A Learning-based Model of Repeated Games with Incomplete Information

Juin-Kuan Chong

The NUS Business School, National University of Singapore
bizcjk@nus.edu.sg

Colin F. Camerer

Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125
camerer@hss.caltech.edu

Teck H. Ho

Haas School of Business, University of California, Berkeley
hoteck@haas.berkeley.edu

Abstract

This paper tests a learning-based model of strategic teaching in repeated games with incomplete information. The repeated game has a long-run player whose type is unknown to a group of short-run players. The proposed model assumes a fraction of 'short-run' players follow a one-parameter learning model (self-tuning EWA). In addition, some 'long-run' players are myopic while others are sophisticated and rationally anticipate how short-run players adjust their actions over time and "teach" the short-run players to maximize their long-run payoffs. All players optimize noisily. The proposed model nests an agent-based quantal-response equilibrium (AQRE) and the standard equilibrium models as special cases. Using data from 28 experimental sessions of trust and entry repeated games, including 8 previously unpublished sessions, the model fits substantially better than chance and much better than standard equilibrium models. Estimates show that most of the long-run players are sophisticated, and short-run players become more sophisticated with experience.

Key words: repeated games, self-tuning experience-weighted attraction learning, quantal response equilibrium

[1] This research was supported by NSF grant SES-0078911. Thanks to John Kagel

1 Introduction

Many transactions in the economy are conducted repeatedly by players who either know the history of behavior by others and anticipate future interactions. Examples include cartels, employment relations, merchant banking relationships, long-standing corporate rivalries, customers who are loyal to retailers, lending to customers with known credit histories, and so forth. Game theorists model these situations as repeated games with incomplete information and study their sequential equilibria (SE).

Two early experimental studies evaluated the accuracy of SE predictions in repeated trust games (Camerer and Weigelt (1988a)) and entry deterrence games (Jung, Kagel, and Levin (1994)). In these games, a long-run player is matched repeatedly with a group of short-run players. The long-run player can be one of the two types (normal or special). The short-run players know the proportions of the two types, but do not know which type of the long-run player they face.

In the trust game, a single borrower B (i.e., the long-run player) wants to borrow money from a series of 8 lenders denoted L_i ($i = 1, \ldots, 8$) (i.e., the short-run players) (cf. Kreps (1990)). A lender makes only a single lending decision (either *Loan* or *No Loan*). The borrower makes a string of decisions, (either *Repay* or *Default*), each time a lender chooses *Loan*.

The entry game deterrence is similar. A series of eight short-run entrants each decide, one at a time, whether to enter or stay out in a series of periods. If the entrant enters in a period, a long-run incumbent decides whether to fight or share.

The payoffs in the trust game imply that if the games were only one stage, the borrower would *Default*; anticipating this, the rational lender would choose *No Loan*. Similarly, in a one-stage entry deterrence game the entrant would enter because she would anticipate that the incumbent would share. The special types of borrowers and incumbents have payoffs which create a preference for repaying or fighting, respectively, rather than defaulting or sharing.

Both experimental studies showed three empirical regularities:

for rapidly supplying data and to Drew Fudenberg, Qi-Zheng Ho, David Hsia, Xin Wang, and four referees, for discussions and help. Useful comments were received from seminar participants at Berkeley, Caltech, Chicago, Harvard, Hong Kong UST, New York University, Pittsburgh, Princeton, and Wharton. We would also like to thank the editors and two anonymous reviewers for their helpful comments and suggestions

(1) The basic patterns predicted by SE occur in the data: In the trust game, borrowers are more likely to default in later rounds than in earlier rounds, and lending rates fall after a previous default. Similarly, incumbents are more likely to share in later rounds, and entry rates increase after sharing.
(2) There are two systematic deviations from the SE predictions: (a) There are too few defaults (by borrowers) and too few fights (by incumbents); and (b) the predicted rates of lending and entering increasing smoothly across rounds, while the SE predicts a step function across periods.
(3) In the experiments, subjects played 50-100 eight period sequences. Equilibration occurred across sequences ("cross-sequence learning") and between experimental sessions (experienced subjects were closer to SE than inexperienced subjects).

Camerer and Weigelt (1988a) and Jung et al. (1994) showed that the SE prediction could be modified to explain both the basic patterns (1) and the deviation (2a) above by assuming that some proportion of normal-type players acted like the special types induced by the experimenter (the "home-made prior").

These early analyses fell short in three ways: First, the prior was inferred from the data rather than measured separately in one-stage games. Second, the SE predictions of trust and entry rely on two different special types of opposite behavioral kinds– one is trustworthy (sacrificing money to help others) and the other is vindictive (sacrificing money to harm others). Third, the modified SE model with a homemade prior cannot explain deviation (2b) and the cross-sequence learning (3). Authors of both studies recognize that the modified SE model cannot explain cross-sequence learning. As (Camerer and Weigelt, 1988a, p. 27-28) note [2]:

> "...the long period of disequilibrium behavior early in these experiments raises the important question of how people learn to play complicated games. The data could be fit to statistical learning models (e.g., Selten and Stoecker (1986)), though new experiments or new models might be needed to explain learning adequately."

Responding to Camerer and Weigelt's call for new learning models, this paper develops and estimates a learning-based model with strategic "teaching". In the model, a fraction of short-run players learn adaptively from experience and the rest are "sophisticated" [3] – they rationally anticipate how the long-

[2] And see (Jung et al., 1994, p. 90)

[3] See Selten (1991), Milgrom and Roberts (1991), and Fudenberg and Kreps (1990) for models of sophistication. Adding sophistication to adaptive learning makes sense because long-run player subjects often have a sense that short-run players are learning. Models with sophistication also predict that players care about the payoffs of others, and how they are matched with partners in the future (adaptive learning

run players learn and behave. Similarly, a fraction of long-run players are sophisticated and the rest are myopic (they act as if they are playing a one-stage game). In repeated games with partner matching, sophisticated long-run players have an incentive to "teach" the short-run learners what to expect. This kind of "strategic teaching" has been proposed as a boundedly rational theory of reputation formation (see Fudenberg and Levine (1989), Watson (1993) and Watson and Battigali (1997)). Camerer et al. (2002) offer the first empirical implementation of such a model using data from repeated trust games. [4]

This general model both extends simple adaptive learning models, by adding sophistication, and weakens equilibrium models, by adding learning. Because the model mixes adaptive and sophisticated types, certain parameter restrictions reduce the model to boundary cases of special interest. Purely adaptive learning is one boundary case. When all players are sophisticated, believe all others are sophisticated, and best-respond, the model reduces to simply another boundary case – Bayesian Nash equilibrium. We study an Agent Quantal Response Equilibrium (AQRE) version of Bayesian-Nash equilibrium. In AQRE, players optimize noisily but update their beliefs using Bayes' rule and anticipate accurately what others will do (McKelvey and Palfrey (1998)).

Adding adaptive players make sense because there are behavioral differences between sophisticated and adaptive players. Consider the lender in the trust games. Adaptive lenders will continue lending until a default occurs, after which later lenders are less likely to lend. A sophisticated lender, in contrast, anticipates default by assessing the probability of the borrower being a normal ("dishonest") type. Hence she will stop lending when the posterior probability of dishonest type is high enough that the expected payoff from lending exceeds not lending. This could happen even without a default in previous rounds. In short, adaptive players *react* to past default behavior but sophisticated players *anticipate* future default behavior.

The general model has the potential to improve the modified SE used in earlier papers, which fell short in three ways:

---

models don't have these properties), consistent with experimental evidence (Partow and Schotter (1993), Mookherjee and Sopher (1994), Cachon and Camerer (1996), Andreoni and Miller (1993)). Models including sophistication have generally fit better in matrix games, signaling games, repeated trust games, and p-beauty contest games (see Stahl (1999), Cooper and Kagel (2001), Camerer, Ho, and Chong (2002), Ho, Camerer, and Weigelt (1998)).

[4] Their model consists solely of adaptive short run players who follow a parametric EWA model (which requires a total of 18 parameters). They do not allow their long run players to be a special home-made type. However, the special home-made type is estimated in the benchmark AQRE model. The models are validated on one trust game dataset from Camerer and Weigelt (1988a).

First, the value of the home-made prior is measured separately, in one-stage experiments where potential reputation effect is absent, rather than estimated from the repeated games.

Second, the model provides a unifying theory of "special types" across different games. Both types of special-type players– trustworthy borrowers and fighting lenders– act like Stackelberg players: They choose the strategy they would commit to, if they could, in order to improve long-run payoffs. (This is the essence of the models of Fudenberg and Levine (1989)) Our model derives the two different types *endogenously* from the game payoff structures and a single common source: Both are special types whose behavior is similar to that of sophisticated long-run players who maximize long-run payoffs. Even though the impact of teaching is quite different between trust and entry games (payback in trust game is mutually beneficial while fighting in entry game is privately beneficial), the model captures both impacts across games with no additional parameter.

Third, although switching from SE to AQRE improves fit and explains the deviation (2b), it cannot explain learning across sequences within an experimental session, and learning across sessions, which the general model can. Cross-sequence learning can be explained by allowing subjects to learn both from previous periods within an eight-period sequence, *and* from previous eight-period sequences.

In this paper, we apply the general model to the 20 experimental sessions published earlier on trust and entry deterrence games, and to 8 brand new sessions. The new data provide additional replication of the basic patterns and give us more statistical power. We estimate that more than 90% of the long-run players in both games are sophisticated. About half of the short-run players are estimated to be sophisticated in sessions with inexperienced subjects, but all the short-run players are estimated to be sophisticated after experience.

To verify that the general model captures the trends in the data and to understand the impact of each feature of the model in tracking the data, we simulate the behavior of the model under various parameter restrictions (i.e., after "disabling" features one at a time) and compare with the data.

The three empirical regularities discussed above can be translated into cross-round trends and cross-sequence trends. We find that the model tracks data well in both cross-round and cross-sequence trends: when there is a significant trend in the data, the model picks up the trend as well.

