Recency, Records and Recaps: The effect of feedback on behavior in a simple decision problem*

Drew Fudenberg[1], Alexander Peysakhovich[2,3]

Current version: Dec. 2013

First version: May 2013

Abstract

Suboptimal behavior can persist in simple stochastic decision problems. This has motivated the development of solution concepts such as cursed equilibrium (Eyster & Rabin 2005) and behavioral equilibrium (Esponda 2008). We experimentally study a simple adverse selection (or "lemons") problem and find that learning models that heavily discount past information (i.e. display recency bias) explain patterns of behavior better than Nash, cursed or behavioral equilibrium. Providing counterfactual information or a record of past outcomes does little to aid convergence to optimal strategies, but providing sample averages ("recaps") gets individuals most of the way to optimality. Thus recency effects are not solely due to limited memory but stem from some other form of cognitive constraints.

* Much of this work was done while Peysakhovich was a graduate student at the Harvard University Department of Economics. We thank Douglas Bernheim, Gary Charness, Ignacio Esponda, Armin Falk, Simon Gaechter, Ed Glaeser, Moshe Hoffman, David Laibson, David K. Levine, Sendhil Mullainathan, Matthew Rabin, David G. Rand, Ran Spiegler, Andrei Shleifer, Josh Schwartzstein, Shoshana Vasserman and Erez Yoeli for helpful comments, the Harvard Program for Evolutionary Dynamics for its hospitality, and NSF grants SES 0951462 and 1258665 for financial support.

[1]Department of Economics, [2]Program for Evolutionary Dynamics, Harvard University Cambridge, MA 02139, [3]Department of Psychology, Yale University, New Haven, CT 06511

Introduction

People make mistakes. When does repeat experience lead individuals to optimal behavior and when can mistakes be persistent? This question lies at the heart of interpreting the potential market impacts of many findings from behavioral economics. Two classes of models allow the persistence of mistakes. In general, Bayesian-based models predict that only mistakes about "off-path" or otherwise unobserved events can persist in the long run. In contrast, many behavioral models assume that even "on-path" mistakes can be equilibrium phenomena (i.e. permanent or at least reasonably persistent even in the presence of feedback).

We steer a middle course and argue that "on path" mistakes do occur and can *sometimes* persist. Importantly, the well-documented tendency to discount older information, called *recency bias* (see e.g. Erev & Haruvy 2013), plays a key role in determining whether mistakes are permanent or temporary. In particular, recency bias can imply persistent suboptimal behavior in stationary stochastic environments.[1] To show this we conduct a series of experiments in which participants face a simple stochastic decision problem. A simple learning model with high recency better organizes behavior across experiments than three well known equilibrium concepts. Additionally, insights gained from this model suggest an intervention that markedly improves payoffs and is generalizable across several decision problems.

Our stochastic decision problem is the simplified version of Akerlof's (1970) classic lemons model introduced in Esponda (2008), the *additive lemons problem* (ALP). In the ALP a seller is endowed with an object with a value known to the seller and unknown to the buyer, and the object is worth a fixed amount more to the buyer than the seller. The buyer makes a single take-it-or-leave-it offer to the seller[2] who accepts or rejects it. Buyers receive full feedback if their offer is accepted but receive no feedback if their offer is rejected.

The ALP with subjects acting as buyers and computers playing sellers makes a useful laboratory model for studying the persistence of mistakes. The calculation of the buyer's optimal strategy requires individuals to use conditional expectations, something

[1] Recency can be a useful heuristic in a non-stationary world where older observations will not necessarily be relevant.
[2] Played in our experiments by a computer who always uses the dominant strategy.

that is generally unintuitive for many individuals, and leads to large deviations from optimal behavior in the first round (Tor & Bazerman 2003). Allowing individuals to play the ALP repeatedly allows us to examine the persistence of these mistakes and the effects of manipulations on convergence to optimality.

Our baseline ALP lets us generate predictions using Nash equilibrium (NE), cursed equilibrium (CE), behavioral equilibrium (BE), and a simple learning model, temporal difference reinforcement learning (TDRL; Sutton & Barto 1998).[3] In our first experiment subjects were randomly assigned to one of four conditions: two payoff structures for the lemons problem (high or low "value added") were crossed with two information conditions. In the informed condition, participants were told the prior distribution of seller valuations; in the uninformed condition they were not given any information about this distribution of values other than its support.[4]

These four conditions were chosen so that Nash, cursed and behavioral equilibrium have clear and distinct predictions: in the "low value added" payoff conditions CE predicts offers above NE offers while in the "high value added" conditions it predicts offers below NE. In contrast, BE predicts possible underbidding in conditions where participants are uninformed of the seller's valuation but coincides with NE predictions in conditions where the subjects were told the distribution of seller values. We use simulation methods to generate predictions for the TDRL model, and find that with relatively high recency the TDRL model predicts offers above NE in "low-added value" conditions but close-to optimal aggregate behavior in "high added value" conditions.

There is no underbidding in any of these conditions, and informing subjects of the distribution of seller values has little effect. Moreover, there is more overbidding than (fully) cursed equilibrium predicts when it predicts overbidding, and almost optimal behavior where cursed equilibrium predicts underbidding. Thus neither Nash equilibrium, behavioral equilibrium, nor cursed equilibrium fit well with the data. In contrast, a very simple learning model with relatively high recency organizes the

[3] To use simulation methods, we need to fix a functional form for the learning model. We choose the TDRL model for its simplicity, but the qualitative predictions come from the assumption of recency, not the chosen functional form.

[4] This latter structure corresponds to the assumptions in Esponda (2008), who argues that it seems a better description of many field settings.

aggregate behavioral patterns relatively parsimoniously. In addition, we find direct evidence for recency: Individuals react strongly to last period outcomes even after experience with the decision problem.

We then consider a succession of treatments with different feedback structures. In a second experiment, we add explicit counterfactual information: buyers are informed of the object's value regardless of whether their bid is accepted or rejected. This allows us to test whether behavior in the main treatment comes from incorrect expectations about the value of the rejected items. Providing this additional information has very little effect and our qualitative findings are unchanged. Because this treatment makes the information subjects receive exogenous to their actions, it also permits a cleaner test of recency effects, which we again confirm.

Recency effects are very powerful in the ALP because a single experience with a strategy contains very little information about whether that strategy is successful. Thus, when participants heavily discount the past's information, they are not able to learn the optimal behavior. In our next experiment we ask whether this discounting of past information is a result of limited memory or of more complicated cognitive constraints.

To answer this we consider two treatments. In the *more information* condition participants play the ALP against 10 sellers simultaneously. Each round buyers make a single offer decision that applied to all 10 sellers. At the end of a round, participants receive feedback about each of the ten transactions: what the seller's value was, whether the offer was accepted and the buyer's profits on that transaction. The *simple information* condition has identical rules. However, instead of receiving fully detailed feedback on each transaction, participants are told their *average* profit out of the 10 transactions and *average* values of the objects they actually purchased.

Providing more information has little effect, but providing the information in the pithy, more readily understood form of averages ("recaps") significantly improves the subjects' payoffs. This suggests that recency effects may not simply be an issue of "memory space" but also the (lack of) computational resources to construct useful summary statistics from multiple pieces of data. Exploring these computational constraints is an important avenue for future research.

Finally, to test the robustness of our findings, we study a slightly modified version of the Monty Hall problem. This problem also involves non-trivial probabilistic reasoning and, again, participants do not solve it correctly on the first try. Allowing for feedback we see that the effectiveness of recaps remains as before. Repeated play of the Monty Hall problem does not lead to optimal behavior but simple recaps lead to better decisions and higher payoffs.

2. Theory

2.1 Nash Equilibrium in the Additive Lemons Problem

To investigate the effects of different information and feedback conditions on learning, payoffs, and convergence or non-convergence of behavior to optimality, we focus on the *additive lemons problem* (ALP) as formulated in Esponda (2008). In this game: there are two players, a buyer and a seller. The seller begins with an object of value v drawn from a uniform distribution between 0 and 10; this value is known to the seller but is unknown to the buyer. The buyer makes a single take-it-or-leave-it offer b to the seller. If the seller accepts this offer, the buyer receives the object and pays b to the seller. The object is worth $v+k$ to the buyer, thus there is a gain from the occurrence of trade.

This game has a unique Nash equilibrium in weakly undominated strategies: It is weakly dominant for the seller to accept all offers below v and reject all offers above v. Because the seller has a dominant strategy, we transform the ALP into a single-person decision for the rest of our study. The buyer's optimization problem is thus

$$\max_b \Pr(v \le b)[E(v+k \mid v \le b) - b] = b[k - b/2].$$

Solving the maximization shows that the optimal bid is b* = k, so buyers offer k every round and sellers accept when $v<k$ and reject if $v>k$.

We chose the ALP for several reasons. First, lemons problems are familiar to economists. Second, the ALP is easy to describe to subjects but also tends to elicit suboptimal first responses due failures of probabilistic reasoning. The key to this failure is that the expectation in the buyer's maximization problem is a *conditional* expectation.

To make an optimal decision the buyer needs to take into account that if a bid of b is accepted the item's value must lie below v. There is a large amount of experimental evidence that individuals frequently fail to this correction in many decisions of interest, including common value auctions (Kagel & Levin 1986), the Monty Hall problem (Krauss & Wang 2003) and strategic voting games (Guarnaschelli et al. 2000). Additionally, the ALP can be played repeatedly in a short amount of time. We will focus on two payoff conditions: a "low added value" condition where $k=3$ and a "high added value" condition where $k=6$.

The ALP is very similar to the Acquire a Company Game (ACG) introduced by Samuelson & Bazerman (1985). The ACG has the same extensive form, but the value to the buyer has the multiplicative form kv instead of the additive form $v+k$ that we consider here. In the ACG, for $k>2$ the optimal bid is 10 and for $k<2$ the optimal bid is 0. There has been a large amount of research on this game which shows that when $k<2$, individuals fail to play the optimal strategy (Samuelson & Bazerman 1985) even with learning opportunities (Ball et al. 1991). However, the fact that the optimal bid is on the boundary is a significant confound here, given the aversion of individuals for corner solutions (Rubinstein et al. 1993). Our specification of the ALP avoids this confound as for any value of k the optimal solution is interior.[5]

2.2 Predicted Play, I: Equilibrium Concepts

Nash equilibrium requires that each player's strategy is a best response to the true distribution of opponents' play, and so implies that the buyers in the ALP should make the optimal bid. Some alternative equilibrium concepts (Battigalli & Guatoli 1997, Dekel et al. 1999, 2004, Fudenberg & Kamada 2013, Fudenberg and Levine 1993a, Kalai & Lehrer 1993, Lehrer 2012) maintain the assumption that players correctly interpret and process the information they receive and best respond to this information, while allowing players to have incorrect beliefs provided those beliefs are consistent with their observations, so that players can only have wrong beliefs "off the equilibrium path." We

[5] One of our treatments demonstrates this confound by restricting the range of allowed bid in the ALP.

focus here on a particular example of such a concept: behavioral equilibrium (Esponda 2008).

A variety of behavioral experiments show that mistakes in probabilistic reasoning are fairly common (Charness & Levin 2009, Kagel & Levin 1986, Krauss & Wang 2003, Samuelson & Bazerman 1985, Ball et al. 1991). This motivates equilibrium concepts that allow or require individuals to make mistakes in updating beliefs about opponents' play and computing the associated best responses. In particular, cursed equilibrium (Eyster &Rabin 2005) allows for a specific type of mistake in computing conditional expectations, without distinguishing between on-path and off-path errors.

Behavioral equilibrium and cursed equilibrium make different predictions of behavior in the ALP. In addition, they suggest different causes for deviations from optimal play. We now discuss these predictions.

Behavioral Equilibrium

Esponda (2008) develops the solution concept of behavioral equilibrium (BE) specifically for the ALP. This concept is meant to model settings where (1) subjects need to learn the distribution of Nature's moves (i.e. values) at the same time that they learn the distribution of opponent's play, and (2) buyers don't see the seller's value when the seller rejects the object.[6]

In our setting BE can be expressed as a two-tuple (p^*, b_{BE}) where p^* is a probability distribution on the interval [0,10]. BE imposes two conditions on this tuple. First, b_{BE} must be optimal for the buyer given distribution p^* and the belief that sellers play optimal strategies. Second, p^* must be consistent with what buyers observe in equilibrium, so that $p^*(A)$ for any subinterval A of the interval [0, b_{BE}] must coincide with the true probability (in this case, uniform) of A.

However, no restrictions are placed on what probabilities p^* may place on of the distribution of values that buyers never actually see. Given these two conditions, BE is a

[6] BE allows for two types of agents: naïve agents whose beliefs are only required to be self-confirming but do not know the distribution of Nature's move (as in Dekel et al. 2004) and sophisticated agents who know the payoff functions of the other players, as in rationalizable self confirming equilibrium (Dekel et al. 1999). In our formulation of the ALP the buyers are told the seller's strategy so the two type of agent are equivalent. When the ALP is formulated as a game, the sophisticated agents deduce that the seller will not accept a price below their value, but the naïve agents need not do so.

set valued solution concept with the property that $b_{BE} \leq b_{NE}$. Thus BE predicts that buyers cannot persistently overbid in either of the payoff conditions we examine in the ALP: Were they were to do so, they would learn that it would be better to make the NE bid instead. However, buyers can persistently underbid if they have overly pessimistic beliefs about the distribution of values above their bid.

