aptara	IERE	iere˙560	Dispatch: August 22, 2009	CE: OIO
	Journal	MSP No.	No. of pages: 24	PE: Martha

INTERNATIONAL ECONOMIC REVIEW
Vol. 50, No. 4, November 2009

SUBJECTIVE PROBABILITIES IN GAMES: AN APPLICATION TO THE OVERBIDDING PUZZLE*

BY OLIVIER ARMANTIER AND NICOLAS TREICH[1]

Federal Reserve Bank of New York, U.S.A., Université de Montréal, CIRANO, and CIREQ, Québec, Canada; Toulouse School of Economics (LERNA, INRA), France

This article illustrates how the joint elicitation of subjective probabilities and preferences may help understand behavior in games. We conduct an experiment to test whether biased probabilistic beliefs may explain overbidding in first-price auctions. The experimental outcomes indicate that subjects underestimate their probability of winning the auction, and indeed overbid. When provided with feedback on the precision of their predictions, subjects learn to make better predictions, and to curb significantly overbidding. The structural estimation of different behavioral models suggests that biased probabilistic beliefs is a driving force behind overbidding, and that risk aversion plays a lesser role than previously believed.

1. INTRODUCTION

It has been well established in the psychology literature on judgment that the subjective beliefs agents hold about probabilities are typically biased. In particular, psychologists have observed that individuals tend to overestimate low probabilities, whereas they often underestimate high probabilities.[2] Likewise, several individual decision experiments suggest that agents exhibit comparable biased probabilistic beliefs when making risky choices.[3] The object of this article is to show how the joint elicitation of subjective probabilities and preferences may help us better understand behavior in a game. More specifically, we conduct an experiment showing that biased probabilistic beliefs may be considered a driving force behind the overbidding puzzle.[4]

* Manuscript received October 2006; revised August 2007.

[1] We would like to thank Paul Pezanis-Christou, Jordi Brandts, as well as seminar participants at the Institut d'Analisi Economica, the University of Auckland, Iowa State University, the 2004 ESEM conference in Madrid, the 2004 ESA conferences in Amsterdam and Tucson, the 2004 Experimental and Behavioral Economics Workshop in Calgary, and the 2005 JEE in Rennes for helpful discussions. Finally, we would like to thank three anonymous referees for the quality of their comments. All remaining errors are ours. The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of New York, or the Federal Reserve System. Please address correspondence to: Olivier Armantier, Research Department, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045. E-mail: *olivier.armantier@ny.frb.org*.

[2] See, e.g., Sanders (1973), and Murphy and Winkler (1984) for probability forecasting in meteorology; Lichtenstein et al. (1978), Viscusi et al. (1997), and Benjamin et al. (2001) for predictions of lethal risks; Viscusi and O'Connor (1984), and Gerking et al. (1988) for perception of job-related hazards; as well as Hurley and Shogren (2005) for evidence in sterile laboratory experiments.

[3] See the numerous references in Camerer (1995) for lottery choices; Ali (1977), and Golec and Tamarkin (1998) for horse-track betting; as well as Wakker et al. (1997) for insurance decisions. Note that it has been suggested that the subjective probabilities in individual decision experiments may reflect both the beliefs agents may hold about the probabilities they face, as well as intrinsic preferences over probabilities when making a risky choice (see, e.g., Kahneman and Tversky, 1979). Throughout the article, we concentrate exclusively on subjective beliefs about probabilities.

[4] A clarification about semantics is in order: throughout the article we characterize the disconnect between objective and subjective probabilities as a bias. We do not, however, take a stand on the source of this disconnect. In particular, we do not suggest that it results from cognitive errors. As suggested by one of the referees, the disconnect may also reflect the agent's lack of rational expectation.

There is a wealth of evidence indicating that subjects in independent first-price private-values auctions tend to bid above the risk-neutral Bayesian Nash equilibrium (hereafter RNBNE).[5] Although often rationalized by risk aversion, there does not seem to be a consensus in the literature around the cause(s) of overbidding.[6] In particular, Goeree et al. (2002) (hereafter GHP) find that a model with a sensible constant relative risk aversion parameter, and a model in which risk-neutral agents have biased beliefs about their probability of winning the auction, both fit observed bids equally well.[7] GHP's experimental data, however, do not allow them to disentangle the two effects.[8] Indeed, they only observe bidding choices, and it is well known that different combinations of beliefs and preferences may generate the same observed behavior.

Following Manski (2002, 2004), we propose to circumvent this identification problem by conducting an experiment in which we jointly elicit choices and subjective probabilities. Indeed, if we observe both choices and subjective probabilities, then we can jointly estimate risk aversion and probabilistic beliefs, and thereby test directly whether one of these two effects is more relevant to explain overbidding. Our approach however, is only valid under two conditions. First, in accordance with standard theory, subjects must hold probabilistic beliefs on which they base their actions. Second, the belief-elicitation technique must be effective and it should not affect bidding behavior.[9] Although we will not be able to test formally whether these conditions hold, we will present supporting evidence.

The experiment we propose consists of a first-price independent private-values auction similar to the one in GHP. In addition, we also ask subjects to predict their probability of winning the auction in order to elicit their probabilistic beliefs. We conduct two informational treatments differentiated by the feedback provided to subjects at the end of each round on the precision of their prediction. No information is revealed in the *no-feedback* treatment, whereas in the *feedback* treatment, subjects are informed of the quality of their prediction. We also conduct three belief-elicitation treatments, each using a different payment scheme.

The experimental outcomes in the *no-feedback* treatment indicate that subjects overbid and underestimate their probability of winning the auction. However, after observing their objective probability of winning, subjects in the *feedback* treatment not only learn to make more accurate predictions, but they also drastically curb their tendency to overbid. In fact, bidding above the RNBNE virtually disappears once subjects have learned to predict correctly their probability of winning. These results therefore illustrate how probabilistic beliefs play a central role in shaping behavior in games.

To explain further the experimental outcomes, we estimate several structural models of noisy behavior. The estimation results suggest that subjects are heterogenous with respect to both probabilistic beliefs and risk preference. In addition, we find that actions seem to be consistent with subjective beliefs, as subjects appear to "best-respond" to their predictions about their

[5] Although the phenomenon has been documented in other contexts (e.g., second price auction), we restrict our attention in the present article to overbidding in independent first-price private-values auctions.

[6] When first documented, overbidding was initially rationalized by risk aversion (see, Cox et al., 1983). Several experiments have since questioned this hypothesis (see, Kagel, 1995 for a survey of these experiments). Other explanations proposed include a "joy of winning" (see, e.g., Cox et al., 1983, 1988), a lack of monetary incentives (the "flat maximum critique" of Harrison, 1989), bidding errors (Kagel and Roth, 1992), and asymmetric costs of deviating (Friedman, 1992). See also the December 1992 issue of the American Economic Review for a flavor of the debate pertaining to the causes of overbidding.

[7] Cox et al. (1985) were in fact the first to propose to study first-price auctions with a utility function exhibiting nonlinearity in the probabilities. The authors however, only consider a power probability weighting function (PWF), and they conclude without further analysis that it is observationally equivalent to a model with risk aversion.

[8] Dorsey and Razzolini (2003) face a similar identification problem. Indeed, their experiment suggests the presence of probability misperception in private-values auctions, but they cannot confirm formally this hypothesis as they do not elicit beliefs.

[9] As we will see, there are in particular two issues we must pay attention to. First, it has been suggested that probing subjects with intrusive elicitation techniques may alter the way they play during the course of a repeated game experiment (Croson, 2000; Camerer et al., 2001; or Rutström and Wilcox, 2005). Second, the elicitation technique should not provide subjects with an incentive to state biased predictions (e.g., due to insurance motives) because of the choices they make in the auction game.

probability of winning. The structural models also suggest that the presence of biased probabilistic beliefs is a main source of overbidding, and that risk aversion may play a lesser role than previously believed. In fact, the estimated average constant relative risk aversion parameter drops from 0.6 to less than 0.2 when one accounts for the subjective probabilities revealed by subjects. These results appear to be robust, as no statistically significant difference may be detected across the three belief-elicitation treatments we conducted.

The article is structured as follows. The experimental design is presented in Section 2, and discussed in Section 3. The experimental outcomes are presented in Section 4. Different noisy models of behavior are estimated and compared in Section 5. Finally, Section 6 concludes.

2. THE EXPERIMENTAL DESIGN

In this section, we present the different experimental treatments. The principal choices made when designing the experiment are then discussed in a subsequent section. The experiment was conducted with volunteers at the State University of New York at Stony Brook. Although in the article we ultimately report the experimental outcomes for a single elicitation procedure, the experiment was in fact conducted under three belief-elicitation treatments (no payment, a scoring rule, and a prediction contest) and two informational treatments (with feedback and without feedback). For each of the six possible combinations of informational and elicitation treatments, we conducted four experimental sessions, each with 10 subjects and 15 rounds. No subject participated in more than one session. At the beginning of a session, players were assigned to an isolated computer. Subjects were told in advance how many rounds would be played, and they knew that the experiment would not exceed 1 hour and 30 minutes. Instructions were then read aloud, followed by participants' questions, and a brief training with the computer software.[10]

As further discussed in Section 3, the experimental design is essentially motivated by the following two objectives: first, the auction has to be similar to the one in GHP; second, the elicitation mechanism should lead subjects to reveal their beliefs as precisely as possible, without affecting the way they behave during the course of the repeated auction experiment.

Before describing how the experiment unfolds, we summarize the discrete auction model in GHP. Two players participate in a sealed bid independent private-values auction. Each player receives a randomly determined prize-value, which is equally likely to be \$0, \$2, \$4, \$6, \$8, or \$11. The players must simultaneously make a sealed bid, which is constrained to be an integer dollar amount. The prize is awarded to the highest bidder (with ties decided by the flip of a coin), for a price equal to his bid. GHP prove that the unique RNBNE is to bid \$0, \$1, \$2, \$3, \$4, and \$5, for values of, respectively, \$0, \$2, \$4, \$6, \$8, and \$11. The equilibrium strategy yields subjects an average profit of \$1.9 per round.

In each round, subjects are randomly matched in pairs. In order to limit reputation building, the subjects are informed that the assignment is such that it is not possible to identify the other member of the pair.[11] The problem in each round may be decomposed in three phases. In phase 1, the GHP's auction is implemented in a strategy form. That is, subjects are asked to identify the bid they would make for each possible prize-value. They are told that a prize-value will be assigned to them at the end of the round, and that their conditional choice in phase 1 corresponding to that value would determine what we called their "effective bid." Subjects are also informed that the outcome of the auction is decided by comparing the effective bids of the two members of the pair. We emphasized in the instructions that subjects should make careful decisions in phase 1, as their effective bid directly influences their auction payoffs.

In phase 2, subjects are asked to predict their own probability of winning the auction for a given list of bids. A coin flip decides whether all subjects must make their predictions for a

[10] The complete list of instructions is available in the Appendix.

[11] The matching protocol, although standard in repeated game experiments (including auction experiments), does not guarantee that subjects will meet at most once. As a consequence, it may be argued that each round does not replicate perfectly a one-shot game environment.

list of even bids (i.e., \$0, \$2, \$4, \$6, \$8, and \$10), or uneven bids (i.e., \$1, \$3, \$5, \$7, \$9, and \$11).[12] Subjects are asked to express as a frequency (i.e., an integer between 0 and 100) their chances of winning the auction for each of the six bids in the list.[13] We carefully explained to the participants that their probability of winning with a given bid is equal to the probability that this bid is higher than the effective bid of a random person in the room. We also made them aware that the experiment was designed in such a way that their choices and payoffs in phase 1 are independent of their choices and payoffs in phase 2 (and vice-versa).

Once every prediction has been submitted, we determine in phase 3 the auction and (when relevant) the prediction payoffs. To do so, we randomly match subjects in pairs. Then, we successively stop by each subject computer station, and roll a six-sided die to assign a prize-value to each participant. As previously mentioned, we determine a subject's effective bid by matching his prize-value with his corresponding conditional choice in phase 1. The effective bids of the members of each pair are then compared, and the auction payoffs are calculated.

As previously mentioned, we experimented with three different mechanisms to elicit subjective probabilities. Under the first elicitation treatment, subjects are not rewarded for the precision of their predictions; under the second elicitation treatment, subjects are rewarded according to a quadratic scoring rule; finally, under the third elicitation treatment subjects compete in a prediction contest. We concentrate in this section on the description of the prediction contest for which we will report the experimental outcomes in the Section 4.[14] In order to determine the prediction payoffs in the contest treatment, a six-sided die is first rolled once in public. The outcome of the draw determines which of the six integer bids in the list given in phase 2 will be used to measure the accuracy of the subjects' predictions. The predictions of both members of a pair are compared with an "objective probability" of winning the auction for the bid randomly selected. In the instructions, we carefully explained that this objective probability is calculated precisely by looking at the bid decisions made in phase 1 of that round by all the participants in the room, other than the members of the pair to which the subject belongs. As a result, the objective probability differs slightly across pairs, but it is the same for both members of a pair. The prediction payoff is \$4 to the member of the pair with the closest prediction to the objective probability for the bid randomly selected. We told subjects that in the event of a tie, the allocation of the \$4 would be decided by the flip of a coin.[15]

At the end of the round, the auction outcomes (i.e., the bids and payoffs of both members) are revealed to each member of the pair. This is the only information revealed in the *no-feedback* treatment (hereafter referred as treatment 1). In the *feedback* treatment (hereafter referred as treatment 2), participants are also informed about the quality of their absolute predictions. More precisely, we presented both on a graph and in a table the subject's own predictions along with her objective probabilities of winning for each of the six integer bids in the list given in phase 2. In order to promote absolute precision, instead of relative precision, we did not reveal to the subjects the predictions payoffs, nor the accuracy of any other participant. In other words, subjects do not know until the end of the session whether their prediction in a given round was more accurate.