We disable four key features of the model one at a time: cross-sequence learning, the proportion of sophisticated lenders, the home-made prior, and the proportion of sophisticated borrowers. And we find that disabling the features has significant impact on the ability of the model to pick up the trends in the data. In some cases, the prediction of the restriction model either does not pick up the trend at all or predicts an opposite trend.

The next section introduces the model of repeated games and reports new experiments that measure the proportions of special types. Section 3 discusses the key differences among equilibrium, AQRE and the proposed model. Section 4 reports estimates of the models on three data sets from repeated trust and entry-deterrence games. Section 5 checks the robustness of the model through simulation. Section 6 concludes.

2 A Model of Repeated Games

We consider any two-player repeated game with incomplete information, where the long-run player can be one of two types (or equivalently, have one of the two induced payoff functions) and short-run player is uncertain about long-run player's type. In repeated borrower-lender trust relationships, a lender is uncertain about whether a borrower is honest or dishonest. In repeated incumbent-entry games, an entrant is uncertain whether an incumbent will always fight entry or not. We refer to the honest borrowers and aggressive borrowers as special types. Standard equilibrium analysis assumes both players are sophisticated and behave according to the prediction of Bayesian Nash equilibrium.

Table 1 shows the various player segments in the proposed model. p fraction of long-run players are induced to be special type and $(1-p)$ fraction to be normal type. Of the normal type players, a fraction θ has an inherent preference for special type's payoff function, a fraction $(1-\theta)\cdot\alpha_B$ are sophisticated and a fraction $(1-\theta)\cdot\alpha_B$ are myopic. A α_L fraction of short-run players are sophisticated and the remaining $1-\alpha_L$ are adaptive. If $\alpha_L=\alpha_B=1$, the model reduces to AQRE. If $\alpha_L=\alpha_B=0$, the model reduces to the self-tuning EWA learning model.

The proposed model allows a fraction θ of long-run player's with normal-type payoff to act like the special-type payoff (this fraction is previously labeled as homemade prior). Along with the fraction p of the borrower players who are induced to behave like special types by the experimenter, the total fraction who actually behave like special types is $p+(1-p)\theta$.

Earlier experiments imputed a value for the homemade prior (Camerer and Weigelt (1988a), Neral and Ochs (1992)) or estimated it from a structural model (Palfrey and Rosenthal (1988), McKelvey and Palfrey (1992)). We *measured* the frequency of homemade prior in two separate experimental sessions with one-shot games with random rematching (using the same subject pools used in the early trust experiments). In these games, there is no reputational incentive for behaving like special type players. The measured rate of those behaviors is then used to constrain their frequency in the repeated game estimation.

In a typical experimental session, subjects are randomly assigned fixed roles of borrower, or lender (e.g., 11 subjects are divided into 3 borrowers and 8 lenders). In a single sequence, a borrower B is randomly chosen to play in all the periods of an eight-period supergame (the other borrowers sit and watch). In addition, the borrower may be payoff-induced to be an honest type with probability p where p is set by the experimenter a priori. A borrower type remains the same for all 8 periods of the sequence. Each lender L_i plays in exactly one of the 8 stage games in each supergame in a random position each time (the position of a particular lender-subject in each sequence is unknown to the borrower). The entire eight-stage supergame is repeated in a series of *sequences* (typically 50 to 100 sequences).

We model the choice probabilities of each segment f (f assumes a value of a for adaptive learner and a value of s for sophisticated player) of players at time t. In specifying the probability, we adopt the logistic approach in which lenders of segment f attach an attraction value $A_L^j(f,k,t)$ to each strategy j in a given round t of a sequence k. Similarly, borrowers of segment f' (f' assumes a value of m for myopic player, a value of s for sophisticated player, and a value of h if the borrower behaves like an honest type) have an attraction value $A_B^j(f',k,t)$ to each strategy j in a given round t of a sequence k. Below, we will discuss how $A_L^j(f,k,t)$ and $A_B^j(f',k,t)(i = L, B)$ are determined for each segment of the players in Table 1.

2.1 Adaptive Lenders

Recall that lenders play only once in each sequence. Yet they clearly respond to the experiences of the other players, which they only observe. So we assume "observational learning": Players can learn from previous rounds in a sequence *and* from previous sequences. Consider round 7 in sequence 14. The round 7 lender who is deciding what to do saw what happened in the previous 6 rounds of sequence 14, and learned about the attractiveness of lending from what happened in those rounds. But the lender also knows what happened in the upcoming (7th) round of the previous sequences 1-13 –a glance at the past– and learned about whether she should loan in round 7 from those previous round 7 experiences. We call the latter effect *cross-sequence learning.*

Within-sequence learning can be modeled by standard learning theories. We use a "self-tuning" EWA model of Ho, Camerer, and Chong (2004) for its parsimony (with only 1 parameter) and versatility (it has predicted reasonably accurately in other games). (Other adaptive models could be used in its place as well).

Returning to our example, the strategy *loan* for a lender before period 7 of sequence 14 is influenced by two sources of experience –the attraction of *loan* after period 6 of sequence 14, and the experience after choosing *loan* in period 7 of sequences 1-13. These influences are captured by differentially updating the attractions of the strategies.

The strength of cross-sequence learning is parameterized by a parameter τ. If $\tau = 0$ there is no cross-sequence learning; if $\tau = 1$ experience in upcoming periods of previous sequences is just as important as experience in the previous period of the current sequence. The data will tell us how strong cross-sequence learning is through the value of τ.

The updating of the attraction for an adaptive lender $A_L^j(a,k,t)$ occurs in 2 steps. The idea is to create an "interim" attraction for round t, $B_L^j(a,k,t)$, based on the attraction $A_L^j(a,k,t-1)$ and payoff from the round t, then incorporate experience in round $t+1$ from previous sequences, transforming $B_L^j(a,k,t)$ into a final attraction $A_L^j(a,k,t)$. The exact specification of the attraction updating is as follows:

(1) Learning across rounds within a sequence:

$$B_L^j(a,k,t) = \frac{\phi(k,t)\cdot N(k,t-1)\cdot A_L^j(a,k,t-1)}{M(k,t)} + \frac{[\delta_j(k,t)+[1-\delta_j(k,t)]\cdot I(j,s_L(k,t))]\pi_L(j,s_B(k,t))}{M(k,t)} \quad (1)$$

$$M(k,t) = \phi(k,t)\cdot N(k,t-1)+1 \quad (2)$$

where $\phi(k,t)$ and $\delta(k,t)$ are functional parameters and $N(k,0) = 1$, as specified in Ho et al. (2004) (see Appendix for further details).[5] The initial attraction $A_L^j(a,k,0) = A_0$(No Loan/Not Enter) for the strategy $j =$ No Loan (Not Enter) is estimated. $I(j,s_L(k,t))$ is the indicator function

[5] In Camerer et al. (2002), adaptive short run players follows a parametric EWA model with a fixed set of parameter estimates. Having a fixed set of learning parameters restricts model flexibility. It seem reasonable to assume that the adaptive player relies less on her past learning experience when she senses that her past experience does not help (lending no longer seems that attractive when more defaults are happening). The reliability of past experience does deteriorate in later rounds when default happens more often. Having fixed parameters hinders this learning flexibility.

that equals 1 if strategy j is the chosen strategy $s_L(k,t)$ of lender L in round t of sequence k, and equals 0 otherwise.

(2) Learning in a coming round from previous sequences:

$$A_L^j(a,k,t) = \frac{\phi(k,t)^\tau \cdot M(k,t) \cdot B_L^j(a,k,t) + \tau \cdot \delta(k,t)\hat{\pi}_L^j(k,t+1)}{N(k,t)} \quad (3)$$

$$N(k,t) = \phi(k,t)^\tau \cdot M(k,t) + \tau \quad (4)$$

where we assume that the learning about an upcoming round t from previous sequences is driven by the average payoff in the round t's in all previous sequences (i.e., $\hat{\pi}_L^j(k,t+1) = \sum_{m=1}^{k-1} \pi_L(j, s_B(m,t+1))/(k-1)$).

The attraction at the end of time period t then determines the predicted adaptive lender's choice probability at $t+1$ according to the logit rule,

$$\hat{P}_L^j(a,k,t+1) = \frac{e^{\lambda_L^a \cdot A_L^j(a,k,t)}}{\sum_{j'} e^{\lambda_L^a \cdot A_L^{j'}(a,k,t)}}. \quad (5)$$

2.2 Sophisticated Lenders

The sophisticated lender rationally anticipates the action of the borrower and maximizes her own expected payoff in each period. Let the lender's belief about the overall fraction of honest types at sequence k and end of time t be $r(k,t)$. Then, the remaining fraction $(1 - r(k,t))$ of borrowers are either myopic or sophisticated. Their combined predicted probability of choosing strategy j' at $t+1$ is as follows: $\hat{P}_B^{j'}(d,k,t+1) = [(1-\alpha_B) \cdot \hat{P}_B^{j'}(m) + \alpha_B \cdot \hat{P}_B^{j'}(s,k,t+1)]$. The expected payoff of lender for choosing strategy j is then given by:

$$A_L^j(s,k,t) = \sum_{j'} \left[(1 - r(k,t))\hat{P}_B^{j'}(d,k,t+1) + r(k,t)\hat{P}_B^{j'}(h)\right] \pi_L(j,j') \quad (6)$$

where $\pi_L(j,j')$ is the lender's payoff for strategy j when borrower chooses j'. $\hat{P}_B^{j'}(h)$ is the probability that an honest borrower chooses strategy j'.[6]

If the sophisticated lender chooses *loan*, she updates her belief in a Bayesian manner at the end of $t+1$ using the borrower's choice probabilities as follows:

[6] Notice that $\hat{P}_B^{j'}(h)$ does not depend on arguments k and t because the probability does not vary across periods.

$$r(k,t+1) = \frac{\hat{P}_B^{j'}(h)\cdot r(k,t)}{\hat{P}_B^{j'}(h)\cdot r(k,t)+\hat{P}_B^{j'}(d,k,t+1)\cdot(1-r(k,t))} \tag{7}$$

where j' is the chosen strategy.

