0.01,0.01,0.02
e_k or e_{kj}, *j=H,M,L,0*	0.285	0.286	0.24,0.43,0.58,0.81	0.50,0.39,0.47,0.85
Optimistic p_k	0.199	0.240	0.000	0.022
z_{kj}, *j=A,0*	—	—	—	1.00,0.00
e_k or e_{kj}, *j=A,0*	0.285	0.286	—	0.29,0.50
L2 p_k	0.344	0.496	0.442	0.441
z_{kj}, *j=H,M,L,0*	—	—	(0.88,0.07,0.02,0.03)	0.87,0.04,0.03,0.06
e_k or e_{kj}, *j=H,M,L,0*	0.233	0.203	0.18,0.35,0.21,0.21	0.25,0.61,0.16,0.22
D1 p_k	0.298	0.175	0.195	0.000
z_{kj}, *j=H,M,L,0*	—	—	(0.44,0.12,0.06,0.38)	—
e_k or e_{kj}, *j=H,M,L,0*	0.276	0.704	0.43,0.63,0.15,1.00	—
D2 p_k	0.000	0.000	0.000	0.000
z_{kj}, *j=H,M,L,0*	—	—	—	—
e_k or e_{kj}, *j=H,M,L,0*	—	—	—	—
Equilibrium p_k	0.160	0.044	0.052	0.000
z_{kj}, *j=H,M,L,0*	—	—	(0.00,0.08,0.75,0.17)	—
e_k or e_{kj}, *j=H,M,L,0*	0.165	0.163	—,0.41,0.00,0.97	—
Sophisticated p_k	0.000	0.000	0.000	0.022
z_{kj}, *j=H,M,L,0*	—	—	—	0.00,0.00,0.71,0.29
e_k or e_{kj}, *j=H,M,L,0*	—	—	—	—,—,0.54,1.00

Table IV. Implications of Types and Aggregate Look-up Measures for TS and Baseline Subjects, and Baseline Subjects by Most Likely Type Estimated from Decisions Alone (— vacuous)

Type	Own Look-Ups	Other Look-Ups	Own String Length	Other String Length	Own Gaze Time	Other Gaze Time	Own Payoff First	Own Payoff Last	Own-Own Trans.	Other-Other Trans.	Own Up-Dn. Trans.	Other L.-Rt. Trans.	Altr. Own-Oth. Trans.
Implications of Types													
Altruistic	=5.8	=5.8	=1.82	=1.82	Long	Long	—	—	=45%	=45%	—	—	=10%
Pessimistic	=3.9	—	=1.82	—	Long	Short	=50%	=50%	=45%	—	=31%	—	=10%
Naïve	=5.8	—	=1.82	—	Long	Short	=50%	=50%	=45%	—	˜31%	—	=10%
Optimistic	=5.8	—	=1.82	—	Long	Short	=50%	=50%	=45%	—	—	—	=10%
L2	=2.4	=5.8	=1.82	=1.82	Long	Long	=50%	=50%	=45%	=45%	=31%	˜31%	=10%
D1	=4.6	=2.0	=1.82	=1.82	Long	Long	—	=50%	=45%	=45%	=31%	=31%	=10%
D2	=4.5	=2.4	=1.82	=1.82	Long	Long	—	=50%	=45%	=45%	=31%	=31%	=10%
Equilibrium	=4.1	=3.6	=1.82	=1.82	Long	Long	—	=50%	=45%	=45%	=31%	=31%	=10%
Sophisticated	=5.8	=4.2	=1.82	=1.82	Long	Long	—	=50%	=45%	=45%	=31%	=31%	=10%
Random	—	—	1.82	1.82	—	—	50%	50%	45%	45%	31%	31%	10%
Aggregate Look-up Measures													
TS	19.0	15.7	6.88	7.33	0.60	0.45	68.3%	83.9%	84.2%	81.6%	63.3%	69.3%	5.1%
All Baseline	16.8	14.6	5.46	5.95	0.67	0.60	72.8%	78.5%	79.7%	77.5%	31.6%	42.9%	8.5%
Altruistic	24.4	26.5	2.20	2.26	0.48	0.44	88.9%	33.3%	33.5%	38.0%	21.0%	60.0%	36.8%
Pessimistic	—	—	—	—	—	—	—	—	—	—	—	—	—
Naïve/Optim.	13.7	8.4	6.76	6.03	0.82	0.69	96.0%	77.3%	84.9%	80.5%	21.1%	48.3%	4.9%
L2	18.0	17.2	5.80	7.13	0.59	0.52	58.5%	87.9%	84.7%	83.0%	39.4%	30.3%	6.2%
D1	14.6	12.8	3.74	3.73	0.81	0.76	70.6%	54.8%	70.4%	68.3%	28.3%	61.7%	14.5%
D2	—	—	—	—	—	—	—	—	—	—	—	—	—
Equilibrium	18.4	13.4	4.05	3.67	0.55	0.51	100.0%	72.2%	72.0%	69.5%	21.5%	79.0%	5.3%
Sophisticated	—	—	—	—	—	—	—	—	—	—	—	—	—

