Exp Econ (2015) 18:569–608
DOI 10.1007/s10683-014-9420-1

ORIGINAL PAPER

Eliciting beliefs in continuous-choice games: a double auction experiment

Claudia Neri

Received: 18 December 2012 / Revised: 23 October 2013 / Accepted: 27 October 2014 / Published online: 7 April 2015
© Economic Science Association 2015

Abstract This paper proposes a methodology to implement probabilistic belief elicitation in continuous-choice games. Representing subjective probabilistic beliefs about a continuous variable as a continuous subjective probability distribution, the methodology involves eliciting partial information about the subjective distribution and fitting a parametric distribution on the elicited data. As an illustration, the methodology is applied to a double auction experiment, where traders' beliefs about the bidding choices of other market participants are elicited. Elicited subjective beliefs are found to differ from proxies such as Bayesian Nash equilibrium beliefs and empirical beliefs, both in terms of the forecasts of other traders' bidding choices and in terms of the best-response bidding choices prescribed by beliefs. Elicited subjective beliefs help explain observed bidding choices better than BNE beliefs and empirical beliefs. By extending probabilistic belief elicitation beyond discrete-choice games to continuous-choice games, the proposed methodology enables to investigate the role of beliefs in a wider range of applications.

Keywords Probabilistic beliefs · Belief elicitation · Private information · Experiments

1 Introduction

Beliefs play an important role in game-theoretic models of strategic decision-making. In equilibrium models, such as the Nash equilibrium, the outcome of a game is

Electronic supplementary material The online version of this article (doi:10.1007/s10683-014-9420-1) contains supplementary material, which is available to authorized users.

C. Neri (✉)
Department of Economics, University of St. Gallen, Varnbüelstrasse 19, 9000 St. Gallen, Switzerland
e-mail: claudia.neri@unisg.ch

interpreted as a steady state where players hold correct beliefs about opponents' behavior and act rationally by best responding to those beliefs. In learning models, such as the belief learning model, which focus on if and how a steady state outcome is reached, players update their beliefs about opponents' actions by observing opponents' past actions and then make their choice by best responding to those beliefs. When game-theoretic models are used to explain empirical or experimental choice data in the absence of beliefs data, proxies are used to represent beliefs. In the equilibrium framework, the proxy is provided by equilibrium beliefs defined in accordance with equilibrium strategies. In the belief learning framework, the proxy is provided by empirical beliefs computed from opponents' past behavior.

In recent years a literature on the probabilistic elicitation of subjective beliefs in games has emerged, with the aim to verify (otherwise non-verifiable) model assumptions and to explore whether elicited beliefs lead to better predictions of choice behavior than alternative proxies do. The literature has focused on normal-form games with a discrete-choice variable, either in a finitely-repeated (Nyarko and Schotter 2002; Rutström and Wilcox 2009; Hyndman et al. 2012; Danz et al. 2012) or in a one-shot setting (Costa-Gomes and Weizsäcker 2008; Rey-Biel 2009).

This paper proposes a methodology to extend probabilistic belief elicitation to games with a continuous choice set, thus allowing for the use of beliefs data in a wider range of applications than the one covered by the existing literature. As an illustration, the methodology is applied to a uniform-price double auction experiment where multiple buyers and sellers with independent private valuations submit their bidding choices and state their probabilistic beliefs about other participants' bidding choices. The elicited beliefs data then allow for the investigation of how beliefs affect bidding choices.[1]

Moving from a discrete- to a continuous-choice setting requires profound modification to the belief elicitation procedure. In a discrete-choice setting, probabilistic beliefs are represented by a discrete probability distribution and are elicited by having participants assign a probability to each possible element in the set. In a continuous-choice setting, instead, probabilistic beliefs are represented by a continuous probability distribution. In this paper, partial information about the subjective distribution is elicited by partitioning the choice set into intervals and requiring each participant to report the probability that opponents' choices fall within each interval. A parametric distribution is then fitted to the elicited beliefs data in order to recover beliefs in the form of a fitted continuous subjective distribution. The procedure builds on previous work on non-strategic decisions and a survey setting (see the Survey of Economic Expectations and Engelberg et al. (2009)), modifying them to be appropriate for strategic decisions and an experimental setting.

Belief elicitation within the double auction experiment enables us to compare elicited subjective beliefs with its proxies, Bayesian Nash equilibrium (BNE) beliefs and empirical beliefs.[2] Subjective beliefs differ from BNE and empirical beliefs not only in terms of the forecasts of other agents' bidding choices but also in terms of the best-response bidding choices prescribed by beliefs. The extent to which elicited

[1] In a uniform-price double auction (also called static double auction, call auction, or clearinghouse auction) all traders trade at the same time, when the market is 'called'.

[2] The equilibrium model for the double auction consists of the BNE.

subjective beliefs help explain observed choices is evaluated comparing the consistency between choice and subjective beliefs to the consistency between choice and BNE or empirical beliefs. Choices are found to be more often consistent with the best response to subjective beliefs than with the best response to BNE or empirical beliefs. Deviation of choice from BNE best response is related to the deviation of subjective beliefs from BNE beliefs. Neither subjective beliefs converge to BNE beliefs nor choice converges to BNE best response. However, evidence suggests that convergence in beliefs may be necessary but not sufficient for convergence in choice, while convergence in choice may be sufficient but not necessary for convergence in beliefs. Despite helping explain observed choices, subjective beliefs have a low accuracy in predicting other subjects' behavior, whether accuracy is measured by a scoring rule or by calibration.

The belief elicitation design is relevant and applicable, beyond the context of a double auction, to settings that involve multiple players (possibly with different roles and types), a continuous choice set, and a payoff function that depends on a (possibly involved) function of players' choices. Relevant settings may include market competition in oligopoly with incomplete information about production costs or bargaining with incomplete information about reservation values.

The remainder of the paper is organized as follows. Section 2 describes the experimental design and the belief elicitation methodology. Section 3 presents the main results. Section 4 concludes.

2 Experimental procedures

A total of 66 subjects, recruited among Northwestern University undergraduate students, participated in a total of 11 auction markets designed as computer lab experiments using the z-Tree software. Four auctions were conducted with eight traders, four with four traders and three with six traders. At the beginning of the experiment participants are randomly matched into groups of traders and interact within the same group over the course of 15 rounds. In the first round each player is randomly assigned the role of either buyer or seller. If a player is initially assigned to the role of buyer (seller), then she is a buyer in rounds 1 through 5, a seller (buyer) in rounds 6 through 10, and a buyer (seller) again in rounds 11 through 15. The information about each participant's current role always appears on the computer screen. The experimental auction follows the rules of the double auction model of Rustichini et al. (1994), which is reviewed in Appendix 2.[3]

In each round, each buyer has \$10 and each seller has one unit of the commodity. Each buyer can purchase a single unit of the commodity from any seller, and each seller

[3] Cason and Friedman (1997) conduct a double auction experiment under the trading rules of the model of Rustichini et al. (1994). Most importantly, their design differs from the one of this paper because beliefs data is not elicited. Moreover, private valuations are drawn from the uniform distribution over \$[0,4.99], participants play 30 trading rounds, profits accumulate throughout the rounds and participants are paid the accumulated amount. They report analogous evidence that bidding choices deviate from the risk-neutral BNE predictions. However, by collecting only bidding choices, they face an identification problem, which I solve, following Manski (2002, 2004), by collecting both choice and beliefs data.

can sell a single unit of the commodity to any buyer. At the beginning of each round each buyer's personal value, v, and seller's personal cost, c, are assigned randomly: the computer determines each of them as an independent random draw of the 1,000 numbers (to the nearest cent) between \$0.00 and \$9.99.[4] After being privately informed about their valuation, buyers and sellers submit their bids and offers.[5] Any non-negative amount, specified up to a precision of cents of a dollar, is allowed. The market price p is set by sorting bids and offers in increasing order and selecting the midpoint between the middle two figures.[6] Buyers who submit a bid larger than or equal to p and sellers who submit an offer smaller than or equal to p trade. If trading at p, a buyer with private value v earns a payoff equal to $v - p$, and a seller with private cost c earns a payoff equal to $p - c$. Participants who do not trade earn zero.

The novelty of the experimental design consists of requiring participants not only to submit a bidding choice, but also to report their probabilistic beliefs about the bidding choices of other market participants.[7] Beliefs about the bids submitted by other buyers and beliefs about the offers submitted by other sellers are elicited separately.

At the end of each round, each participant receives feedback about the results for that round: a graph of the bids and offers submitted by market participants (i.e., the demand and supply curves), the market price, her trading status (i.e., whether she traded), her profit, the accuracy of her stated beliefs, and the corresponding monetary reward.[8] At the end of the last round, the computer randomly draws one of the rounds and the participants are paid according to their performance in that round only. On average, subjects earned \$10.26. In addition to \$5 for attendance, they earned on average \$0.82 in the trading task (\$2.08 if concluding a trade), \$2.24 in the forecasting task, and \$2.20 in a comprehension quiz administered at the beginning of the session.[9] The large relative size of forecasting earnings (and comprehension earnings) compared to trading earnings could limit the generalizability of the results.[10] Supplementary Material contains additional material, including the experiment instructions and a description of the participant pool.

---

4 Thus, in each round a new set of observations of bidding choices, market price and trades is collected.

5 The personal values and costs of other buyers and sellers are never revealed. After submitting their choices, players are not allowed to modify them.

6 In a market with n buyers and n sellers, after sorting all $2n$ bids and offers as $\psi_{(1)} \leq \psi_{(2)} \leq \cdots \leq \psi_{(2n)}$, the market price is set at $p = 0.5\psi_{(n)} + 0.5\psi_{(n+1)}$. $\psi_{(n)}$ denotes the n-th order statistic among all 2 n submitted bidding choices.