If the lender chooses *noloan*, then $r(k,t+1) = r(k,t)$. Each lender starts at round 1 with the prior P(Honest), or $r(k,1) = p+(1-p)\theta$.

Updating the belief $r(k,t)$ changes the attractions $A_L^j(s,k,t)$ and captures learning. The updated attraction determines the sophisticated lender's choice probability according to the logit rule,

$$\hat{P}_L^j(s,k,t+1) = \frac{e^{\lambda_L^s\cdot A_L^j(s,k,t)}}{\sum_{j'} e^{\lambda_L^s\cdot A_L^{j'}(s,k,t)}}. \tag{8}$$

2.3 Honest Borrowers

Honest borrowers always earn more from repaying (by definition). They choose according to the stage game payoffs of honest type conditional on loan by the lender using a logit rule as follows:

$$\hat{P}_B^j(h) = \frac{e^{\lambda_B^h\cdot\pi_H(j,Loan)}}{\sum_{j'} e^{\lambda_B^h\cdot\pi_H(j',Loan)}}. \tag{9}$$

2.4 Myopic Borrowers

Since borrowers move after the lenders do, there is nothing for a borrower to learn. We call the borrowers who care only about immediate payoff "myopic". The attractions of *repay* and $default$ are simply the stage-game payoffs of dishonest type conditional on *loan* by the lender. They choose between those strategies using a logit rule:

$$\hat{P}_B^j(m) = \frac{e^{\lambda_B^m\cdot\pi_D(j,Loan)}}{\sum_{j'} e^{\lambda_B^m\cdot\pi_D(j',Loan)}} \tag{10}$$

2.5 Sophisticated Borrowers

The sophisticated borrower maximizes the total expected payoff from all remaining periods. The payoff from choosing strategy j in round t of sequence k is as follows:

$$A_B^j(s,k,t|r(k,t)) = \pi_B(j,j') + V_B(k,t+1|r(k,t+1)) \tag{11}$$

where $V_B(k,t+1|r(k,t+1))$ refers to the *ex ante* value of the borrower for all remaining rounds after $t+1$ of the game given the lender's posterior belief $r(k,t+1)$ at time $t+1$. Note that $r(k,t+1)$ above is determined by $r(k,t)$ and the probability of strategy j using equation 7.

The *ex ante* value of the borrower for future rounds can be specified recursively as follows:

$$\begin{aligned} V_B(k,t|r(k,t)) = \max_{J_t}(\sum_{j'} \bar{P}_L^{j'}(d,k,t|J_t) \cdot \sum_{j} \hat{P}_B^j(s,k,t) \cdot \\ [\pi_B(s_B^j, s_L^{j'}) + V_B(k,t+1|r(k,t+1))]); \end{aligned} \tag{12}$$

where J_t is the sequence of future actions by the sophisticated borrower from round $t+1$ until the end of the game sequence,[7] $J_t \equiv \{j_t, j_{t+1}, \ldots, j_{T-1}, j_T\}$. The lender's probability given a future path J_t is given by $\bar{P}_L^{j'}(d,k,t|J_t) = [(1-\alpha_L) \cdot \bar{P}_L^{j'}(a,k,t|J_t) + \alpha_L \cdot \bar{P}_L^{j'}(s,k,t|J_t)]$ where α_L is the proportion of sophisticated lender.

This future payoff term in 12 gives an incentive for the sophisticated borrower to influence the lender. That is, even if $\pi_B(repay,j')$ is lower than $\pi_B(default,j')$, the attraction $A_B^{repay}(s,k,t|r(k,t))$ may be higher because of the consequences of choosing *Repay* for $V_B(k,t+1|r(k,t+1))$. The idea is to make the lender want to lend in future rounds. This is accomplished by repaying in earlier rounds (with a view to teach the lenders), so that the sophisticated lender revises upward her prior on the borrower (through equation 7; cf Kalai and Lehrer (1993)) and the adaptive lender improves the attraction of the lending strategy (through equation 1).

[7] Technically, computing the teaching borrower attractions requires evaluating all the paths J_t to find the maximum. This is computationally cumbersome in early periods (e.g., in period 8 there are $2^8 = 64$ paths). To simplify computation, we maximize only over paths of future borrower actions that never have repayment following default, because repayments following default are rare and usually yield lower payoffs. (This reduces the number of paths to only nine – always repaying, plus repayment followed by defaulting in period t then always defaulting, for each t from 1 to 8.)

The final attraction then determines the lender's choice probability according to the logit rule,

$$\hat{P}_B^j(s,k,t+1) = \frac{e^{\lambda_B^s \cdot A_B^j(s,k,t)}}{\sum_{j'} e^{\lambda_B^s \cdot A_B^{j'}(s,k,t)}}. \tag{13}$$

2.6 Likelihood and Estimation

The models are estimated using 3 datasets: two trust game datasets from Camerer and Weigelt (1988a,b) and one entry deterrence dataset from Jung et al. (1994). All experimental sessions within a dataset are restricted to have a common set of parameters (except for the scale sensitivity parameters λs where each session has its own). Maximum likelihood estimation (MLE) was used to calibrate the model on 70% of the sequences in each experimental session, then forecast behavior in the remaining 30% of the sequences in that session. If the model fits in-sample purely by overfitting, it will perform surprisingly *poorly* out-of-sample.

The likelihood function used in estimation consists of three parts:

(1) The likelihood of observing the data of the lenders is as follows:

$$L_L = [(1-\alpha_L) \cdot \Pi_k \Pi_t \hat{P}_L^{s_L(k,t)}(a,k,t) + \alpha_L \cdot \Pi_k \Pi_t \hat{P}_L^{s_L(k,t)}(s,k,t)] \tag{14}$$

where $s_L(k,t)$ is the strategy actually chosen by lender L at time t in sequence k.

(2) For the sequences where an honest-type borrower is drawn (with probability p), the likelihood of observing the data of the borrowers is as follows:

$$L_B^H = \Pi_{k"} \Pi_t \hat{P}_B^{s_B(k",t)}(h) \tag{15}$$

where k" are the sequences with honest types and $s_B(k",t)$ is the strategy chosen by the borrower at time t in sequence k"

(3) For the sequences where an dishonest-type is drawn, the likelihood of observing the data of the borrowers is as follows:

$$L_B^D = \theta \cdot \Pi_{k'} \Pi_t \hat{P}_B^{s_B(k',t)}(h) + (1-\theta) \cdot \Big[(1-\alpha_B) \cdot \Pi_{k'} \Pi_t \hat{P}_B^{s_B(k',t)}(m) + \alpha_B \cdot \Pi_{k'} \Pi_t \hat{P}_B^{s_B(k',t)}(s,k',t)\Big] \tag{16}$$

where k' are the sequences with dishonest-type draws and $s_B(k',t)$ is the strategy chosen by the borrower at time t in sequence k'.

Finally, the total likelihood of observing all the data is given by $L_L \cdot L_B^H \cdot L_B^D$.

2.7 Measuring the Homemade Prior θ Experimentally

Earlier trust and entry experiments showed that even when the induced fraction of honest borrowers or fighter types is zero, there is a substantial repayment and fighting in finite games (even in the last period). Inspired by the "gang of four" model of Kreps and Wilson (1982), Camerer and Weigelt (1988a) suggested this was due to the presence of an endogenous fraction of subjects who, despite monetary incentives to default, simply preferred to act reciprocally and repay– a "homemade prior" of reciprocal types. Palfrey and Rosenthal (1988) used the same idea to explain contribution in public goods games. [8]

In Camerer et al. (2002), the homemade prior θ is estimated from the data as part of fitting a QRE model. The resulting estimates were high– from .5 to 1– compared to the values around .1-.2 suggested by early experiments. This probably means the QRE model needs to overestimate θ in order to make up for some other basic misspecification.

Since the homemade prior is intimately tied to the extent of repaying or fighting, it is important to estimate it precisely and plausibly. By definition, honest or aggressive types will repay or fight even in one-shot games (their behavior springs from preferences, not strategy). Therefore, we recently measured θ by conducting two experimental sessions of one-shot games, reproducing the original experimental conditions [9] from repeated games as closely as possible while generating enough data for a reliable estimate.

One session used the most common payoff structure in trust games and the other session used the most common structure in entry games (see Table 2). [10] Each session used 12 subjects playing two blocks of 6 rounds in a fixed-role protocol (as in the original experiments). In each block of six rounds, each borrower was matched with each lender once in a "zipper" design. Each bor-

[8] More recently, this intuition has been formalized in models of social preference used to explain contribution (and punishment) in public good games, reciprocity, rejections of ultimatum offers, and so forth (e.g., Fehr and Gächter (2000) and Camerer (2003, chap. 2)).
[9] The original experiments were run in 1986 and 1990 respectively.
[10] The lender's payoff used was -50 when the borrower reneges. This payoff is identical to trust data sessions 6-8 where $p = 0.1$ and new trust data sessions 1-7 where $p = 0.1$. The entrant's payoff used was 80 when the weak monopolist fights in market-entry games. This corresponds to market entry game sessions 1-3 (inexperienced) and 6 (experienced) where $p = 1/3$.

rower therefore plays the same lender twice, but never knows which lender she is playing. A total of 72 single-shot games were played in each experimental session.

Since the crucial behavior is repayment by borrowers, we used the "strategy method" in which borrowers chose whether to repay or default *before* knowing whether they received a loan. (Otherwise, repayment decisions are only observed when lenders lend, which severely limits the number of such decisions.)

Dollar payments were those used in the original experiments, adjusted upward for inflation.[11] In trust games there were 17 repayments (26%) and in entry games there were 11 fight choices out of 72 (18%). These percentages are close to the 17% figure originally imputed by Camerer and Weigelt (1988a).