Table V
Aggregate Rates of Compliance with Types' Occurrence and Adjacency for TS and Baseline Subjects, and for Baseline Subjects by Most Likely Type Estimated from Decisions Alone, in percentages (— vacuous)

Treatment (# subjects)	Altruistic *J*=H,M,L,0	Pessimistic *i*=H,M,L,0	Naïve *i*=H,M,L,0	Optimistic *i*=A,0	L2 *i*=H,M,L,0	D1 *i*=H,M,L,0	D2 *i*=H,M,L,0	Equilibrium *i*=H,M,L,0	Sophisticated *i*=H,M,L,0
TS (12)	**3.10.50.27**	**44.7.36.13**	**83.2.0.15**	**86.14**	**76.2.0.22**	**92.3.1.5**	**92.3.1.5**	**96.1.1.3**	**75.1.1.24**
Baseline (45)	**14,11,51,24**	**74,2,11,14**	**78,4,4,14**	**85,15**	**67,14,5,14**	**52,19,15,14**	**50,19,15,14**	**42,23,19,16**	**39,21,20,21**
***Altruistic* (2)**	**78.6.11.6**	**56.8.33.3**	**53.3.42.3**	**97.3**	**47.8.39.6**	**36.6.56.3**	**33.8.56.3**	**31.11.56.3**	**28.14.56.3**
***Pessimistic* (0)**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**
Naïve/Optim.	**9,5,53,33**	**85,1,9,5**	**89,5,3,4**	**96,4**	**42,24,3,31**	**45,22,20,13**	**43,18,23,16**	**26,24,28,23**	**23,23,27,27**
***L2* (23)**	**8,12,58,22**	**72,2,9,17**	**78,3,0,18**	**80,20**	**85,6,3,6**	**57,20,9,15**	**54,21,10,15**	**49,24,12,15**	**46,22,12,20**
***D1* (7)**	**23,21,26,29**	**59,3,16,23**	**63,7,6,23**	**77,23**	**53,21,6,21**	**48,17,14,20**	**45,19,15,21**	**42,20,17,21**	**38,14,21,27**
***D2* (0)**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**
Equilibrium	**6,8,86,0**	**100,0,0,0**	**97,3,0,0**	**100,0**	**64,36,0,0**	**69,17,14,0**	**67,19,14,0**	**56,25,19,0**	**53,19,28,0**
Sophisticated	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**	**—,—,—,—**

References

Algaze [Croson], Rachel (1990): "A Test of Presentation Effects on Strategy Choice," B. A. Honors Thesis, University of Pennsylvania.

Aumann, Robert, and Adam Brandenburger (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161-1180.

Beard, T. Randolph, and Richard Beil (1994): "Do People Rely on the Self-interested Maximization of Others?: An Experimental Test," *Management Science*, 40, 252-262.