7 The elicitation task is presented after each subject has submitted her bid or offer. Other studies where beliefs are elicited after choices are submitted include Bellemare et al. (2010) and (2012). Eliciting beliefs after choices are submitted has the advantage that choices are unaffected by the belief elicitation but raises the concern that participants state beliefs to rationalize choices. Bellemare et al. (2012) compare the beliefs in a benchmark treatment (with participants who make choices and then state beliefs) to the beliefs in a treatment with observer-participants (who only state their beliefs but do not make any choice). They find that stated beliefs are not significantly affected by choices.

8 If a player does not trade, her initial cash (if a buyer) or the unsold unit of the commodity (if a seller) does not count towards profits.

9 The maximum amount a participant could earn in the comprehension quiz was \$5. The maximum amount a participant could earn for forecasting was \$4. According to the rules, any negative trading profit would be deducted from the total payments. However, no participant received a negative payoff at the end of the session.

2.1 Belief elicitation

I now turn to characterize the role of subjective beliefs in the decision problem of market participants and describe how belief elicitation is carried out.

According to the auction rules, the relationship between an individual's bidding choice and the choices of other participants determines what her payoff will be, depending on whether she trades and, if so, at what price. Lets consider buyer i and seller j in a market with n buyers and n sellers. Sorting all other $2n-1$ bids and offers as $\zeta_{(1)} \leq \zeta_{(2)} \leq \cdots \leq \zeta_{(2n-1)}$, a buyer i trades if her bid $b > \zeta_{(n)}$ and the trade occurs at price $p = 0.5\zeta_{(n)} + 0.5b$ if $\zeta_{(n)} < b < \zeta_{(n+1)}$ or $p = 0.5\zeta_{(n)} + 0.5\zeta_{(n+1)}$ if $b > \zeta_{(n+1)}$.[11] Analogously, a seller j trades if her offer $a < \zeta_{(n)}$ and the trade occurs at price $p = 0.5a + 0.5\zeta_{(n)}$ if $\zeta_{(n-1)} < a < \zeta_{(n)}$ or $p = 0.5\zeta_{(n-1)} + 0.5\zeta_{(n)}$ if $a < \zeta_{(n-1)}$.

Buyer i (seller j) chooses bid b (offer a) without knowing the other $2n$-1 bids and offers but holding subjective probabilistic beliefs about them. Assuming an agent behaves as a risk-neutral subjective expected utility maximizer, she submits a choice that maximizes her expected payoff given her beliefs about $\zeta_{(n)}$ and $\zeta_{(n+1)}$ (if buyer) or about $\zeta_{(n-1)}$ and $\zeta_{(n)}$ (if seller). Since bids and offers can be any non-negative number, the choice space is continuous and beliefs about other agents' choices are represented by a continuous probability distribution. A buyer's beliefs are represented by the joint density of $\zeta_{(n)}$ and $\zeta_{(n+1)}$, denoted $f_{(n),(n+1)}$, and a seller's beliefs are represented by the joint density of $\zeta_{(n-1)}$ and $\zeta_{(n)}$, denoted $f_{(n-1),(n)}$.

Eliciting beliefs in the form of $f_{(n),(n+1)}$ and $f_{(n-1),(n)}$ is likely to prove very cumbersome. First, assessment of the joint probability of events is usually not coherent.[12] Second, $f_{(n),(n+1)}$ and $f_{(n-1),(n)}$ are joint densities of order statistics, and the concept of order statistic is likely to produce confusion and misunderstanding. The approach explored in this paper circumvents these obstacles taking advantage of the fact that, as shown in Appendix 2, $f_{(n),(n+1)}$ and $f_{(n-1),(n)}$ are a function of the beliefs about other buyers' bids (represented by the probability distribution with density y_b and cdf Y_b) and the beliefs about other sellers' offers (represented by the probability distribution with density y_s and cdf Y_s). Thus, the experiment implements the elicitation of beliefs about other buyers' bids and about other sellers' offers, leaving the joint densities $f_{(n),(n+1)}$ and $f_{(n-1),(n)}$ to be computed according to the formulas in Appendix 2.[13]

In order to elicit the continuous distribution Y_b (Y_s), first partial information is obtained by eliciting a discrete distribution over several intervals in the support, and thus the value of Y_b (Y_s) at several points in the support. Then, a parametric

[10] For a discussion on the effects of the size of monetary incentives in belief elicitation see Armantier and Treich (2013).

[11] $\zeta_{(n)}$ is the n-th order statistic among the $2n-1$ choices submitted by the other agents.

[12] For example, the assessment can be affected by the so-called conjunction fallacy: the assessment of the joint probability is higher than the probabilities of the constituent events.

[13] It is worth pointing out that in games with incomplete information, such as the double auction, beliefs about other players' actions depend on beliefs about the distribution of players' preferences (i.e., types, or private valuations) and beliefs about other players' strategies.

distribution is fitted on the elicited data and a fitted continuous distribution $\hat{y}_b$ ($\hat{y}_s$) is obtained. Both steps are described below in detail.

The choice set \$ $[0, \infty)$ is divided into six intervals: \$[0,2], [2.01,4], [4.01,6], [6.01,8], [8.01,10] and [10.01,∞).[14] I place the cutoff at \$10 because realizations of private values and costs are below \$10. I choose a width equal to \$2 for the intervals and, therefore, a total number of six intervals to allow for a balance between the goal of eliciting multiple pieces of information regarding subjective beliefs and the goal of keeping the task manageable and fairly quick for the respondents. The respondents are required to assign to each interval $j = 1, 2, \ldots, 6$ a probability y_j that represents their beliefs that the variable of interest (bids or offers) falls within the interval.[15] The exact wording is reported in "Belief elicitation" section in Supplementary Material.

Belief elicitation is incentivized by means of a monetary reward, determined by a quadratic scoring rule.[16] If a subject reports beliefs about the choices of other N subjects by assigning probabilities $y = (y_1, \ldots, y_6)$ to the intervals $j = 1, 2, \ldots, 6$, and $(N_1, \ldots, N_6)$ are the actual numbers of choices falling within each interval, then the quadratic scoring rule determines that the reward is equal to

[14] The choice set is unbounded since a bidding choice can be any non-negative number, rounded up to the second decimal. Beliefs are a continuous probability distribution defined over a potentially unbounded support. Thus, respondents are not artificially led to believe that other subjects' choices are restricted to be within a certain upper bound.

[15] Previous 2-player games, in which Player 1 states probabilistic beliefs about randomly-matched Player 2's choice, use either a probability format (Nyarko and Schotter 2002; Costa-Gomes and Weizsäcker 2008; Rutström and Wilcox 2009) or a proportion format (Rey-Biel 2009; Blanco et al. 2010; Bellemare et al. 2011). The proportion format has been supported by evidence that people are better at working with natural frequencies than with percent probabilities (Hoffrage et al. 2000). In the proportion format, Player 1 states the 'number of Player 2s out of ...' that will choose each alternative choice. The base ('out of ...'), set by the experimenter, is often 100, which facilitates the correspondence between probabilities and proportions. In this paper, several characteristics of the auction experiment make the probability format more appropriate. The auction is a multi-player game with of 4, 6 or 8 participants, half buyers and half sellers. Thus, each subject faces 1, 2 or 3 other subjects with the same role and 2, 3 or 4 other subjects with the opposite role, and states separately beliefs about bids and beliefs about offers. On the one hand, a proportion format with base 100 could confuse subjects since the actual number of other buyers (sellers) is 1, 2, 3 or 4 (depending on market size and one's own role). On the other hand, a proportion format with base equal to the actual number of other buyers (sellers) has two limitations. First, in a markets with four participants the wording 'out of 1' would appear when a subject states beliefs about the choice of the only other subject with the same role. Second, different wording would appear across markets (since markets differ in size) and across questions (since the numbers of other buyers and other sellers differ). It is out of the scope of this paper to investigate the effects of differences in wording.

[16] The quadratic scoring rule is an incentive-compatible mechanism, which provides subjects with an incentive to report their beliefs truthfully, provided that subjects are risk neutral and do not distort probabilities. However, rewarding subjects for the accuracy of their beliefs and using the quadratic scoring rule to determine the reward may have several limitations. Rutström and Wilcox (2009) and Palfrey and Wang (2009) report evidence on the effects that the elicitation of beliefs may have on choice behavior, including the possibility of more strategic behavior, lower risk aversion and overconfidence. Blanco et al. (2010) investigate whether subjects employ the belief elicitation task in order to hedge against adverse outcomes in the choice task. They find no evidence of hedging being a major problem in belief elicitation. Armantier and Treich (2013) report that increasing the payment in a proper scoring rule induces more biases towards reporting uniform probabilities.

$\$2 - \frac{1}{N}\sum_{j=1}^{6} N_j\left[(y_j - 1)^2 + \sum_{h\neq j} y_h^2\right]$.[17] Subjects are rewarded separately for the accuracy of their beliefs about bids and for the accuracy of their beliefs about offers. For each task, the rule generates a reward bounded in \$[0,2].[18]

From knowledge of the probabilities $(y_1, \ldots, y_6)$ assigned to the intervals, the value of the cdf Y at the right endpoints of the intervals, $(r_1, \ldots, r_6) = (2, 4, 6, 8, 10, \infty)$, is easily computed.[19] Fitting a parametric distribution over the data $Y(r_1), \ldots, Y(r_6)$ allows to obtain the fitted density and cdf of other subjects' choices, $\hat{y}$ and $\hat{Y}$. This method can be used separately to obtain the fitted density and cdf of other subjects' bids ($\hat{y}_b$ and $\hat{Y}_b$), and the fitted density and cdf of other subjects' offers ($\hat{y}_s$ and $\hat{Y}_S$).