Because these samples are modest in size, θ may not be estimated too precisely. Therefore, the estimation below restricts θ, as estimated in the repeated games, to lie in a 95% confidence interval of the values measured in the one-shot experiments. These intervals are (.19,.29) for trust and (.11,.20) for entry.

3 Special Cases

To provide a context on which the empirical results can be discussed, we first contrast several key characteristics of the SE AQRE, adaptive learning, and the general models.

The delicate logic of the repeated-game equilibrium can be illustrated with the trust game. Table 2 shows payoffs in the Camerer-Weigelt repeated trust game. Recall that a single borrower is drawn to play an 8-period sequence. Her type (either honest or dishonest) is drawn randomly using a commonly-known prior and communicated only to the borrower herself. The borrower then plays a sequence of stage games with eight lenders who play once each in random order.

In each stage game, the lender can choose not to lend (then both earn 10 currency units) or can choose to lend. Lenders prefer to lend if the borrower will repay, yielding 40 for the lender. But if the borrower defaults the lender earns -100.[12] A dishonest borrower earns 60 if she repays, and 150 if she defaults. Honest-type borrowers have the same payoffs except a default pays 0. Note that in the subgame after receiving a loan, the myopic dishonest borrower prefers to default while the honest borrower prefers to repay. The

[11] The original experiments were in 1986 and 1990, so we adjusted payments by the GDP deflator, increasing them by 50% and 23% respectively.
[12] In sessions 4-6, the lender's default payoff was -50. In sessions 7-8, it was -75.

probability that an borrower had honest-type payoffs in a particular sequence was varied across experimental sessions from 0.33 to 0.

The SE is computed from the last period backward (see Camerer et al. (2002) for details). In the last period, risk-neutral lenders lend if their perceived P(Honest) is above a threshold $\gamma = .79$. Anticipating this, normal borrowers mix in period 7 by repaying with enough probability to make the lender's updated P(Honest)=.79 in period 8, which makes lenders indifferent. Guessing accurately how borrowers will mix the lender's P(Honest) threshold in period 7 is γ^2. The same argument works by induction back to period 1. In each period the lender has a threshold of perceived P(Honest) which makes her indifferent between lending and not lending. The path of these threshold P(Honest) values is simply γ^{9-n} in period n. When the updated P(Honest) in period t is above the threshold in the period $t+1$, the lender always lend and normal borrowers always repay in period t. After that phase, lenders mix and borrowers default with increasing probability if they get a loan.

Besides this sharp restriction on equilibrium lending and default, Bayesian updating and optimization impose two more very strong restrictions: Since only normal borrowers default, after a default the borrower's type is revealed and players should neither lend nor repay after that. And after a later period in which there is no loan, the borrower misses an opportunity to improve her reputation, so players should neither lend nor repay after that period.

Jung et al. (1994) ran a 'chain-store' entry deterrence game with payoffs as shown in Table 2. With these parameters, the sequential equilibrium is very much like the one in the trust game: Fighting for a couple of periods (and entrants wisely staying out) followed by mixing, with an increasing tendency to share toward later periods.

The SE predicts that many events have zero probability (e.g., lending after a default). But these events are actually observed occasionally, so the likelihood function blows up unless *some* notion of error or trembling is added to the model. AQRE (McKelvey and Palfrey (1998)) adds trembling toward better-responses (noisily best-respond), and assumes that the agents understand the likely trembling that other players are doing. A homemade prior θ is also included into the AQRE model (because this proved useful in fitting data in the earlier analyses).

The AQRE model is implemented with four parameters– three different response sensitivities (λ's) for sophisticated lenders, honest borrowers, and sophisticated borrowers (since there are no adaptive lenders and myopic borrowers), and a perceived prior belief of lenders about P(honest) or P(aggressive) (restricted to be within the confidence interval determined by the new one-shot game data mentioned above). AQRE is a plausible benchmark and fits many other data sets well (e.g., McKelvey and Palfrey (1998), Goeree and Holt (1999), Ho et al. (1998)). However, it is noteworthy that even in the agent-based form, AQRE estimation in these data is much more computationally challenging than in any previous applications to extensive-form games, which have all used much simpler games with fewer nodes (see Camerer et al. (2002) for details).[13]

The general model relaxes the key AQRE assumptions that all players Bayesian-update belief and predict accurately the likely actions of others. The model allows for the existence of the strategically *un*-sophisticated players (those who learn adaptively or respond myopically). If $\alpha_L = \alpha_B = 1$, the general model reduces to AQRE.

The general model also nests a self-tuning version of the experience-weighted attraction model that has been used to fit and predict a wide range of experimental data (Camerer and Ho (1999), Ho et al. (2004)). If $\alpha_L = \alpha_B = 0$, all lenders are adaptive and borrowers are either honest or myopic. We do not report results of this special case because all three datasets use a fixed matching protocol for the long-run player and the parameters α_L and α_B are generally greater than 0.5, indicating the existence of a significant portion of sophisticated players.

4 Data and Results

This paper fits the general and AQRE models to experimental data from three sources. The first is eight experimental sessions of a repeated borrower-lender trust game reported by Camerer and Weigelt (1988a). The second is a previously unpublished sample of eight more sessions of the same game (with prior P(honest), $p = .10$) in which players also report beliefs about whether the borrower will default if there is a loan (Camerer and Weigelt (1988b)).[14] These data are called "new trust" games. The third is 12 sessions of an entry-deterrence game from Jung et al. (1994). Eight of the sessions use inexperienced subjects (participating in that particular game for the first time) and four use experienced subjects who returned for a second session

[13] QRE is computationally nightmarish with 64 strategies because solving it requires solving a system of 64 simultaneous equations with Bayesian updating between nodes. Rather than using a distribution over all $2^8 = 64$ supergame strategies, players choose a distribution of strategies at each node (as if each node is controlled by a separate "agent"); hence the modified AQRE. The model can then be approximated by computing beliefs and expected payoffs at each node using backward induction, which is a convenient shortcut.

[14] See Camerer and Weigelt (2005) for discussion of the beliefs.

playing the same game. There are a total of 28 experimental sessions, roughly 2,000 8-period sequences and 26,000 choices.[15]

4.1 Trust games

Typical patterns in the old trust data can be seen in Figures 1a-b (pooling across all sessions to reduce sampling error). Periods 1,...,8 denote periods in each sequence. The figures show relative frequencies of not lending (all data) and default (conditional on lending, for dishonest borrowers only), assuming there was no default earlier in the sequence. Sequences are combined into ten-sequence blocks (denoted "sequence" in the figures) and average frequencies are reported from those blocks.

Two patterns in the data are of primary interest. First, what is the rate of lending across periods (and how does it change across sequences)? Second, how do borrowers respond to loans in different periods (and how do these responses vary across sequences)? Figure 1a-b show that lenders start by generally making loans (i.e., low frequency of no-loan) in early periods, then learn to rarely lend (i.e., high frequency of no-loan) in periods 7-8. Borrowers rarely default in early periods, but frequently default in periods 7-8. The pattern of increasing default in later periods is particularly dramatic in later sequences so there is cross-sequence learning. Figures 2a-b show frequencies for the eight new trust sessions. The general pattern of results is similar to that in Figure 1 although $NoLoan$ and $Default$ choices are more common in earlier periods.

How well the models capture these patterns can be judged in two ways: 1) overall statistics measuring fit (log likelihood); and 2) reported parameter values. Table 3 summarizes log likelihoods (LL) for in-sample calibration and out-of-sample validation. The general model performs significantly better in-sample and out-of-sample than the AQRE in both old (1988a)and new (1988b)trust data. The fact that the per-period log likelihoods are similar in calibration and validation suggests that the general model does not overfit (if it did, validation would have a larger negative log likelihood). The general model seems to improve a little on AQRE by allowing a sizeable fraction of adaptive learners,

[15] Subjects in the trust games were either MBA students at NYU (in the original data) or undergraduates at the University of Pennsylvania (in the new trust data). They were paid an average of $18 for a 2-1/2 hour session. Instructions are available in Camerer and Weigelt (1988a). Subjects in the entry-deterrence games were University of Pittsburgh undergraduates. See Jung et al. (1994) for design details. Each session had 48-101 eight-period sequences. In each trust session, there were 11 subjects, three borrowers and eight lenders. In each entry-deterrence session, there were 7 subjects, three monopolists and four entrants.

which AQRE does not. [16]

Table 4 gives estimated parameter values. [17] The estimated percentages of sophisticated lenders α_L are 43% and 63%, respectively, for old and new trust data. The corresponding percentages of sophisticated borrowers are 100% and 95% for old and new trust data, suggesting that virtually all the long-run borrowers are teaching.

4.2 Entry-deterrence games

Now we turn to the Jung et al. data on entry-deterrence. Since they ran experiments both with inexperienced subjects and experienced subjects, we can see whether subjects grow more sophisticated when they repeat an entire experimental session.

Equilibrium predicts rates of entry and sharing to start low and rise as the end of a sequence draws near. Actual entry and sharing by inexperienced subjects are far too frequent in early periods but there is some convergence toward early entry-deterrence across the experimental session (see Figures 3a-b). Inexperienced entrants just didn't quite figure out how much it pays to fight entry in early periods.

Figures 4a-b show data from experienced subjects. The correspondence of behavior to equilibrium is much more dramatic. In the first sequence block, players often enter in the first 3 periods, but they quickly learn early entry is rarely met with sharing, and they stay out in early periods of later sequences.

[16] Note also that the model is almost as accurate when all sessions are pooled, with common parameters, as when fit statistics from session-specific estimation are totaled up, although 40 fewer parameters are estimated when data are pooled. (See our 2004 working paper for details of session-by-session estimation). This is a big hint that the parameter estimates are quite stable across sessions for the teaching model. Our earlier working paper reports two other comparisons. Allowing ϕ, δ, and κ to be free parameters (common within each data set) and estimating them, rather than deriving them from functions as self-tuning EWA does, improves out-of-sample accuracy slightly in hit rate and likelihood in most data sets. Fixing the homemade prior θ to zero hurts the likelihood substantially in two data sets and gives hit rates less than chance (below 50%) except in the inexperienced entry game data.