Brandts, Jordi, and Charles Holt (1995): "Limitations of Forward Induction and Dominance: Experimental Evidence," *Economics Letters*, 49, 391-95.

Cachon, Gerard, and Colin Camerer (1996): "Loss Avoidance and Forward Induction in Experimental Coordination Games," *Quarterly Journal of Economics*, 111, 165-194.

Camerer, Colin, Eric Johnson, Talia Rymon, and Sankar Sen (1993): "Cognition and Framing in Sequential Bargaining for Gains and Losses," 27-47 in Kenneth Binmore, Alan Kirman, and Piero Tani (editors), *Frontiers of Game Theory*. Cambridge: MIT Press.

Card, S. K., T. P. Moran, and A. Newell (1983): *The Psychology of Human-Computer Interaction.* Hillsdale N. J.: Lawrence Erlbaum.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross (1990): "Selection Criteria in Coordination Games: Some Experimental Results," *American Economic Review*, 80, 218-233.

Costa-Gomes, Miguel, Vincent Crawford, and Bruno Broseta (1998): "Cognition and Behavior in Normal-Form Games: an Experimental Study," Discussion Paper 98-22, University of California, San Diego (available as a pdf file at http://weber.ucsd.edu/~vcrawfor/).

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

El-Gamal, Mahmoud, and David Grether (1995): "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association*, 90, 1137-1145.

Gilboa, Itzhak, Ehud Kalai, and Eitan Zemel (1993): "The Complexity of Eliminating Dominated Strategies," *Mathematics of Operations Research*, 18, 553-565.

Harless, David, and Colin Camerer (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62, 1251-1289.

Harless, David, and Colin Camerer (1995): "An Error Rate Analysis of Experimental Data Testing Nash Refinements," *European Economic Review*, 39, 649-660.

Haruvy, Ernan, Dale Stahl, and Paul Wilson (1999): "Evidence for Optimistic and Pessimistic Behavior in Normal-Form Games," *Economics Letters*, 63, 255-259.

Ho, Teck Hua, and Keith Weigelt (1996): "Task Complexity, Equilibrium Selection, and Learning: An Experimental Study," *Management Science*, 42, 659-679.

Ho, Teck Hua, Colin Camerer, and Keith Weigelt (1998): "Iterated Dominance and Iterated Best Response in Experimental 'P-Beauty Contests'," *American Economic Review*, 88, 947- 969.

Holt, Debra (1999): "An Empirical Model of Strategic Choice with an Application to Coordination Games," *Games and Economic Behavior*, 27, 86-105.

Johnson, Eric, Colin Camerer, Sankar Sen, and Talia Rymon (1996): "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining," manuscript, The Wharton School, University of Pennsylvania.

Kagel, John, and Alvin Roth, editors (1995): *Handbook of Experimental Economics*. Princeton University Press: Princeton, New Jersey.

Knuth, Donald, Christos Papadimitriou, and John Tsitsiklis (1988): "A Note on Strategy Elimination in Bimatrix Games," *Operations Research Letters*, 7, 103-107.

McKelvey, Richard and Thomas Palfrey (1992): "An Experimental Study of the Centipede Game," *Econometrica*, 60, 803-836.

Nagel, Rosemarie (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85, 1313-1326.

Payne, John, James Bettman, and Eric Johnson (1993): *The Adaptive Decision Maker*. Cambridge, U.K.: Cambridge University Press.

Roth, Alvin (1987): "Bargaining Phenomena and Bargaining Theory," 14-41 in Alvin Roth (ed.), *Laboratory Experimentation in Economics: Six Points of View*. New York: Cambridge University Press.

Roth, Alvin and Michael Malouf (1979): "Game-Theoretic Models and the Role of Information in Bargaining," *Psychological Review*, 86, 574-594.

Roth, Alvin and J. Keith Murnighan (1982): "The Role of Information in Bargaining: An Experimental Study," *Econometrica*, 50, 1123-1142.