Depending on (a) the number of intervals where positive probability is placed, and (b) whether the intervals, where positive probability is placed, are adjacent to each other, I fit the data using a unimodal beta distribution, a triangular distribution, the union of two or three triangular distributions, or the union of a unimodal beta and a triangular distribution.[20] Appendix 3 contains an overview of the fitting methods, which build on procedures introduced by Engelberg et al. (2009). After obtaining $\hat{y}_b$, $\hat{Y}_b$, $\hat{y}_s$ and $\hat{Y}_s$, the joint densities $f_{(n),(n+1)}$ and $f_{(n-1),(n)}$ can be computed and used in turn to compute expected payoffs, as shown in Appendix 2.

2.2 Remarks

In this section I discuss the main features of the belief elicitation design and its relevance beyond the context of a double auction.

The design implements a probabilistic belief elicitation, in which respondents report their beliefs as a probabilistic (or distributional) forecast. An alternative approach is non-probabilistic belief elicitation, in which respondents report their beliefs as a non-probabilistic (or deterministic, or point) forecast.[21] Probabilistic

[17] It is worth pointing out that the rule is not necessarily incentive-compatible if the participant is rewarded for predicting the choice frequencies of a finite number of opponents. In such a setting, the rule is incentive-compatible if his beliefs corresponds to one of the outcomes that his opponents' choices could generate, but it is not incentive-compatible for all possible beliefs that he could hold. Consider a subject facing two other buyers and reporting probabilistic beliefs $(p_1, p_2, p_3, p_4, p_5, p_6)$ about bids. The rule is incentive-compatible only if his beliefs match one of the 21 possible vectors $(f_1, f_2, f_3, f_4, f_5, f_6)$ of empirical fractions where $f_j = \{0, \frac{1}{2}, \frac{2}{2}\}$ for $j = 1, 2, \ldots, 6$ and $\sum_{j=1}^{6} f_j = 1$. Costa-Gomes and Weizsäcker (2008) makes an analogous argument.

[18] Subjects are provided with a simplified explanation of the rule in the instructions. At the end of each round subjects receive feedback including a comparison between the probabilities assigned to each interval and the number of other bids and offers falling within each interval.

[19] $Y(r_1) = y_1$, $Y(r_2) = y_1 + y_2$, and so on.

[20] Approximately 90 % of the beliefs are fitted with a unimodal beta distribution.

[21] Several papers elicit traders' non-probabilistic beliefs about future prices in a double auction experiment (Smith et al. 1988; Sonnemans et al. 2004; Hommes et al. 2005; Haruvy et al. 2007). The definition of beliefs as market price forecasts is certainly appropriate in the investigation of the effect of past prices on traders' beliefs about future price movements, which is the objective of the above-mentioned papers. However, such a definition is less appropriate in the investigation of how beliefs affect strategic decision-making, which is the objective of this paper. The market price is determined jointly by

belief elicitation has two main advantages. First, there is evidence that agents, when reporting a non-probabilistic forecast, reveal different features of their subjective distribution (mean, median, or other quantiles), which makes it difficult for researchers to interpret point forecasts.[22] Second, probabilistic belief elicitation provides information on the uncertainty faced by an agent, which can be used to study choice under uncertainty where features of the subjective distribution play a role.[23] The use of non-probabilistic beliefs (either elicited as such or computed as a summary statistic of elicited probabilistic beliefs) would also limit the scope of the analysis of subjective beliefs, preventing us, for example, from investigating whether beliefs' accuracy is poor on specific portions of the support.

The belief elicitation design is relevant and applicable, beyond the context of a double auction, to settings that involve multiple players (possibly with different roles and types), a continuous choice set, and a payoff function that depends on a (possibly involved) function of players' choices.[24] Relevant settings may include market competition in oligopoly with incomplete information about production costs or bargaining with incomplete information about reservation values.[25] Such settings differ non-trivially from the one where belief elicitation is usually implemented: a 2-player matrix game in which Player 1 (2) assigns a discrete probability vector to Subject 2 (1)'s alternative choices. As a consequence, elicitation of probabilistic beliefs (possibly involving joint densities, order statistics, etc.) also differs non-trivially.

The approach involves three steps: *(a)* the characterization of subjective beliefs as (a function of) continuous probability distributions, *(b)* the elicitation of partial information about the continuous probability distributions via a discrete format, and *(c)* the parametric fitting of the data. Within this general approach, specific aspects may need to be modified when moving from an auction market to other games. Regarding *(a)*, while the approach suggests, independently of the setting, elicitation

Footnote 21 continued
the choice made by a subject holding beliefs and the choices made by other subjects and it is therefore the *result* of strategic interaction.

22 See Manski (2004) for a review, Engelberg et al. (2009) and Manski and Neri (2013) for a comparison of probabilistic and non-probabilistic elicitation.

23 In the double auction, as discussed above, the ordering of choices determines the auction outcome and the relevant beliefs about the ordering are represented by a joint density of two order statistics (n-th and $(n+1)$-th for a buyer, $(n-1)$-th and n-th for a seller). Since beliefs about the ordering of choices cannot be revealed by a best-point forecast of choice, the design implements the elicitation of a distributional forecast. An alternative, non-probabilistic approach, not explored in the paper, is to take the order statistics of choice as object of forecast, and elicit a best-point forecast of the order statistics.

24 In the double auction, for example, there are $N > 2$ players with different roles and types (buyers and sellers with independent private valuations), the choice set is $[0, \infty)$, and an individual's payoff depends on a function of other players' ordered choices.

25 For example, consider a Cournot market of incomplete information with each seller's constant marginal cost drawn from a commonly known distribution in $[c_L, c_H]$. Quantity-setting sellers act as strategic decision-makers who choose a non-negative level of production from a continuous and possibly unbounded choice set, best-responding to their beliefs about their rivals' production. Belief elicitation consists of dividing the choice set into J intervals and having each seller assign a vector $y = (y_1, \ldots, y_J)$ of subjective probabilities to the J intervals. Then the relation between observed production choices and subjective beliefs can be investigated.

of marginal densities of individual variables instead of joint densities of aggregate variables, the specification of the densities will depend on the setting. Regarding *(b)*, the intervals over which probabilities are assigned should cover the choice set without restricting it arbitrarily and allow for eliciting several points over the subjective cdf, while keeping the task simple. The number and width of intervals used in the paper were chosen to satisfy these requirements, but are not meant as a general specification.[26] For a discussion on *(c)*, see Appendix 3.

3 Results

3.1 Subjective beliefs, BNE beliefs and empirical beliefs

Table 7 reports the sample distribution of the elicited subjective probabilities y_j assigned to each interval $j = 1, \ldots, 6$. There is substantial heterogeneity across subjects both in the beliefs about bids and in the beliefs about offers: probabilities range from 0 to 100 % for most intervals.[27]

The experimental design restricts bidding choices to be non-negative but does not prevent subjects from submitting choices above the highest possible valuation nor from assigning positive probability to such choices. Empirically, while choices above \$10 are negligible, beliefs assigning a positive (despite small) probability to such choices are not. However, the percentage of subjects with such beliefs decreases sharply from the first to the last round (from 35 to 9 % for bids and from 36 to 14 % for offers) and the positive probabilities still assigned in the last round are mostly smaller than 10 %, suggesting that initially too high forecasts of the likelihood of such choices are corrected after repeated observation of their low empirical frequency. The accuracy of subjective beliefs is discussed in detail in Sect. 3.2.2.[28]

Before turning to investigate the role of subjective beliefs in explaining observed choice behavior, it is useful to consider what alternatives would be available if subjective beliefs were not elicited. If that were the case, choice behavior could be studied assuming that the beliefs that subjects hold can be proxied with BNE beliefs or with empirical beliefs.[29] Therefore, I now define BNE beliefs and empirical

[26] Future work should evaluate alternative methods to choose the intervals in order to provide more accurate fitting and/or require less elicitation questions.

[27] The sample distribution is reported for the first and the last round separately.

[28] There are only 4 observations out of 990 of a bid or offer above \$10: buyer 49 in round 12 bids \$10.60, seller 21 in round 9 offers \$11, seller 7 in round 2 offers \$15, seller 29 in round 2 offers \$18.50. Across all subjects, the probabilities assigned in round 1 to a bid above \$10 are 0 % (65 % of subjects), 1–10 % (29 % of subjects), 11–20 % (3 % of subjects), 21–30 % (3 % of subjects). In round 15, the probabilities are 0 % (91 % of subjects), 1–10 % (9 % of subjects). Across all subjects, the probabilities assigned in round 1 to an offer above \$10 are 0 % (64 % of subjects), 1–10 % (24 % of subjects), 11–20 % (5 % of subjects), 31–60 % (7 % of subjects). In round 15, the probabilities are 0 % (86 % of subjects), 1–10 % (8 % of subjects), 11–20 % (5 % of subjects), 21–30 % (1v of subjects).

[29] This paper focuses on two benchmarks: the beliefs learning model using empirical beliefs and the BNE. Differently than in beliefs learning, in reinforcement learning (Erev and Roth 1998) players learn by observing what was successful in the past and play more often strategies that paid off relatively well in the past. The experience-weighted attraction (EWA) learning model (Camerer and Ho 1999) is a hybrid

beliefs. For a review of the BNE bidding strategies in the double auction see Rustichini et al. (1994) and Appendix 4.