Cursed Equilibrium

Individuals often fail to deal correctly with conditional probabilities. Eyster & Rabin (2005) build this assumption into the equilibrium concept of fully cursed equilibrium (CE). Formally, CE assumes that when individuals in a Bayesian game optimize they completely ignore the correlation between their other players' types and their strategies.[7]

As a simple example, consider a game with two players, 1 and 2 and two states of Nature, A and B, with probabilities *p* and *1*-p respectively. The states are observed by player 2 but not by player 1. Suppose player 2 plays action *s* when the state is A and plays *r* when the state is B. Then when player 1 sees action *s*, they should infer that the state is A. However, in CE player 1 ignores this information, and so optimizes against beliefs of the form "player 2 plays *s* with probability *p,* and *r* with probability *1-p* regardless of the state of the world"

Following Esponda (2008), we adapt CE to the additive lemons problem by supposing that the buyer's maximization problem replaces the conditional expectation of the value *v* with its *unconditional expectation*

$$\max_b \Pr(v \leq b) E(v - b) = \max_b (b/10)[5 + k - b].$$

so the cursed equilibrium bid here is

$$b_{CE} = (5+k)/2.$$

[7] In applying cursed equilibrium to the lemons problem Eyster & Rabin use a refinement to restrict off path play that is analogous to our assumption that the sellers do not use weakly dominated strategies. Eyster & Rabin also propose the notion of partially cursed equilibrium, in which beliefs are a convex combination of the fully cursed beliefs and those in the Nash equilibrium.

Note that this leads to overbidding (relative to the best response) if k < 5 and underbidding when k > 5. Thus CE predicts overbidding in the low added value conditions (*k=3*) and underbidding in high added value conditions (*k=6*).

As noted above, the predictions of CE do not depend on whether or not players are told the distribution of Nature's moves or on the sort of feedback they receive in the course of repeated trials. Another property of CE is that in many games, including the ALP, the payoff that players expect to receive in equilibrium does not match the actual payoffs they will receive. Thus, to the extent that CE is meant to describe behavior that persists when subjects have experience (as the "equilibrium" part of its name suggests), it implies that individuals have *permanently* incorrect yet stable beliefs about their expected payoffs.

2.3: Predicted Play, II: Learning Dynamics with Recency

A common argument given for the use of equilibrium analysis is that equilibrium arises as the long-run result of a non-equilibrium learning process (Fudenberg & Kreps 1995, Fudenberg & Levine 1998, Samuelson 1997). However, there is a substantial amount of evidence both from the lab (Camerer 2003, Erev & Haruvy 2013) and the field (Agarwal et al. 2008, Malmandier & Nagel 2011) that individuals react strongly to recently experienced outcomes and discount past information. Individuals who display such "recency effects" will not converge to using a single strategy in a stochastic environment, and so will be poorly described by an equilibrium model. Thus it is interesting to explore the use of learning dynamics to generate predictions in place of an equilibrium concept. This method is less concise, and will typically require the use of simulations, so it does not lead to as sharp predictions. However, it has the additional benefit of providing predictions over intermediate time scales as well as in the long run (Roth & Erev 1995, Erev & Roth 1998).

Recency has been incorporated into both belief-based and reinforcement-based models of learning, by adding a parameter that controls the speed of informational discounting (see e.g. Cheung & Friedman 1997, Fudenberg & Levine 1998, Sutton & Barto 1998, Camerer & Ho 1999, Benaim et al. 2009). Recency effects have also been

modeled by supposing that individuals "sample" a set of experiences either with all experiences in the recent past weighted equally (Young 1993) or with more recent experiences being more likely to be sampled (Nevo & Erev 2012). Most of these learning models converge to an ergodic distribution. In general the details of those distributions depend on the specifics of the model, but it is easier to characterize the effect of recency in some limit cases. At one extreme, with very little recency, each of a large number of past outcomes has approximately equal weight, so in a stationary decision environment we expect each individual to obtain close to the optimal payoff.[8] On the other hand, the most extreme case of recency is to play a best response to last period's information. In the ALP if the seller's value today is expected to be exactly the same as yesterday's, then the optimal bid equals yesterday's value; this implies that for both the $k=3$ and $k=6$ versions of the ALP the population average bid will be the unconditional expectation of the seller's value, which is 5.[9]

In practice we do not expect observed behavior to correspond to either of these limits but instead to reflect an intermediate weight on recency, so we would like to know the aggregate implications of such intermediate weights in our two conditions. To get a sense of this we now specialize to a specific model that is easy to work with: the temporal difference reinforcement learning model (TDRL). This model has a single parameter that controls the rate at which information from past observations is discounted. Although more complex learning models fit various data better, variations of TDRL have been shown to fit human and animal learning behavior reasonably well (Glimcher et al. 2008) we believe that the qualitative effect of recency on the aggregate distribution of play will be roughly the same for many of the alternative models.

[8] This assumes either that the agents play all actions with positive probability, as in smooth fictitious play, or that they receive counterfactual information about the payoffs of the actions they did not use. Benaim et al. (2009) prove a continuity result that relates the limit of the ergodic distribution of recency-weighted smooth fictitious play to the long run outcomes of smooth fictitious play without any recency bias at all.

[9] Note that if bids are restricted to be integers, the optimal response is to bid the smallest integer larger than the realized computer value; this predicts an average bid of 5.5 in the ALP.

TDRL works as follows: for each action *a* the agent begins at time 1 with a valuation $v_1(a)$ which we assume is chosen randomly.[10] In each period, individuals use a logit choice function, so they choose action *a* with probability

$$P(s_t = a) = \frac{\exp(\gamma v_t(a))}{\sum_{a' \in A} \exp(\gamma v_t(a'))}$$

Here γ represents the degree of maximization; note that as γ goes to infinity, the probability of the action with the highest value goes to 1, so the choice function is approximates maximization, while as γ goes to 0 all actions are chosen with equal probability.

After each choice, individuals receive feedback and update their valuations. In the case of the ALP, individuals receive different feedback depending on whether their offer is accepted or not. We deal with these cases in turn.

First, suppose that the individual's offer is accepted. The individual then sees the seller's valuation for the object in that round. In the TDRL model individuals update their valuations for action *a* according to

$$v_t(a) = v_{t-1}(a) + \alpha\left(\pi_t(a) - v_{t-1}(a)\right)$$

where $\pi_t(a)$ is the *payoff* from choosing action *a* in that round. The basic idea is simple, the function v(*a)* measures the value assigned to action *a*. The term in parentheses represents the *prediction error*. If it is positive, this means *a* did better than expected and conversely if it is negative, then *a* did worse than expected; the value $v_t(a)$ is then incremented upward or downward accordingly. Parameter α is the *learning rate* – the higher it is, the more responsive individuals are to recent rounds.

Note that this formula requires individuals to be able to compute the counterfactual payoffs when information sets are censored. How will updates occur when offers are rejected and individuals are not informed of the seller's valuation? For this

[10] A more realistic model would include a generating process for the initial valuations, but because the initial values are swamped by data fairly quickly we leave this out.

model, we assume that if individuals bid *a* and are rejected they correctly infer that the computer's value *v* was above their bid *a*, draw a random value *v* from the interval [*a*, 10] and update their valuations as if this hypothetical *v* was the true computer value.[11]

We simulate N=1000 agents playing 30 rounds of the ALP. Figure 1 shows the average simulated behavior in the final round for two different *k* values as well as different values of α and γ =1.The simulation results are qualitatively robust to variations in γ. We note that even at α= .9 we still see average offers very close to optimal for *k=6* but in the *k=3* condition low levels of α are required for approximately optimal play.

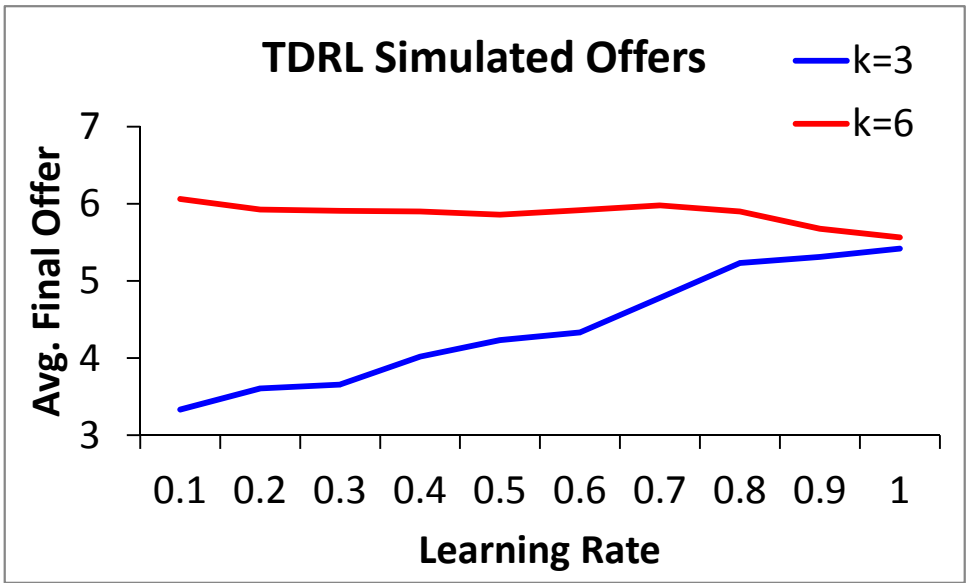

Figure 1: Simulation results for TDRL.

3. General Experimental Design

Subject Recruitment

All of the experimental participants were recruited online using the labor market Amazon's Mechanical Turk (MTurk); we had 617 subjects in various treatments of the ALP and another 150 in the Monty Hall problem described in experiment 4. Although one might worry about the lack of control in an online platform, a number of studies have

[11] Note that this assumes individuals are drawing from the correct distribution for counterfactual payoffs. We will see in experiment 2 that providing the subjects with the full counterfactual information does not change our results..

demonstrated the validity of psychological and economics experiments conducted on the MTurk platform; see for example Ouss & Peysakhovich (2013), Peysakhovich & Karmarkar (2013), Horton et al. (2011), Simons & Chabris (2012), Amir & Rand (2012), Peysakhovich & Rand (2013), Rand et al. (2013). All recruited subjects were US based (had US bank accounts).

In all of the following studies individuals read the instructions for the games and answered a comprehension quiz. Individuals who failed the comprehension quiz were not allowed to participate in the study; reported participant numbers are for those who passed the quiz. Across all of our studies, the failure rate on the quizzes was approximately 25%, which is slightly higher than the rate in the studies above (typically 10-20%). However, our game is much more complicated than simple games such as the one shot Public Goods Game (see Appendix for instructions and example comprehension quizzes).

All experiments were incentive compatible: participants earned points during the course of the experiment that were converted into USD. Participants earned a show-up fee of 50 cents and could earn up to $2 extra depending on their performance. All games were played for points and participants were given an initial point balance to offset potential losses.[12] Experiments took between 10 and 17 minutes thus this amounted to, on average, approximately $9 per hour. Each subject participated in only one experiment in this series.

Experiment 1: Baseline

Design

We recruited N=190 participants to play 30 rounds of the additive lemons problem. In each round participants made a bid to a computerized seller who played the dominant strategy. Participants were informed of the seller's strategy in the instructions.
If a participant's bid was accepted, they received full feedback about the round including the value of v and their payoff. If a participant's bid was rejected, they were informed about this and received no additional information.

[12] In the ALP experiments participants began with a reserve of 50 points; – thus even if participants offered 9 points per round in the $k=3$ condition (where subjects could overpay the most) they would still, on average, not go broke over the course of the experiment.

We varied two parameters to form 4 conditions. First, we varied the level of *k*, setting it equal either to 3 or 6. Second, we varied whether participants were informed of the distribution of seller values. In one case, they were told that *v* is distributed uniformly between 0 and 10. In the other, they were informed that there was a distribution, but not what it was. This gave us 4 conditions, which let us "score" the fit of each of the theories discussed above. Individuals were randomized into a single condition.

The theories discussed above give clear hypotheses about what should happen in each of these treatments. CE predicts overbidding when k=3 and underbidding when k=6 (CE predicts bids of 4 and 5.5 respectively). BE is a set-valued solution concept; it rules out overbidding, allows underbidding in the "uniformed" condition, and predicts optimal bids in the informed treatments. Finally, simulations of the TDRL model with high recency higher than CE-overbidding in aggregate when k=3 and almost NE behavior when k=6 treatments. Table 1 summarizes these hypotheses.