Schotter, Andrew, Keith Weigelt, and Charles Wilson (1994): "A Laboratory Investigation of Multiperson Rationality and Presentation Effects," *Games and Economic Behavior*, 6, 445-468.

Selten, Reinhard (1998): "Features of Experimentally Observed Bounded Rationality," *European Economic Review*, 42, 413-436.

Stahl, Dale (1999): "Evidence Based Rule Learning in Symmetric Normal Form Games," *International Journal of Game Theory*, 28, 111-130.

Stahl, Dale, and Paul Wilson (1994): "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior and Organization*, 25, 309-327.

Stahl, Dale, and Paul Wilson (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10, 218-254.

Van Huyck, John, Ray Battalio, and Richard Beil (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80, 234-248.

Van Huyck, John, John Wildenthal, and Ray Battalio (2000): "Tacit Cooperation, Strategic Uncertainty, and Coordination Failure: Evidence from Repeated Dominance Solvable Games," *Games and Economic Behavior*, in press.

Appendix A: Explanation of Information Search Measures in Table IV

This appendix explains how to derive our types' implications for the 13 information search measures in Table IV under simple assumptions about cognition and information search.

The implications about the minimal numbers of own and other's look-ups follow by enumeration from Occurrence. Some of their relationships are easily understood from general principles. *Naïve*, *Optimistic*, and *Pessimistic* have no implications for other's look-ups because other's payoffs are irrelevant for them. *Altruistic*, *Naïve*, *Optimistic*, and *Sophisticated* must always look at all own payoffs (with a minor exception involving bounds for *Optimistic*), and *Altruistic* and *L2* must always look at all other's payoffs; their common "= 5.8" entries are based on the numbers of own or other's payoffs, averaged over games. The minimal number of own look-ups for *Pessimistic* is somewhat lower because it can avoid the need for some look-ups more often, using bounds as for *Optimistic*. *Equilibrium* requires still fewer own look-ups because it can sometimes avoid the need for many look-ups by identifying dominance for its partner. (*Sophisticated*, by contrast, must always check for dominance relationships among own decisions when forming beliefs, even when its partner has a dominant decision.) Both *Sophisticated* and *Equilibrium* require fewer other's than own look-ups, on average, because identifying an own dominant decision makes other's look-ups unnecessary, but not vice versa.[76]

D1 and *D2* require fewer own look-ups than *Altruistic*, *Naïve*, *Optimistic*, and *Sophisticated* because identifying a dominant decision for their partner sometimes allows them to avoid some own look-ups. However, *D1* and *D2* require more own look-ups than *Equilibrium* because identifying their naive best responses sometimes requires more own look-ups than identifying equilibrium decisions. *D1* and *D2* require fewer other's than own look-ups because after checking for dominance, they must compare the expected payoffs of own but not other's decisions. *L2*, by contrast, requires fewer own than other's look-ups—fewer own look-ups than any other type—because it starts by identifying other's *Naïve* decision, which requires all other's payoffs, and then best responds to it, which requires own payoffs for only one other's decision. The large difference in *D1*'s and *L2*'s search implications is surprising, because both respect two rounds of iterated dominance, and therefore make the same

[76]Random look-ups have no implications for the minimal numbers of look-ups, because they take the total numbers as given. Adjacency implies no further restrictions on minimal numbers of look-ups.

decisions in all but two of our games. This separation plays an important role in Section 4.D's econometric analysis.

Table IV's implications for average string lengths follow from the assumption that a type's relevant comparisons are more frequently represented by adjacent pairs in the look-up sequence than in a random sequence with the same total number of look-ups. This implies average string lengths at least as long as random for *Pessimistic*, *Naïve*, *Optimistic*, *L2*, *D1*, *D2*, *Equilibrium*, and *Sophisticated*, for which no relevant comparisons cross the boundary between own and other's payoffs; and at most as long as random for *Altruistic*, for which all relevant comparisons do so. Average string length would approach 2.0 for long sequences of random look-ups if transitions to the same payoff were as likely as to other payoffs, because there are as many own as other's payoffs, on average. However, our subjects hardly ever returned immediately to the same payoff.[77] If we elevate this empirical regularity to an assumption, average string length approaches a limit less than 2.0, which depends on the numbers of decisions, as the total number of look-ups increases. An easy calculation, assuming equal numbers of look-ups in each game and averaging over games, yields a limiting average string length of 1.82 for random look-ups. Our subjects' look-up sequences were long enough to make this limit an appropriate benchmark.