[17] Keep in mind that in self-tuning EWA ϕ, δ and κ are *not* estimated, they are functions of the data. The averaged functional values of ϕ, δ and κ are quite consistent across sessions. They are also in the ballpark of the values estimated in parametric EWA (see Ho et al. (2004)), except that the functional ϕ is always too high (.76-.77 compared to unconstrained estimates of .45 and .25). The fact that pooling across sessions degrades overall fit only a little, and parameters are consistent across the new and old trust data sets, is encouraging. See our working paper for details.

Summary statistics in Table 3 shows that the general teaching and AQRE models are about equally accurate for experienced subjects. With 100% of the borrowers and lenders sophisticated ($\alpha_L = \alpha_B = 1$), the general model reduces to AQRE. For inexperienced subjects, the general model is much more accurate than AQRE, reflecting the presence of adaptive lenders.

Table 4 shows estimated parameter values. The estimated fractions of sophisticated players are smaller for inexperienced subjects ($\alpha_L = .67$, $\alpha_B = .91$) than for experienced subjects ($\alpha_L = \alpha_B = 1$). This increase in sophistication is also observed by Stahl (1999) in matrix games, Camerer et al. (2002) in a dominance-solvable (p-beauty contest) game, and Cooper and Kagel (2001) in signaling games. This seems to be a robust finding, and a sensible one– players come to realize how others are learning after they play the same game in two consecutive sessions.

A challenging test for both the general and AQRE models is whether similar parameter values can be used to explain behavior in trust games and entry-deterrence games. These games are opposite in incentive structure in the sense that special type behavior (repaying or fighting) is mutually-beneficial in trust games but only privately-beneficial in entry-deterrence. If the same general model structure and parameters can explain both games that shows some robustness which encourages broader application. In fact, the trust and inexperienced entry data give similar values of self-tuning EWA parameters of ϕ (.76-.78) and δ (.15-.19). This is an encouraging first step towards a *general* learning-based theory of different repeated games.

Many results are consistent across both games. The AQRE model predicts rather well, but it is helped substantially by allowing the constrained homemade prior above zero. Restricting $\theta = 0$ degrades fit of AQRE a lot. In terms of overall out-of-sample fit, the general model is always a little better than AQRE. The key difference between the two models is that some lenders and entrants learn in the general model but they always anticipate what borrowers and incumbents will do in AQRE. The fact that the general model generally fits and predicts better than AQRE means that weakening sophistication of 'short-run' players has some is empirical value. However, the two models are equivalent with experienced entry-game subjects, which shows the power of experimental experience to create full sophistication.

5 Model Robustness

We subject the general model to a stress test by checking the robustness of the model using simulation (c.f. Cooper, Garvin, and Kagel (1997)). The idea of this robustness check is: 1) to verify if the general model is able to reproduce the empirical trends exhibited by the data and 2) to assess the impact of the features of the model in tracking data.

Four trends in the data emerge from the visual inspection of the figures 1-4. There are two cross-round trends: 1. the frequency of $NoLoan$ or $Entry$ increases across rounds; and 2. the frequency of $Default$ or $Share$ increases across rounds. The other two are cross-sequence trends: 3. the frequency of $NoLoan$ or $Entry$ decreases across sequence in early rounds but increases across sequence in later rounds; and 4. the frequency of $Default$ or $Share$ decreases across sequences in early rounds but increases across sequences in later rounds. We check for the statistical significance of these visual observations.

To confirm the significance of cross-round trends in the data, we run the following regressions on each of the four datasets across round t:

$$Prob_{kt} \text{ (No Loan/Entry)} = a_{Lt} + b_{Lt} \cdot t, \tag{17}$$

$$Prob_{kt} \text{ (Default/Sharing Given Loan/Entry)} = a_{Bt} + b_{Bt} \cdot t, \tag{18}$$

where t indexes round and k indexes the sequence blocks in Figures 1-4. Both b_{Lt} and b_{Bt} are significantly positive, confirming our visual observation of trends 1 and 2.

To check the significance of *cross-sequence* trends in the data, we partition each dataset into the first 4 rounds and last 4 rounds and run separate regressions for each partition across sequences k:

$$Prob_{kt} \text{ (No Loan/Entry)} = a_{LkR} + b_{LkR} \cdot k, \tag{19}$$

$$Prob_{kt} \text{ (Default/Sharing Given Loan/Entry)} = a_{BkR} + b_{BkR} \cdot k \tag{20}$$

where $R = 1$ represent the first 4 rounds and $R = 2$ represent the last 4 rounds.

We expect $b_{Lk1} < 0$ and $b_{Lk2} > 0$ for a significant trend 3 and $b_{Bk1} < 0$ and $b_{Bk2} > 0$ for a significant trend 4. We are only able to find partial confirmation of the cross-sequence trends. Specifically, we find strong evidence of cross-sequence trend in lender's behavior for both the Trust data and the inexperienced entry data in the first 4 rounds of the game ($b_{Lk1} < 0$). We also find strong evidence of cross sequence trend in incumbent's behavior in the first 4 rounds of the game ($b_{Bk1} < 0$). But the second-half cross-sequence coefficients b_{Lk2} and b_{Bk2} are not significant.

Next, we investigate if these significant trends in the data are replicated by the model prediction. We first generate the prediction of the general model by simulating choices using the parameter estimates of the general model from Table 4. For each 8-round sequence, we produce 1000 simulated choice paths of both lenders (or entrants) and borrowers (or incumbents). The 1000 simulated paths are averaged to produce the simulated probabilities for each sequence. These simulated probabilities are subjected to the same significance tests across periods and sequences we run for the data; slope coefficients for cross-round and cross-sequence trends for data and the simulated model are reported below the figures. The results for the Trust data and the inexperienced entry data are presented in Figures 5 and 6.[18] Judging from the plots and the significant estimates, the general model is able to reproduce the significant trends in the data fairly accurately (although the estimated $b_{.}$ trend coefficients are a little smaller in magnitude).

Is each key feature of the model necessary to capture the trends? We find out by "disabling" each feature individually and simulating choice using the restricted models with each feature disabled separately. There are four key features of the model: cross-sequence learning, the proportion of sophisticated lenders, the home-made prior, and the proportion of sophisticated borrowers. The first two are parameters driving the lender's choices and the last two are driving the borrower's choices. The details of the simulation and the trend significance regression analysis are the same as before. The plots and regression results are reported in Figures 7 and 8. The figures combine all sequences so several parameter configurations can be put on a single 3-D graph. The general findings benchmarked against data are:

(1) disabling cross-sequence learning results in less *Loan* initially and more *Loan* in later rounds (the cross-round trend has a lower slope b_{Lt} and higher intercept a_{Lt} than the general model). There is obviously more learning in the data than within-sequence learning. We also see more default in later rounds in response to this lender behavior (I.e., the slope b_{Bt} when $\tau = 0$ is lower than in the general model). In entry game, there is more entry and less sharing in later rounds (i.e., the slope b_{Bt} when $\tau = 0$ is lower than the general model).
(2) disabling sophisticated lenders ($\alpha_L = 0$) results in a flatter rise in $NoLoan$ frequency across rounds (i.e., b_{Lt} with $\alpha_L = 0$ is lower than the general model) and a steep drop in entry rate across rounds (which is going against data trend). The restricted model predicts a lower default and more sharing pattern than the data. This shows that allowing no sophisticated lenders harms fit.
(3) disabling the sophisticated borrowers ($\alpha_B = 0$) results in a flat default and sharing rate. This suggests that there are *non*-myopic borrowers. It

[18] In addition, the significant cross-round effects for Camerer and Weigelt (1988b) are $b_{Lt} = 0.08, b_{Bt} = 0.06$ for data and $b_{Lt} = 0.05, b_{Bt} = 0.04$ for the model. The significant cross-round effects for the experienced subjects in Jung et al. (1994) are $b_{Lt} = 0.09, b_{Bt} = 0.11$ for data and $b_{Lt} = 0.07, b_{Bt} = 0.07$ for the model. All are significant at 5%.

predicts a flatter rise in no loan across rounds than the data (b_{Lt} is too low).

(4) disabling the home-made prior ($\theta = 0$) results in substantially overprediction of default and sharing rate (They are higher than both the data and the general model prediction in every round, i.e. a_{Bt} is higher.). This provide strong evidence of the Special home-made type. It results in a flatter rise in no loan rate (b_{Lt} is lower) and an almost flat entry rate.

In general, disabling each of these four features, one at a time, shows that each of the features contribute to the ability of the general model to describe subjects' behavior. The simulations also indicate that each feature indeed captures the right part of the behavior that feature is meant to capture.

6 Conclusion

Many empirical learning models implicitly assume that players do not realize others are learning. This paper adds "sophisticated" players who *do* realize others are learning, in repeated games with incomplete information. Sophisticated players who know they are playing a repeated game have an incentive to take actions which are costly in the short-run, but which "teach" learners what to expect, in a way that benefits the teachers. Including teaching effects extends learning models to the many domains in which economic relationships are long-lasting.

This paper applies a precise model of sophisticated teaching to finitely-repeated experimental games of trust and entry-deterrence, with incomplete information about players types (some are induced to be honest or to fight entry). Earlier experiments have shown that some features of behavior in these games are approximated by very complex and delicate equilibria (Camerer and Weigelt (1988a), Jung et al. (1994)). But it is unlikely that players approximate the equilibria by introspection, and their comparative static predictions are often wrong (Neral and Ochs (1992)). A boundedly rational model of learning is one answer to the question of how people can approximate hyperrational Bayesian equilibria.