3.1.1 BNE beliefs

I denote with $\hat{v}_b$ ($\hat{v}_s$) the continuous probability distribution representing the beliefs about bids (offers) held by an agent in the BNE. In the BNE, beliefs are correct. Thus $\hat{v}_b$ and $\hat{v}_s$ correspond to the BNE bid and offer density, which can be computed given knowledge of the BNE strategies b_{BNE} and a_{BNE}. Thus, the bid density evaluated at a bid b is $f_b(v)/b'_{BNE}(v)$, where $b_{BNE}(v) = b$ and $f_b(v)$ is the density of private values at v. Analogously, the offer density evaluated at an offer a is $f_s(c)/a'_{BNE}(c)$, where $a_{BNE}(c) = a$ and $f_s(c)$ is the density of private costs at c.[30] Figure 1 illustrates BNE beliefs. Given $\hat{v}$, $v = (v_1, \ldots, v_6)$ is the vector of probabilities that in the BNE an agent assigns to the intervals $j = 1, \ldots, 6$.

3.1.2 Empirical beliefs

Following Cheung and Friedman (1997), subject i's γ-weighted empirical beliefs that another subject chooses a bid (offer) in interval j in round t+1 is defined as $\xi_{i,t+1,j} = \left(N_{t,j} + \sum_{u=1}^{t-1} \gamma_i^u N_{t-u,j}\right) / \left(N_t + \sum_{u=1}^{t-1} \gamma_i^u N_{t-u}\right)$, where $N_{t,j}$ is the number of other players who chose a bid (offer) in interval j in period t, γ_i^u is the weight given to the observation of a bid (offer) in interval j being chosen by another buyer (seller) in round $t - u$, and N_t is the total number of buyers (sellers) whose behavior is observed.[31] Then γ^*-weighted empirical beliefs is defined as $\eta_{i,t+1,j} = \left(N_{t,j} + \sum_{u=1}^{t-1} \gamma_i^{*u} N_{t-u,j}\right) / \left(N_t + \sum_{u=1}^{t-1} \gamma_i^{*u} N_{t-u}\right)$, i.e. where γ_i^* minimizes the absolute squared difference between subjective beliefs y and γ-weighted empirical beliefs ξ, $\sum_{t=1}^{15} \sum_{j=1}^{6} |y_{i,t,j} - \xi_{i,t,j}|^2$.[32] In what follows, I refer to γ^*-weighted empirical beliefs simply as empirical beliefs.

Footnote 29 continued
model that encompasses beliefs and reinforcement learning models as special cases. Nyarko and Schotter (2002) find that the belief learning model using elicited subjective beliefs outperforms both a reinforcement and EWA model. Alternatively to BNE, non-equilibrium models such as level-k and cognitive hierarchy (CH) could serve as benchmark. While the belief elicitation methodology described in the paper is readily applicable in the context of a comparison between elicited beliefs and level-k /CH beliefs, the feedback design should be modified, since level-k /CH models are most appropriate in one-shot games with no feedback regarding opponents' choices or one's own payoff, in order to minimize learning effects. While these extensions go beyond the scope of this paper, I consider them an interesting topic for further research.

[30] Given private values and costs distributed uniformly over $[0, 10]$, $f_b(v) = f_s(c) = 0.1$.

[31] Specific cases of empirical beliefs include Cournot beliefs, defined as opponents' last-period actions, and fictitious-play beliefs, defined as the average of opponents' past actions. They are obtained by setting $\gamma = 0$ or $\gamma = 1$, respectively. See also Fudenberg and Levine (1998). Note that, since in the experiment each subject changes role in round 6 and 11, N_t is not constant.

[32] The median estimated γ^* is 1.02 for beliefs about bids and 0.99 for beliefs about offers. For round 1, I assume that γ-weighted empirical beliefs prescribe $\xi_{i,1,j} = 1/6$ for $j = 1, \ldots, 6$. For previous work using γ^*-weighted empirical beliefs see Nyarko and Schotter (2002).

Can subjective beliefs be modeled by BNE or empirical beliefs? In order to answer this question I compute for each participant i and for each round t the average absolute differences (AAD) between the subjective beliefs $y_{i,t}$ and the empirical beliefs $\eta_{i,t}$, and between the subjective beliefs $y_{i,t}$ and the BNE beliefs $v_{i,t}$, which I define respectively $AAD_{i,t}(y,\eta) = \frac{1}{6}\sum_{j=1}^{6} \left| y_{i,t,j} - \eta_{i,t,j} \right|$ and $AAD_{i,t}(y,v) = \frac{1}{6}\sum_{j=1}^{6} \left| y_{i,t,j} - v_{i,t,j} \right|$. In what follows I restrict attention to beliefs about bids.[33]

Table 1 reports summary statistics of $AAD(y,\eta)$ and $AAD(y,v)$ in round 5, 10 and 15. Subjective beliefs differ from empirical and BNE beliefs. The median of $AAD(y,\eta)$ ranges between 12 and 8 % age points and the median of $AAD(y,v)$ is 8 % points. $AAD(y,\eta)$ decreases across rounds while $AAD(y,v)$ displays no significant differences across rounds.[34] In round 5 and 10 $AAD(y,v)$ is significantly smaller than $AAD(y,\eta)$, indicating that subjective beliefs are closer to BNE beliefs than to empirical beliefs. The difference is not significant in round 15.[35] I draw two conclusions. First, neither BNE nor empirical beliefs provide an appropriate proxy for subjective beliefs, since the differences between beliefs are not zero. Second, if we wish to use a proxy for subjective beliefs, empirical beliefs are a poorer proxy than BNE beliefs in earlier rounds, and become as good as BNE beliefs in later rounds, as information about more previous rounds becomes available.[36]

BNE or empirical beliefs, despite differing from subjective beliefs, may still prescribe the same best response as subjective beliefs. If that were the case, eliciting subjective beliefs to determine which best response they prescribe, in order to later investigate choice behavior, could be avoided and a simple computation of BNE or empirical beliefs would suffice.

Table 2 shows how best responding to subjective beliefs differ from best responding to BNE or empirical beliefs, by reporting summary statistics of the absolute difference between best responses in round 5, 10 and 15.[37,38] In round 15

[33] Tables 8 and 9 report results separately on beliefs about bids and beliefs about offers.

[34] Defining as pairs of comparison round 5 vs. round 10 and round 5 vs. round 15, Wilcoxon signed rank test finds no significant differences in $AAD(y,v)$ (round 5 vs. 10 $p = 0.1770$, round 5 vs. 15 $p = 4188$). Wilcoxon signed rank test finds significant differences in $AAD(y,\eta)$ (round 5 vs. 10 $p = 0.0000$, round 5 vs. 15 $p = 0000$).

[35] One-side Wilcoxon signed rank test against no difference between $AAD(y,\eta)$ and $AAD(y,v)$ within the same round. In round 5 $p<0.0000$, round 10 $p = 0.0093$, round 15 $p = 0.1339$.

[36] Additional analysis is reported in Tables 8 and 9 replicate Table 1 by market size. Tables 10 and 11 show that differences between subjective and BNE beliefs, and between subjective and empirical beliefs do not vary with private valuations.

[37] Computing the best responses prescribed by beliefs requires using beliefs $\hat{y}$, $\hat{v}$, $\hat{\eta}$. $\hat{\eta}$ is the continuous distribution obtained applying to the discrete distribution η the fitting methods described in Sect. 2 (already employed to define $\hat{y}$). First the fitted density and cdf of other subjects' bids, $\hat{\eta}_b$ and $\hat{H}_b$, and the fitted density and cdf of other subjects' offers, $\hat{\eta}_s$ and $\hat{H}_S$, are obtained. Then $\hat{\eta}_b$, $\hat{H}_b$, $\hat{\eta}_s$ and $\hat{H}_s$ are used to compute the fitted joint densities of the relevant order statistics.

[38] It is worth pointing out that while there exists a unique BNE best response (as discussed in Appendix 4 and in Cason and Friedman (1997)), there may be multiple best responses to subjective beliefs and multiple best responses to empirical beliefs. In the case of a range of best responses, the midpoint of the range is used for comparison in Table 2. Best responses are expressed up to cents of a dollar, so as are bids and offers in the experimental design.

Table 1 Average absolute difference between subjective and BNE beliefs and between subjective and empirical beliefs

	ADD subj. vs. BNE beliefs			ADD subj. vs. empirical beliefs		
	Round 5	Round 10	Round 15	Round 5	Round 10	Round 15
Mean	0.09	0.08	0.09	0.13	0.09	0.09
SD	0.05	0.05	0.06	0.04	0.04	0.05
Min	0.02	0.00	0.01	0.05	0.04	0.02
Q1	0.05	0.05	0.05	0.10	0.07	0.05
Median	0.08	0.08	0.08	0.12	0.09	0.08
Q3	0.10	0.11	0.12	0.15	0.11	0.12
Max	0.25	0.26	0.26	0.25	0.21	0.23

Beliefs about bids

Table 2 Absolute difference (in dollars) between the best response to subjective beliefs and the best response to BNE beliefs (column 1–3) and absolute difference between the best response to subjective beliefs and the best response to empirical beliefs (column 4–6)

	‖BR to subj. beliefs − BR to BNE beliefs‖			‖BR to subj. beliefs - BR to emp. beliefs‖		
	Round 5	Round 10	Round 15	Round 5	Round 10	Round 15
Mean	0.25	0.31	0.35	0.53	0.33	0.40
SD	0.31	0.62	0.52	0.71	0.59	0.56
Min	0.00	0.00	0.00	0.00	0.00	0.00
Q1	0.06	0.07	0.06	0.08	0.06	0.06
Median	0.15	0.15	0.12	0.23	0.16	0.15
Q3	0.33	0.26	0.41	0.77	0.30	0.53
Max	1.64	3.81	3.03	3.24	3.75	2.92

Summary statistics for round 5, 10 and 15

the absolute difference between best response to subjective beliefs and best response to BNE beliefs ranges between \$0 and \$3.03 (mean \$0.35 and median \$0.12). Analogously, the absolute difference between best response to subjective beliefs and best response to empirical beliefs ranges between \$0 and \$2.92 (mean \$0.40 and median \$0.15).[39]

[39] Moreover, the best responses prescribed by subjective, BNE or empirical beliefs may generate different expected payoffs. Table 12 reports summary statistics of the absolute difference in expected payoffs between best responses for round 5, 10 and 15. In round 15 the absolute difference in expected payoff between best response to subjective beliefs and best response to BNE beliefs ranges between \$0 and \$2.30 (mean \$0.44 and median \$0.12). Analogously, the absolute difference in expected payoff between best response to subjective beliefs and best response to empirical beliefs ranges between \$0 and \$2.38 (mean \$0.41 and median \$0.16).