Table 1: Hypotheses by condition in the Baseline Experiment

K=3, Informed			*K=3, Uninformed*		
CE	BE	TDRL	CE	BE	TDRL
> NE	NE		> NE	≤ NE	> NE
K=6, Informed			*K=6, Uninformed*		
CE	BE	TDRL	CE	BE	TDRL
< NE	NE	Close to NE	< NE	NE	Close to NE

Results

Figure 2 shows time courses of average offers binned by 5 round blocks. There is no effect of being informed about the true distribution of values either in the averages (regressions show an insignificant effect of being informed; Table 2) or in the full distribution of behavior in the last round (pooled across *k* conditions, Kolmogorov-Smirnov test p=.68).[13] Bids are higher in the *k=6* condition, as would be expected (Table

[13] One possible explanation of this is that subjects who are not told the true distribution expect it to be uniform distribution. In principle we could test this by using different value distributions, we chose not to because it is harder to describe other distributions to subjects, and the effect of described probabilities is not our main focus.

2). In particular this is true in the first round (col. 3; p<.001), so participants do condition their initial play on this payoff relevant parameter. To increase power, we now pool across informed and uninformed conditions.

We first focus on the *k=3* condition. By the last 10 rounds aggregate behavior appears to have converged: in regressions the significance of round number on bid disappears when we restrict the sample to the last 1/3 of the game (Table 3; mean first 5 rounds = 5.59, mean bid in last 10 rounds = 5.148). In addition, the distribution of bids in the first 10 rounds is significantly different from the distribution of bids in the last 10 rounds (Kolmogorov-Smirnov p<.001), but the distribution of behavior in rounds 11-20 is not significantly different from the distribution of behavior in rounds 21-30 (Kolmogorov-Smirnov p=.43). The average bid in the last 10 rounds is significantly above the optimal bid of 3 (95% confidence interval [4.88, 5.41]).

In the *k=6* condition we see no experience effects. For symmetry with *k=3*, we focus on the last 10 rounds. Here bids are much closer to optimality and are not significantly different from the optimal bid of 6 (mean bid in last 10 rounds = 6.02 with subject-level clustered 95% confidence interval given by [5.72, 6.27]).

To check whether average behavior is the correct statistic to represent the behavior of the population (or, for example, if using the average masks an underlying bimodal distribution), we show the histogram of bids in the last 10 rounds by condition in Figure 3. The bid distributions look well centered, and the aggregate overbidding in *k=3* conditions reflects overbids by many subjects. The figure also shows that behavior does not correspond to a mixture of some subjects optimizing and others choosing at random due to inattention.

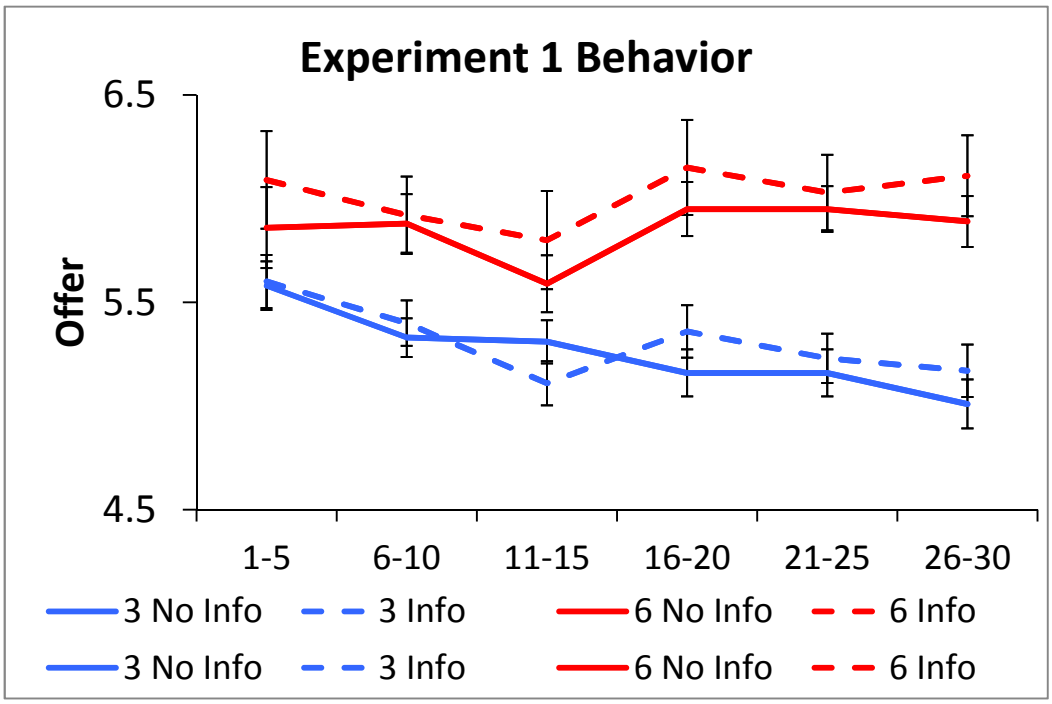

Figure 2: Average bids by condition smoothed at 5 round blocks. Error bars represent standard errors clustered at the participant level.

Table 2: Effects of treatments on behavior.

	bid	bid	First round
1=(k=6)	0.652	0.594	0.948
	(0.164)***	(0.228)*	(0.199)***
1=informed	0.103	0.052	0.135
	(0.163)	(0.223)	(0.197)
round	-0.007	-0.007	
	(0.003)*	(0.003)*	
(k=6)Xround		0.109	
		(0.327)	
constant	5.341	5.368	5.080
	(0.129)***	(0.146)***	(0.156)***
R^2	0.04	0.04	0.11
N	5,700	5,700	190

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

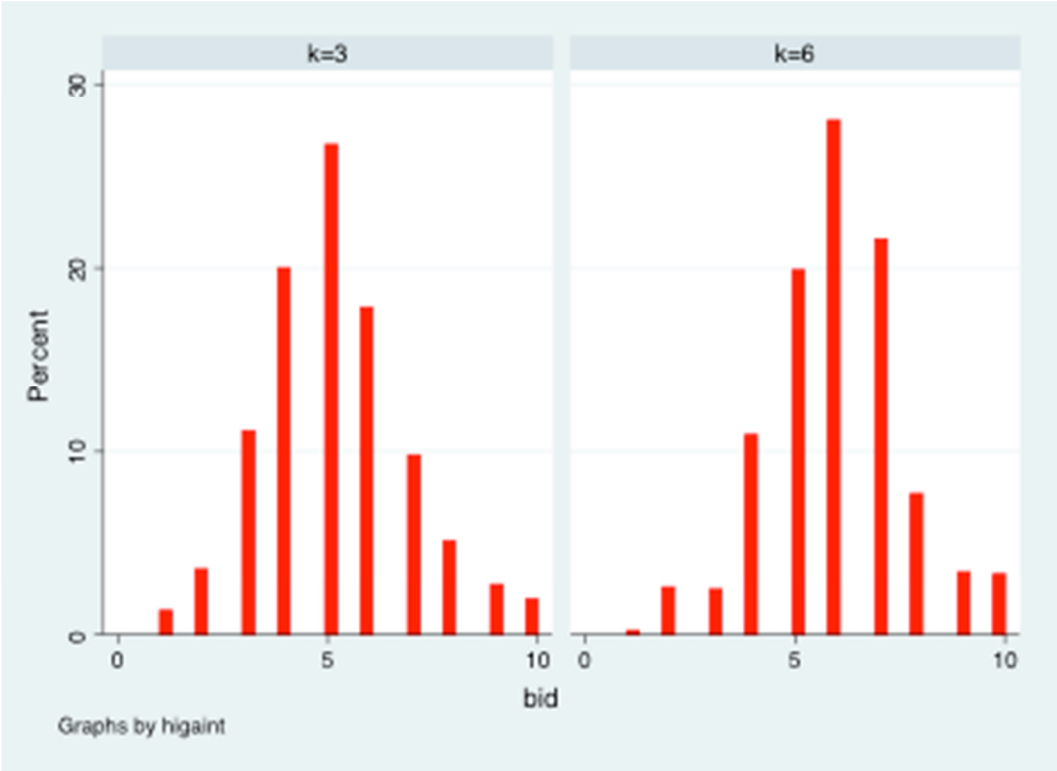

Figure 3: Bids in the last 10 rounds of experiment 1 by condition

Table 3: Experience effects in Experiment 1.

	All rounds	round > 20
K=6 dummy	0.337	0.476
	(0.167)*	(0.599)
round	-0.016	-0.011
	(0.005)***	(0.016)
(k=6)Xround	0.020	0.015
	(0.006)**	(0.022)
constant	5.545	5.423
	(0.115)***	(0.441)***
R^2	0.04	0.06
N	5,700	1,900

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Participants behave nearly optimally in *k=6* conditions but even towards the end of the experiment they behave quite suboptimally in the *k=3* conditions. To show that the misoptimization is economically significant we look at the payoff consequences of these decisions. We define the efficiency of an individual decision as the expected payoff as a percentage of the expected payoff of the optimal strategy. Figure 4 shows that this misoptimization does affect earned payoffs substantially: average efficiency in the last 1/3rd of the game is only approximately 10% in the *k=3* conditions.

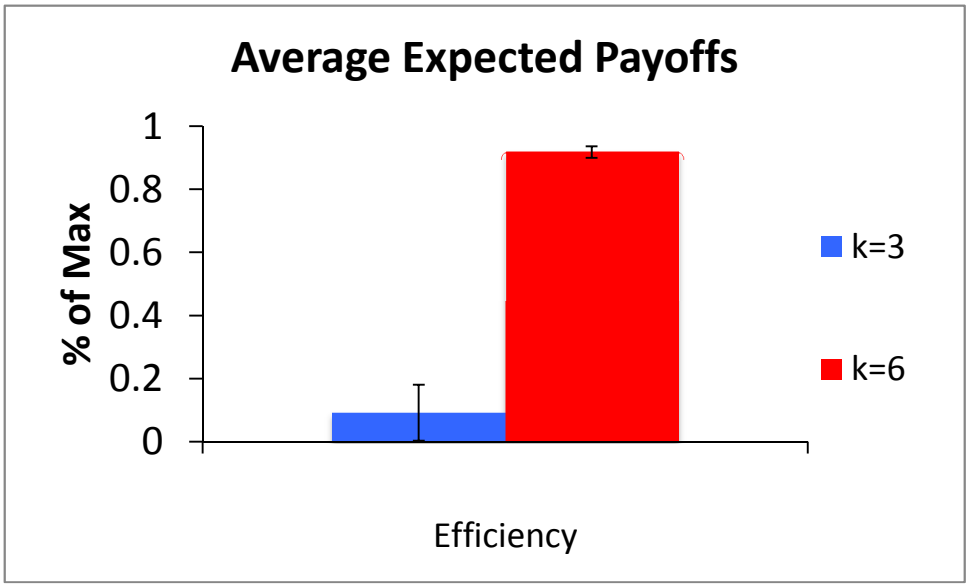

Figure 4: Mean efficiency of offers in last 10 rounds in experiment 1 by condition. Error bars represent standard errors clustered at participant level.

We now turn to the performance of the theories. There is a substantial amount of variance in the last 10 rounds (Figure 3) so we look at the average bids. In the averages, we see overbidding at k=3 and no underbidding at k=6, and we do not see any differences between informed and uninformed conditions. Thus, BE does not fit well with our data, despite its substantial intuitive appeal. The substantial overbidding in the *k=3* conditions is qualitatively consistent with CE, but the overbidding is even higher than that CE predicts, and significantly so (mean bid in last 10 rounds = 5.148, 95% confidence interval clustered at participant level [4.88, 5.41]). Eyster & Rabin (2005) find a similar effect when trying to fit CE models to some experimental data. In addition we do not see the underbidding in the *k=6* conditions that CE predicts (mean bid in last 10 rounds = 6.02 with subject clustered 95% confidence interval given by [5.72, 6.27]).

Finally, we turn to the TDRL model. We first discuss whether the model matches patterns in the aggregate data: as in TDRL simulations with high recency we see that aggregate behavior exhibits extreme overbidding in the *k=3* conditions and optimal behavior in the *k=6* conditions.

Next we look at the dynamics of behavior. Because both CE and BE are equilibrium concepts, they make predictions about aggregate behavior once subjects have enough experience/feedback for equilibrium to roughly approximate their behavior. However, these models do not make predictions about how behavior should change between rounds before the equilibrium is reached, and predict little change in play once subjects have enough experience. In contrast, any learning model with a high weight on recent outcomes predicts there should be *non-random* changes in individual behavior between rounds and that this non-stationarity should continue even when individuals have played and received feedback on a substantial number of trials.

To look for this individual-level effect, we define a variable called Δbid as the offer in round *t* minus the offer in round *t-1*. We then look at how Δbid is affected by what happens in round *t-1,* with the prediction that good outcomes of accepted bids should lead individuals to revise their bid upward, bad outcomes should lead individuals to revise their bids downward and rejections (which indicate that the computer had a high value that round) should lead individuals to (on average) revise their bid upward. Again, we restrict this analysis to the last 1/3 of all rounds, where *aggregate* behavior has converged.

Figure 5 shows Δbid as a function of outcomes in a last round. We look at three bins: when an individual's bid was accepted and earned a positive profit, when bids were accepted and yielded a loss, and when bids were rejected. The figure shows that there is strong relationship between the previous period's outcome and Δbid.