By including several learning types, the model both adds sophistication to adaptive learning, and adds learning to an AQRE model. Lenders and entrants either learn adaptively (using a self-tuning functional EWA rule) or sophisticatedly anticipate what borrowers and incumbents will do. Borrowers and incumbents are either myopic, always behave in a special way (trustworthy or fighting entry), or teach strategically.

A key difference between the proposed model and equilibrium is that the *play-*

ers build reputations in AQRE ("this guys seems honest"), but the learners' *strategies* have reputations ("entry is dangerous") in the proposed model. The proposed model is also a *partial equilibrium* one, because some adaptive players do not fully anticipate what others will do.

The general model was fit to 28 sessions of data from both repeated trust and entry games. Both models reproduce most of the basic trends in the data, particularly increasing default and market-sharing in later periods of a sequence, and some cross-sequence learning. The key parameter in the general model is the fraction of strategic teachers, α_B. This figure is reliably estimated to be about .91 for inexperienced entry-game subjects and close to 1 for trust games and experienced entry-game data. The fact that α_B rises with subject experience corroborates other findings. The AQRE model generally fits reliably worse than the general model, which is an indication that adding a learning component to equilibrium models helps explain how people behave.

A key point is that the same model can account for quite different behavior in these games: Borrowers in trust games behave in an honest way that is mutually-beneficial, while aggressive incumbents benefit only themselves. The same model explains both because the two behaviors emerge endogenously from the same kind of interaction between teaching and payoffs.

Finally, we subject the general model to a stress test to check the model robustness. Relying only on parameter estimates, the general model is able to simulate behaviors that match the significant trends found in the data. Furthermore, we show that this descriptive ability degrades significantly when any of the four key features of the model is turned off.

In future research, it would be useful to endogenize some of the parameters of most interest (particularly the rate of sophistication α_B). The model could also be applied to games and markets where the interaction of sophistication and adaptive learning is interesting (e.g. inflation-setting, see Sargent (1999); and price bubbles, see Smith, Suchanek, and Williams (1988), Camerer and Weigelt (1990) and Lei, Noussair, and Plott (2001)).

References

Andreoni, J., Miller, J., 1993. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. Economic Journal 103, 570–585.

Cachon, G. P., Camerer, C. F., 1996. Loss-avoidance and forward induction in experimental coordination games. Quarterly Journal of Economics 111, 165–194.

Camerer, C. F., 2003. Behavioral Game Theory: Experiments on Strategic Interaction, 1st Edition. Princeton University Press, Princeton, N.J.

Camerer, C. F., Ho, T.-H., 1999. Experience-weighted attraction learning in normal-form games. Econometrica 67, 827–874.

Camerer, C. F., Ho, T.-H., Chong, J.-K., 2002. Sophisticated EWA learning and strategic teaching in repeated games. Journal of Economic Theory 104 (1), 137–188.

Camerer, C. F., Weigelt, K., 1988a. Experimental tests of a sequential equilibrium reputation model. Econometrica 56, 1–36.

Camerer, C. F., Weigelt, K., 1988b. Measuring beliefs in reputation games, presentation to the Economic Science Association.

Camerer, C. F., Weigelt, K., 1990. Inducing bubble expectations in experimental asset markets, unpublished manuscript.

Camerer, C. F., Weigelt, K., 2005. Belief calibration, QRE and learning in repeated trust games. In: Lupia, A., Alt, J. (Eds.), Essays in Honor of Richard McKelvey.

Cooper, D., Kagel, J., 2001. Learning and transfer in signaling games, working Paper, Department of Economics, Case Western Reserve University.

Cooper, D. J., Garvin, S., Kagel, J. H., 1997. Signalling and adaptive learning in an entry limit pricing game. RAND Journal of Economics 28 (4), 662–683.

Fehr, E., Gächter, S., 2000. Fairness and retaliation: The economics of reciprocity. Journal of Economic Perspective 14, 159–181.

Fudenberg, D., Kreps, D., 1990. A theory of learning, experimentation and equilibrium in games, working Paper, Stanford University.

Fudenberg, D., Levine, D., 1989. Reputation and equilibrium selection in games with a patient player. Econometrica 57, 759–778.

Goeree, J. K., Holt, C. A., 1999. Stochastic game theory: For playing games, not just for doing theory, proceedings of the National Academy of Sciences.

Ho, T.-H., Camerer, C. F., Chong, J.-K., May 2004. The economics of learning models: A self-tuning theory of learning in games, working Paper, University of California Berkeley.

Ho, T.-H., Camerer, C. F., Weigelt, K., 1998. Iterated dominance and iterated best-response in p-beauty contests. American Economic Review 88, 947–969.

Jung, Y. J., Kagel, J. H., Levin, D., 1994. On the existence of predatory pricing: An experimental study of reputation and entry deterrence in the chain-store game. RAND Journal of Economic 25, 72–93.

Kalai, E., Lehrer, E., 1993. Rational learning leads to nash equilibrium. Econometrica 61 (5), 1019–1045.

Kreps, D., 1990. Corporate culture and economic theory. In: Alt, J., Shepsle, K. (Eds.), Perspectives on Positive Political Economy. Cambridge University Press, Cambridge.

Kreps, D., Wilson, R., 1982. Reputation and imperfect information. Journal of Economic Theory 27 (1), 253–279.

Lei, V., Noussair, C., Plott, C., 2001. Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. Econometrica 69 (4), 831–859.

McKelvey, R. D., Palfrey, T. R., 1992. An experimental study of the centipede game. Econometrica 60, 803–836.

McKelvey, R. D., Palfrey, T. R., 1998. Quantal response equilibria for extensive form games. Experimental Economics 1, 9–41.

Milgrom, P., Roberts, J., 1991. Adaptive and sophisticated learning in repeated normal form games. Games and Economic Behavior 3, 82–100.

Mookherjee, D., Sopher, B., 1994. Learning behavior in an experimental matching pennies game. Games and Economic Behavior 7, 62–91.

Neral, J., Ochs, J., 1992. The sequential equilibrium theory of reputation building: A further test. Econometrica 60, 1151–1169.

Palfrey, T. R., Rosenthal, H., 1988. Private incentives in social dilemmas: The effects of incomplete information and altruism. Journal of Public Economics 35, 309–332.

Partow, Z., Schotter, A., 1993. Does game theory predict well for the wrong reasons? an experimental investigation, working Paper 93-46, C.V. Starr Center for Applied Economics, New York University.

Sargent, T., 1999. The Conquest of American Inflation, 1st Edition. Princeton University Press, Princeton.

Selten, R., 1991. Anticipatory learning in 2-person games. In: Selten, R. (Ed.), Game Equilibrium Models. Vol. I. Springer-Verlag, Berlin, pp. 98–154.

Selten, R., Stoecker, R., 1986. End behavior in sequences of finite prisoner's dilemma supergames: A learning theory approach. Journal of Economic Behavior and Organization 7, 47–70.

Smith, V. L., Suchanek, G., Williams, A., 1988. Bubbles, crashes and endogenous expectations in experimental spot asset markets. Econometrica 56, 1119–1151.

Stahl, D. O., 1999. Sophisticated learning and learning sophistication, working Paper, University of Texas at Austin.

Watson, J., 1993. A 'reputation' refinement without equilibrium. Econometrica 61, 199–205.

Watson, J., Battigali, P., 1997. On 'reputation' refinements with heterogeneous beliefs. Econometrica 65, 363–374.

7 Appendix: Self-tuning EWA Model Specification

Ho et al. (2004) developed a *Self-tuning* EWA model in which the fixed parameters ϕ and δ are replaced by functions of data which self-adjust across games and over time. These functions determine parameter values for each player, each round and each sequence, which are then plugged into the EWA updating equation to determine attractions.

The function $\phi(k,t)$ is designed to detect change in the learning environment.

It takes the differences in corresponding elements of two frequency vectors, squares them, and sums the squares over strategies. The change-detection function $\phi(k,t)$ is

$$\phi(k,t) = 1 - .5(\sum_j [\frac{\sum_{\sigma=t-W+1}^{t} I(s_B^j, s_B(k,\sigma))}{W} - \frac{\sum_{\sigma=1}^{t} I(s_B^j, s_B(k,\sigma))}{t}]^2)$$

where W is the minimal number of equilibrium strategies. s_B^j denotes the j-th strategy of borrower and the term $\frac{\sum_{\sigma=t-W+1}^{t} I(s_B^j, s_B(k,\sigma))}{W}$ is the j-th element of a vector that simply counts how often strategy j was played by the borrower in the W periods from $t-W+1$ to t, and divides by W. The term $\frac{\sum_{\sigma=1}^{t} I(s_B^j, s_B(k,\sigma))}{t}$ is the relative frequency count of the j-th strategy over all t periods.

The parameter δ is the weight on foregone payoffs. Presumably this is tied to the attention subjects pay to alternative payoffs, ex-post. Subjects who have limited attention are likely to focus on strategies that would have given higher payoffs than what was actually received, because these strategies present missed opportunities. To capture this property, define

$$\delta_j(k,t) = \begin{cases} 1 \text{ if } \pi_L(s_L^j, s_B(k,t)) > \pi_L(k,t), \\ 0 \text{ otherwise.} \end{cases}$$

where $\pi_L(k,t)$ is the actual payoff lender received in round t of sequence k. Subjects reinforce chosen strategies (where the top inequality necessarily binds) and all unchosen strategies with better payoffs (where the inequality is strict) with a weight of one. They reinforce unchosen strategies with equal or worse payoffs by zero.