3.2 Subjective beliefs and choice

3.2.1 Is choice consistent with the best response to subjective beliefs?

Let us now turn to investigate the consistency between observed bidding choices and elicited subjective beliefs by analyzing whether choices are consistent with the best response to subjective beliefs. Best responding to beliefs is defined as making a choice that gives the highest expected payoff given the beliefs. Denoting the best response to subjective beliefs as $b_{BR,\hat{y}}$, choice b is consistent with (the best response to) subjective beliefs if $b \in b_{BR,\hat{y}}$.[40]

Since BNE and empirical beliefs are alternative proxies for subjective beliefs, the extent to which subjective beliefs help explain choices can be evaluated comparing the consistency between choice and subjective beliefs to the consistency between choice and BNE or empirical beliefs. Denoting the best response to empirical beliefs as $b_{BR,\hat{\eta}}$ and the best response to BNE beliefs as $b_{BR,\hat{v}}$, choice b is consistent with (the best response to) empirical beliefs if $b \in b_{BR,\hat{\eta}}$ and consistent with (the best response to) BNE beliefs if $b \in b_{BR,\hat{v}}$.[41]

Since assessing consistency involves computing expected payoff, it is useful pointing out that expected payoff is more sensitive to bidding choice for mid-range private valuations than for extreme ones. On the one hand, buyers with a low value and sellers with a high cost face a very small probability of trading and, provided they bid at or below value or offer at or above cost, the largest payoff they can expect is most likely approximately zero. On the other hand, buyers with a high value and sellers with a low cost are likely to trade at a price determined by the choices of other subjects. Therefore, evidence that choice is consistent with best response to subjective beliefs provides stronger support that subjective beliefs help explain choice if found when expected payoff is more sensitive to choice.

Table 3 reports the percentage of observations for which observed choice is consistent with the best response to subjective, empirical, or BNE beliefs. The three cases are not mutually exclusive, since choice may be consistent with the best response to only one, only two, or all three types of beliefs, nor jointly exhaustive, since choice may not be consistent with the best response to any of the beliefs.[42] Results are reported separately over the entire sample, over the subsample with mid-range valuations ($[3,7]), and over the subsamples with extreme valuations.[43]

[40] Given subjective beliefs $\hat{y}$, a participant expects to receive payoff $\pi(b|\hat{y})$ by choosing b. If $b \in b_{BR,\hat{y}}$, then b and $b_{BR,\hat{y}}$ generate, given subjective beliefs $\hat{y}$, equal expected payoffs $\pi(b|\hat{y}) = \pi(b_{BR,\hat{y}}|\hat{y})$.

[41] Given empirical beliefs $\hat{\eta}$, a participant expects to receive payoff $\pi(b|\hat{\eta})$ by choosing b. Given BNE beliefs $\hat{v}$, a participant expects to receive payoff $\pi(b|\hat{v})$ by choosing b. If $b \in b_{BR,\hat{\eta}}$ then $\pi(b|\hat{\eta}) = \pi(b_{BR,\hat{\eta}}|\hat{\eta})$, and if $b \in b_{BR,\hat{v}}$ then $\pi(b|\hat{v}) = \pi(b_{BR,\hat{v}}|\hat{v})$.

[42] Choice is not consistent with the best response to any belief in 53 % of the observations.

[43] In a setting in which valuations are drawn independently from a uniform distribution over $[0,10], the range $[3,7] exclude those valuations for which the BNE model predicts that buyers with a low value and sellers with a high cost will have a zero probability of trading. See Rustichini et al. (1994).

Table 3 Percentage of observations for which choice is consistent with best response to subjective beliefs, empirical beliefs, or BNE beliefs

Obs.	No.	Choice consistent with best response to		
		Subjective beliefs	Empirical beliefs	BNE beliefs
All	990	36 (12.6)	32 (10)	34 (9)
Value or cost in [3,7]	372	20.2 (9.4)	13.7 (5.7)	14.8 (4)
Value <3 or cost >7	315	69.5 (18.5)	69.5 (17.7)	71 (16.2)
Value >7 or cost <3	303	19.5 (10.4)	16.2 (7.3)	18.5 (7.5)

The percentages, which would be generated assuming that subjects choose randomly, are reported in parenthesis

Choice is consistent with the best response to subjective, empirical, and BNE beliefs in 36, 32 and 34 % of the observations, respectively.[44] The same results could not be obtained under the assumption that subjects make random choices, since the probability that a choice consistent with the best response to subjective, empirical or BNE beliefs would be picked by chance is 12.6, 10 or 9 %, respectively.[45] Restricting to the subsample with mid-range valuations, choice is consistent with the best response to subjective, empirical or BNE beliefs in 20.2, 13.7 and 14.8 % of the observations, respectively.[46] While these percentages are smaller compared to the ones computed over the entire sample, subjective beliefs perform relatively better than empirical and BNE beliefs in explaining observed choice over the subsample for which the comparison is most relevant.[47]

[44] Table 13 reports the results across market-size treatments and separately for first and last round. Within each market-size treatment, there are no significant differences in the consistency of choice with subjective beliefs between first and last round (Wilcoxon signed rank test against no difference in consistency across pairs: $p = 0.999, p = 0.999, p = 0.8238$, for market-size 4, 6, 8 respectively). Across market-size treatments, the hypothesis that the rate of consistency of choice with subjective beliefs is the same for market size 4, 6 and 8 cannot be rejected either for round 1 (two-sided Fisher exact test, $p = 0.052$) or for round 15 ($p = 0.254$).

[45] The probability is defined as the ratio between the number of choices that generate an expected payoff equal to the one generated by the best response and the number of all possible alternatives from which a subject can randomly choose. Since in the experiment choices and payoffs are defined to the nearest cent, numbers are expressed in cents when computing the probability. The possible alternatives from which to randomly choose are restricted to the values between $0 and $10.

[46] Choice is not consistent with the best response to any belief in 69 % of the observations.

[47] How do the subsamples with extreme valuations compare with the one with mid-range valuations? Among observations with a low value or a high cost, consistency of choice with the best response to subjective, empirical or BNE beliefs is 69.5, 69.5 and 71 %, respectively. Within this subsample, consistency artificially increases due to the extremely low sensitivity of expected payoff to choice. Among observations with a high value or a low cost, consistency of choice with the best response to subjective, empirical or BNE beliefs is 19.5, 16.2 and 18.5 %, respectively. For both subsamples with extreme valuations there is no evidence of a better performance of subjective beliefs, compared to empirical or BNE beliefs, in explaining observed choice.

Table 4 Relative frequency of choice and/or best response to subjective beliefs consistent with the best response to BNE beliefs

Choice and best response to BNE beliefs	Best response to subjective beliefs and best response to BNE beliefs				
	Consistent		Not consistent		All
Consistent	286	0.85	49	0.15	335
	0.44		0.14		
Not consistent	363	0.55	292	0.45	655
	0.56		0.86		
All	649		341		

The discussed evidence highlight how subjective beliefs differ from BNE beliefs (Sect. 3.1) and how choice differ from best response to BNE beliefs (Sect. 3.2.1). Since the definition of the BNE requires convergence of both beliefs and choices to the BNE, it is not surprising that there is no evidence of convergence.[48] Thus, it is not possible to perform a comparison between convergent and non-convergent cases, which could provide insight into the conditions leading to convergence itself. However, a simple inspection of the frequency with which choice is consistent with the BNE best response and the best response to subjective beliefs is consistent with the BNE best response, can provide an insight into assessing the sufficiency and/or necessity of the two conditions for equilibrium: convergence in choices and convergence in beliefs.

Table 4 shows that the frequency with which choice is consistent with the BNE best response increases from 14 to 44 % when the best response to subjective beliefs becomes consistent with the BNE best response. The frequency with which the best response to subjective beliefs is consistent with the BNE best response increases from 55 to 85 % when choice becomes consistent with the BNE best response. The evidence suggests that (i) convergence in beliefs is necessary but not sufficient for convergence in choice, and that (ii) convergence in choice is sufficient but not necessary for convergence in beliefs.

3.2.2 How accurate are subjective beliefs?

In the experiment the accuracy of subjective beliefs in predicting other subjects' choices is measured and rewarded according to a quadratic scoring rule.[49] The

[48] Convergence in choices to the BNE requires choices to be consistent with the best response prescribed by BNE beliefs from a certain point in time onwards without deviation. Similarly, convergence in beliefs to the BNE requires subjective beliefs to prescribe a best response consistent with the best response prescribed by BNE beliefs from a certain point in time onwards without deviation.

[49] Recall from Sect. 2 that subjects are rewarded based on the comparison of their subjective beliefs to the bidding behavior of other participants in the same auction market.