Following GHP, subjects were paid half of their accumulated earnings in cash at the end of the session. The auction earnings in treatment 1 are comparable to those in GHP (respectively, \$10.58 versus \$10.7 in GHP), but auction payoffs in treatment 2 are slightly higher (\$11.49). In both treatments, the auction profits are smaller than the RNBNE expected profit of \$14.25.

[12] GHP also conduct a "low-cost" treatment with prize values of \$0, \$2, \$4, \$6, \$8, and \$12. Although this "low-cost" treatment was likely to generate slightly more overbidding, it would have required to ask predictions for either six uneven bids, or seven even bids. In order to avoid such asymmetry, we preferred to adopt GHP's "high-cost" treatment.

[13] In a series of pilot experiments, we experimented with different mechanisms for subjects to express their beliefs (e.g., a scale, a pie, a comparison with other random events). No substantial difference was identified in the way subjects reported their beliefs (i.e., the distributions of the predictions in the various treatments were not statistically different).

[14] A detailed description of the other two elicitation techniques may be found in a separate document (Addendum 1) on one of the author's web site at http://www.sceco.umontreal.ca/liste_personnel/armantier/index.htm.

[15] Note that, to promote a similar level of introspection when selecting bids and making predictions, we have arranged so that the number of choices and the expected payoffs are roughly the same in phases 1 and 2.

3. COMMENTS ON THE EXPERIMENTAL DESIGN

In this section, we discuss the principal features of the experimental design (for additional details, see Armantier and Treich, 2006). The first objective of the experiment is to evaluate whether the beliefs subjects hold about their probability of winning the auction are biased. To do so, we compare a subject's prediction with her objective probability, which can only be derived if we observe her (potential) opponents full-bidding strategy. Two features of the design allow us to infer the subjects' entire strategies: first, the number of possible private-values is finite; second, the auction is implemented in its strategic form.[16] It has been argued that the strategy method may affect behavior, as subjects may regret their choices once they observe their actual private-values (see e.g., Roth, 1995). In order to prevent such a problem, we emphasized to subjects that they had to select their bids carefully in phase 1, as one of them would directly influence their auction payoffs. Moreover, we will see in the Section 4 that a comparison with the experimental outcomes obtained by GHP under the extensive implementation of the game does not reveal any significant treatment effect.[17]

There is no consensus across fields on an appropriate methodology to measure probabilistic beliefs: most psychologists (and some economists) simply ask subjects for their best estimates;[18] statisticians (and some experimental economists) typically rely on "proper" scoring rules in an attempt to promote truthful revelation;[19] whereas prediction contests are often conducted in real-life applications.[20] Two desired properties are difficult to reconcile when devising an elicitation mechanism: it should be transparent, and it should encourage truthfulness. Asking subjects for their best estimates is clearly the simplest approach, but economists are in general concerned it may not incite subjects to report their true probabilities. In contrast, proper scoring rules are maximized by truthful revelation, but they have been criticized for their lack of transparency and their possible impact on behavior in repeated games.[21] Prediction contests, although essentially used in real-life contexts, may be an interesting compromise between the two previous approaches. Indeed, they provide salient financial incentives through a natural mechanism.[22] In the absence of a firm methodological foundation, we decided to adopt an agnostic approach, and to conduct our experiment separately under the three elicitation methods. As we shall see, these different elicitation methods produced essentially similar experimental outcomes.

Finally, a key feature of the experimental design is that the subjects' predictions should only depend on their beliefs, not on their bidding decisions. This feature derives primarily from the

[16] The strategy method also allows us to collect a much larger sample of data in each round. This additional information will prove very helpful when we estimate various behavioral models.

[17] Selten and Buchta (1999), as well as Pezanis-Christou and Sadrieh (2003) adopted an essentially comparable strategy method in an auction experiment. In accordance with our results, they did not identify the presence of a significant treatment effect.

[18] See e.g., Lichtenstein et al. (1978), Viscusi et al. (1997), as well as Manski (2004) and the references therein.

[19] See e.g., De Finetti (1965), Murphy and Winkler (1970), Savage (1971).

[20] Examples of practical implementations of this method include the National Collegiate Weather Forecasting Contest, the Earthquake Prediction Contest, the Federal Forecasters Forecasting Contest, or the Wall Street Journal semi-annual forecasting survey.

[21] Note that proper scoring rules are not necessarily incentive compatible. This is not the case for instance when expected utilities are nonlinear in payoffs and/or in probabilities, as it will turn out to be the case in our experiment. For evidence that quadratic scoring rules may affect the way subjects play in a repeated game (see Croson, 2000; Camerer et al., 2001; Rutström and Wilcox, 2005). Finally, observe that experimental economists do not rely exclusively on proper scoring rules. Indeed, while still rewarding precision, some authors prefer to give a greater weight to transparency (see, e.g., Dufwenberg and Gneezy, 2000; Croson, 2000; Wilcox and Feltovich, 2000; Charness and Dufwenberg, 2006).

[22] Note that the exact theoretic properties of such a contest depend on how one models the prediction game. For instance, if it is assumed that a subject believes that the true probability is a deterministic number P, then independently of his opponent's beliefs, this subject will maximize his chances of winning the contest by announcing P. Truthful revelation is therefore a dominant strategy in that case. Alternatively, one may model the prediction contest as a game of an incomplete information. Probabilistic beliefs are then interpreted as random signals informative about the unknown objective probability. In this case, the prediction contest is incentive compatible under some specific conditions on the subjects' prior beliefs (see Ottaviani and Sørensen, 2003 for details).

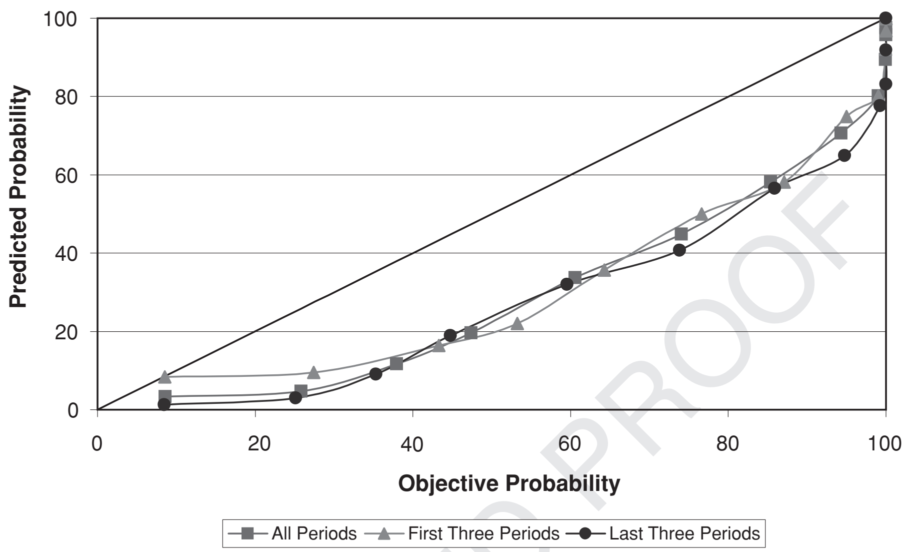

FIGURE 1

PREDICTIONS IN TREATMENT 1

choice of objective probability to which a subject's prediction is compared. Indeed, recall that the characterization of the objective probability for a given pair is based on the actions of all participants except the members of that pair. As a result, the prediction payoffs obtained by both members of a pair cannot be affected by the bids they each submit in phase 1. Likewise, the auction payoffs obtained by both members of a pair are independent of the predictions they make in phase 2. As a result, subjects should make their choices separately in phases 1 and 2, as they have no incentive to tie their bidding and prediction decisions.[23]

4. EXPERIMENTAL OUTCOMES

In order to shorten the article, we concentrate in this section on the experimental outcomes obtained under the prediction contest mechanism. Indeed, although the nature of the results presented below remains unaffected by the belief-elicitation method adopted, we find that the prediction contest produced slightly more homogenous experimental outcomes.[24]

4.1. *Treatment 1: No-Feedback*

4.1.1. *Probability predictions.* In Figure 1, we display on the x-axis the subjects average objective probabilities of winning the auction, and on the y-axis the subjects average predictions in treatment 1. We plotted three curves differentiated by the number of periods considered when

[23] In contrast, alternative characterizations of the objective probability might lead subjects to tie their bidding and prediction decisions. For instance, if one compares the predictions to the average probability of winning based on the bids submitted by every subject in that round, then a subject could potentially select bids (e.g., bid 0 for all prize-values) in phase 1 in order to skew knowingly the average probability of winning, and improve his chances of winning the prediction reward in phase 2. Likewise, suppose that one calculates each subject's objective probability of winning based on the bid function submitted by the other member of the pair to which he belongs. Then, the auction and prediction payoffs of the subject both depend on the bids of his opponent. As a result, a risk-averse subject could select jointly his bids and predictions in order to insure himself against his opponent bidding behavior.

[24] We reach this conclusion on the basis of an econometric analysis indicating that predictions and bidding behavior are slightly more homogenous across periods and across subjects under the prediction contest mechanism. In addition, we will see that the prediction contest procedure is further validated by the fact that the predictions made by subjects are consistent with their actions. The lack of major differences between the experimental outcomes obtained under the two other elicitation techniques may be appreciated in a separate document (Addendum 1) available on one of the author's web-site at http://www.sceco.umontreal.ca/liste_personnel/armantier/index.htm.

averaging. Each curve actually consists of 12 consecutive dots representing the predictions for integer bids ranging from \$0 to \$11. For instance, the first dot on the left corresponds to the predicted probability of winning the auction with a bid $b = 0$. The diagonal has also been plotted to guide the eye. If subjects make unbiased predictions, then their stated probabilities should fall around the diagonal. Figure 1 indicates that subjects underestimate their probability of winning for any bid between \$0 and \$11. The underestimation is in fact quite substantial for a large number of bids. For instance, throughout the 15 periods, subjects believe that their probability of winning with a bid $b = 3$ and $b = 5$ (the RNBNE bids for a private-value of $v = 6$ and $v = 11$) are on average 19.66% and 44.86%, whereas the actual probabilities are 47.38% and 74.08%.

It is interesting to note that, although we find evidence of overestimation for the lowest bid (i.e., $b = 0$) during the first three periods, subjects' predictions did not produce the traditional inverse S-shape pattern (i.e., overestimation of small probabilities and underestimation of high probabilities) commonly observed in the psychology literature (see e.g., Camerer, 1995). This observation is consistent with GHP whose results also suggest that the subjective probability function may be convex everywhere. In other words, it appears that subjects in our experiment are pessimists, and do not exhibit the traditional biases. Similar evidence has been detected in other experimental studies (e.g., Schotter and Sopher, 2001 identify pessimistic beliefs in an ultimatum game experiment), as well as field studies (Giordani and Söderlind, 2006 show that professional forecasters have been historically pessimistic in their predictions of GDP).[25] Our experimental results may therefore raise the question of the relevance of the inverse S-shape pattern to apply to private-values auctions, and maybe more generally to game situations.

No evidence of learning may be detected in Figure 1. Indeed, the predictions in the first and last three periods appear indistinguishable. In order to confirm this observation statistically, we estimate a model of the form:

$$\hat{P}_{i,b,t} = \delta_b^0\left(1 + \delta_b^5 \mathbb{I}_{[t>5]} + \delta_b^{10} \mathbb{I}_{[t>10]} + \eta_i\right) + u_{i,b,t}, \tag{1}$$

where $\hat{P}_{i,b,t}$ is the prediction of player $i = 1, \ldots, 10$ in period $t = 1, \ldots, 15$ regarding his probability of winning the auction with a bid equal to $b = 0, \ldots, 11$, $\mathbb{I}_{[.]}$ is the indicator function satisfying $\mathbb{I}_{[x]} = 1$ when x is true and $\mathbb{I}_{[x]} = 0$ otherwise; $(\delta_b^0, \delta_b^5, \delta_b^{10})$ are parameters to be estimated; η_i is an individual random effect with mean zero and variance σ_η^2; finally, $u_{i,b,t}$ is a mean zero error term with a standard deviation accounting for possible heteroskedasticity across predictions for different bid values and/or different time periods: $Std(u_{i,b,t}) = \sigma|\gamma_1 - b|^{\gamma_2} t^{\gamma_3}$.