Short-Run Player (e.g., Lenders, Entrants)	Proportion	Specification of Behavior	**Long-Run Player** (e.g., Borrowers, Incumbents)	Proportion	Specification of Behavior
Adaptive	$1-\alpha_L$	Section 2.1	Special (Induced) (e.g., Honest. Aggressive)	p	Section 2.3
Sophisticated	α_L	Section 2.2			
			Normal (Induced) (Dishonest, Non-aggressive)	1-p	
			Special (Home-made)	θ	Section 2.3
			Normal	$1-\theta$	
			Myopic	$1-\alpha_B$	Section 2.4
			Sophisticated	α_B	Section 2.5

Table 1: A Model of Repeated Games with both Adaptive and Sophisticated Players

Table 2: Payoffs for the Borrower-lender Trust Games and the Entry-deterrence Games

Payoffs in the borrower-lender trust game, Camerer and Weigelt [1988a]

lender strategy	borrower strategy	payoffs to lender	payoffs to borrower normal (X)	payoffs to borrower honest (Y)
loan	default	-100^*	150	0
	repay	40	60	60
no loan	no choice	10	10	10

Payoffs in the entry-deterrence game, Jung, Kagel and Levin [1994]

entrant strategy	incumbent strategy	payoffs to entrant	payoffs to incumbent normal (X)	payoffs to incumbent fighter (Y)
in	fight	80	70	160
	share	150	160	70
out	no choice	95	300	300

Note: * Loan-default lender payoffs were -50 in sessions 6-8 and -75 in sessions 9-10.

Table 3: In-sample and Out-of-sample Performance of the General Model and AQRE Model

Dataset	Camerer and Weigelt (1988a)	Camerer and Weigelt (1988b)	Jung, Kagel and Levin (1994) Inexperienced Subjects	Jung, Kagel and Levin (1994) Experienced Subjects
In-sample Calibration [1]				
Sample size	5757	3820	5847	2232
Log-likelihood				
The General Model	-2919.43	-2007.20	-2246.97	-1345.76
AQRE	-3218.52	-2094.17	-2418.43	-1345.76
Log-likelihood Ratio [2]	598.18	173.93	342.91	0.00
Average Probability				
The General Model	0.60	0.59	0.68	0.55
AQRE	0.57	0.58	0.66	0.55
Out-of-sample Validation				
Sample size	2894	1882	2866	1072
Log-likelihood				
The General Model	-1425.16	-947.25	-1341.15	-553.11
AQRE	-1525.69	-989.44	-1425.53	-553.11
Average Probability				
The General Model	0.61	0.60	0.63	0.60
AQRE	0.59	0.59	0.61	0.60

Note 1: Calibrated on all observations for 70% of the subjects instead of 70% observations of all subjects.

Note 2: Threshold for significant χ^2 test is 12.59 at 5% for 6 degrees of freedom

Table 4: Parameter Estimates [1]

Dataset	Camerer and Weigelt (1988a)	Camerer and Weigelt (1988b)	Jung, Kagel and Levin (1994) Inexperienced Subjects	Jung, Kagel and Levin (1994) Experienced Subjects
The General Model				
Adaptive Lender				
Functional ϕ	0.76	0.77	0.78	0.76
Functional δ	0.15	0.16	0.19	0.34
τ	0.94	0.68	0.35	0.12
A0 (No Loan/Not Enter)	-2.09	-1.52	-1.63	0.39
λ^a_L	11.40	5.89	3.90	3.70
Sophisticated Lender				
α_L	**0.43**	**0.63**	**0.67**	**1.00**
λ^s_L	7.75	8.19	5.41	6.53
Myopic Borrower				
λ^m_B	2.66	3.76	5.62	3.10
Sophisticated Borrower				
α_B	**1.00**	**0.95**	**0.91**	**1.00**
λ^s_B	5.87	8.45	2.90	1.36
θ	0.28	0.28	0.18	0.17
Honest Borrower				
λ^h_B	27.32	24.76	5.70	6.30
Agent-based Quantal Response Equilibrium (AQRE)				
Sophisticated Lender				
α_L	1.00	1.00	1.00	1.00
λ^s_L	6.31	5.37	4.76	6.53
Sophisticated Borrower				
α_B	1.00	1.00	1.00	1.00
λ^s_B	3.84	4.76	3.28	1.36
θ	0.27	0.28	0.16	0.17
Honest Borrower				
λ^h_B	26.70	25.89	2.75	6.30

Note 1: All sessions in a dataset are pooled to produce a common set of estimates except for the scale parameters like the λs which are session-specific. The estimates reported here are averages. Standard errors are reported in our working paper.

Figures 1a-b: Frequency Plots for the Trust Data from Camerer and Weigelt [1988a]

Figure 1a: Empirical Frequency for No Loan

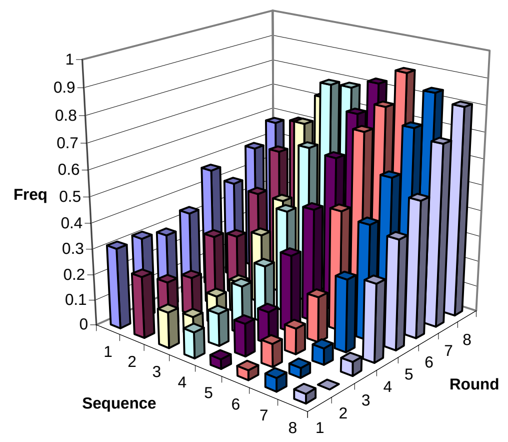
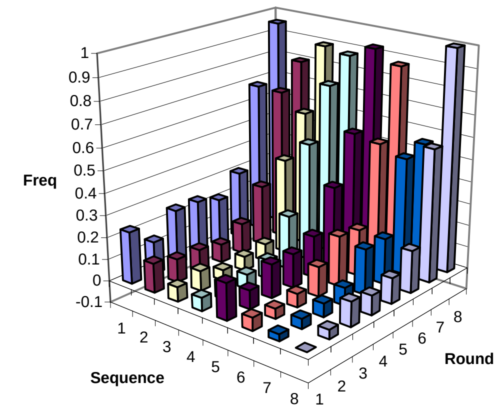


Figure 1b: Empirical Frequency for Default conditional on Loan (Dishonest Borrower)



Figures 2a-b: Frequency Plots for the Unpublished Trust Data from Camerer and Weigelt [1988b]

Figure 2a: Empirical Frequency for No Loan

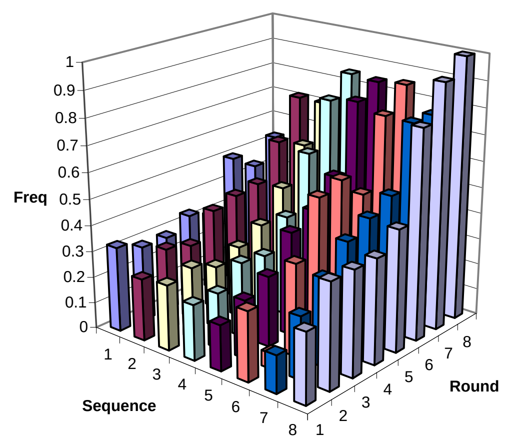
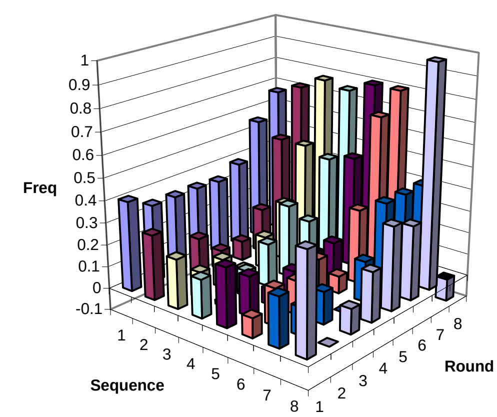


Figure 2b: Empirical Frequency for Default conditional on Loan (Dishonest Borrower)



Figures 3a-b: Frequency Plots on Inexperienced Subjects from the Entry Data from Jung, Kagel and Levin [1994]

Figure 3a: Empirical Frequency for Entry

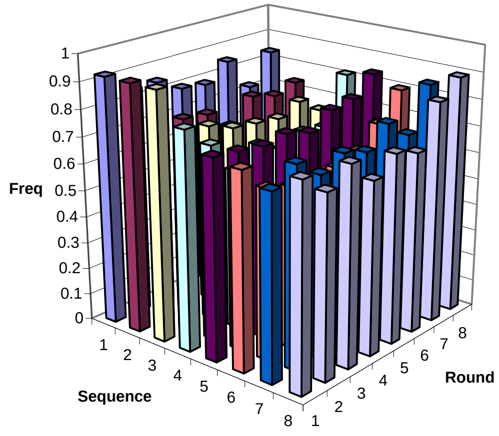
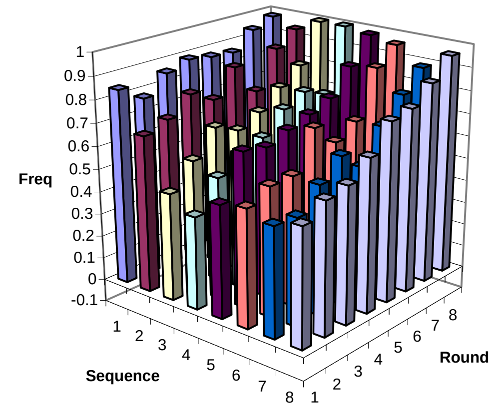


Figure 3b: Empirical Frequency for Sharing conditional on Entry (Weak Incumbent)



Figures 4a-b: Frequency Plots on Experienced Subjects from the Entry Data from Jung, Kagel and Levin [1994]