We can also use a linear regression of Δbid on outcomes in rounds where offers were accepted to estimate the average slope of the relationship between last rounds' outcome and bid.[14] Our regressions show a positive sign (Table 4). Additionally, we can

[14] As one might expect, the average response is much steeper in rounds where participants realize a loss compared to those where they realize a gain (Table 4, col. 3; slope in loss = .41, slop in gain = .14).

look at what happens when an offer is rejected: individuals raise their offer by .405 points (95% confidence interval [.302, .508]) next round.

Although the TDRL model with high recency describes first-order patterns in the data well, a high recency parameter implies a very strong behavioral response in the next round's offer (1 for 1 in the limit case of extreme recency), and we do not see such a strong response in the individual-level regressions. We could improve the fit of TDRL by adding additional parameters, but we are content to sacrifice in-sample fit for portability and simplicity. TDLR does better than either of the equilibrium concepts at organizing the general patterns in our experiments, and can provide intuition about the effects of recency bias on the ALP and other learning scenarios.

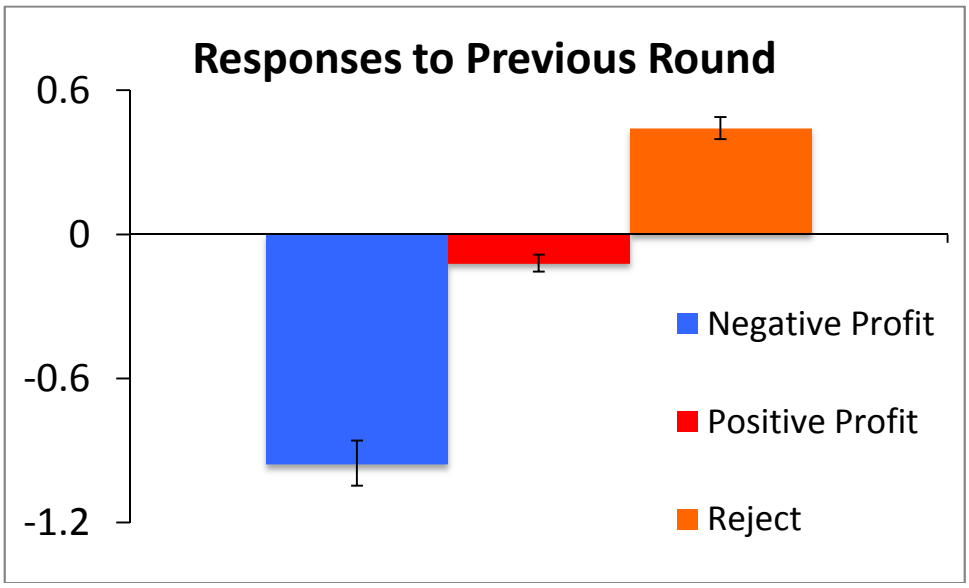

Figure 5: Relationships between outcomes in a round and changes in bid in the last 10 rounds of experiment 1. Error bars represent standard errors clustered at participant level.

Table 4: Effects of information on Δbid (last 10 rounds)

	Δbid	Δbid	Δbid
roundprofit	0.191	0.340	0.143
	(0.026)***	(0.057)***	(0.039)***
1=(K=6)		0.028	-0.221
		(0.143)	(0.101)*
(K=6)Xroundprofit		-0.207	
		(0.068)**	
1=Loss Round			-0.038
			(0.164)
(LossRound)XProfit			0.276
			(0.137)*
constant	-0.576	-0.547	-0.299
	(0.063)***	(0.076)***	(0.079)***
R^2	0.08	0.11	0.10
N	1,014	1,014	1,014

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

To test whether this pattern is driven by a small subset of individuals or is representative, we define a *step* as moving a bid up or down 1 point. We then look at the number of steps that individuals take in the last 10 rounds of the ALP (Figure 6). If the recency results were driven by a small number of individuals then we should expect to see a large mass of individuals at 0. If the results are representative, we should expect to see a smaller mass at 0 and most people taking multiple steps.

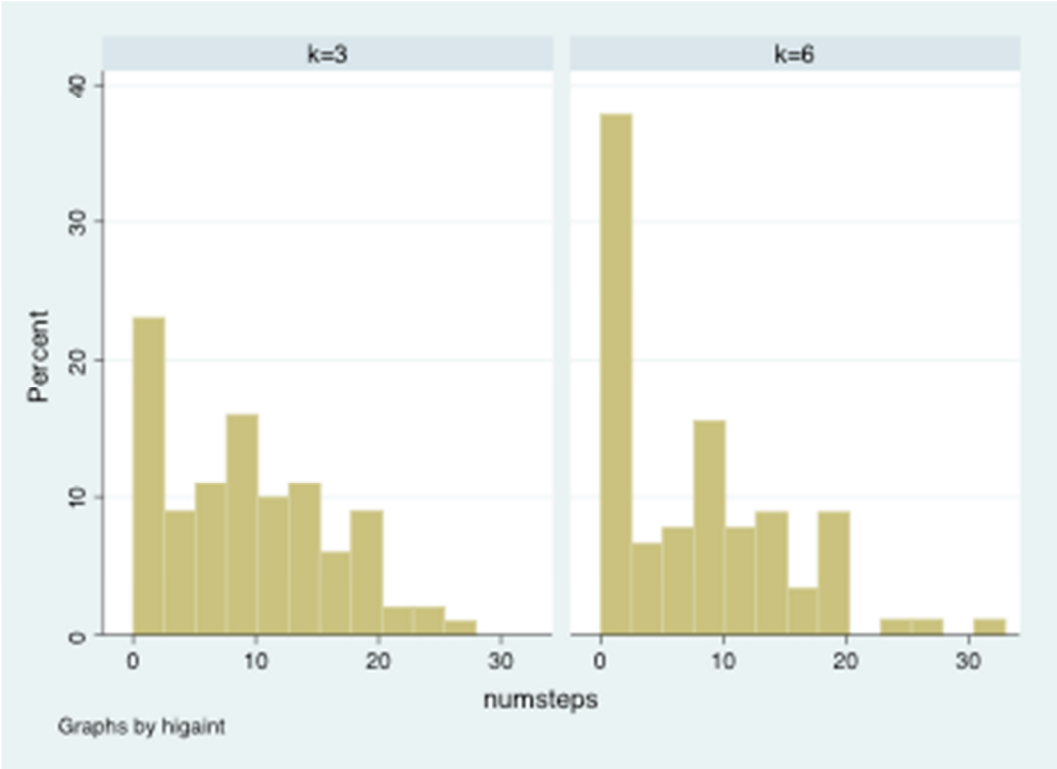

Figure 6: Number of steps taken by individuals. A majority of individuals change their offer decisions between rounds.

Between 65% (*k=6*) and 80% (*k=3*) of participants' offer behavior exhibits persistent variance, even in the last 1/3 of experimental rounds. This finding is hard to reconcile with any sort of equilibrium analysis.[15] We also examine the behavior of the 20-30% of individuals who make no changes to their strategy in the last 10 rounds. Figure 7 shows the bids that these individuals have converged to. The fact that these individuals behavior is constant is consistent with some equilibrium notion, but even

[15] The amount of this variance does not appear to be decreasing during the course of the experiment. The average absolute value of Δbid is 1.02 in the first 10 rounds and .98 in the last 10 rounds and this difference is not significant (two-sided clustered t-test p=.58).

among this group there is substantial overbidding in the k=3 condition. Thus the assumption of rational Bayesians is a poor fit among this group as well.

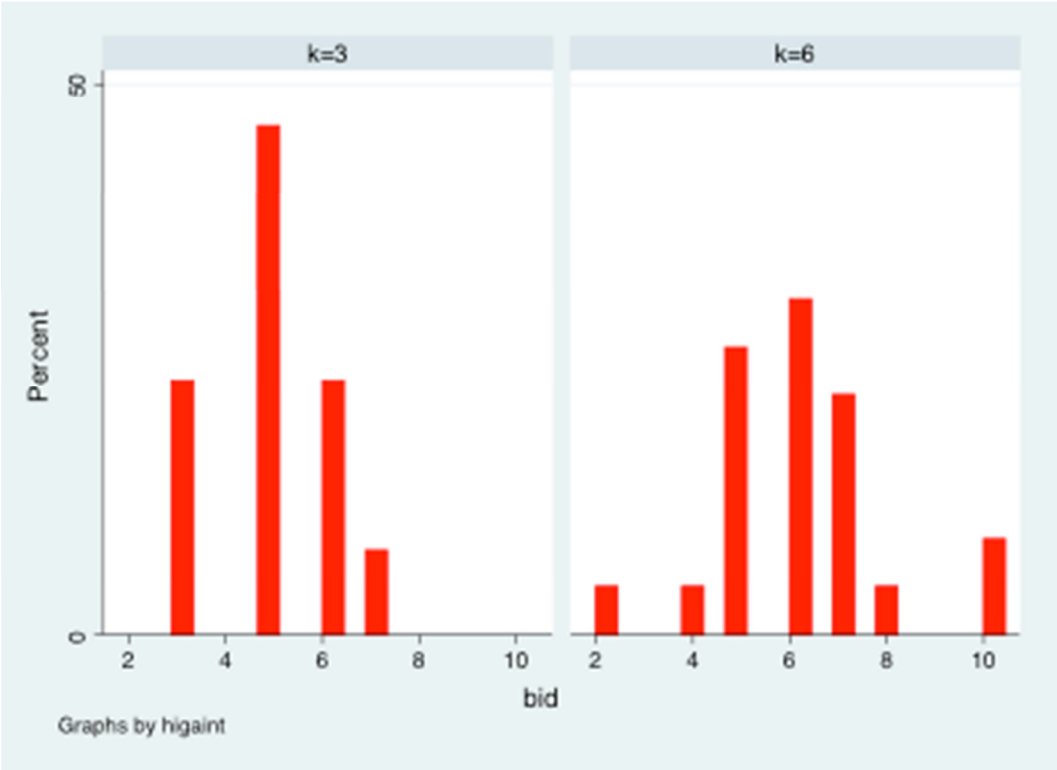

Figure 7: Bids of individuals whose strategies are constant over the last 10 rounds.

Thus we see that most subjects have a high responsiveness to recent outcomes and that a simple model incorporating this recency effect does better at predicting aggregate behavior than the solution concepts of BE and CE. It is important to note that we do think these concepts are useful for some purposes. First, we believe that CE captures a real and important psychological phenomenon. However, we suggest that in repeated contexts CE is most useful as a guide to initial play before subjects get much feedback, while learning models are better for making medium and long-run predictions of most subjects' behavior. Even in this longer run, a concept like CE may act as a useful prediction for some fraction of the population, as shown by the 20-30% of our subjects who do not appear to adjust their behavior in response to feedback. Second, we find it plausible that BE does occur in the field, and other experiments provide evidence for related forms of

self-confirming equilibrium (Fudenberg & Levine 1997) even though there is no evidence for them in the data we consider here.

In experiment 3 we show that our learning model is also useful in that it suggests particular interventions that can lead individuals closer to optimal behavior. Before turning to this, we present another experiment that tests the robustness of our results and further demonstrates the prevalence of recency-based learning.

Experiment 2: Counter Factual Information

The next experiment is designed to control for a potential confound in experiment 1: we saw that the average Δbid in a round in which an offer was accepted was -.270, so the observed overbidding primarily occurs due to individuals moving their bid upward after a rejected offer. One potential explanation for this is misperceptions about the value of v conditional on rejection. To check for this, as well as replicate our original results, we performed a second experiment.

Design

We recruited 75 new participants to play $k=3$ and $k=6$ conditions with one twist: whereas in experiment 1 participants simply received a *rejected* message if their offer was not accepted, participants now received full feedback about the seller's value v whether their offer was accepted or not

Results

Comparing the data from experiment CFI ("counter-factual information") to the behavior from experiment 1, we see little difference between behaviors of individuals who have counterfactual information vs. those who do not (Figure 8). If anything, the individuals with counterfactual information do slightly worse (overbid more) in the $k=3$ condition, but this difference is not significant (Table 5).

As before, the aggregate outcomes are not driven by outliers. Figure 9 shows a histogram of bids in the last 10 rounds of the CFI experiment by condition – the data appear to be well centered. Thus, our results replicate and are the overbidding in $k=3$ conditions are not driven by the lack of information in rejection rounds.

We can also investigate whether most individuals appear to change their behavior from round to round in the CFI experiment. We define a step as in our baseline experiment and plot the number of steps individual behavior takes in the last 10 rounds. As in experiment 1 we find that approximately 80% of individuals indeed change their behavior from round to round (Figure 10).