Before moving to the estimation results, let us briefly point out some of the features of the econometric model. First, note that under the specification in (4.1), δ_b^0 may be interpreted as the initial average prediction for a bid b, whereas δ_b^5 and δ_b^{10} capture how this prediction changed (in percentage) after, respectively, 5 and 10 periods.[26] Second, the specification of the error term's standard deviation is quite flexible and it allows in particular (i) for subjects to learn to make less-noisy predictions with time (i.e., $\gamma_3 < 0$); and (ii) for predictions to be less variable for low and high bids (for which the probabilities of winning the auction are close to 0 and 100) than for intermediate bids (i.e., $\gamma_1 \in [0, 11]$ and $\gamma_2 < 0$ in which case the standard deviation has an inverse U-shape). Finally, the econometric model captures potential systemic effects across individuals by exploiting the panel structure of the data to account for possible correlations between the predictions made by a subject over time.[27]

The results of the feasible least square estimation presented in Table 1 indicate no systematic prediction adjustment, as most estimated parameters $\left(\hat{\delta}_b^5, \hat{\delta}_b^{10}\right)$ are not significantly different

[25] Note that evidence of overconfidence or optimism has also been detected in experimental economics (see e.g., Palfrey and Rosenthal, 1991; or Camerer and Lovallo, 1999).

[26] We opted for the specification in (1) because the parameters are easily interpretable and comparable across regressions. In particular, it enables to compare the speed of learning between early and late periods. In addition, we will be able to compare the speed with which subjects adjust their bids, and the speed with which they adjust their predictions. Different specifications of the model (e.g., including a time trend, the subject's prediction in the previous

TABLE 1
EVOLUTION OF PREDICTIONS

	Treatment 1			Treatment 2			Treatment Effect♣		
	$\hat{\delta}_b^0$	$\hat{\delta}_b^5$	$\hat{\delta}_b^{10}$	$\hat{\delta}_b^0$	$\hat{\delta}_b^5$	$\hat{\delta}_b^{10}$	$\Delta\hat{\delta}_b^0$	$\Delta\hat{\delta}_b^5$	$\Delta\hat{\delta}_b^{10}$
$b = 0$	4.753*	−0.572*	−0.221*	7.908*	0.081*	−0.005	2.540	0.483*	0.204*
	(0.948)	(0.056)	(0.072)	(0.792)	(0.037)	(0.026)	(1.695)	(0.022)	(0.044)
$b = 1$	9.661*	−0.466*	−0.150*	15.179*	0.293*	0.098*	3.356	0.593*	0.235*
	(1.569)	(0.039)	(0.063)	(2.685)	(0.028)	(0.019)	(2.114)	(0.024)	(0.035)
$b = 2$	15.312*	−0.198*	−0.068	23.432*	0.465*	0.116*	5.759*	0.604*	0.170*
	(2.341)	(0.024)	(0.037)	(2.913)	(0.020)	(0.017)	(2.098)	(0.016)	(0.020)
$b = 3$	25.219*	−0.120*	0.008	30.265*	0.409*	0.124*	2.315	0.428*	0.131*
	(3.082)	(0.026)	(0.029)	(3.460)	(0.021)	(0.016)	(1.637)	(0.021)	(0.021)
$b = 4$	37.478*	−0.053*	−0.004	42.152*	0.336*	0.133*	3.252	0.310*	0.123*
	(4.110)	(0.022)	(0.025)	(4.689)	(0.019)	(0.015)	(1.857)	(0.019)	(0.019)
$b = 5$	48.276*	−0.040	−0.004	51.386*	0.347*	0.122*	1.564	0.319*	0.110*
	(5.275)	(0.023)	(0.020)	(5.951)	(0.019)	(0.013)	(1.488)	(0.017)	(0.014)
$b = 6$	57.132*	−0.031	0.004	62.300*	0.296*	0.069*	1.182	0.281*	0.062*
	(6.413)	(0.019)	(0.018)	(7.763)	(0.019)	(0.016)	(1.693)	(0.020)	(0.013)
$b = 7$	70.111*	−0.011	0.002	75.709*	0.208*	0.024*	2.225	0.192*	0.019*
	(6.929)	(0.015)	(0.015)	(6.768)	(0.016)	(0.010)	(1.580)	(0.022)	(0.008)
$b = 8$	82.168*	−0.002	−0.001	87.989*	0.093*	0.013	3.851	0.084*	0.011
	(7.516)	(0.009)	(0.012)	(7.890)	(0.011)	(0.011)	(2.233)	(0.016)	(0.007)
$b = 9$	92.172*	−0.002	0.001	95.743*	0.057*	0.003	1.946	0.047*	−0.002
	(8.251)	(0.006)	(0.010)	(8.513)	(0.008)	(0.008)	(1.904)	(0.018)	(0.005)
$b = 10$	96.317*	−0.002	−0.002	97.552*	0.020*	0.001	0.200	0.027*	0.001
	(5.950)	(0.005)	(0.009)	(6.292)	(0.006)	(0.005)	(1.262)	(0.012)	(0.004)
$b = 11$	98.964*	−0.002	0.000	99.216*	0.009*	0.001	−0.116	0.010	0.000
	(3.357)	(0.002)	(0.005)	(2.874)	(0.003)	(0.002)	(1.006)	(0.009)	(0.003)

Variance Parameters

Treatment 1					Treatment 2				
$\hat{\sigma}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\sigma}_\eta$	$\hat{\sigma}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\sigma}_\eta$
7.892*	6.331*	−0.629*	0.018	0.130*	8.109*	5.478*	−0.541*	−0.091*	0.088*
(0.557)	(0.726)	(0.170)	(0.041)	(0.043)	(0.743)	(0.877)	(0.183)	(0.039)	(0.035)

*Parameter larger than zero at a 5% significance level. Numbers in parenthesis refer to the standard deviations of the estimates

♣In order to test for the presence of a treatment effect, we estimate $\Delta\hat{\delta}_b^\tau = \hat{\delta}_b^\tau(T_2) - \hat{\delta}_b^\tau(T_1)\ \forall\tau \in \{0, 5, 10\}$, where $\hat{\delta}_b^\tau(T_2)$(respectively, $\hat{\delta}_b^\tau(T_1)$) is the parameter estimated with the data collected in Treatment 2 (respectively, Treatment 1).

from zero.[28] Subjects however, appear to learn to lower their predictions for small bids after the first five periods, since $\hat{\delta}_5$ is significantly less than zero for $b = 0, \ldots, 4$. In fact, the reduction in the stated probabilities can be quite substantial. For instance, when asked to predict their probability of winning for a bid $b = 0$ (respectively, $b = 1$), subjects lower their predictions by 57.2% (respectively, 46.6%) after the first five periods, and an additional 22.1% (respectively,

round, or a dummy variable identifying the session in which the subjects participated) do not change the nature of the results presented.

[27] Observe that because of time and individual effects, the two samples consisting of predictions collected during the first and last periods of a sessions are not independent. Therefore, we cannot compare these two samples with traditional nonparametric tests such as the Mann–Whitney test. In contrast, such a comparison may be conducted with our parametric approach, as the time and individual effects are explicitly modeled.

[28] Throughout the article the terms "significant" and "significantly" are used in their statistical sense, to indicate whether a hypothesis may be accepted or rejected at a 5% significance level. Note also that distributions of the test statistics, and the standard deviations of the estimates have been evaluated by bootstrap, in order to control for the finiteness of the sample (see Shao and Tu, 1995).

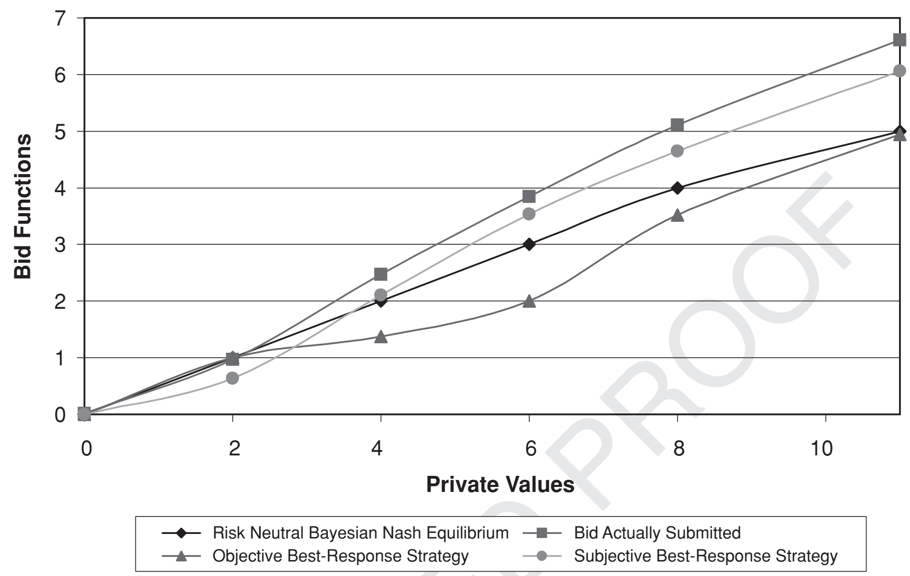

FIGURE 2

BID FUNCTIONS IN TREATMENT 1

15.0%) after period 10.[29] In other words, the only evidence of adjustment we can identify is that subjects rapidly reduce their predictions for small bids.

We find that the variance of the predictions exhibits a nearly symmetrical inverse U-shape. Indeed, we can see at the bottom of Table 1 that $\hat{\gamma}_1$ is almost centrally located in [0, 11], while $\hat{\gamma}_2$ is significantly lower than zero. In other words, predictions are less variable for bids with which bidders are almost certain to loose or to win the auction (i.e., low and high bids). The variance of the predictions, however, does not seem to contract with time, as $\hat{\gamma}_3$ is found to be insignificant. Finally, we identify substantial heterogeneity within the subjects' pool as $\hat{\sigma}_\eta$, the standard deviation of the individual random effect, is found to be significant and relatively large. For instance, a subject with an individual effect equal to one standard deviation ($\eta_i = \hat{\sigma}_\eta$) will submit predictions 13.0% larger than the average bidder (for whom $\eta_i = 0$).

4.1.2. *Bidding behavior.* Let us now turn to the bids submitted in treatment 1. In Figure 2, we compare the average bids actually submitted for each possible private-value (identified by squares), with the RNBNE (identified by diamonds). The figure confirms that except for low private-values (i.e., $v = 0$ and $v = 2$) subjects tend to bid significantly above the RNBNE. This result however, does not necessarily imply that subjects are either risk averse or do not have rational expectations. Indeed, even though subjects do not submit the RNBNE, it is entirely possible that they best-respond to their opponents actions, or that their choices are consistent with their beliefs. In order to explore these hypotheses, two additional bid functions are plotted in Figure 2. The set of triangles represents a subject's risk-neutral objective best-response, that is what a risk-neutral agent should bid if he knew or could infer correctly the other participants actions. The set of disks represents a subject's risk-neutral subjective best-response, that is what a risk-neutral utility maximizer should bid conditional on her stated beliefs regarding her

Q1

[29] We conjecture that this decrease may be explained in part by the fact that most subjects did not witness any auction won by a bidder submitting a bid of 0 or 1 during the experiment. Indeed, out of the 300 auctions conducted during the four sessions of treatment 1, only two were won by a bidder submitting a bid of 0, and 9 by a bidder submitting a bid of 1.

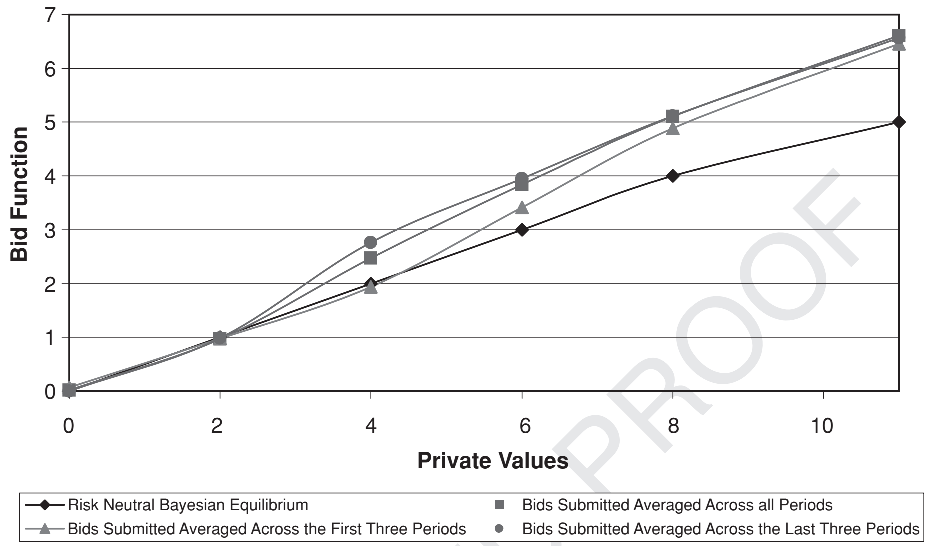

FIGURE 3

EVOLUTION OF BIDS IN TREATMENT 1

probability of winning the auction.[30] Figure 2 indicates that subjects overbid compared to their objective best-responses. Their actions however, appear to be quite consistent with their beliefs. Indeed, the bids submitted for the different private-values are only slightly higher (on average) than the subjective best-responses.[31]

To summarize, subjects cannot be assimilated to risk-neutral agents with rational expectations because they do not comply with either the RNBNE, or their objective risk-neutral best-responses. Subjects, however, may not be far from risk-neutral utility maximizers, because their actions are close to their subjective risk-neutral best-responses. Nevertheless, the slight overbidding remaining may still be rationalized by risk aversion, although it would suggest that risk aversion may play a lesser role than previously believed. This conjecture will need to be confirmed statistically in the Section 5 when we estimate noisy behavioral models.