Table 5 Accuracy of beliefs. Medians in round 1 and round 15. Comparison of subjective beliefs, BNE beliefs, and empirical beliefs according to different measures of accuracy: quadratic score, linear score and *AAD*)

	Subjective beliefs	BNE beliefs	Empirical beliefs	Perfect foresight
Round 1				
Quadratic score ($)	2.33	2.43*	2.33	3.12***
Linear score ($)	0.82	0.85	0.67†††	2.24***
AAD	0.39	0.36	0.44†††	0***
Round 15				
Quadratic score ($)	2.30	2.40*	2.40***	3.06***
Linear score ($)	0.80	0.83	0.86**	2.12***
AAD	0.40	0.41*	0.39*	0***

As a benchmark, the accuracy of perfect foresight is the one that would be achieved by correctly forecasting the actual behavior of other market participants. One-side Wilcoxon signed rank test against no difference in accuracy across pairs of beliefs. Pairs are defined as subjective beliefs versus BNE beliefs, empirical beliefs, or perfect foresight

† † †, ††, † Denotes $p<0.001$, $p<0.01$, $p<0.05$ when the alternative hypothesis is that the accuracy of subjective beliefs is the highest in the pair

***, **, * Denotes $p<0.001$, $p<0.01$, $p<0.05$ when the alternative hypothesis is that the accuracy of subjective beliefs is the lowest in the pair

higher the accuracy, the higher the quadratic score. The same criterion can be used to compare the relative accuracy of subjective, BNE and empirical beliefs. Table 5 reports the median quadratic score for each type of beliefs in the first and last round.[50] Subjective beliefs are significantly less accurate than BNE beliefs both in round 1 ($2.33 vs. $2.43) and in round 15 ($2.30 vs. $2.40). Instead, subjective beliefs are significantly less accurate than empirical beliefs in round 15 ($2.30 vs. $2.40) but not in round 1 ($2.33 vs. $2.33).[51]

The assessment could depend on the use of the quadratic score. As alternative measures of accuracy, Table 5 reports also the linear score and the average absolute difference *AAD*. The higher the accuracy, the higher the linear score and the lower the *AAD*.[52] According to both the linear score and the *AAD*, subjective beliefs are significantly more accurate than empirical beliefs in round 1, but less accurate in round 15. According to the *AAD*, subjective beliefs are significantly less accurate

[50] In the experiment subjects report separately their beliefs about the bids chosen by buyers and their beliefs about the offers chosen by sellers. Therefore, the accuracy of beliefs about bids and the accuracy of beliefs about offers can be analyzed separately or jointly. In the analysis that follows, I consider a joint measure of accuracy. Each measure in Table 5 is the sum of the measure for beliefs about bids and the measure for beliefs about offers. Table 14 is analogous to Table 5 but reports results separately by market size.

[51] Analogous evidence of subjective beliefs' ability to explain observed choices despite low predictive accuracy is reported by Nyarko and Schotter (2002), who elicit beliefs in a 2x2 game and compare subjective beliefs with empirical beliefs.

[52] Linear score is defined as $\$2 - \sum_{j=1}^{6} n_j \left[(1-y_j) + \sum_{h\neq j} y_h\right]$ and *AAD* is defined as $\frac{1}{6}\sum_{j=1}^{6} |y_j - n_j|$, where y_j is the probability assigned to an opponent submitting a choice within interval j, and n_j is the fraction of opponents submitting a choice within interval j.

than BNE beliefs in round 15. Therefore, evidence suggests that subjective beliefs are less accurate than BNE, while they are first more and later less accurate than empirical beliefs. Finally, subjective beliefs do not appear to become more accurate in later rounds: the median accuracy in round 15 is not significantly different than in round 1, irrespective of using the quadratic score, the linear score or the *AAD*.[53]

To interpret the extent to which beliefs differ in accuracy, the last column of Table 5 reports as benchmark the maximum score (quadratic and linear) and the minimum *AAD* that would be achieved by perfect foresight, i.e. by correctly forecasting the actual behavior of other market participants. The accuracy achieved by perfect foresight is markedly higher than the accuracy achieved by any of the beliefs. For example, according to the quadratic score in round 1, while the accuracy of BNE beliefs is 4 % higher than the one of subjective beliefs, the accuracy of perfect foresight is 34 % higher.

In the above discussion accuracy is measured by comparing each forecast of an event with its corresponding observed realization, i.e. comparing the probabilities assigned to each interval of choices with the observed realizations of choices falling in each interval. An alternative measure of accuracy is calibration (Seidenfeld 1985), which instead measures the degree of correspondence between the probabilities assigned to the choices and the empirical frequency of those choices, pooling all instances when those probabilities were stated.[54]

Table 15 reports the empirical frequency of choices falling in each interval, by subjective beliefs. Among all stated subjective beliefs, only those with more than 10 observations are displayed. Calibration is good over the upper-extreme interval $[10.01,\infty)$, poor over the intervals \$[0,2] and \$[8.01,10], and better over the mid-intervals.[55]

Individual calibration scores can be computed taking participants as unit of observation, instead of pooling data. The calibration score is measured by the weighted average of the squared difference between the vector of subjective probabilities and the vector of empirical choice frequencies, for all instances in which those subjective probabilities were stated.[56] The smaller the calibration

[53] Two-side Wilcoxon signed rank test cannot reject the null hypothesis of no difference in accuracy between round 1 and 15 (quadratic rule $p = 0.8043$, linear rule $p = 0.7982$, *AAD* $p = 0.6147$).

[54] By Seidenfeld's definition, a set of probabilistic predictions are calibrated if p percent of all predictions reported at probability p are true. In the experiment participants forecast the fraction of choices falling in each interval \$[0,2], [2.01,4], [4.01,6], [6.01,8], [8.01,10] and $[10.01,\infty)$. Beliefs are perfectly calibrated if for all the instances when the probabilities $(0.25, 0.5, 0.25, 0, 0, 0)$ are assigned, choices fall 25 % of the time in \$[0,2], 50 % of the time in \$[2.01,4], and 25 % of the time in \$[4.01,6]; for all the instances when the probabilities $(0.4, 0.4, 0.2, 0, 0, 0)$ are assigned, choices fall 40 % of the time in \$[0,2], 40 % of the time in \$[2.01,4], and 20 % of the time in \$[4.01,6]; and so on. Calibration adopts a frequentist approach to measure forecast accuracy, while proper scoring rules adopt a Bayesian approach.

[55] The most-often stated beliefs (0.2,0.2,0.2,0.2,0.2,0) are the best calibrated.

[56] Denote with $y_t = (y_1, \ldots, y_6)$ the vector of probabilities that subject i assigns in round t to choices falling in the intervals $(1, \ldots, 6)$. Subject i submits T vectors, one in each round. Of the T vectors, there are K distinct ones, identified as $y_k = (y_{k1}, \ldots, y_{k6})$ with $k = 1, .., K$. Each distinct y_k occurs in m_k instances. Denote with $n_k = (n_{k1}, \ldots, n_{k6})$ the relative frequency of choices falling in the intervals in the m_k instances when y_k is submitted. Weighting the squared difference between y_k and n_k by m_k, the Brier-

Table 6 Calibration scores

Object of forecast	Role of forecaster	No.	Calibration score						
			Mean	SD	Min	Q1	Median	Q3	Max
Bids	Buyer	66	0.39	0.24	0.02	0.23	0.33	0.52	1.04
Offers	Buyer	66	0.27	0.17	0.00	0.16	0.25	0.40	0.72
Bids	Seller	66	0.25	0.15	0.00	0.14	0.26	0.34	0.63
Offers	Seller	66	0.38	0.24	0.01	0.20	0.36	0.53	0.99

Summary statistics across participants

score is, the better the calibration. Table 6 reports summary statistics of the calibration score. Calibration is relatively poor compared to that found by Camerer et al. (2002) but better than that found by Costa-Gomes and Weizsäcker (2008).[57] Table 6 distinguishes object of forecast (bids or offers) and role of forecaster (buyer or seller). Calibration is poorer when the forecaster's role coincides with the role of those subjects whose choices are forecasted. A participant playing as buyer (seller) scores better at forecasting offers (bids). Analogously, a participant forecasting bids (offers) scores better when playing as seller (buyer).[58] The causes of this asymmetry, on which I do not conjecture at this stage, are an interesting question for further research .

Finally, the results help reconcile evidence in Armantier and Treich (2009) and Kirchkamp and Reiß (2011). Within a first-price independent-private-values auction experiment, Armantier and Treich (2009) find that probabilistic beliefs of winning the auction for a given list of bids are inaccurate, while Kirchkamp and Reiß (2011)

Footnote 56 continued
calibration score is $\frac{1}{T}\sum_{k=1}^{K} m_k(y_k - n_k)(y_k - n_k)'$. It ranges in [0,2], 0 corresponding to perfect calibration.

[57] In a set of one-shot games with no feedback, Costa-Gomes and Weizsäcker (2008) find evidence of poor calibration (calibration scores of 0.432 and 0.374 for Rows and Columns players, respectively). In repeated games, Camerer et al. (2002) find that beliefs are very well calibrated (calibration scores of 0.006 and 0.011 for Borrower and Lender players, respectively). Camerer et al. (2002) find that Borrower beliefs and Lender beliefs are similarly accurate at forecasting the behavior of Borrowers.