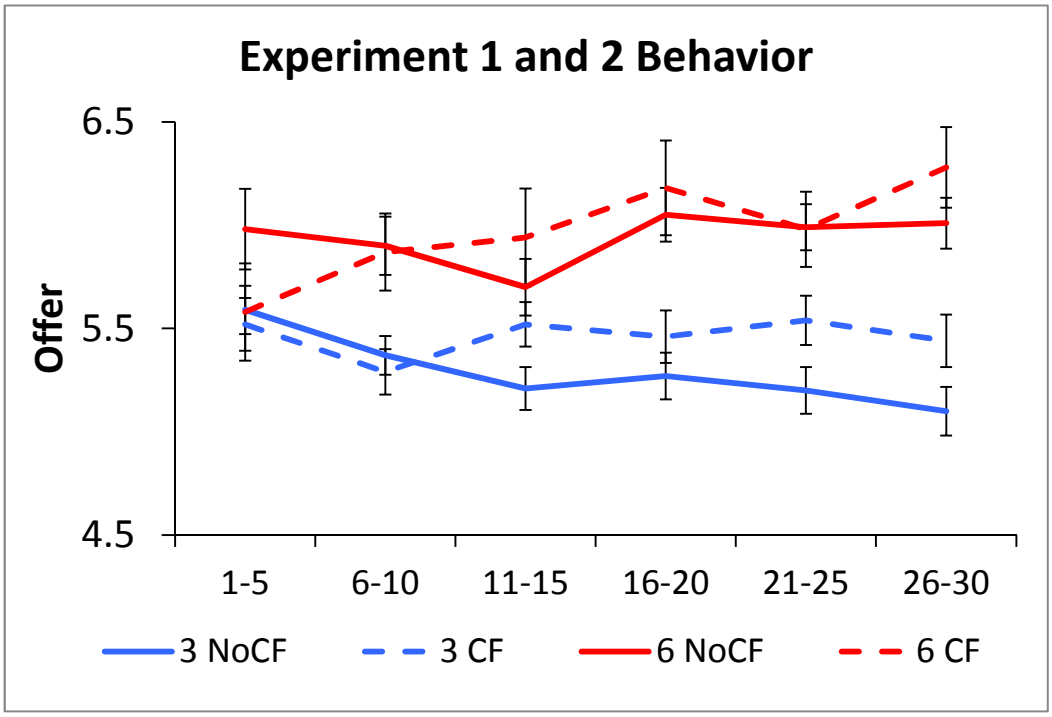

Figure 8: Counterfactual information does not help individuals optimize.

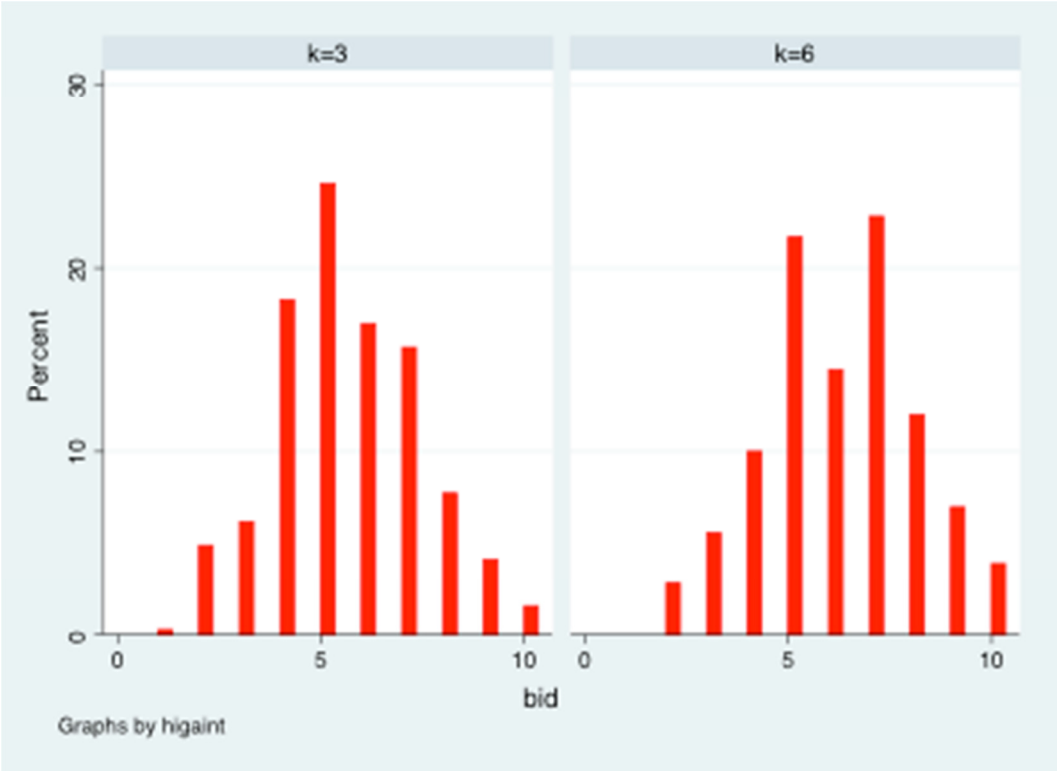

Figure 9: Bids in the last 10 rounds in CFI.

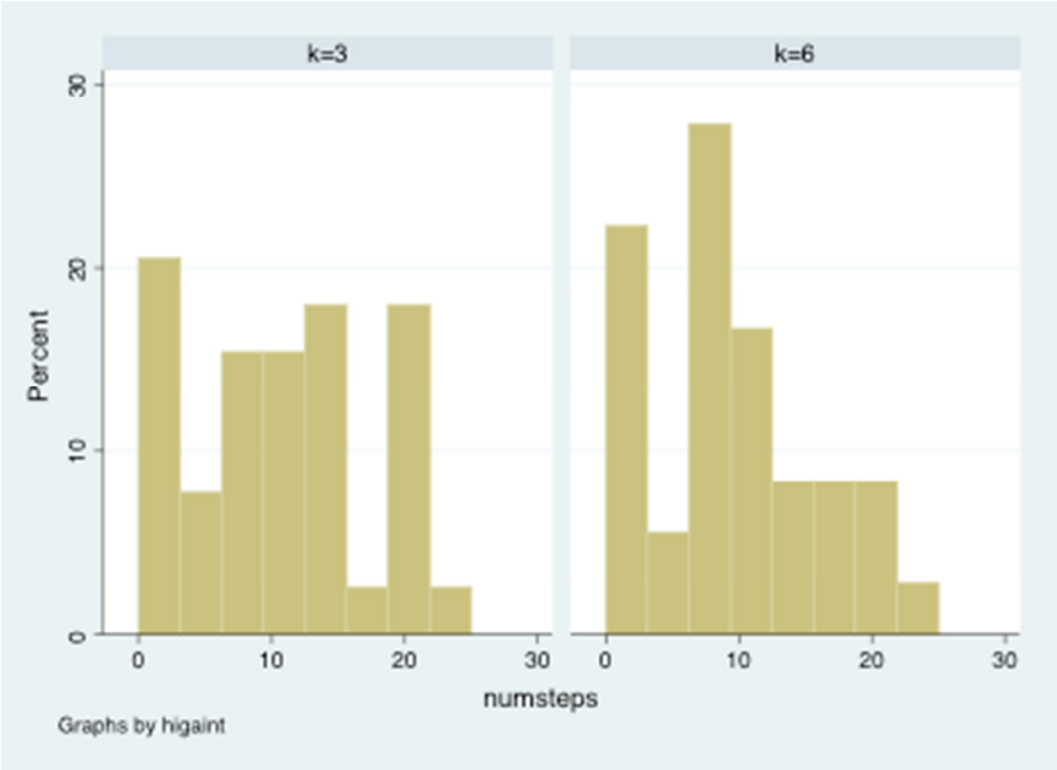

Figure 10: Number of steps taken by individuals in the last 10 rounds in CFI .

Table 5: OLS Regressions pooling data from experiments 1 and 2.

	bid	bid	bid
1=CFI experiment	0.108	0.108	0.174
	(0.148)	(0.148)	(0.182)
round	-0.002	-0.012	-0.012
	(0.003)	(0.004)**	(0.004)**
1=+6	0.614	0.290	0.329
	(0.137)***	(0.145)*	(0.172)
+6Xround		0.021	0.021
		(0.006)***	(0.006)***
+6XCFI			-0.138
			(0.299)
constant	5.335	5.489	5.470
	(0.108)***	(0.110)***	(0.118)***
R^2	0.03	0.04	0.04
N	7,950	7,950	7,950

Coefficients from OLS regressions.
Standard errors in parentheses clustered at participant level.
* $p<0.05$; ** $p<0.01$; *** $p<0.001$

The CFI experiment lets us perform a reduced form test of recency effects. In the baseline experiment, information that individuals received was partially endogenous (high bids were much more likely to get accepted). However, with counterfactual information, the computer's value v acts like an exogenous shock in round t. In a monotone learning model, higher values of v increase potential valuations of higher bids and low values of v decrease valuations of higher bids. Thus, we expect a monotone relationship between bids at time t and histories of observed computer values v.

We can see a recency effect very starkly even in the last 10 rounds. We first split realized computer values into very low (values of 3 or below, bottom 30% of realizations) or very high (values of 7 or bigger, top 30% of realizations). We then take

the average bid of each individual over the last 10 rounds and set that as the individual's "baseline". We then look for the effect of observing a very high or very low value in round *t* on round *t+1, t+2* and *t+3* deviations from this average bid. Here, a positive deviation represents a higher than average (for that individual) bid and a negative deviation represents a lower than average bid. Figure 11 shows there is a large effect on behavior in the $t+1^{st}$ round, and no statistically appreciable effect on the $t+2^{nd}$ or $t+3^{rd}$ round.

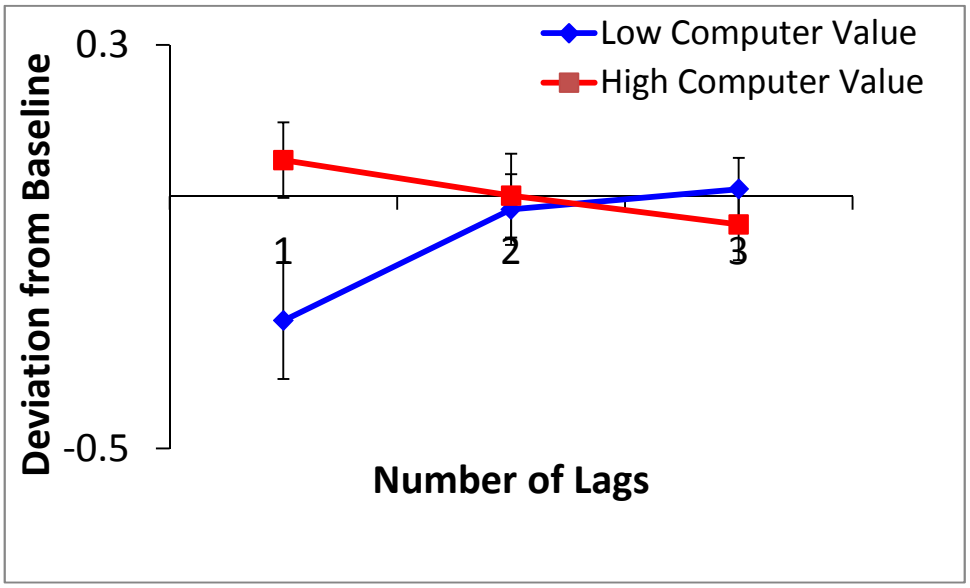

Figure 11: Individual behavior in round *t* conditional on experiencing high/low outcomes in rounds *t-1* or *t-2* or *t-3* in the last 10 rounds.

Figure 11 pools the recency effect across conditions. In our regressions below, we see that when *k=6* the effect of observed computer values is much less pronounced. We attribute this to the relatively flat payoff profile of the bids: Even if the computer's value is extremely low, most bids will make similar (and positive) levels of profits.

An alternative way to quantify recency effects is to regress bids at time *t* on lagged experiences then use a model selection criterion to choose an "optimal" (from a model fit point of view) number of lags. If there are strong recency effects, the selected model should use a relatively small number of lags.

We use the Bayesian Information Criterion (BIC) to select the number of lags to include. The BIC of a fixed model is given by

$$-2\ln(L) + m*\ln(n)$$

where L is the likelihood of the model given the data, n is the number of observations (which we treat as number of subjects) and m is the number of model parameters. Thus, the BIC decreases in variance explained (model fit) and increases in number of free parameters (m*ln(n)).[16]

Table 6 shows the results of this analysis restricted to the second half of the experimental rounds.[17] We use bids pooled across conditions as our dependent variable, add subject level fixed effects and use BIC to select the optimal number of lagged experiences. To account for the fact that the same computer value v may have different effects on behavior in $k=3$ and $k=6$ conditions we also include lagged interaction terms. As predicted by the recency hypothesis, the linear regression model selected by the BIC involves only a single lag.[18]

In addition to recency effects, research on learning and memory also identifies a *primacy effect*: first or initial experiences tend to be recalled more vividly and hence have a large impact on behavior later (Erev & Haruvy 2013). To test for the effects of primacy on behavior we remove the subject level fixed effects from our regressions and add a term for "computer value observed in first round" to the regression, while as above retaining one lagged computer value. Doing so we find that the effect of first round experience is insignificant (p=.943; Table 7 col. 2) so primacy effects do not seem to be important here.

Finally, we turn to individual level heterogeneity. Dropping subject level fixed effects from the one-lag regression we see the R^2 drop dramatically from .49 to .05 (Table 7 col. 1). Thus there is a large amount of individual-level heterogeneity in our data. If a major source of heterogeneity in bids was caused by differences in

---

[16] See Hastie et al. 2009 for a description and derivation of the BIC.