Figure 3 indicates that, if anything, subjects learn to slightly increase their bids over time. This observation is confirmed by the estimation of an econometric model similar to (1):

$$B_{i,v,t} = \delta_v^0\left(1 + \delta_v^5 \mathbb{I}_{[t>5]} + \delta_v^{10} \mathbb{I}_{[t>10]} + \eta_i\right) + u_{i,v,t}, \tag{2}$$

where $B_{i,v,t}$ is the bid submitted by player $i = 1, \ldots, 10$ in period $t = 1, \ldots, 15$ for a private-value $v \in \{0, 2, 4, 6, 8, 11\}$. The remaining parameters and variables are defined as in Equation (1), except for the standard deviation of $u_{i,v,t}$ which is modeled as $Std(u_{i,v,t}) = \sigma|\gamma_1 - v|^{\gamma_2} t^{\gamma_3}$.

The estimation outcomes in Table 2 indicate that, except for private-values equal to $v = 4$ and $v = 6$, subjects in treatment 1 do not learn to decrease their bids significantly. In other words, we cannot find conclusive evidence of systematic strategy adjustment. The estimates of the standard deviation parameters indicate that (i) the variance of the bids increases with the private-value v (i.e., $\hat{\gamma}_1$ is insignificant whereas $\hat{\gamma}_2$ is positive and significant), (ii) bidding behavior does not

[30] Unlike previous private-values experiments, we can compare subjects' choices with their objective and subjective best-responses. Indeed, two features specific to our design enable the calculation of these best-responses: first, the auction is implemented in strategic form; second, we elicit subjects' beliefs about their probability of winning.

[31] This observation is consistent with Nyarko and Schotter (2002), as well as Bellemare et al. (2005) who find that subjects' actions appear to be more compatible with their subjective instead of their objective best-responses. This result also provides some support to one of the two hypotheses underlying our approach. Namely, agents appear to have probabilistic beliefs on which they base their actions.

TABLE 2

EVOLUTION OF BIDDING BEHAVIOR

	Treatment 1			Treatment 2			Treatment Effect♣		
	$\hat{\delta}_v^0$	$\hat{\delta}_v^5$	$\hat{\delta}_v^{10}$	$\hat{\delta}_v^0$	$\hat{\delta}_v^5$	$\hat{\delta}_v^{10}$	$\Delta\hat{\delta}_v^0$	$\Delta\hat{\delta}_v^5$	$\Delta\hat{\delta}_v^{10}$
$v = 0$	0.024	−0.003	0.000	0.026	−0.000	0.000	0.001	0.001	0.000
	(0.017)	(0.005)	(0.003)	(0.014)	(0.004)	(0.004)	(0.046)	(0.003)	(0.001)
$v = 2$	1.024*	−0.016	0.000	1.036*	−0.022*	−0.000	0.007	−0.004	−0.000
	(0.131)	(0.012)	(0.007)	(0.106)	(0.008)	(0.005)	(0.054)	(0.005)	(0.002)
$v = 4$	2.285*	0.154*	0.038	2.393*	−0.104*	−0.100*	0.128	−0.231*	−0.124*
	(0.331)	(0.029)	(0.025)	(0.375)	(0.022)	(0.022)	(0.089)	(0.019)	(0.017)
$v = 6$	3.655*	0.121*	0.012	3.766*	−0.116*	−0.109*	0.077	−0.224*	−0.102*
	(0.558)	(0.021)	(0.027)	(0.395)	(0.027)	(0.021)	(0.085)	(0.022)	(0.016)
$v = 8$	4.962*	0.027	−0.007	5.183*	−0.137*	−0.096*	0.065	−0.140*	−0.080*
	(0.763)	(0.020)	(0.016)	(0.684)	(0.021)	(0.020)	(0.101)	(0.022)	(0.017)
$v = 11$	6.592*	0.019	−0.008	6.710*	−0.091*	−0.142*	0.108	−0.091*	−0.117*
	(0.688)	(0.015)	(0.013)	(0.653)	(0.016)	(0.020)	(0.094)	(0.021)	(0.018)

Variance Parameters

Treatment 1					Treatment 2				
$\hat{\sigma}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\sigma}_\eta$	$\hat{\sigma}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\sigma}_\eta$
0.246*	0.752	0.883*	−0.008	0.153*	0.196*	0.444	0.956*	−0.051*	0.120*
(0.044)	(1.676)	(0.214)	(0.025)	(0.051)	(0.040)	(1.171)	(0.237)	(0.020)	(0.045)

*In order to test for the presence of a treatment effect, we estimate $\Delta\hat{\delta}_v^\tau = \hat{\delta}_v^\tau(T_2) - \hat{\delta}_v^\tau(T_1)$ $\forall\tau \in \{0, 5, 10\}$, where $\hat{\delta}_v^\tau(T_2)$(respectively, $\hat{\delta}_v^\tau(T_1)$) is the parameter estimated with the data collected in Treatment 2 (respectively, Treatment 1).

♣Parameter larger than zero at a 5% significance level. Numbers in parenthesis refer to the standard deviations of the estimates.

TABLE 3

FREQUENCY WITH WHICH THE CORRECT RNBNE BID OR STRATEGY IS PLAYED

	Frequency of RNBNE Bid Played for Value Equals to						Frequency of RNBNE Strategy
	$v = 0$	$v = 2$	$v = 4$	$v = 6$	$v = 8$	$v = 11$	
Treatment 1							
All periods	98.67%	94.67%	38.33%	15.33%	12.17%	5.67%	1.00%
First three periods	93.33%	87.50%	46.67%	26.67%	18.33%	6.67%	0.83%
Last three periods	100.00%	98.33%	24.17%	13.33%	10.00%	5.83%	4.17%
Treatment 2							
All periods	99.17%	95.67%	67.83%	51.33%	39.17%	24.17%	12.50%
First three periods	95.83%	87.50%	34.17%	14.17%	17.50%	6.67%	0.00%
Last three periods	100.00%	99.17%	95.83%	88.33%	70.00%	50.00%	41.67%

become less variable with time (i.e., $\hat{\gamma}_3$ is insignificant), and (iii) there is significant heterogeneity across subjects in their bid selection (i.e., $\hat{\sigma}_\eta$ is significantly greater than zero).

The frequency with which subjects played the correct RNBNE bid for a given private-value is reported in Table 3. This table shows that although subjects select the RNBNE bid 98.67% and 94.67% of the time for the lowest private-values of $v = 0$ and $v = 2$, they rarely play the RNBNE bid for the highest private-values of $v = 8$ and $v = 11$ (respectively, 12.17% and 5.67%). These percentages vary only slightly between the first and last three periods, which further confirms that subjects do not learn to play the RNBNE bids. The last column of Table 3 reports the frequency with which subjects submitted the entire RNBNE bid function (i.e., $B(0) = 0$, $B(2) = 1$, $B(4) = 2$, $B(6) = 3$, $B(8) = 4$ and $B(11) = 5$). The correct RNBNE strategy has been chosen only 1.00% of the time overall, but there seems to be a slight increase between the first and last three periods from 0.83% to 4.17%.

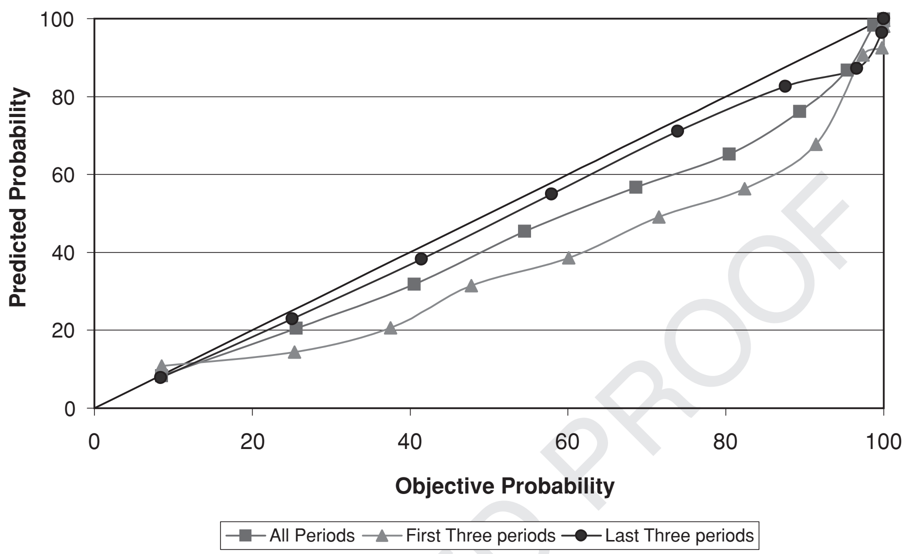

FIGURE 4

PREDICTIONS IN TREATMENT 2

Finally, note that the auction outcomes in treatment 1 are both qualitatively and quantitatively consistent with those in GHP. This conjecture is confirmed by a formal statistical analysis that provides no evidence that bidding behavior is significantly different in treatment 1 and in GHP's experiment.[32] In other words, implementing the auction in strategic form, and eliciting beliefs with a prediction contest do not appear to have introduced any significant treatment effect. This result is consistent with Rutström and Wilcox (2005) who suggest that, in contrast to a scoring rule, eliciting beliefs with a nonintrusive procedure does not significantly affect the way subjects play during the course of a repeated game experiment (see also Nyarko and Schotter, 2002).

4.2. *Treatment 2: Feedback*

4.2.1. *Probability predictions.* Figure 4 shows that providing feedback has a dramatic effect on the accuracy of subjects' predictions. Indeed, although predictions in the first three periods of treatments 1 and 2 are roughly similar, subjects in treatment 2 make nearly unbiased predictions during the last three periods of treatment 2. In order to confirm the presence of learning, we re-estimate the model in (1) with the data collected in treatment 2. The results in Table 1 indicate that subjects' predictions for all possible bids increase rapidly after the first five periods, and keep increasing, although at a slower pace, after period 10. For instance, when asked to predict their probability of winning for a bid $b = 2$, subjects increase their predictions by 46.5% after the first five periods, and an additional 11.6% after period 10.

In other words, from the feedback provided, subjects in treatment 2 appear to learn to make better predictions over time.[33] This result seems to contrast sharply with the outcomes in treatment 1. In order to confirm statistically the presence of a treatment effect, we estimate $\Delta\hat{\delta}_b^\tau = \hat{\delta}_b^\tau(T_2) - \hat{\delta}_b^\tau(T_1)$ for any $\tau \in \{0, 5, 10\}$, where $\hat{\delta}_b^\tau(T_2)$ (respectively, $\hat{\delta}_b^\tau(T_1)$) is the parameter estimated with the data collected in treatment 2 (respectively, in treatment 1). The results reported in the last three columns of Table 1 indicate that the subjects' predictions in the early

[32] The outcome of this statistical analysis may be found in a separate document (Addendum 2) available on one of the author's web-site at http://www.sceco.umontreal.ca/liste_personnel/armantier/index.htm.

[33] Further econometric analyses indicate that subjects do not simply copy the feedback provided by the experimenter, but instead use this information to learn to make better predictions (see Armantier and Treich, 2006).