[58] Participants play several rounds as buyer and several rounds as seller. In each round, independently of their role, participants state their beliefs about bids and their beliefs about offers. This allows us to make two pair-wise comparisons for each individual, either keeping the role fixed (Does an individual playing as buyer/seller score better at forecasting bids or offers?) or keeping the object of forecast fixed (Does an individual forecasting bids/offers score better when playing as buyer or seller?). I use a one-side Wilcoxon signed rank test against no difference in calibration score across pairs of observations. A participant forecasting bids scores better when playing as seller than buyer ($p < 0.0000$). A participant forecasting offers scores better when playing as buyer than seller ($p = 0.0005$). A participant playing as buyer scores better at forecasting offers than bids ($p < 0.000$). A participant playing as seller scores better at forecasting bids than offers ($p < 0.000$).

find that non-probabilistic beliefs about the opponent's bidding strategy are fairly accurate. The belief of winning the auction for some bid depends on (i) the beliefs about others' *bidding strategies*, (ii) the beliefs about others' *private valuations*, (iii) how (i)–(ii) combine into the beliefs about others' *bids*, and (iv) how the beliefs about others' bids combine into the belief of winning.[59] Despite the different auction format, by eliciting probabilistic beliefs about others' bidding choices, I show that inaccuracies arise as soon as in step (iii).[60]

4 Conclusion

This paper proposes a methodology to implement probabilistic belief elicitation in continuous-choice games. Representing beliefs as a continuous subjective probability distribution, the methodology involves eliciting partial information about the subjective distribution and fitting a parametric distribution on the elicited data.

As an illustration, the methodology is applied to a double auction experiment, where traders' beliefs about the bidding choices of other market participants are elicited. Elicited subjective beliefs are found to differ from proxies such as BNE beliefs and empirical beliefs, both in terms of the forecasts of other traders' bidding choices and in terms of the best-response bidding choices prescribed by beliefs. Moreover, elicited subjective beliefs are found to explain observed bidding choices better than BNE beliefs and empirical beliefs.

By extending probabilistic belief elicitation to continuous-choice games, the proposed methodology enables to investigate the role of beliefs in a wide range of applications. Thus, the use of beliefs data can be applied to investigations in which traditionally only choice data are studied.

Acknowledgments I had the opportunity to present earlier versions of this work at the 2010 International Meeting of the Economic Science Association, and at Northwestern University, University of St.Gallen, University of Zurich, LUISS, Universidade Nova de Lisboa, CSEF, Bank of Italy. I am grateful to seminar participants and to Charles Manski, Marco Ottaviani, Brian Rogers, Charles Bellemare, and Jacob Goeree for their comments. A previous version of this work circulated as 'Strategic Thinking and Subjective Expectations in a Double Auction Experiment'.

[59] Kirchkamp and Reiß (2011) argue that 'one reason for erroneous best replies is connected to the handling of probabilities, particularly when transforming the expected bidding strategies used by others (together with the underlying distribution of valuations) into the probability distribution of winning bids'.

[60] In Armantier and Treich (2009) and Kirchkamp and Reiß (2011) bidding strategies are recorded by means of the strategy method.

Appendix 1: Figures and Tables

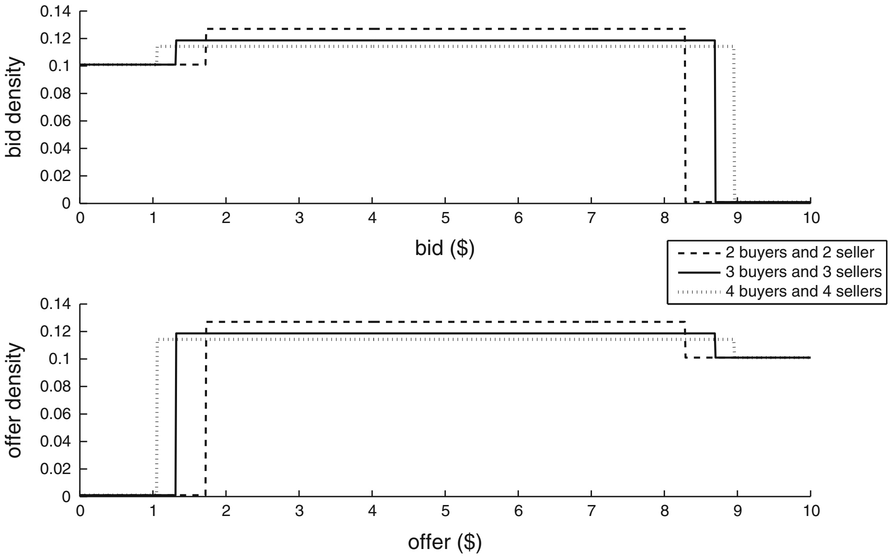

Fig. 1 BNE beliefs: bid and offer density functions. According to the BNE strategies, the probability of trading is zero for buyers with a very low value and sellers with a very high cost. Therefore, their misrepresentation of private values and costs cannot be bounded. In a market with four buyers and four sellers, buyers with value lower than \$1.05 and sellers with cost higher than \$8.95 can never trade. In a market with three buyers and three sellers, buyers with value lower than \$1.31 and sellers with cost higher than \$8.69 can never trade. In a market with two buyers and two sellers, buyers with value lower than \$1.72 and sellers with cost higher than \$8.28 can never trade

Table 7 Sample distribution (in percent) of the subjective probabilities assigned to each interval

Subj. prob. %	Beliefs about other buyers' choices						Beliefs about other sellers' choices					
	Intervals						Intervals					
	1st $[0,2]$	2nd $(2,4]$	3rd $(4,6]$	4th $(6,8]$	5th $(8,10]$	6th $(10,\infty)$	1st $[0,2]$	2nd $(2,4]$	3rd $(4,6]$	4th $(6,8]$	5th $(8,10]$	6th $(10,\infty)$
Beliefs in the first round ($t = 1$)												
0	30.3	7.58	4.55	9.09	18.18	65.15	24.24	12.12	7.58	10.61	22.73	63.64
1	1.52					1.52	1.52					
2	1.52	1.52			1.52			1.52	1.52		1.52	1.52
3				1.52			1.52					
5	7.58		1.52	3.03	15.15	19.7	12.12			6.06	9.09	15.15
6										1.52		
8		1.52										
10	18.18	4.55	4.55	13.64	25.76	7.58	19.7	12.12	9.09	10.61	21.21	7.58
15	1.52	9.09	7.58	3.03	3.03		1.52	6.06	7.58	1.52	6.06	
16	1.52				1.52	1.52	1.52					1.52
17		3.03	1.52	1.52	1.52			3.03	1.52	1.52	1.52	
20	18.18	42.42	24.24	31.82	24.24	1.52	22.73	24.24	28.79	36.36	25.76	3.03
23			1.52									
25	1.52	10.61	13.64	10.61	3.03	1.52	1.52	10.61	9.09	10.61	1.52	
30	1.52	7.58	18.18	13.64	4.55	1.52	1.52	10.61	15.15	10.61	1.52	
33										1.52	1.52	
34												1.52
35	3.03	1.52	1.52	1.52			1.52		4.55	1.52	1.52	
40	1.52	7.58	6.06	6.06			3.03	15.15	6.06	1.52	1.52	1.52
45	3.03		3.03				1.52		1.52			

Table 7 continued

Subj. prob. %	Beliefs about other buyers' choices						Beliefs about other sellers' choices					
	Intervals						Intervals					
	1st $[0,2]$	2nd $(2,4]$	3rd $(4,6]$	4th $(6,8]$	5th $(8,10]$	6th $(10,\infty)$	1st $[0,2]$	2nd $(2,4]$	3rd $(4,6]$	4th $(6,8]$	5th $(8,10]$	6th $(10,\infty)$
50	3.03	3.03	10.61	3.03	1.52		4.55	4.55	7.58	4.55	3.03	3.03
60	3.03			1.52			1.52					
80	1.52		1.52							1.52		1.52
100	1.52										1.52	
Beliefs in the last round ($t = 15$)												
0	31.82	15.15	3.03	15.15	36.36	90.91	27.27	10.61	4.55	21.21	39.39	86.36
1												
2					1.52		1.52					
3								1.52				
5	4.55	1.52			3.03	3.03	4.55	1.52		1.52	3.03	3.03
6												
8	1.52											
10	18.18	4.55	4.55	6.06	13.64	6.06	13.64	4.55	6.06	7.58	16.67	4.55
12					3.03		1.52					
15	4.55	4.55		1.52	3.03		3.03	4.55	4.55	3.03	7.58	
16												
17								1.52				
20	22.73	30.3	27.27	40.91	30.3		25.76	27.27	25.76	30.3	25.76	4.55
22	1.52	1.52	3.03	1.52				1.52	1.52	1.52	1.52	
23												
24			1.52	1.52								

Table 7 continued

Subj. prob. %	Beliefs about other buyers' choices						Beliefs about other sellers' choices					
	Intervals						Intervals					
	1st $[0, 2]$	2nd $(2, 4]$	3rd $(4, 6]$	4th $(6, 8]$	5th $(8, 10]$	6th $(10, \infty)$	1st $[0, 2]$	2nd $(2, 4]$	3rd $(4, 6]$	4th $(6, 8]$	5th $(8, 10]$	6th $(10, \infty)$
25	4.55	15.15	15.15	10.61	4.55		6.06	13.64	13.64	13.64	3.03	1.52
28		1.52										
30	4.55	13.64	15.15	12.12			6.06	18.18	18.18	12.12	1.52	
33							1.52		1.52	1.52		
34								1.52				
35			4.55						1.52			
36								1.52				
40	3.03	6.06	9.09	6.06			3.03	3.03	7.58	3.03		
45												
50	3.03	3.03	9.09	4.55	3.03		6.06	7.58	9.09	1.52	1.52	
55			1.52									
60		1.52	1.52		1.52			1.52		1.52		
64									1.52			
80									1.52			
100		1.52	4.55						3.03	1.52		