[17] Here we use the last 15 rounds instead of the last 10 as in the rest of our analyses to increase power. Restricting to the last 10 rounds gives qualitatively similar (but less significant) results and BIC continues to select a single lag.

[18] Note that this does not mean that individuals literally discount all information beyond the last outcome. Rather this result means that the additional predictive power from looking at longer individual-level histories is small.

risk attitudes or some other aspect of preferences we would expect first round bids to help explain the subsequent heterogeneity, so we add it as a regressor in Table 7 col. 3. We see that first round bids is not a significant predictor of behavior in the last 10 rounds (p=.167) so we hypothesize that the individual level heterogeneity is driven by differences in learning processes. However, further exploration of the causes of this heterogeneity is beyond the scope of this paper.

As a final robustness check, we can consider another method of model selection from the literature on machine learning: L1 regularized linear regression, i or "lasso." [19] In this procedure the linear regression coefficients β are chosen to minimize the sum of the squared residuals plus a coefficient λ times the sum of the absolute values of the coefficients (ie. a penalty for model complexity). We choose the penalty coefficient λ using cross-validation: splitting the sample into subsets, training a series of models on some of the data and using fit on the held-out set to gauge performances. As in our BIC analysis, lasso selects a model involving only a single lag of information and subject-level fixed effects as the best model for predicting individual bid behavior in the last 10 rounds. We describe the procedure and results more fully in the online appendix.

[19] See Tibshirani (1996) or the textbook by Hastie et al (2009).

Table 6: BIC-based model selection. Regressions include subject-level fixed effects.

	bid	bid	bid	bid
L1.v	0.078	0.075	0.078	0.079
	(0.042)*	(0.041)*	(0.040)*	(0.038)**
L1.(k=6)Xv	-0.092	-0.090	-0.093	-0.096
	(0.046)**	(0.045)**	(0.044)**	(0.043)**
L2.v		0.017	0.012	0.012
		(0.022)	(0.022)	(0.022)
L2.(k=6)Xv		-0.013	-0.008	-0.009
		(0.029)	(0.030)	(0.030)
L3.v			0.030	0.029
			(0.019)	(0.018)
L3.(k=6)Xv			-0.030	-0.027
			(0.028)	(0.027)
L4.v				0.004
				(0.028)
L4.(k=6)Xv				-0.017
				(0.037)
constant	5.615	5.568	5.495	5.519
	(0.123)***	(0.161)***	(0.159)***	(0.174)***
R^2	0.49	0.49	0.49	0.49
N	1,200	1,200	1,200	1,200
BIC	**4063.22**	**4071.17**	**4077.78**	**4086.01**

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 7: Checking for primacy and effect of first round behavior on bids in last 10 rounds.

	bid	bid	bid
L1.v	0.105	0.101	0.104
	(0.053)*	(0.055)*	(0.053)*
1=(k=6)	1.467	1.468	1.433
	(0.448)***	(0.448)***	(0.441)***
L1.vX(k=6)	-0.152	-0.152	-0.151
	(0.063)**	(0.064)**	(0.063)**
First Experience		-0.009	
		(0.060)	
First Bid			0.072
			(0.094)
Constant	4.926	4.996	4.542
	(0.339)***	(0.529)***	(0.621)***
R^2	0.04	0.04	0.05
N	750	750	750

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

A Cautionary Note

We noted before that we chose the ALP to avoid the extremeness of optimal offers that obtains in similar games such as the ACG. When optimal offers are on the boundary of the feasible set, censoring and corner aversion can each generate misleading inferences from the observed data.

Censoring is a mechanical confound present when optimal play is implemented with errors. Recall that in our experimental data we see that distributions of behavior appear to be symmetric and centered on an interior point of the strategy space. Thus, purely idiosyncratic noise should "wash out" in our estimations. When the optimal offer is on the lower boundary of the feasible set (as it is in the ACG) white noise in responses would push the observed mean of decisions higher than the optimum even if each agent's uncensored strategy was to pick the optimum plus a mean-zero noise term Similarly, when $k=6$ in the ALP, censoring would lead to apparent underbidding by optimizing agents if the highest allowed bid were close to 6.

A second potential confound is corner aversion. Existing work shows that individuals are prone to 'avoid the edges' in some strategic settings (Rubinstein et al. 1993). Such a bias would make optimal behavior in the ACG harder to learn than in the ALP. To test the effect of the feasible set on bids, we performed another experiment.

Design

We recruited an additional N=79 individuals to participate in 30 rounds of our baseline ALP conditions with offers restricted to the integers 0-6. This restriction makes the optimal offer in the *k=6* condition on the edge of the interval and the optimal offer in the *k=3* condition squarely in the middle. We see that this lowers offers in the final 10 rounds from 5.09 to 4.29 when *k=3* (clustered s.e. = .209) and from 5.98 to 4.21 (clustered s.e. = .185) when k=*6*. Both of these changes are statistically significant (clustered t-test both p<.01).

We can test whether this decrease is driven purely by censoring of bids or if corner aversion also appears to play a role. To do so, we take combine this data with the results of experiment 1. We then manually censor the distributions of bids in experiment 1. If the difference in outcomes is driven purely by censoring, we should see that the manually censored experiment 1 distribution is identical to the one from the current experiment. If corner aversion plays a role, we should see a difference in distributions. Figure 12 shows the histogram produced by this exercise. The distributions are different across experiments (Kolmogorov-Smirnov tests p<.01 for both *k* conditions) and especially so in the *k=6* condition where optimal bids lie on the border of the strategy space. This shows that both censoring and corner aversion can act as confounds when optimal bids are on the edge of the interval.

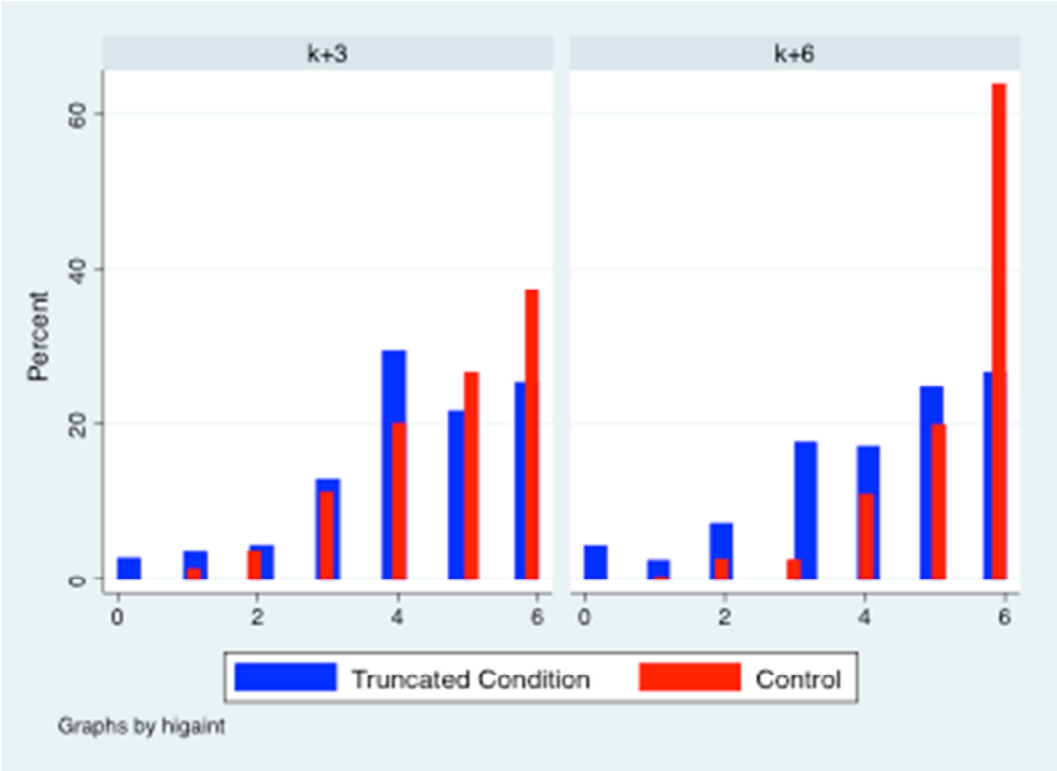

Figure 12: Histogram of bids in last 10 rounds of the restricted strategy space experiment overlaid with experiment 1 bids censored at 6.

Experiment 3: Recaps

We now show that an explicitly dynamic viewpoint is useful in not just describing the data but also in designing interventions to help individuals make better decisions. Thus, considering the dynamics of learning delivers insights that equilibrium models did not.

Why does suboptimal behavior persist at the aggregate level in our experiments? The ALP's feedback structure is such that a relatively small sample of outcomes typically doesn't reveal the optimal bid. Thus high recency acts as a barrier towards learning optimal behavior in this setting. This suggests a prescription for intervention: increasing the number of outcomes subjects observe simultaneously should help them make better decisions. To test this hypothesis, we performed another experiment.

Design

We recruited N=273 more participants. In experiment 3 participants were assigned to one of 3 ALP conditions all with $k=3$. The *control* condition simply replicated the *k=3* condition from experiment 2. In the *more information* condition

participants played the ALP against 10 sellers simultaneously. Sellers' object values were determined independently. Each round (of 30) buyers made a single offer decision that applied to all 10 sellers (who, as before played the optimal strategy). Participants were informed of all this. At the end of a round, participants received feedback about each of their transactions simultaneously: what the seller's value was, whether the offer was accepted and the buyer's profits on that transaction.

There is much existing evidence that in addition to having limited memory, individuals also have limited computational resources (Miller 1956). Thus one may expect that more information is only useful if it is in easily "digestible" form. To look for evidence of computational constraints we added a *simple information* condition. This condition was almost identical to the *more information* condition; individuals played 30 rounds with 10 sellers simultaneously and made a single offer that applied to each seller. However, instead of receiving fully detailed feedback on each transaction, participants received pithy recaps: they were told their *average* profit out of the 10 transactions and *average* values of the objects they actually purchased (see online appendix section 4 for examples of feedback screens).

Results

Figure 13 shows the average offers in the experiment binned in 5 round increments. We see that the addition of more information doesn't seem to help individuals converge to optimal behavior (round 21-30 mean offer in control = 5.02, mean offer in *more info* = 5.22).[20] However, simple information in the form of pithy recaps does appear to be useful (rounds 21-30 mean offers = 4.02), which suggests that the learning problems caused by recency do not stem solely from limited memory.[21] Table 8 confirms the statistical significance of these results. This latter finding is consistent with that of Bereby-Myer & Grosskopf (2008), who find that recaps are

[20] We also replicate our recency results: BIC analysis on the baseline condition of experiment 3 again yields selection of a single lag.(see web appendix).

[21] Sampling -based learning models such as Erev & Haruvy (2013) would also predict that recaps are more useful than records.

helpful in their version of the ACG, though they do not compare recaps with the more-information condition.

Though individuals in the simple information condition still offer above the NE offer in the final round (one sided t-test p<.01) they perform significantly better in terms of efficiency (Figure 14) than in the control and more information treatments. (Table 8; col. 2).

In addition to comparing sample averages we can also see what effect the simple information condition has on the full distribution of behavior. Figure 15 shows a histogram of bids in the last 10 rounds by condition. As the figure suggests, the simple information condition doesn't just affect a subset of the population but rather seems to drive the whole distribution of bids towards the optimum as well as decreasing the variance (test for equal variance in last 10 rounds of baseline vs. simple info. condition rejects equal variance, p<.001).

We also use the control condition to test whether many participants don't understand and/or are not paying much attention to the experiment. We see an effect of the payoff parameter *k* in each of our experiments, even in the first round, and these effects go in the expected direction (higher *k* leads to higher bids) so at least some of our respondents did pay attention. We can further investigate whether inattention is driving our results by using data on response times. (We did not include a response time measure for each individual decision in experiments 1 and 2, but we do include one here.) In the control condition the average participant spends 4.21 seconds on each decision screen (bottom decile of participants spend on average 2.42, top decile of participants spend on average 5.8). We also see no correlation between average response time and bids, earnings or bid variance across experiment (linear regression clustered on participant gives p=.7, .55, .8 respectively). Note that this also implies that response time does not predict extent of recency bias as higher recency bias mechanically correlates negatively with earned payoff.