The implications for look-up durations, or *gaze times*, follow immediately from the assumption that a type's relevant look-ups have longer average gaze times than other look-ups.[78]

The implications about the frequencies of inspecting own payoffs first and last follow from the assumptions that first and last look-ups are more likely than not to be relevant, and that a type's relevant look-ups appear more frequently, on average, than in a random look-up sequence, for which first and last look-ups are equally likely to be of own and other's payoffs. For *Altruistic* these assumptions imply no restrictions on first or last look-ups. For *Pessimistic*, *Optimistic*, and *Naïve* only own payoffs are relevant, hence first and last look-ups are more likely to be of own payoffs. For *L2* more other's than own payoffs are relevant, hence first look-ups are more likely to be of other's payoffs; while for *D1*,

[77]We can distinguish returning from staying because clicking is required to close as well as open boxes.
[78]Standard decision-theoretic notions have no implications for gaze time because they focus on the information look-ups reveal, which is independent of gaze time provided that (as here) it suffices for comprehension. Subjects seem to make many irrelevant look-ups out of curiosity, and these may have shorter gaze times than relevant look-ups; but they could also have longer gaze times if subjects make relevant comparisons via brief, frequently repeated look-ups.

D2, *Equilibrium*, and *Sophisticated*, own and other's payoffs are usually both relevant, hence there is no presumption about first look-ups. For *Equilibrium* the last payoff relevant to identifying its decision via iterated dominance is an own payoff, while other methods are neutral on this point; and for *D1*, *D2*, *L2*, and *Sophisticated* the last relevant information is always from an expected-payoff or dominance comparison of own decisions. Thus for these five types, last look-ups are more likely than not to be of own payoffs.

Table IV's implications for the frequencies of transitions from own to own and other's to other's payoffs, of up-down transitions in own payoffs and left-right transitions in other's payoffs, and of altruistic own-to-other's transitions, all follow from the assumption that a type's relevant comparisons are more frequently represented by adjacent pairs in the look-up sequence than in a random sequence with the same total number of look-ups. For random look-ups the expected frequencies of those transitions, averaged across games, are 45.0%, 45.0%, 30.6%, 30.6%, and 10.0%, respectively, again assuming that subjects never return immediately to the same payoff.

Because the associated comparisons are all more likely than not to be relevant, the frequencies of own to own and other's to other's payoff transitions for *L2*, *D1*, *D2*, *Equilibrium*, and *Sophisticated*; of up-down in own and left-right in other's payoff transitions for *D1*, *D2*, *Equilibrium*, and *Sophisticated*; of up-down in own payoff transitions for *L2*; of own to own transitions for *Naïve*, *Optimistic*, and *Pessimistic*; and of Altruistic own to other's transitions for *Altruistic* should all be at least random. Because comparisons associated with left-right in own transitions are more likely than not to be relevant for *Pessimistic*, its frequency of up-down in own transitions should be at most random. Because comparisons associated with own to other's transitions are irrelevant for all types other than *Altruistic*, their frequencies should be at most random for those types. Because *Naïve* decisions (either for *Naïve* or for *L2*'s partner) can be identified equally well by left-right or up-down comparisons in own payoffs, the frequencies of up-down in own transitions for *Naïve* and left-right in other's transitions for *L2* should be approximately random. Finally, because there are no relevant comparisons associated with up-down in own transitions for *Optimistic*; with left-right in other's transitions for *Naïve*, *Optimistic*, or *Pessimistic*; or with up-down in own and left-right in other's transitions for *Altruistic*, there is no presumption about the frequencies of such transitions.