Figure 4a: Empirical Frequency for Entry

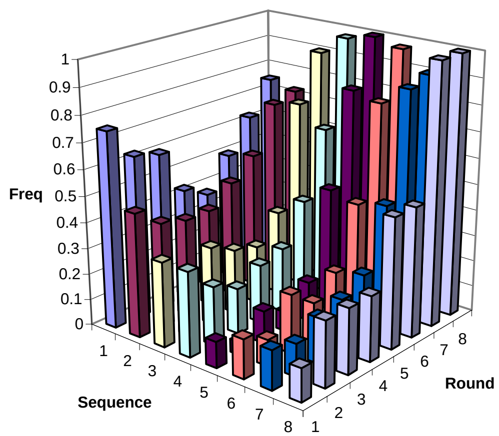
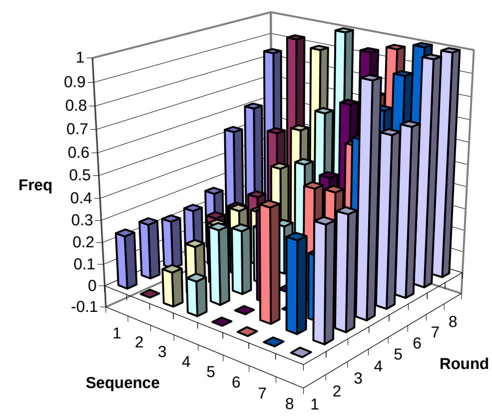


Figure 4b: Empirical Frequency for Sharing conditional on Entry (Weak Incumbent)



Figures 5a-d: Cross-sequence and Cross-round Effects for the Trust Data from Camerer and Weigelt [1988a]

Figure 5a: The General Model Simulated Frequency for No Loan

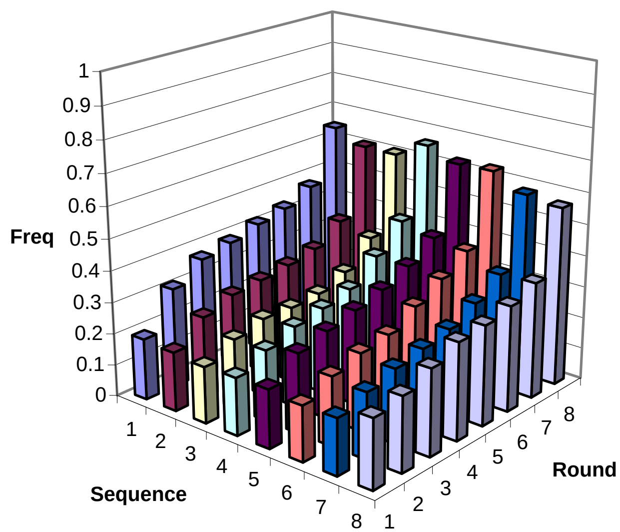
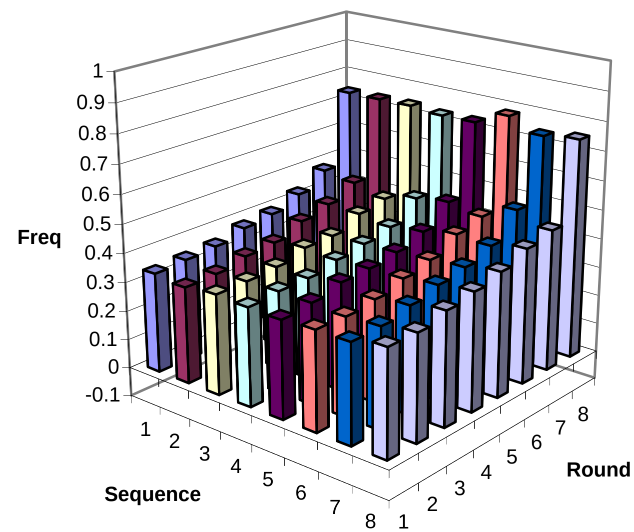


Figure 5b: The General Model Simulated Frequency for Default conditional on Loan (Dishonest Borrower)



Cross-sequence Trends: b_{Lk1} (data) = -0.02, b_{Lk1} (model) = -0.01. Both are significant at 5%

Cross-round Trends: b_{Lt} (data) = 0.10, b_{Lt} (model) = 0.05, b_{Bt} (data) = 0.10, b_{Bt} (model) = 0.04. All are significant at 5%

Figures 6a-b: Cross-sequence and Cross-round Effects for Inexperienced Subjects from the Entry Data from Jung, Kagel and Levin [1994]

Figure 6a: The General Model Simulated Frequency for Entry

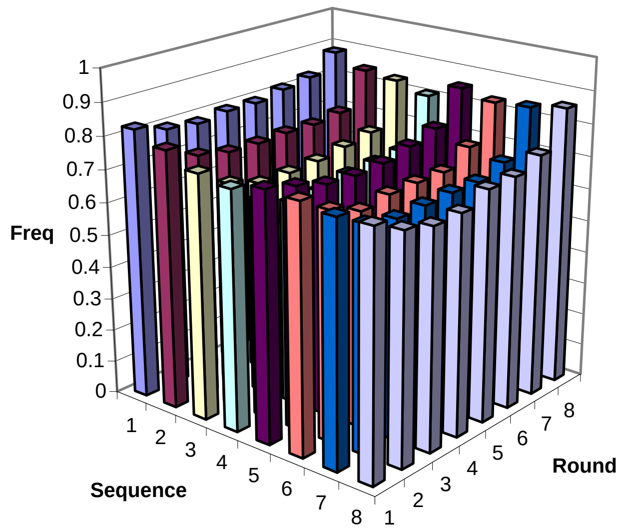
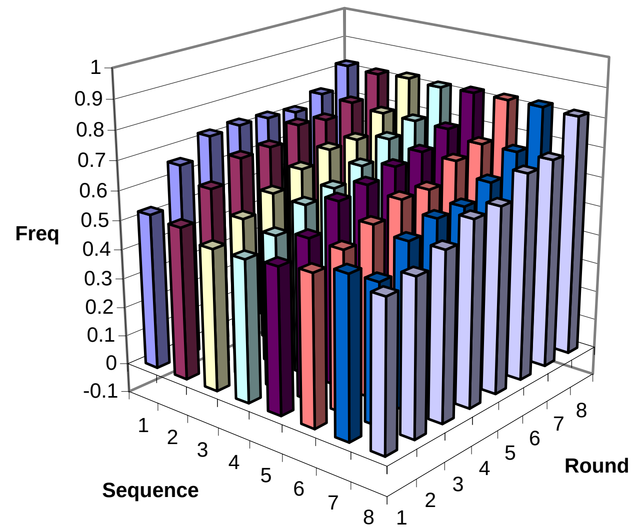


Figure 6b: The General Model Simulated Frequency for Sharing conditional on Entry (Weak Incumbent)



Cross-sequence Trends: b_{Lk1} (data) = -0.02, b_{Lk1} (model) = -0.01, b_{Bk1} (data) = -0.03, b_{Bk1} (model) = -0.01. All are significant at 5%

Cross-round Trends: b_{Bt} (data) = 0.05, b_{Bt} (model) = 0.04. All are significant at 5%

Figures 7a-b: Parameter Restrictions for the Trust Data from Camerer and Weigelt [1988a]

Figure 7a: Average Probability of No Loan

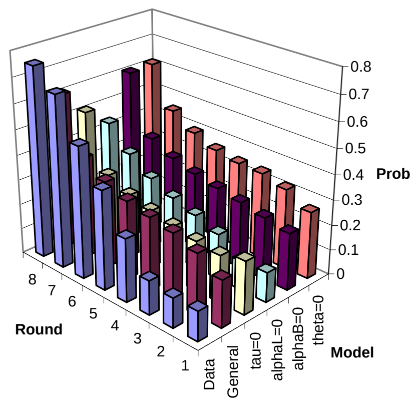
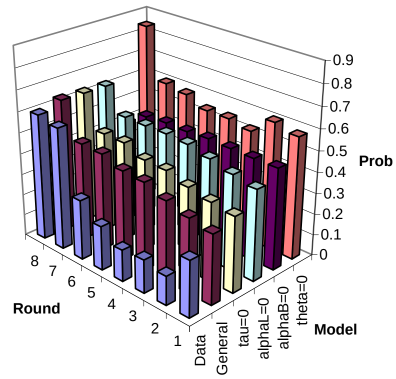


Figure 7b: Average Probability of Default Given Loan (Dishonest Borrower)



Lender: b_{Lt} (data) = 0.10, b_{Lt} (model) = 0.05, b_{Lt} (τ=0) = 0.03, b_{Lt} (α_L=0) = 0.04, b_{Lt} (α_B=0) = 0.04, b_{Lt} (θ=0) = 0.04.
Borrower: b_{Bt} (data) = 0.10, b_{Bt} (model) = 0.04, b_{Bt} (τ=0) = 0.03, b_{Bt} (α_L=0) = 0.04, b_{Bt} (α_B=0) = 0.00, b_{Bt} (θ=0) = 0.05

Figures 8a-b: Parameter Restrictions for Inexperienced Subjects from the Entry Data from Jung, Kagel and Levin [1994]

Figure 8a: Average Probability of Entry

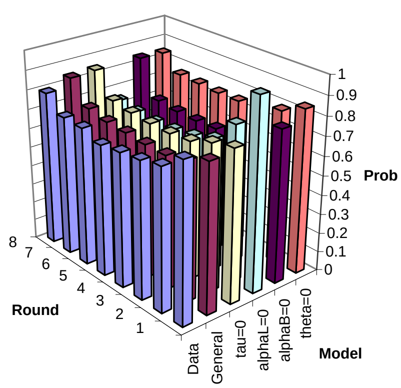
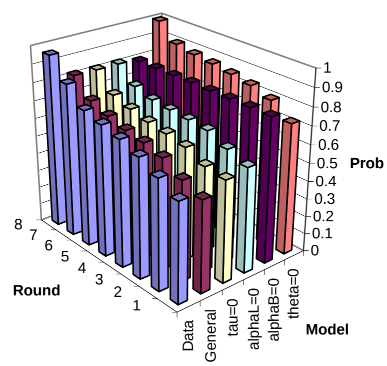


Figure 8b: Average Probability of Sharing Given Entry (Weak Incumbent)



Incumbent: b_{Bt} (data) = 0.05, b_{Bt} (model) = 0.04, b_{Bt} (τ=0) = 0.03, b_{Bt} (α_L=0) = 0.03, b_{Bt} (α_B=0) = 0.00, b_{Bt} (θ=0) = 0.03