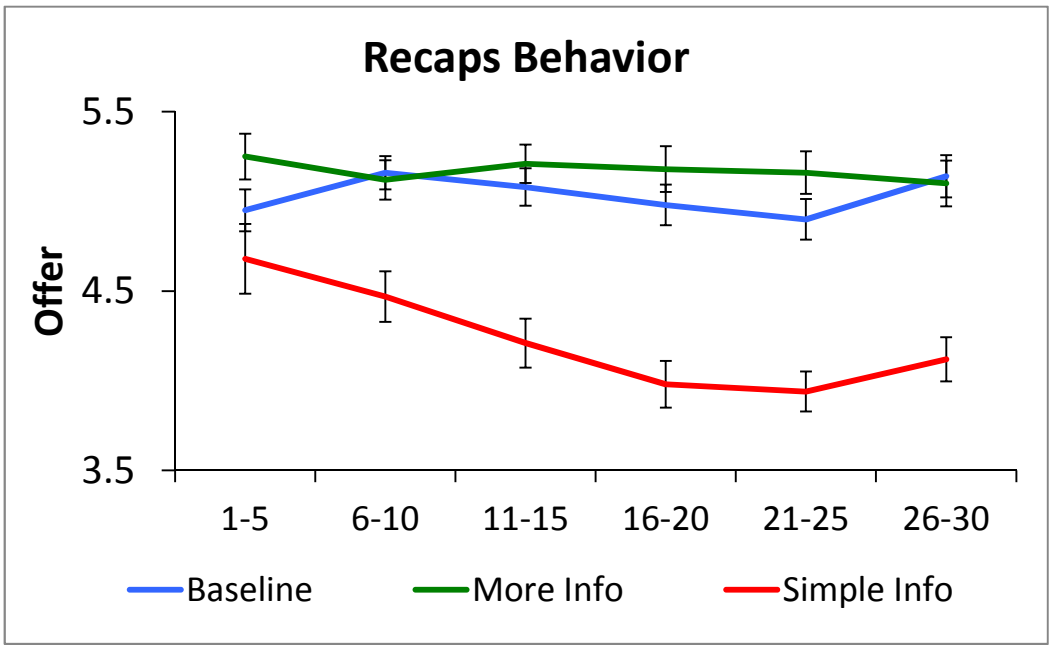

Figure 13: Bids (*k=3*) by 5 round blocks in experiment 3. Error bars represent standard errors clustered at participant level.

Table 8: The effect of condition on bids and efficiency.

	bid	efficiency
1=Extra Info	0.213	-0.213
	(0.263)	(0.171)
1=Simple Info	-0.812	0.489
	(0.192)***	(0.111)***
constant	5.046	0.130
	(0.162)***	(0.104)
R^2	0.05	0.06
N	5,730	5,730

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

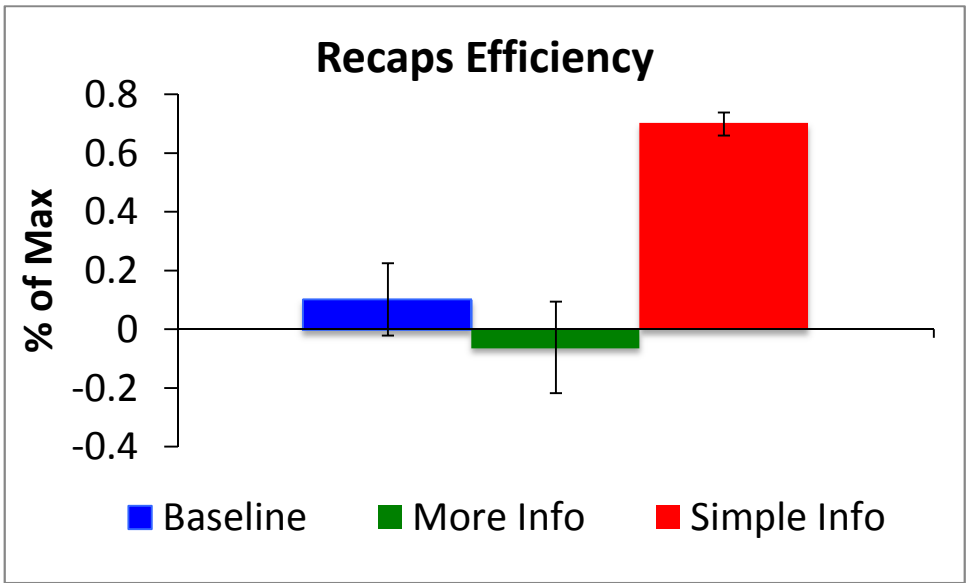

Figure 14: Efficiency of strategies chosen in recaps experiment (last 10 rounds). Error bars represent standard errors clustered at participant level.

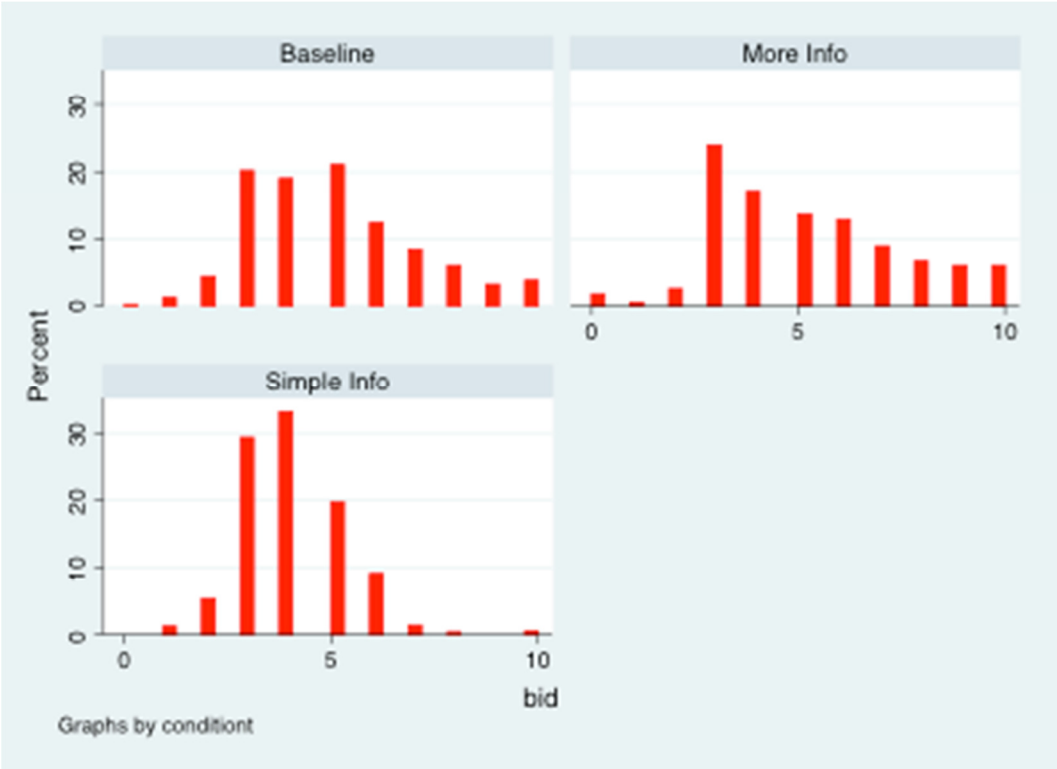

Figure 15: Bids in the last 10 rounds of the recaps experiment, by condition.

Experiment 4: Monty Hall

Our experiments so far have dealt with a specific stochastic decision problem. To test whether our recency-based intuitions extend beyond the ALP we performed another experiment.

Design

We recruited N=139 more participants[22] to participate in 30 rounds of the Monty Hall problem: participants are faced with 3 doors with a prize randomly hidden (by Monty Hall) behind one of the doors. The participants choose a door and Monty opens one of the other doors that does not hide the prize. Participants are then asked if they want to stick to their original choice or switch to the other unopened door.

Note that because Monty always removes a non-winning door, the probability of winning the prize from switching is 66% while the probability of winning from your initial choice is 33%. As in the ALP, to find the optimal strategy participants need to correctly compute conditional probabilities, so we again expect mistakes. Moreover, past work on Monty Hall games has concluded that most individuals will not learn the optimal strategy, even with experience and some sorts of "nudges" (Friedman 1998, Kraus & Wang 2003, Tor & Bazerman 2003, Palacio-Huerta 2003).

Our subjects play a modified version of this game. First, we added a small cost of switching choices (the prize was 10 points and the cost of switching was 1 point). Without this modification CE predicts that participants are indifferent between switching and keeping (both having a probability of 1/3 of winning in CE). With the addition of this small cost the unique CE prediction is for participants to choose to keep their original choice.

Our experiment included two conditions. One control condition where participants played 30 rounds of the Monty Hall game with full feedback and one simple information condition where participants played 30 rounds against 30 computers per round and received feedback in a manner similar to the simple information condition of the ALP (see online Appendix for screenshots of feedback screens).

[22] A total of 150 participants finished the experiment. Due an error in our experimental program, choices for some rounds were not recorded for 11 participants; we drop these participants in what follows. Including them in the analysis does not change the results.

Because the Monty Hall Problem has received attention in popular culture, we framed it in terms of choosing one of three cards (see the online Appendix for screenshots of the experiment/full experimental instructions). At the end of the experiment, participants were asked whether they were familiar with this game or any game like it: approximately 9% of our participants responded affirmatively (and had a 98% switch rate during their trials). We restrict our analysis to the naïve subjects (N=126).

Results

Figure 16 shows the aggregate behavior by condition. As in the ALP we found that with only standard feedback subjects did quite poorly. In the last 10 rounds of the control condition subjects played the optimal strategy only 37% of the time. However, in the simple information condition subjects played the optimal strategy (always switch) 77% of the time in the last 10 rounds. This difference is statistically significant (Table 9 col. 1).

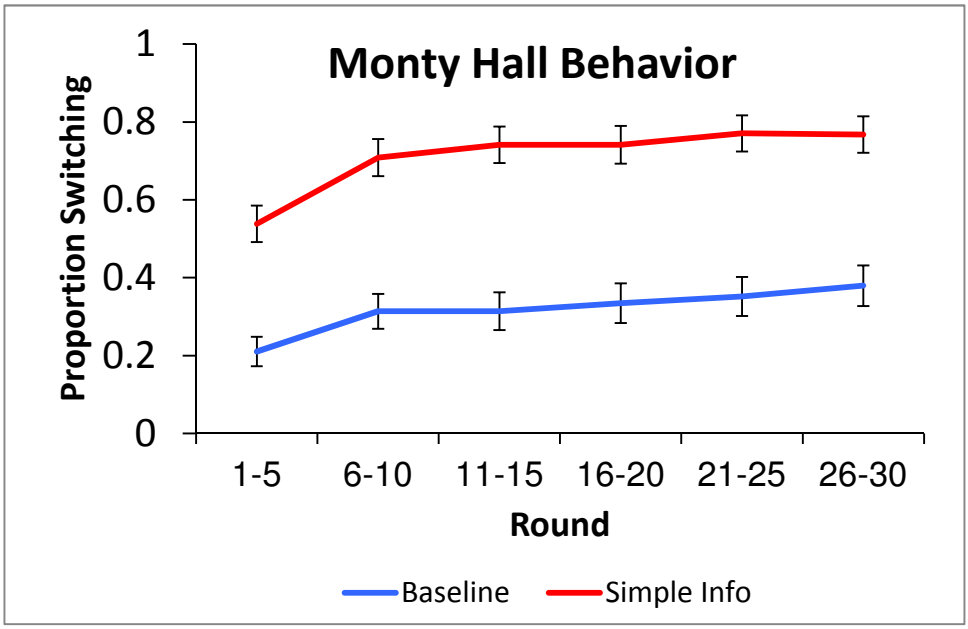

Figure 16: Monty Hall behavior by condition, by five round blocks. Error bars represent standard errors clustered at participant level.

There is an experience effect in the control condition. But this effect disappears in the last 10 rounds. To confirm this statistically, we use a linear probability model to

predict choosing to switch or not. We include a round indicator, a condition dummy and an interaction term. The coefficient on round is significant and positive in the first 10 rounds (as is the interaction with condition; Table 9 col. 2) but it is much smaller and not significant in the last 10 rounds (Table 9 col. 3). This is again consistent with high recency - the more recency a learning model displays, the faster it discounts initial conditions and the faster aggregate behavior converges to the ergodic distribution.

To make the finding even more stark, Figure 17 shows the number of switches individuals make in the last 10 rounds by condition. Since it is always optimal to switch, making 10 switches implies that an individual played optimally in each of the last 10 rounds. Many more individuals act in an optimal manner in the last 10 rounds of the simple information condition (57% vs. 10%; Ranksum test $p<.001$).

In the ALP we explicitly checked for recency by looking at the effects of last period's computer value on current bid. In the Monty Hall problem, it is less clear how to check for recency in a reduced form specification. The simplest approach is to suppose that subjects use the binary strategy space "stay" or "switch." Replicating our regression analyses using this strategy space reveals no relationship between information and behavior. However, it is not clear that this is the appropriate strategy space; alternatives include a preference for ending up with a given card, and the 6- element set (3 choices for initial card followed by “switch” or “stay”). This problem is confounded further by the fact that different individuals may use different strategy spaces. As a consequence we do not have an explicit demonstration that recency matters in the ALP, but the fact that the simple recaps that helped subjects in the ALP help here as well is suggests that a similar learning process may be involved.

Table 9: Linear probability model predicting switching.

	All Rounds	Early Rounds	Late Rounds
1=Simple Info	0.394	0.282	0.506
	(0.061)***	(0.066)***	(0.207)*
round		0.022	0.004
		(0.006)***	(0.007)
(Simple)Xround		0.014	-0.004
		(0.008)	(0.008)
constant	0.317	0.144	0.270
	(0.042)***	(0.042)***	(0.184)
R^2	0.15	0.16	0.17
N	3,780	1,260	1,260

Coefficients from OLS regressions.