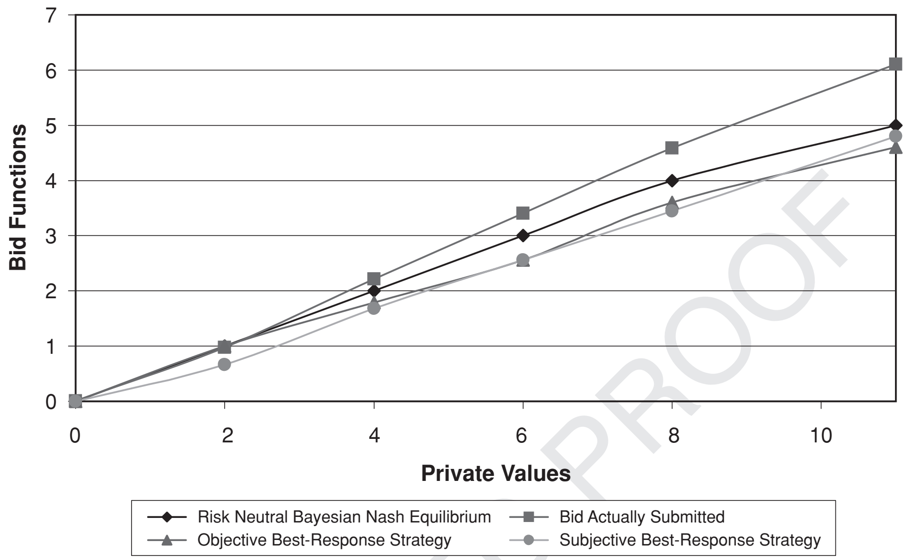

FIGURE 5

BID FUNCTIONS IN TREATMENT 2

periods of treatments 1 and 2 are nearly indistinguishable. Indeed, except for $b = 2$, the parameters $\Delta\hat{\delta}_b^0$ are not significantly different from zero. With time however, subjects in treatment 2 learn to make substantially higher predictions than subjects in treatment 1, because most of the parameters $\Delta\hat{\delta}_b^5$ and $\Delta\hat{\delta}_b^{10}$ are significantly greater than zero.[34]

4.2.2. *Bidding behavior.* Let us now turn to the bidding behavior of subjects in treatment 2. Figure 5 indicates that on average subjects still bid above the RNBNE, although overbidding is slightly less prominent than in treatment 1. Figure 5 however, only tells one part of the story. Indeed, Figure 6 shows a dramatic reduction of overbidding over time. Although the bids submitted during the first three periods are comparable in treatments 1 and 2, subjects in treatment 2 learned to drastically reduce their bids. In fact, during the last three periods of treatment 2 only the bids submitted for high private-values exceed slightly (on average) the RNBNE. For instance, for the highest private-value $v = 11$, subjects submitted instead of the RNBNE $b = 5$, an average bid of 5.29 during the last three periods of treatment 2, compared to 6.57 in treatment 1. This treatment effect is confirmed statistically in Table 2. Indeed, while all $\Delta\hat{\delta}_v^0$ are insignificant, most of the estimated differences $\Delta\hat{\delta}_v^5$ and $\Delta\hat{\delta}_v^{10}$ are significantly lower than zero. In other words, early bidding is statistically indistinguishable, but subjects in treatment 2 quickly learned to make significantly lower bids than subjects in treatment 1.

Some of the economic implications of these strategy adjustments are quite substantial. In particular, the expected profit of a subject during the last three periods jumps from \$1.34 in treatment 1, to \$1.79 in treatment 2, a 33.5% increase. In addition, subjects leave less money on the table by the end of treatment 2. Indeed, during the last three periods, the average foregone profit (i.e., the difference between a subject's actual expected profit, and the expected profit generated by his risk-neutral objective best-response) drops from \$0.21 (or 15.41%) in treatment 1, to \$0.05 (or 3.08%) in treatment 2.

A comparison of Tables 1 and 2 also indicates that subjects in treatment 2 adjust their strategy at a slower, but more constant pace than their predictions. Indeed, we can see in Table 1 that

[34] These results are also confirmed by a series of Mann–Whitney tests comparing the samples of predictions collected during the first (respectively, last) period of treatments 1 and 2. Observe however, that this nonparametric test cannot be extended to compare more than one period because of individual and time effects.

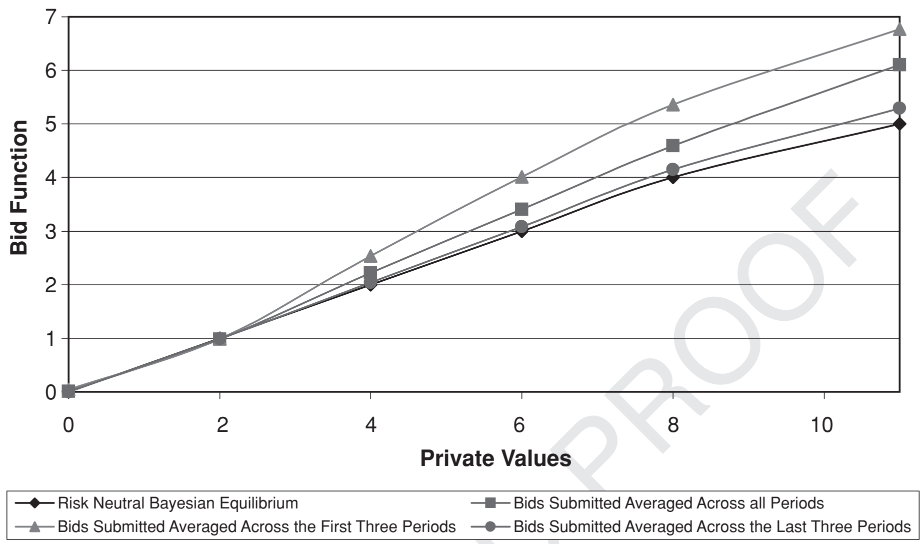

FIGURE 6

EVOLUTION OF BIDS IN TREATMENT 2

subjects in treatment 2 essentially adjust their predictions within the first five periods (the average percentage increase in the predictions is 21.7% after the first five periods and an additional 5.8% during the last 5 periods). In contrast, Table 2 indicates that the magnitude of the strategy adjustment is basically comparable after 5 and 10 periods (the average percentage decrease in the bids submitted is 7.8% after the first five periods and an additional 7.4% during the last 5 periods). In other words, subjects tend to first learn to make better predictions, and then adjust their bidding behavior accordingly.[35] The asymmetry between the prediction and strategy learning speeds may explain why, in contrast to treatment 1, subjects do not appear to best-respond to their beliefs in treatment 2. Indeed, Figure 5 indicates that on average subjects in treatment 2 bid above their risk-neutral subjective best-response. After the adjustment period (i.e., by the end of the session) subjects' actions, subjective and objective risk-neutral best-responses become consistent in treatment 2, because predictions are almost unbiased whereas behavior nearly conforms with the RNBNE.

Finally, one of the most striking differences between treatments 1 and 2 may be found in Table 3. Indeed, we can see that during the last three periods of treatment 2, subjects submitted the RNBNE bids for each possible private-value much more frequently. For instance, subjects submitted the RNBNE bid $b = 3$ for a private-value $v = 6$, 88.33% of the time in treatment 2, versus only 13.33% of the time in treatment 1. Even more remarkable, subjects precisely complied with the RNBNE bid function 41.67% of the time in the last three periods of treatment 2, versus 4.17% of the time in treatment 1.

To summarize, subjects overbid and underestimate their probability of winning in treatment 1. From the feedback provided in treatment 2, subjects first learn to make better predictions, and then nearly eliminate their tendency to overbid. This treatment effect is not as surprising as it may first appear. Indeed, it seems natural that the feedback and the financial stimulus lead subjects to make better predictions. Then, we can reasonably expect a subject to lower his bids, once he realizes that his probability of winning with any given bid is higher than he previously believed. For instance, when a subject realizes that he is virtually guaranteed to win the auction with a bid of $b = 5$ or $b = 6$, he should be very unlikely to submit a higher bid.

[35] This conjecture is also supported by a series of statistical tests indicating that the null hypothesis $\{\hat{\delta}_b^5 > \hat{\delta}_v^5\}$ cannot be rejected at the 5% significance level $\forall b \in \{1, \ldots, 7\}$ and $\forall v \in \{1, \ldots, 11\}$.

5. QUANTAL RESPONSE MODELS WITH HETEROGENOUS AGENTS

The experimental outcomes just presented suggest that risk aversion may play a lesser role than previously believed. In order to provide statistical support to this conjecture, we estimate a behavioral model. We adopt the Quantal Response Equilibrium concept (QRE hereafter) developed by McKelvey and Palfrey (1995) to model "noisy" decision making.[36] This equilibrium concept has been shown to be powerful to organize behavior in numerous experimental settings.[37] In particular, GHP find that the QRE model performs very well at tracking bidding behavior in our discrete auction model.

In short, the QRE approach essentially relies on two key principles: behavior is random, and the probability to choose an action increases with the expected utility this action may yield. In other words, agents are not expected to select their best-responses systematically, but they play these strategies with a higher frequency. The probabilistic choice rule is typically characterized by a "noise parameter" $\mu > 0$ reflecting the sensitivity of the choice probabilities to expected utilities. A large μ yields essentially random behavior, whereas μ close to zero implies Nash-like behavior. A condition on the consistency of actions and beliefs is then imposed to determine the QRE choice probabilities.

In order to explain overbidding, the leading risk aversion model, the so-called CRRAM (see e.g., Cox et al., 1988), assumes that agents may have unobserved heterogenous preferences. In addition, we have seen in Section 4 that subjects in our experiment exhibit significant heterogeneity with respect to probabilistic beliefs. Therefore, we develop a generalized QRE model accounting for unobserved heterogeneity with respect to both risk preferences and probabilistic beliefs.[38] We only summarize in this section the main features of the heterogenous QRE model for our auction game. A more detailed presentation may be found in the Appendix.

Agent i is endowed with a utility function of the form: $U(x) = (1 - r_i)^{-1}(x)^{1-r_i}$, where r_i is the agent's coefficient of constant relative risk aversion. In addition, to model the probabilistic beliefs of agent i, we generalize the PWF proposed by Prelec (1998): $\hat{P}_i = \Phi_i(P) = \exp(-\beta_i(-\ln(P))^{\alpha_i})$, where $\hat{P}_i$ and P are, respectively, the subjective and objective probability, and $(\alpha_i, \beta_i) \in \mathbb{R}^2_{++}$. Note that $(\alpha_i, \beta_i) = (1, 1)$ yields the identity function, in which case subjective and objective beliefs are equivalent.

Following Cox et al. (1988), agent i is assumed to be endowed with a vector of parameters (r_i, α_i, β_i), but she only knows the distribution from which her opponents' parameters are generated. In order to complete the specification of the model, we assume that (r_i, α_i, β_i) is identically and independently distributed across subjects from a trivariate normal distribution in which the last two components are truncated on the left at zero. Finally, (r_0, α_0, β_0) will denote the mean of (r_i, α_i, β_i), and Σ its variance–covariance matrix.[39] We propose in the Appendix a novel approach to estimate $(\mu, r_0, \alpha_0, \beta_0, \Sigma)$, the vector of structural parameters characterizing the QRE model under unobserved heterogeneity. We now present different estimated models to illustrate how taking into consideration elicited probabilities, in addition to observed actions, can improve inferences about preferences.

5.1. *Restricted QRE Estimated with Observed Bids.*

We start by estimating two restricted versions of the heterogenous QRE model: a QRE model with risk aversion and objective beliefs (i.e., we impose $(\alpha_i, \beta_i) = (1, 1)\ \forall i$), and a QRE model with subjective beliefs and risk neutrality (i.e., we impose $r_i = 0\ \forall i$). In order to estimate these models, we follow GHP in using only the

[36] The GHP discrete auction model does not have the Bayesian Nash equilibrium in a pure strategy for every possible coefficient of relative risk aversion. Therefore, it will not be possible to estimate the Bayesian Nash equilibrium model under risk aversion, as a possible alternative to the QRE.

[37] See, e.g., McKelvey and Palfrey (1995, 1998), Capra et al. (2002), Anderson et al. (2002), Yi (2003), as well as GHP.

[38] To the best of our knowledge, this is the first generalization of the QRE approach to unobserved heterogeneity.

[39] The results presented in this section appear to be robust to the parametric specification adopted for the joint distribution of (r_i, α_i, β_i). In particular, the truncations at zero do not seem to be binding. In fact, the results remain essentially unchanged when (r_i, α_i, β_i) is assumed to follow either a multivariate normal or a gamma distribution.

TABLE 4
CONSTRAINED QUANTAL RESPONSE EQUILIBRIUM MODELS
(ESTIMATED WITH OBSERVED BIDS ONLY)

	Model with Risk Aversion and Objective Beliefs			Model with Risk Neutrality and Subjective Beliefs					
	μ	r_0	σ_r	μ	α_0	β_0	σ_α	σ_β	$\rho_{\alpha,\beta}$
Treatment 1	0.063*	0.673*	0.119*	0.136*	1.012*	2.281*	0.057*	0.174*	0.108
	(0.004)	(0.006)	(0.014)	(0.007)	(0.014)	(0.028)	(0.022)	(0.040)	(0.124)
Treatment 2	0.102*	0.521*	0.100*	0.172*	0.790*	2.165*	0.093*	0.220*	0.026
	(0.006)	(0.007)	(0.019)	(0.015)	(0.015)	(0.029)	(0.035)	(0.046)	(0.148)

*Parameter larger than zero at a 5% significance level.
Numbers in parenthesis refer to the standard deviations of the estimates.

bids collected during the experiment, thereby ignoring for the moment the predictions made by the subjects.[40] These two restricted models will serve as benchmarks to test whether our understanding of behavior improves when the analysis accounts for both actions and subjective probabilities. In addition, this will give us an opportunity to compare our results to those of GHP.[41]

We first concentrate on the estimation of the restricted QRE models with the data collected in treatment 1. The estimates of the model with risk aversion and objective beliefs presented in Table 4 are consistent with GHP's results. Indeed, we estimate the noise and average risk aversion parameters to be $(\hat{\mu}, \hat{r}_0) = (0.063, 0.673)$, versus (0.08,0.55) in GHP.[42] These results are also compatible with recent estimates of the relative risk aversion parameter in private-value auctions.[43] Observe also that, in accordance with GHP, we find that the average risk aversion parameter is significantly greater than zero, thereby rejecting the risk-neutrality hypothesis. Finally, the estimation of the model with risk-neutrality and subjective beliefs yields an average PWF of the form $\Phi(P) = P^{2.281}$, closely matching the nearly quadratic PWF obtained by GHP. In fact, no significant difference may be detected when we estimate either of the two restricted QRE models with the data collected by GHP and the data collected in treatment 1.[44] This result therefore reinforces the conjecture that implementing the auction in a strategic form, and asking subjects to make predictions did not introduce any behaviorally significant treatment effect.

When we estimate the two restricted models with the data collected in treatment 1, we find the standard deviation parameters $(\sigma_r, \sigma_\alpha, \sigma_\beta)$ to be significantly greater than zero and substantial compared to their respective means. In other words, our estimates suggest significant heterogeneity across subjects in terms of risk aversion and probabilistic beliefs. The first part of this result is consistent with a number of experimental analyses of first-price auctions in which heterogeneity in risk preferences is often identified (see e.g., Cox et al., 1988; Cox and Oaxaca, 1996; or Chen and Plott, 1998). To the best of our knowledge however, the presence of

[40] As illustrated in the Appendix, we cannot estimate a more general QRE model including simultaneously subjective probability and risk aversion when using only the observed bids. Indeed, the average coefficient of relative risk aversion r_0 is only identified separately from the parameters of the PWF (α_0, β_0) when one relies on both the auction data and the predictions.

[41] The results obtained with our heterogenous models are not directly comparable with those in GHP because they assume that subjects are homogenous. For a direct comparison, we refer the reader to an earlier working paper (Armantier and Treich, 2006), where we estimated exactly the same models as GHP. As we will see however, the estimates under the homogeneity and heterogeneity assumptions are consistent, which makes the comparison with GHP's results relevant.

[42] Under homogeneity, we find $(\hat{\mu}, \hat{r}) = (0.067, 0.611)$ (Armantier and Treich, 2006).

[43] See e.g., Cox and Oaxaca (1996), Chen and Plott (1998), Campo et al. (2000), as well as Pezanis-Christou and Romeu (2003).

[44] A formal comparison of the QRE models estimated with the data collected by GHP and the data collected in treatment 1 may be found in a separate document (Addendum 2) available on one of the author's web-site at http://www.sceco.umontreal.ca/liste_personnel/armantier/index.htm.

heterogeneity in probabilistic beliefs has never been previously detected in experimental economics. As we will see, these results remain valid when we estimate the full QRE models.

We now turn to the estimation of the restricted models with the data collected in treatment 2. As indicated in Table 4, the estimated results vary slightly compared to treatment 1, but they remain within the same order of magnitude. For instance, the noise parameter in the model with risk aversion and objective beliefs increases significantly in treatment 2 (0.102), whereas the average risk aversion parameter drops slightly (0.521). These differences may be partially explained by the dual adjustment process taking place in treatment 2, as subjects first learned to make better predictions, and then to adjust their strategies.

5.2. *Full QRE Model Estimated with Observed Bids and Predictions.* We now consider the full QRE model combining probabilistic beliefs and risk aversion. This model is estimated with both the bids and the predictions made by subjects during the experiment. As previously shown however, subjects in treatment 2 learned to make better predictions over time. Therefore, the shape of the PWF is likely to differ significantly between the first and last periods. In order to account for this possibility, the model has been successively estimated with the data collected during the entire experiment (i.e., 15 periods), and with the data collected during the last five periods only.

The estimation results are presented in Table 5. In both treatments, the parameters (α_0, β_0) are significantly lower than (1, 2), which implies that, although of similar shape, the PWF is different from the one suggested by GHP. As expected, the PWF estimated with the last five periods of treatment 2 is flatter, i.e., $(\hat{\alpha}_0, \hat{\beta}_0)$ are significantly closer to (1, 1). This result reflects the fact that by the end of the session, subjects have learned to make nearly unbiased estimates. The estimated parameters however, are significantly different from $(\alpha_0, \beta_0) = (1, 1)$. In other words, although nearly unbiased, subjects' predictions in the last five periods of treatment 2 remain imperfect on average.

The estimations of the average risk aversion parameter decrease sharply compared to Table 4. For instance, the average risk aversion parameter estimated with the entire sample collected in treatment 1 drops from 0.673 to 0.328 when we account for the stated predictions. Even more remarkably, this parameter is estimated at 0.196 during the last five periods of treatment 2. These results however, do not imply that subjects behaved as if risk neutral. Indeed, r_0 remains significantly greater than zero in both treatments. Finally, observe that the estimated noise parameters μ are not significantly different in Table 5, except for the last five periods of treatment 2

TABLE 5
FULL QUANTAL RESPONSE EQUILIBRIUM MODEL
(ESTIMATED WITH BIDS AND PREDICTIONS)

Periods	μ	r_0	α_0	β_0	σ_r	σ_α	σ_β	$\rho_{r,\alpha}$	$\rho_{r,\beta}$	$\rho_{\alpha,\beta}$
					Treatment 1					
All	0.096*	0.328*	0.691*	1.745*	0.073*	0.043*	0.166*	0.219	0.679*	0.070
	(0.007)	(0.007)	(0.009)	(0.036)	(0.011)	(0.019)	(0.033)	(0.194)	(0.203)	(0.122)
Last five	0.094*	0.342*	0.702*	1.787*	0.068*	0.048*	0.181*	0.294	0.510*	0.034
	(0.008)	(0.009)	(0.014)	(0.037)	(0.018)	(0.021)	(0.037)	(0.202)	(0.218)	(0.182)
					Treatment 2					
All	0.091*	0.375*	0.736*	1.240*	0.068*	0.059*	0.177*	0.148	0.578*	0.034
	(0.005)	(0.008)	(0.017)	(0.028)	(0.012)	(0.026)	(0.037)	(0.193)	(0.184)	(0.168)
Last five	0.042*	0.196*	0.787*	1.107*	0.063*	0.049	0.136*	0.252	0.533*	0.053
	(0.006)	(0.011)	(0.019)	(0.031)	(0.015)	(0.027)	(0.041)	(0.206)	(0.221)	(0.204)

*Parameter larger than zero at a 5% significance level.
Numbers in parenthesis refer to the standard deviations of the estimates.

TABLE 6
QUANTAL BEST-RESPONSE MODEL
(ESTIMATED WITH BIDS AND PREDICTIONS)

	Model Estimated with All Periods			Model Estimated with Last Five Periods		
	μ	r_0	σ_r	μ	r_0	σ_r
Treatment 1	0.058*	0.233*	0.061*	0.055*	0.202*	0.069*
	(0.005)	(0.008)	(0.010)	(0.006)	(0.012)	(0.016)
Treatment 2	0.079*	0.265*	0.056*	0.047*	0.189*	0.048*
	(0.007)	(0.010)	(0.011)	(0.010)	(0.012)	(0.015)

*Parameter larger than zero at a 5% significance level.
Numbers in parenthesis refer to the standard deviations of the estimates.

where it becomes closer to zero. This result reflects the fact that, as previously mentioned, subjects' behavior nearly conforms with the RNBNE during the last periods of treatment 2.

Interestingly, we find a correlation between risk preferences and probabilistic beliefs. Indeed, $\rho_{r,\beta}$, the coefficient of correlation between an individual's risk aversion parameter r_i and her PWF parameter β_i, is significantly greater than zero. Therefore, although observationally equivalent, biased subjective probabilities and risk aversion do not appear to act as substitutes for an individual. Instead, our estimation results suggest that they complement each other, as highly risk-averse subjects also appear to be more pessimistic or less accurate when predicting probabilities.[45] This result is somewhat consistent with Bellemare et al. (2005) who also identify a correlation between preferences and beliefs.[46]

Finally, we test which of the full QRE models in Table 5 or the restricted QRE models in Table 4 are better able to organize the experimental data collected over the 15 periods. To do so, we perform a series of likelihood-ratio tests for model selection proposed by Vuong (1989). We find that we can clearly reject at the usual significance levels both of the restricted specifications in favor of the full QRE model.[47] In other words, taking into consideration the subjects' predictions, in addition to their bids, improves significantly the fit of the QRE model.

5.3. *Quantal Best-Response Model.* Finally, we relax the assumption of equilibrium behavior. Instead, inspired by the results in the Section 4, we assume that subjects best-respond to the predictions they made about their probability of winning the auction. In other words, we now compare the subjects' actions in the experiment with their subjective best-responses. In order to account for noisy behavior, we develop a "Quantal Best-Response" model. In accordance with the basic principles underlying the traditional QRE approach, behavior remains random, but the probability to choose a given action now increases with its subjective expected utility, where the latter is calculated with the subjects predictions. We refer the reader to the Appendix for further details on the Quantal Best-Response model.

The estimates of the structural parameters under the Quantal Best-Response model are reported in Table 6.[48] Observe that the average risk aversion parameters are significantly lower

[45] Given the estimated values of (α_0, β_0), β_i essentially controls the degree of convexity of the PWF. Therefore, an individual with a high level of risk aversion r_i is more likely to have a large β_i which corresponds to a more severely convex PWF.

[46] In an analysis of the ultimatum game, Bellemare et al. (2005) find that proposers who are optimistic about the acceptance rates of responders also tend to have significantly higher levels of inequity aversion.

[47] When testing the full model against the restricted model with risk aversion and objective beliefs, we find an adjusted test statistic of 6.017 in treatment 1 and 5.776 in treatment 2, which correspond to P-values of 9.506×10^{-10} and 3.853×10^{-9}. When testing the full model against the restricted model with probabilistic beliefs and risk neutrality, we find an adjusted test statistic of 5.293 in treatment 1 and 5.082 in treatment 2, which correspond to P-values of 4.543×10^{-8} and 1.637×10^{-7}. Note also that the P-values are estimated by Bootstrap, and therefore they differ slightly from the asymptotic P-values.

[48] Note that the Quantal Best-Response model can only explain bidding behavior, not predictions. Therefore, the only parameters estimated are the noise and risk aversion coefficients.

than previously estimated. Indeed, we now find that r_0, although still significant, lies between 0.189 and 0.265, instead of the 0.5 to 0.6 range suggested by Table 4, as well as GHP and other studies relying solely on observed bids. Note also that our estimates of the risk aversion parameter do not appear to be sensitive to the belief-elicitation technique employed. Indeed, we find that r_0 lies between 0.197 and 0.292 in the *no-payment* treatment, and between 0.217 and 0.274 in the *scoring-rule* treatment.[49] Finally, pairwise comparisons on the basis of the Vuong specification test for nonnested hypothesis systematically lead to reject the models in Tables 4 and 5 in favor of the model in Table 6.[50] These tests therefore indicate that bidding behavior observed during the experiment may be best explained by a model in which subjects tend to best-respond to their probabilistic beliefs.

To summarize, the structural estimations suggest that subjects in our experiment may be best characterized as utility maximizers with heterogenous probabilistic beliefs and levels of risk aversion. Moreover, the risk aversion parameters we estimated are far too low to explain alone the subjects' tendency to bid above the RNBNE. Instead, the structural estimations confirm that biased probabilistic beliefs is one of the main determinant of overbidding in our experiment.

6. SUMMARY AND DISCUSSION

We conducted a two-treatment auction experiment in which subjects were asked, in addition to bidding, to predict their probability of winning the auction. The experimental outcomes in treatment 1 indicate that subjects overbid, and underestimate their probability of winning. In treatment 2, we find that providing feedback on the accuracy of their predictions leads subjects to make better predictions, and then to curb drastically their tendency to overbid. The estimation of noisy behavioral models suggests that subjects (i) tend to best-respond to their predictions, and (ii) are heterogenous with respect to probabilistic beliefs and risk aversion. We also find that accounting for elicited probabilities in addition to observed choices improves substantially the fit of the model and yields significantly lower estimates of risk aversion. Finally, we identify a correlation between beliefs and preferences, as we find that highly risk-averse bidders are more likely to underestimate their probability of winning the auction. Our results therefore confirm that biased probabilistic beliefs is one of the main source of overbidding. In contrast, although still necessary to rationalize fully behavior, risk aversion is found to play a lesser role than previously believed. These findings tend to support the views of Kagel and Roth (1992) who argue that risk aversion may be one of the determinant of overbidding, but not necessarily the most important one.

It has to be noted however, that the conclusions of this article should be interpreted with caution, as we need to acknowledge at least three limitations of our experiment. First, to infer the subjects' subjective probabilities, we had to conduct a discrete auction that does not necessarily possess the Bayesian Nash equilibrium in pure strategy for any risk aversion coefficient. This prevented us from directly comparing the biased subjective probabilities hypothesis with the leading risk aversion model, i.e., the Bayesian Nash equilibrium model with constant relative risk aversion (see e.g., Cox et al., 1988). Second, in the absence of a firm methodological foundation, we must be cautious when interpreting the predictions we elicited as the true beliefs on which the subjects based their bidding decisions. Our experimental results, however, provide some support to this hypothesis. Indeed, the beliefs we elicited from the subjects are found to

[49] For additional and more detailed comparisons of the results obtained under the three different belief-elicitation techniques, see Addendum 2, available on one of the author's web-site at http://www.sceco.umontreal.ca/liste_personnel/armantier/index.htm.