Standard errors in parentheses clustered at participant level.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

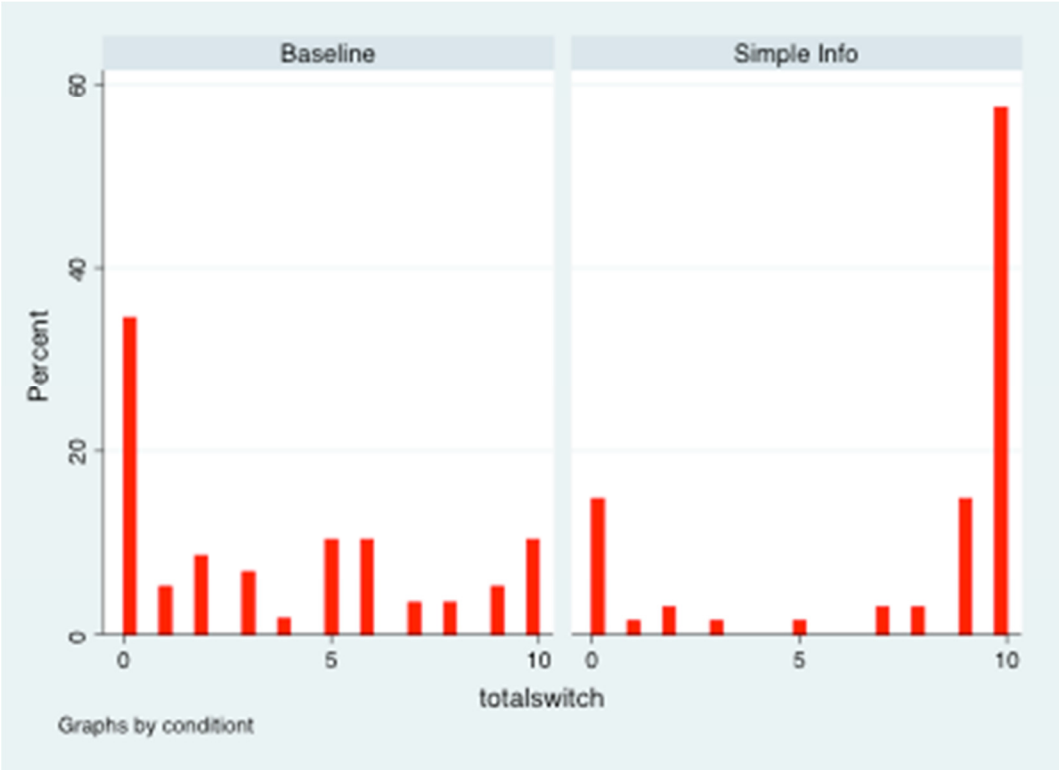

Figure 17: Number of switches made in the last 10 rounds of the Monty Hall problem, by condition.

Conclusion

Although suboptimal behavior *can* persist in the play of simple Bayesian decision problems, this depends crucially on the structure of the feedback that individuals receive. Importantly, we find that just giving participants more information does not aid convergence to optimal behavior, but recaps can be highly effective. We argued that the superiority of recaps comes from economizing on costly cognition. Table 10 summarizes our interpretations of our experimental findings.

Table 10: Recap of experimental findings

Experiment	**N**	**Findings**
Baseline	190	Overbidding in *k=3*, optimal bidding in *k=6*, no effect of being informed of distribution, evidence of recency effects
Counterfactual	75	Baseline results not driven by lack of counterfactual info, more evidence of recency effects, no evidence of primacy, large individual-level heterogeneity
Truncated	79	Corner aversion can act as a confound when optimal bids are extreme
Recaps	273	More information doesn't help but simple recaps drive behavior closer to optimality
Monty Hall	139	Recaps help in other confusing decision problems
TOTAL	**767**	Learning models good at describing behavior in stochastic decision-problems, recency is not just a memory limitation, recaps useful for bringing behavior closer to optimal

Our results demonstrate that explicitly dynamic models of behavior can yield insights in ways that equilibrium models cannot. None of the equilibrium concepts (NE/CE/BE) we consider are able to capture the full variation of behavior in the ALP. By contrast, a learning model with high recency fits aggregate behavior across treatments well. In addition, thinking dynamically gives us intuition about interventions via feedback structure to help nudge individual behavior closer to optimum.

Our experiments show that computational, not just memory, constraints may contribute to the persistence of suboptimal behavior. This reinforces earlier arguments that recency effects are in part driven by computational constraints (Fiedler 2000; Hertwig & Pleskac, 2010), and suggests that here at least memory load is not a primary

driver of recency, in contrast to Barron & Erev (2003). Our results thus support incorporating more accurate representations of computational limits and other forms of bounded rationality into existing learning models.

In addition, there is a debate in the literature about whether findings from learning experiments such as ours can be applied to understand behavior in the field. Individuals may have computational constraints, but in the field they often have access to technological aids. This is argued strongly in Levine (2012):

> *"Even before we all had personal computers, we had pieces of paper that could be used not only for keeping track of information – but for making calculations as well. For most decisions of interest to economists these external helpers play a critical role…"*

Such technologies can provide recaps and thus help guide individuals towards optimal decisions. On the other hand, there is evidence of significant economic costs due to incomplete learning and recency bias in contexts such as credit card late fees (Agarwal et al. 2008), stock market participation (Malmandier & Nagel 2011) and IPO investment (Kaustia & Knupfer 2008), which suggests that even when recaps and record-keeping devices are available they may not be utilized. Recently Hanna et al. (2013) studied the effect of what we call recaps on seaweed farmers in Bali. Seaweed farmers "plant" pods and later harvest them. Two choice variables affect the production of crop: the size of the pods planted and the spacings between pods. Hanna et al. (2013) find that even experienced farmers typically used very suboptimal pod weights and spacings, and that providing a summary table on the returns from different methods that highlighted which method had the highest yields led to substantial improvements in yield and income. They argue that the effect of their summaries comes from focusing farmers' attention on the size of pod used. We note that the effect of their summaries is also consistent with our explanation based on limited cognition. Using their data alone, it seems hard to discriminate between an attention-based explanation and our claim that recency comes from some sort of computational constraint but the focusing explanation seems less suited to our findings in the ALP. At this point, though, the case seems far from settled. Further studies of recaps in the lab and the field could have both scientific and social benefit.

References

Agarwal, S., Driscoll, J. C., Gabaix, X., & Laibson, D. (2008). Learning in the credit card market: National Bureau of Economic Research.

Akerlof, G. A. (1970). The market for" lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84: 488-500.

Amir, O., & Rand, D.G. (2012). Economic games on the internet: The effect of $1 stakes. *PLoS One, 7*

Ball, S. B., Bazerman, M. H., & Carroll, J. S. (1991). An evaluation of learning in the bilateral winner's curse. *Organizational Behavior and Human Decision Processes, 48*:1-22.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of behavioral decision making, 16*: 215-233

Battigalli, P. & D. Guatoli (1997). "Conjectural Equilibria and Rationalizability in a Game with Incomplete Information" in *Decisions, Games and Markets*, Kluwer Academic Publishers, Norwell, MA.

Benaïm, M., J. Hofbauer, & E. Hopkins (2009). "Learning in Games with Unstable Equilibria," *Journal of Economic Theory* 144: 1694-1709.

Bereby - Meyer, Y., & Grosskopf, B. (2008). Overcoming the winner's curse: an adaptive learning perspective. *Journal of Behavioral Decision Making, 21*: 15-27.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton: Princeton University Press.

Camerer, C., & Ho, T. (1999). Experience - weighted Attraction Learning in Normal Form Games. *Econometrica, 67:* 827-874.

Charness, G., & Levin, D. (2009). The origin of the winner's curse: a laboratory study. *American Economic Journal: Microeconomics,* 1: 207-236.

Cheung, Y. & D. Friedman (1997). "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior* 19: 46-76

Crawford, V. P., & Iriberri, N. (2007). Level - k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private - Value Auctions? *Econometrica, 75*: 1721-1770.

Dekel, E., Fudenberg, D., & Levine, D. K. (1999). Payoff information and self-confirming equilibrium. *Journal of Economic Theory, 89*: 165-185.

Dekel, E., Fudenberg, D., & Levine, D. K. (2004). Learning to play Bayesian games. *Games and Economic Behavior, 46*: 282-303.

Erev, I. & Haruvy, E. (2013). Learning and the economics of small decisions. *Forthcoming in The Handbook of Experimental Economics, Vol. 2* Roth, A.E. & Kagel, J. Eds.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88: 848-881.

Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. *American Economic Review*, *98*: 1269-1291.

Eyster, E., & Rabin, M. (2005). Cursed equilibrium. *Econometrica, 73*: 1623-1672.

Fiedler K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107*: 659-676

Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *American Economic Review, 8*8: 933-946.

Fudenberg, D. & Y. Kamada (2013). Rationalizable Partition-Confirmed Equilibrium, mimeo.

Fudenberg, D. & D. Kreps (1995). Learning in Extensive Games, I: Self-Confirming Equilibria," *Games and Economic Behavior,* 8: 20-55.

Fudenberg, D., & Levine, D. K. (1993a). Self-confirming equilibrium. *Econometrica 61*: 523-545.

Fudenberg, D., & Levine, D. K. (1993b). Steady state learning and Nash equilibrium. *Econometrica 61*: 547-573.

Fudenberg, D., & Levine, D. K (1997). Measuring Player's Losses in Experimental Games," *Quarterly Journal of Economics*: 112, 479-506.

Fudenberg, D. & Levine, D.K. (1998). *The theory of learning in games* MIT press, Cambridge MA

Glimcher, P. W., Fehr, E., Camerer, C., & Poldrack, R. A. (2008). *Neuroeconomics: Decision making and the brain*: Access Online via Elsevier.

Guarnaschelli, S., McKelvey, R. D., & Palfrey, T. R. (2000). An experimental study of jury decision rules. *American Political Science Review*, 94:407-423.

Hanna, R., S. Mullainathan, and J. Schwartzstein (2013) Learning through Noticing: Theory and Experimental Evidence in Farming," mimeo.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. *Springer,* N.Y.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition, 115*: 225-237

Horton, J., Rand, DG & Zeckhauser, RJ. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics, 14*: 399-425.

Kagel, J. H., & Levin, D. (1986). The winner's curse and public information in common value auctions. *American Economic Review*, 76:894-920.

Kaustia, M., & Knüpfer, S. (2008). Do investors overweight personal experience? Evidence from IPO subscriptions. *The Journal of Finance, 63* :2679-2702.

Krauss, S., & Wang, X. (2003). The psychology of the Monty Hall problem: discovering psychological mechanisms for solving a tenacious brain teaser. *Journal of Experimental Psychology: General, 13*2:3-22.

Lehrer, E. (2012): "Partially-Specified Probabilities: Decisions and Games," *American Economic Journal: Microeconomics*, 4: 1-33.

Levine, D. K. (2012). *Is Behavioral Economics Doomed?: The Ordinary Versus the Extraordinary*: Open Book Publishers.

Malmendier, U., & Nagel, S. (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? . *Quarterly Journal of Economics, 126*: 373-416.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*: 81-97.

Nevo, I., & Erev, I. (2012). On surprise, change, and the effect of recent outcomes. Frontiers in Cognitive Science, 3.

Ouss, A., & Peysakhovich. A. (2013). When Punishment Doesn't Pay: 'Cold Glow' and Decisions to Punish. *SSRN Working Paper Series*.

Palacio-Huerta, I. (2003), Learning to Open Monty Hall's Doors. Experimental Economics, vol. 6: 235-251.

Peysakhovich, A. & Karmarkar, U. (2013). Asymmetric effects of favorable and unfavorable information on decision-making under ambiguity. *Mimeo*.

Peysakhovich, A. & Rand, D. (2013). Habits of virtue: creating norms of cooperation and defection in the laboratory. *SSRN Working Paper Series*.

Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology, 53*: 793-808.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G., Newman, G., Wurzbacher, O., Nowak, M.A. & Greene, J. (2013). Intuitive Cooperation and the Social Heuristics Hypothesis: Evidence from 15 Time Constraint Studies. *SSRN Working Paper Series*.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior, 8*:164-212.

Rubinstein, A., Tversky, A., & Heller, D. (1993). Naive strategies in zero-sum games: Sackler Institute of Economic Studies.

Samuelson, L. (1997) *Evolutionary Games and Equilibrium Selection,* MIT Press, Cambridge, MA.

Samuelson, W. F., & Bazerman, M. H. (1985). Negotiation under the winner's curse. *Research in experimental economics, 3*: 105-138.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*: 1-27.

Simons, D. J., & Chabris, C. F. (2012). Common (Mis) Beliefs about Memory: A Replication and Comparison of Telephone and Mechanical Turk Survey Methods. *PLoS One*: *7* e51876.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* .Cambridge Univ Press, Cambridge UK.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*58:267-288.

Tor, A., & Bazerman, M. H. (2003). Focusing failures in competitive environments: Explaining decision errors in the Monty Hall game, the acquiring a company problem, and multiparty ultimatums. *Journal of Behavioral Decision Making, 16*: 353-374.

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