[50] In treatment 1, when testing the Quantal Best-Response model against the full QRE model we find an adjusted test statistic of 3.067 with the data collected during the last five periods and 4.200 with the data collected during all 15 periods, which correspond to P-values of 1.351×10^{-3} and 9.698×10^{-6}. In treatment 2, when testing the Quantal Best-Response model against the full QRE model we find an adjusted test statistic of 3.128 with the data collected during the last five periods and 3.952 with the data collected during all 15 periods, which correspond to P-values of 9.235×10^{-4} and 3.884×10^{-5}. When testing the Quantal Best-Response model against the restricted QRE models in Table 4 we find P-values ranging from 5.916×10^{-10} to 7.370×10^{-34}.

be consistent with their bidding decisions. In addition, the experimental outcomes remained essentially unaffected when we employed different belief-elicitation techniques. This last result is particularly important, as it points to the robustness of our results. Third, we have assumed throughout the article that subjective probabilities reflected only the beliefs agents hold about probabilities. We have ignored in particular the decision weights that may reflect agents' preferences over probabilities when facing uncertain decisions (see e.g., Kahneman and Tversky, 1979). It is unclear at this point whether such decision weights would complement or partially offset the effect of risk aversion and of biased probabilistic beliefs in explaining overbidding. We found it nontrivial to modify our experiment to account for decision weights. The design of such an experiment remains an open question left for future research.

Finally, we conclude with some general comments. Although the existence of biased probabilistic beliefs is a well-established phenomenon in the psychology literature, its implications have rarely been explicitly analyzed in a strategic environment. As one of the first attempt to fill this void, our article illustrates how such an analysis may help us better understand strategic behavior. In that respect, two points are worth emphasizing. First, our experimental analysis shows why it may of importance to disentangle biased probabilistic beliefs from other behaviorally equivalent hypotheses such as risk aversion. Indeed, risk aversion is usually assumed to be an intrinsic individual characteristic that is time and context independent. In contrast, our experiment shows that probabilistic beliefs may be corrected, in which case individual behavior may be significantly affected. Second, our article illustrates how the elicitation of subjective probabilities may provide new insights into preferences. Indeed, preferences are usually estimated after imposing some beliefs (e.g., rational or naive) on all agents. This procedure may be flawed, as an incorrect specification of the beliefs may lead to a misrepresentation of preferences. Consistent with Manski (2004), we find that the combination of observed choices with data on subjective probabilities mitigates this identification problem, and improves our understanding of preferences. Indeed, our estimates of risk aversion based on choices and elicited probabilities differ markedly from previous studies relying solely on choice data.

APPENDIX. QRE MODEL WITH UNOBSERVED HETEROGENEITY

Consider a discrete private-values auction with two bidders *i* and *j*. Let us denote v_i and v_j their respective private signals, and b_i and b_j their respective bids. In GHP's auction, $(v_i, v_j) \in \{0, 2, 4, 6, 8, 11\}^2$ and $(b_i, b_j) \in \mathbb{N}^2$. In what follows, the indices *i* and *j* are exchangeable.

In order to account for heterogenous preferences, we follow Cox et al. (1988) in assuming that the utility of bidder *i* may be written as

$$U_i\left(b_i, v_i\right) = \frac{\left(v_i - b_i\right)^{1-r_i}}{1 - r_i} \mathbb{I}_{[i \text{ wins}]},$$

where r_i is agent *i* Arrow–Pratt coefficient of relative risk aversion, and $\mathbb{I}_{[.]}$ is the indicator function satisfying $\mathbb{I}_{[i\text{wins}]} = 1$ when bidder *i* wins the auction (i.e., $b_i > b_j$ or $b_i = b_j$ and *i* wins the coin toss), and $\mathbb{I}_{[i\text{wins}]} = 0$ otherwise.

In order to model heterogenous probabilistic beliefs, we generalize the PWF proposed by Prelec (1998)

$$\Phi_i\left(P\right) = \exp\left(-\beta_i\left(-\ln\left(P\right)\right)^{\alpha_i}\right),$$

where *P* is an objective probability, and the parameters $(\alpha_i, \beta_i) \in \mathbb{R}^2_{++}$ are specific to bidder *i*. Note that this PWF can display the typical inverse S-shape pattern for specific values of (α_i, β_i). In addition, agent *i* subjective and objective beliefs are equivalent when $(\alpha_i, \beta_i) = (1, 1)$. Following Cox et al. (1988), we assume that an agent faces uncertainty about her opponent risk preference and subjective probability. In other words, bidder *i* observes $\theta_i = (r_i, \alpha_i, \beta_i)$, but she only knows the distribution from which θ_j is generated.

Under the QRE approach, stochastic behavior is represented by a probabilistic choice rule. In the present context, the choice rule of bidder i specifies the probability $P_i(b_i, v_i, \theta_i)$ that bidder i submits a bid b_i when endowed with a private-value v_i and a type θ_i. The exclusion of strictly dominated strategies (i.e., bidding above one's own private-value) implies that $P_i(b_i, v_i, \theta_i) = 0$ when $b_i > v_i$. In addition, for the choice rule to be a probability function we must have $\sum_{0 \le b_i \le v_i} P_i(b_i, v_i, \theta_i) = 1$ for any (v_i, θ_i). Finally, let $\bar{P}_i(b_i, v_i) = E_{\theta_i}[P_i(b_i, v_i, \theta_i)]$ denotes the average choice rule that specifies the probability that bidder i submits a bid b_i for a private-value v_i, unconditional on her type θ_i.

We can now define $P_i^W(b_i, \bar{P}_j)$, the objective probability that bidder i wins the auction with a bid b_i when facing an opponent relying on an average choice rule $\bar{P}_j$. In the discrete auction model, this objective probability may be written as

$$P_i^W(b_i, \bar{P}_j) = \frac{1}{6} \sum_{v_j \in \{0,2,4,6,8,11\}} \sum_{b_j < b_i} \bar{P}_j(b_j, v_j) + \frac{1}{12} \sum_{v_j \in \{0,2,4,6,8,11\}} \bar{P}_j(b_i, v_j),$$

where the first term represents the average probability that the opponent of bidder i submits a bid strictly below b_i when receiving one of the six possible private-values, and the second term represents a coin flip favorable to bidder i in the event of a tie. When facing an opponent relying on a choice rule P_j, the expected utility of a bidder i conditional on (b_i, v_i, θ_i) may then be written as

$$\tilde{U}_i(b_i, v_i, \theta_i, P_j) = U_i(b_i, v_i)\Phi_i\big(P_i^W(b_i, \bar{P}_j)\big).$$

In order to find a QRE, we must first specify how the choice probabilities depend on the expected utility. Following GHP, we adopt the power function probabilistic choice rule

$$\text{(A.1)} \qquad P_i(b_i, v_i, \theta_i) = \frac{\{\tilde{U}_i(b_i, v_i, \theta_i, P_j)\}^{1/\mu}}{\sum_{b=0}^{v_i} \{\tilde{U}_i(b, v_i, \theta_i, P_j)\}^{1/\mu}},$$

where $\mu > 0$ is a "noise parameter" reflecting the sensitivity of the choice probabilities to subjective expected utilities. A large μ yields essentially random behavior, whereas μ close to zero implies Nash-like behavior because best-response strategies are chosen with a probability close to one.

We can now try to solve the model, i.e., find the QRE choice probabilities. Observe first that, although agents are endowed with different θ_i, the model is ex-ante symmetric, and therefore, the QRE choice probabilities $P^*(b, v, \theta)$ are also symmetric across bidders. Despite symmetry, however, we are unable to solve the model. Indeed, the QRE choice probabilities are a function of a continuous random vector θ. Consequently, unless we impose arbitrary parametric restrictions on the shape of the choice probabilities as a function of θ, we are unable to solve (even numerically) the QRE model under unobserved heterogeneity.

Fortunately, we do not have to solve the model fully to estimate the structural parameters. Indeed, observe first that unlike $P^*(b, v, \theta)$, the average QRE choice probabilities $\bar{P}^*(b, v)$ can only take a finite number of values, because b and v are both discrete. In addition, note that (A.1) yields

$$\text{(A.2)} \quad \bar{P}_i(b_i, v_i) = E_{\theta_i}\left[\frac{\{\tilde{U}_i(b_i, v_i, \theta_i, P_j)\}^{1/\mu}}{\sum_{b=0}^{v_i}\{\tilde{U}_i(b, v_i, \theta_i, P_j)\}^{1/\mu}}\right] = E_{\theta_i}\left[\frac{\big\{U_i(b_i, v_i)\Phi_i\big(P_i^W(b_i, \bar{P}_j)\big)\big\}^{1/\mu}}{\sum_{b=0}^{v_i}\big\{U_i(b, v_i)\Phi_i\big(P_i^W(b, \bar{P}_j)\big)\big\}^{1/\mu}}\right],$$

which only involves the unconditional choice probabilities $\bar{P}_i$ and $\bar{P}_j$. Therefore, we can find the average QRE choice probabilities by replacing both $\bar{P}_i$ and $\bar{P}_j$ by $\bar{P}^*$ in (A.2), and solving numerically the resulting fixed-point problem.[51]

Once the vector of average QRE choice probabilities $\bar{P}^*$ has been calculated, we can write the likelihood of the QRE model conditional on the bidding decisions as a function of the structural parameters

$$L(\mu, r_0, \alpha_0, \beta_0, \Sigma) = \prod_{i,v,t,b} [\bar{P}^*(b, v)]^{\mathbb{I}[B_{i,v,t}=b]} \quad , \tag{A.3}$$

where $B_{i,v,t}$ is the bid actually submitted by subject i for the private-value v at period t. In other words, the likelihood function in (A.3) compares the subjects' observed actions with their corresponding average QRE probabilities of choice.

As mentioned earlier however, the distribution of the risk-aversion parameters r_i cannot be identified separately from the distribution of the PWF parameters (α_i, β_i) when one relies only on observed actions.[52] In order to identify fully the model, the estimation procedure must account for the subjects' predictions in addition to their bids. To do so, we transform the theoretic PWF into an estimable econometric model by adding an exogenous error term. In other words, we assume that $\hat{P}_{i,b,t} = \Phi_i(P^W_{i,b,t}) + \epsilon_{i,b,t}$, where $\hat{P}_{i,b,t}$ is the prediction submitted by subject i for bid b in round t, $P^W_{i,b,t}$ is the corresponding objective probability calculated with the bids submitted by the other subjects in that round, and $\epsilon_{i,b,t}$ is a normally distributed error term. The likelihood of the PWF model conditional on the observable $P^W_{i,b,t}$ may then be obtained after taking the expectation of $\Phi_i(P^W_{i,b,t})$ with respect to the latent parameters (α_i, β_i).[53] Finally, by combining the likelihood functions of the QRE and PWF models, we can estimate with the method of simulated maximum likelihood (see Gourieroux and Monfort, 2002) the structural parameters of the resulting nonlinear panel data model with unobserved individual effects.

Finally, we conclude this Appendix with a description of the Quantal best-response model. The basic idea is that subjects tend to best-respond to the predictions they made about their probability of winning. In order to account for noisy behavior, we assume that subjects do not necessarily select their subjective best-responses with probability one. Instead, in accordance with the basic principles underlying the QRE approach, we assume that the probability that subject i selects a bid $b_i = 0, \ldots, v_i$ when receiving a private-value v_i at period t is

$$P_{i,t}(b_i, v_i, \theta_i) = \frac{\{\hat{E}_{i,t}[U_i(b_i, v_i)]\}^{1/\mu}}{\sum_{b=0}^{v_i} \{\hat{E}_{i,t}[U(b, v_i)]\}^{1/\mu}} \quad \text{where} \quad \hat{E}_{i,t}[U_i(b_i, v_i)] = \frac{(v_i - b_i)^{1-r_i}}{1 - r_i} \hat{P}_{i,b,t},$$

and $\hat{P}_{i,b,t}$ is subject i's stated probability of winning with a bid b in period t. Recall that we observe $\hat{P}_{i,b,t}$ for each agent and each period. Therefore, the only structural parameters to estimate in this model are (μ, r_0, σ_r), the noise and risk-aversion parameters.

REFERENCES

ALI, M., "Probability and Utility Estimates for Racetrack Bettors," *Journal of Political Economy* 85 (1977), 803–15.

ANDERSON, S., J. GOEREE, AND C. HOLT, "The Logit Equilibrium: A Perspective on Intuitive Behavioral Anomalies," *Southern Economic Journal* 69 (2002), 21–47.

[51] The expectation in (A.2) does not have a closed-form solution, and it must be evaluated numerically. In order to speed-up the numerical integration, we rely on a quasi-Monte Carlo sampling method consisting in generating extensible lattice points modified by the "Baker's transformation" (see Hickernell et al., 2000).

[52] In particular, observe that the two sets of parameters $(\mu, r_i, \alpha_i, \beta_i)$ and $(k\mu, 1 - k(1 - r_i), \alpha_i, k\beta_i)$ (where $k > 0$) yield the same choice probabilities in (A.1) although the vectors of random variables (r_i, α_i, β_i) and $(1 - k(1 - r_i), \alpha_i, k\beta_i)$ have different distributions.

[53] Again, this expectation is calculated numerically using quasi-Monte Carlo sampling techniques.

Armantier, O., and N. Treich, "Probability Misperception in Games: An Application to the Overbidding Puzzle," Mimeo, Université de Montréal, 2006.

Bellemare, C., S. Kröger, and A. van Soest, "Actions and Beliefs: Estimating Distribution-Based Preferences Using a Large Scale Experiment with Probability Questions on Expectations," Mimeo, Tilburg University, 2005.

Benjamin, D., W. Dougan, and D. Buschena, "Individuals' Estimates of the Risks of Death: Part II–New Evidence," *Journal of Risk and Uncertainty* 22 (2001), 35–57.

Camerer, C., "Individual Decision Making," in J. H. Kagel, and A.E. Roth, eds., *Handbook of Experimental Economics* (Princeton, NJ: Princeton University Press, 1995), 587–703.

——, and D. Lovallo, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review* 89 (1999), 306–18.

——, T. Ho, and J. Chong, "Behavioral Game Theory: Thinking, Learning, and Teaching," Paper presented at the Nobel Symposium. in S. Huck, ed., *Essays in Honor of Werner Guth*, 2001.

Campo, S., I. Perrigne, and Q. Vuong, "Asymmetry in First-Price Auctions with Affiliated Private Values," *Journal of Applied Econometrics* 2 (2001), 179–207.

Capra, M., J. Goeree, R. Gomez, and C. Holt, "Learning and Noisy Behavior in an Experimental Study of Imperfect Price Competition," *International Economic Review* 43 (2002), 613–36.

Charness, G., and M. Dufwenberg, "Promises & Partnership," *Econometrica* 74 (2006), 1579–1601.

Chen, K.Y., and C. Plott, "Nonlinear Behavior in Sealed Bid First-Price Auctions," *Games and Economic Behavior* 25 (1998), 34–78.

Costa-Gomes, M., and G. Weizsäcker, "Stated Beliefs and Play in Normal-Form Games," Mimeo, Harvard University, 2004.

Cox, J., and R. Oaxaca, "Is Bidding Behavior Consistent with Bidding Theory for Private Value Auctions," in R. M. Isaac, ed., *Research in Experimental Economics*, Vol. 6, (Greenwich, Conn.: JAI Press, 1996), 131–48.

——, V. Smith, and J. Walker, "Tests of a Heterogeneous Bidder's Theory of First-Price Auctions," *Economics Letters* 12 (1983), 207–12.

——, ——, and ——, "Experimental Development of Sealed-Bid Auction Theory: Calibrating Controls for Risk Aversion," *American Economic Review* 75 (1985), 160–5.

——, ——, and ——, "Theory and Individual Behavior of First-Price Auctions," *Journal of Risk and Uncertainty* 1 (1988), 61–99.

Croson, R., "Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play," *Journal of Economic Behavior and Organization* 41 (2000), 299–314.

De Finetti, B., "Methods of Discriminating Levels of Partial Knowledge Concerning a Test Item," *British Journal of Mathematical and Statistical Psychology* 18 (1965), 87–123.

Dorsey, R., and L. Razzolini, "Explaining Overbidding in First Price Auctions using Controlled Lotteries," *Experimental Economics* 6 (2003), 123–40.

Dufwenberg, M., and U. Gneezy, "Measuring Beliefs in an Experimental Lost Wallet Game," *Games and Economic Behavior* 30 (2000), 163–82.

Friedman, D., "Theory and Misbehavior of First-Price Auctions: Comment," *American Economic Review* 82 (1992), 1374–8.

Gerking, S., M. DeHaan, and W. Schulze, "The Marginal Value of Job Safety: A Contingent Valuation Study," *Journal of Risk and Uncertainty* 1 (1988), 18–99.

Giordani, P., and P. Söderlind, "Is there Evidence of Pessimism and Doubt in Subjective Distributions? Implications for the Equity Premium Puzzle," *Journal of Economic Dynamics and Control* 30 (2006), 1027–43.

Goeree, J., C. Holt, and T. Palfrey, "Quantal Response Equilibrium and Overbidding in Private-Value Auctions," *Journal of Economic Theory* 104 (2002), 247–72.

Golec, J., and M. Tamarkin, "Bettors Love Skewness, Not Risk, at the Horse Track," *Journal of Political Economy* 106 (1998), 205–25.

Gourieroux, C., and A. Monfort, "Simulation-Based Econometric Methods," New York: Oxford University Press, 2002.

Harrison, G., "Theory and Misbehavior of First-Price Auctions," *American Economic Review* 79 (1989), 749–62.

——, "Risk Attitudes in First-Price Auction Experiments: A Bayesian Analysis," *Review of Economics and Statistics* 72 (1990), 541–46.

Hickernell, F. J., H. S. Hong, P. L'Ecuyer, and C. Lemieux, "Extensible Lattice Sequences for Quasi-Monte Carlo Quadrature," *SIAM Journal on Scientific Computing* 22 (2000), 1117–38.

Hurley, T., and J. Shogren, "An Experimental Comparison of Induced and Elicited Beliefs," *Journal of Risk and Uncertainty* 30 (2005), 169–88.

Isaac, M., and J. Walker, "Information and Conspiracy in Sealed Bid Auctions," *Journal of Economic Behavior and Organization* 6 (1985), 139–59.

Kagel, J., "Auctions: A Survey of Experimental Research," in J.H. Kagel, and A.E. Roth, eds., *Handbook of Experimental Economics* (Princeton, NJ: Princeton University Press, 1995), 501–86.

——, AND A. ROTH, "Theory and Misbehavior in First-Price Auctions: Comment," *American Economic Review* 82 (1992), 1379–91.

KAHNEMAN, D., AND A. TVERSKY, "Prospect Theory: an Analysis of Decision under Risk," *Econometrica* 47 (1979), 263–91.

LICHTENSTEIN, S., P. SLOVIC, B. FISCHHOFF, B. LAYMAN, AND B. COMBS, "Judged Frequency of Lethal Events," *Journal of Experimental Psychology: Human Learning and Memory* 4 (1978), 551–78.

MANSKI, C. F., "Identification of Decision Rules in Experiments on Simple Games of Proposal and Response," *European Economic Review* 46 (2002), 880–91.

——,"Measuring Expectations," *Econometrica* 72 (2004), 1329–76.

MCKELVEY, R., AND T. PALFREY, "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior* 10 (1995), 6–38.

—— AND ——,"Quantal Response Equilibria in Extensive Form Games," *Experimental Economics* 1 (1998), 9–41.

MURPHY, A., AND R. WINKLER, "Scoring Rules in Probability Assessment and Evaluation," *Acta Psychologica* 34 (1970), 273–386.

—— AND ——,"Probability Forecasting in Meteorology," *Journal of the American Statistical Association* 79 (1984), 489–500.

NEUGEBAUER, T., AND R. SELTEN, "Individual Behavior of First-Price Sealed-Bid Auctions: The Importance of Information Feedback in Experimental Markets," Mimeo, University of Bonn, 2003.

NYARKO, Y., AND A. SCHOTTER, "An Experimental Study of Belief Learning Using Elicited Beliefs," *Econometrica* 70 (2002), 971–1005.

OTTAVIANI, M., AND P. SØRENSEN, "Forecasting and Rank-Order Contests," Mimeo, London School of Economics, 2003.

PALFREY, T. R., AND H. ROSENTHAL, "Testing Game-Theoretic Models of Free Riding: New Evidence on Probability Bias and Learning," in T. Palfrey, ed., *Contemporary Laboratory Research in Political Economy* (Ann Arbor: University of Michigan Press), 1991.

PEZANIS-CHRISTOU, P., AND A. ROMEU, "Structural Inferences from First-Price Auction Experiments," Mimeo, Institut d'Anàlisi Econòmica, 2003.

——, AND A. SADRIEH, "Elicited Bid Functions in (A)Symmetric First-Price Auctions," Mimeo, Institut d'Anàlisi Econòmica, 2003.

PRELEC, D., "The Probability Weighting Function," *Econometrica* 66 (1988), 497–527.

ROTH, A., "Bargaining Experiments," in J.H. Kagel, and A.E. Roth, eds., *Handbook of Experimental Economics* (Princeton, NJ: Princeton University Press, 1995), 253–348.

RUTSTRÖM, E., AND N. WILCOX, "Learning, Belief Elicitation and Heterogeneity," Mimeo, University of Houston, 2005.

SANDERS, F., "Skill in Forecasting Daily Temperature and Precipitation: Some Experimental Results," *Bulletin of the American Meteorological Society* 54 (1973), 1771–9.

SAVAGE, L., "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association* 66 (1971), 783–801.

SCHOTTER, A., AND B. SOPHER, "Advice and Behavior in Intergenerational Ultimatum Games: An Experimental Approach," Mimeo, New–York University, 2001.

SELTEN, R., AND J. BUCHTA, "Experimental Sealed Bid First Price Auctions with Directly Observed Bid Function," in D. Budescu, I. Erev, and R. Zwick, eds., *Games and Human Behavior—Essays in Honor of A. Rapoport* (Mahwa, NJ, USA, Lawrence Erlbaum Associates, Inc., 1999).

SHAO, J., AND D. TU, *The Jackknife and Bootstrap* (New York: Springer-Verlag, 1995).

VISCUSI, K., J. HAKES, AND A. CARLIN, "Measures of Mortality Risks," *Journal of Risk and Uncertainty* 14 (1997), 213–33.

——, AND C. O'CONNOR, "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?," *American Economic Review* 74 (1984), 942–56.

VUONG, Q., "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica* 57 (1989), 307–33.

WAKKER, P., R. THALER, AND A. TVERSKY, "Probabilistic Insurance," *Journal of Risk and Uncertainty* 15 (1997), 7–28.

WILCOX, N., AND N. FELTOVICH, "Thinking Like a Game Theorist: Comment," Mimeo, University of Houston, 2000.

YI, K.-O., "Quantal-Response Equilibrium Models of Order-Statistics Games," *Journal of Economic Behavior and Organization* 51 (2003), 413–25.

Query

Q1 Author: Figures have been renumbered to make citation in sequential order. Please check.

WILEY-BLACKWELL

350 Main Street, Malden, MA 02148-9933 USA
Phone: (781) 388-8200; Fax: (781) 388-8257

www.blackwellpublishing.com

Recommend a Subscription to Your Library

Dear Contributor:

Together with the Editorial office, our goal is to make sure the journal is easily accessible to all researchers and students around the world. Please help us achieve that goal by recommending that your library subscribe to this journal. Please complete the Library Recommendation Form provided below, and forward it to your librarian or library committee member at your earliest convenience.

Online Access: Subscription to most of our journals entitles all members of the institution to full online access. For details, please visit: www.blackwellpublishing.com/journals

E-mail Alerts: When you register to receive E-mail Alerts, you will receive a complete Table of Contents of the current issue as it publishes. Sign up today at: http://www.blackwell-synergy.com/

Thank you for your contribution to this journal.

Yours sincerely,

Wiley-Blackwell

P.S. If you do not currently have a personal subscription to this journal, and would like to find out more about personal subscriptions, please call us toll-free at 800-835-6770 (outside of the U.S., dial 781-388-8200) or visit our website at http://www.blackwellpublishing.com

✂---

Library Recommendation Form

To: Librarian / Library Acquisition Committee /________________________
From: ____
Department: ____
Please include the **journal**: ______________________ **ISSN** __________

I recommend the journal for the following reasons (please check all that apply):

- ☐ **REFERENCE**: I will refer to this journal frequently for new research articles related to my work.
- ☐ **STUDENT READINGS**: I will be referring my students to this journal regularly to assist their studies.
- ☐ **BENEFIT FOR LIBRARY**: This journal will add to the library's success in fulfilling department, faculty, and student needs
- ☐ **MY PERSONAL AFFILIATION**: I am a contributor to/advisory board member of the journal. I support it strongly and use it frequently in my work. I will regularly recommend articles to colleagues and students.

Your Signature:____Date:________________

This journal is published by Wiley-Blackwell Publishing, 350 Main St., Malden, MA 02148, USA.
Tel: 800-835-6770 in the U.S., outside the U.S. (781) 388-8200; Fax: (781) 388-8232, Email: subscrip@bos.blackwellpublishing.com

Proof Correction Marks

Please correct and return your proofs using the proof correction marks below. For a more detailed look at using these marks please reference the most recent edition of The Chicago Manual of Style and visit them on the Web at: http://www.chicagomanualofstyle.org/home.html

Instruction to typesetter	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	(stet)
Insert in text the matter indicated in the margin	∧	∧ followed by new matter
Delete	through single character, rule or underline or through all characters to be deleted	
Substitute character or substitute part of one or more word(s)	/ through letter or ⊢—⊣ through characters	new character / or new characters /
Change to italics	— under matter to be changed	(ital)
Change to capitals	≡ under matter to be changed	(caps)
Change to small capitals	= under matter to be changed	(sc)
Change to bold type	~ under matter to be changed	(bf)
Change to bold italic	under matter to be changed	(bf+ital)
Change to lower case		(lc)
Insert superscript	∨	∨ under character e.g.
Insert subscript	∧	∧ over character e.g.
Insert full stop	⊙	⊙
Insert comma		
Insert single quotation marks		
Insert double quotation marks		
Insert hyphen	=	=
Start new paragraph	¶	¶
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	#	#
Reduce space between characters or words		