Table 8 Average absolute difference between subjective and BNE beliefs

Round	No.	Beliefs about bids						Beliefs about offers					
		Mean	Min	Q1	Q2	Q3	Max	Mean	Min	Q1	Q2	Q3	Max
Market size 4													
5	16	0.09	0.05	0.07	0.08	0.10	0.25	0.10	0.01	0.07	0.09	0.12	0.25
10	16	0.09	0.00	0.05	0.08	0.10	0.25	0.10	0.01	0.05	0.07	0.12	0.25
15	16	0.11	0.05	0.07	0.09	0.12	0.25	0.09	0.01	0.05	0.07	0.12	0.25
Market size 6													
5	18	0.08	0.04	0.06	0.08	0.10	0.18	0.09	0.04	0.05	0.07	0.11	0.22
10	18	0.08	0.02	0.05	0.07	0.11	0.18	0.10	0.03	0.05	0.08	0.15	0.23
15	18	0.09	0.02	0.04	0.08	0.14	0.18	0.11	0.04	0.05	0.11	0.15	0.23
Market size 8													
5	32	0.09	0.02	0.04	0.08	0.12	0.19	0.08	0.02	0.05	0.07	0.12	0.19
10	32	0.08	0.03	0.04	0.08	0.11	0.26	0.09	0.03	0.04	0.08	0.12	0.26
15	32	0.08	0.01	0.03	0.08	0.10	0.26	0.10	0.01	0.04	0.07	0.15	0.26
All markets													
5	66	0.09	0.02	0.05	0.08	0.10	0.25	0.09	0.01	0.05	0.07	0.12	0.25
10	66	0.08	0.00	0.05	0.08	0.11	0.26	0.09	0.01	0.05	0.08	0.12	0.26
15	66	0.09	0.01	0.05	0.08	0.12	0.26	0.10	0.01	0.05	0.08	0.15	0.26

Beliefs about bids and beliefs about offers

Table 9 Average absolute difference between subjective and empirical beliefs

Round	No.	Beliefs about bids						Beliefs about offers					
		Mean	Min	Q1	Q2	Q3	Max	Mean	Min	Q1	Q2	Q3	Max
Market size 4													
5	16	0.15	0.06	0.12	0.14	0.18	0.25	0.15	0.04	0.11	0.15	0.20	0.23
10	16	0.09	0.05	0.07	0.09	0.11	0.12	0.09	0.05	0.06	0.09	0.12	0.17
15	16	0.11	0.04	0.08	0.10	0.13	0.22	0.09	0.03	0.06	0.08	0.12	0.21
Market size 6													
5	18	0.13	0.06	0.10	0.13	0.14	0.21	0.14	0.03	0.09	0.15	0.19	0.25
10	18	0.09	0.04	0.07	0.09	0.11	0.17	0.12	0.05	0.09	0.12	0.13	0.25
15	18	0.09	0.02	0.05	0.07	0.14	0.19	0.10	0.05	0.07	0.09	0.14	0.20
Market size 8													
5	32	0.12	0.05	0.09	0.11	0.14	0.19	0.11	0.04	0.08	0.09	0.14	0.23
10	32	0.09	0.04	0.06	0.09	0.10	0.21	0.10	0.03	0.07	0.09	0.12	0.24
15	32	0.08	0.02	0.05	0.08	0.09	0.23	0.09	0.03	0.05	0.07	0.11	0.29
All markets													
5	66	0.13	0.05	0.10	0.12	0.15	0.25	0.13	0.03	0.08	0.12	0.16	0.25
10	66	0.09	0.04	0.07	0.09	0.11	0.21	0.10	0.03	0.07	0.10	0.12	0.25
15	66	0.09	0.02	0.05	0.08	0.12	0.23	0.10	0.03	0.06	0.08	0.11	0.29

Beliefs about bids and beliefs about offers

Table 10 Average absolute difference between subjective beliefs and BNE beliefs

Valuation	No.	Beliefs about bids						Beliefs about offers					
		Mean	Min	Q1	Q2	Q3	Max	Mean	Min	Q1	Q2	Q3	Max
Buyers													
[0, 2]	114	0.09	0.01	0.05	0.07	0.12	0.26	0.10	0.01	0.04	0.07	0.12	0.31
[2.01, 4]	95	0.09	0.02	0.05	0.08	0.11	0.26	0.09	0.03	0.05	0.09	0.12	0.26
[4.01, 6]	92	0.10	0.02	0.05	0.09	0.12	0.26	0.10	0.02	0.05	0.09	0.14	0.26
[6.01, 8]	91	0.08	0.01	0.05	0.08	0.11	0.26	0.08	0.01	0.04	0.07	0.11	0.26
[8.01, 10]	103	0.09	0.01	0.05	0.08	0.12	0.26	0.08	0.01	0.04	0.07	0.12	0.26
All	495	0.09	0.01	0.05	0.08	0.12	0.26	0.09	0.01	0.05	0.08	0.12	0.31
Sellers													
[0, 2]	110	0.08	0.01	0.04	0.07	0.10	0.26	0.10	0.01	0.05	0.09	0.14	0.25
[2.01, 4]	104	0.09	0.01	0.06	0.08	0.11	0.26	0.10	0.01	0.05	0.09	0.14	0.26
[4.01, 6]	97	0.09	0.01	0.05	0.08	0.11	0.26	0.10	0.02	0.05	0.09	0.12	0.27
[6.01, 8]	79	0.09	0.02	0.04	0.08	0.11	0.26	0.10	0.01	0.05	0.09	0.13	0.27
[8.01, 10]	105	0.09	0.00	0.05	0.08	0.11	0.26	0.10	0.01	0.05	0.08	0.13	0.26
All	495	0.09	0.00	0.05	0.08	0.11	0.26	0.10	0.01	0.05	0.09	0.13	0.27

Beliefs about bids and beliefs about offers

Table 11 Average absolute difference between subjective beliefs and empirical beliefs

Valuation	No.	Beliefs about bids						Beliefs about offers					
		Mean	Min	Q1	Q2	Q3	Max	Mean	Min	Q1	Q2	Q3	Max
Buyers													
[0, 2]	114	0.11	0.00	0.07	0.10	0.15	0.30	0.11	0.00	0.07	0.10	0.14	0.33
[2.01, 4]	95	0.12	0.02	0.07	0.10	0.15	0.33	0.10	0.02	0.06	0.09	0.13	0.29
[4.01, 6]	92	0.13	0.03	0.09	0.12	0.17	0.33	0.13	0.03	0.08	0.12	0.16	0.33
[6.01, 8]	91	0.11	0.02	0.07	0.09	0.14	0.30	0.10	0.01	0.07	0.10	0.13	0.28
[8.01, 10]	103	0.12	0.03	0.07	0.11	0.16	0.30	0.11	0.03	0.07	0.10	0.14	0.30
All	495	0.12	0.00	0.07	0.11	0.16	0.33	0.11	0.00	0.07	0.10	0.14	0.33
Sellers													
[0, 2]	110	0.11	0.01	0.07	0.10	0.13	0.33	0.12	0.01	0.07	0.11	0.15	0.33
[2.01, 4]	104	0.12	0.02	0.08	0.10	0.15	0.30	0.12	0.03	0.08	0.11	0.15	0.33
[4.01, 6]	97	0.11	0.02	0.07	0.10	0.14	0.33	0.12	0.02	0.07	0.10	0.16	0.33
[6.01, 8]	79	0.11	0.02	0.07	0.10	0.13	0.27	0.13	0.03	0.08	0.11	0.16	0.33
[8.01, 10]	105	0.11	0.02	0.07	0.11	0.14	0.28	0.13	0.03	0.08	0.11	0.17	0.31
All	495	0.11	0.01	0.07	0.10	0.14	0.33	0.12	0.01	0.07	0.11	0.16	0.33

Beliefs about bids and beliefs about offers

Table 12 Absolute difference (in dollars) between the expected payoff generated by the best response to subjective beliefs and the best response to BNE beliefs (column 1–3) and absolute difference between the expected payoff generated by the best response to subjective beliefs and the best response to empirical beliefs (column 4–6)

	Absolute difference in expected payoffs BR to subj. beliefs versus BR to BNE beliefs			Absolute difference in expected payoffs BR to subj. beliefs versus BR to emp. beliefs		
	Round 5	Round 10	Round 15	Round 5	Round 10	Round 15
Mean	0.45	0.34	0.44	0.54	0.29	0.41
SD	0.57	0.45	0.62	0.67	0.44	0.56
Min	0.00	0.00	0.00	0.00	0.00	0.00
Q1	0.01	0.00	0.01	0.02	0.01	0.02
Median	0.25	0.11	0.12	0.34	0.07	0.16
Q3	0.63	0.63	0.57	0.62	0.43	0.61
Max	2.13	1.69	2.30	3.05	1.77	2.38

Summary statistics for round 5, 10 and 15

Table 13 Percentage of observations for which choice is consistent with the best response to subjective beliefs, empirical beliefs, or BNE beliefs

	Choice consistent with best response to		
	Subjective beliefs	Empirical beliefs	BNE beliefs
All markets			
Round 1	38	24	39
Round 15	36	29	27
Markets with four players			
Round 1	25	25	38
Round 15	25	19	19
Markets with six players			
Round 1	22	17	28
Round 15	28	28	17
Markets with eight players			
Round 1	53	28	47
Round 15	47	34	38

Table 14 Accuracy of beliefs. Medians in round 1 and round 15. Comparison of subjective beliefs, BNE beliefs, and empirical beliefs according to different measures of accuracy: quadratic score, linear score and average absolute difference (*AAD*)

	Subjective beliefs	BNE beliefs	Empirical beliefs	Perfect foresight
Markets with 4 players				
Round 1				
Quadratic score ($)	2.31	2.47	2.33	3.50***
Linear score ($)	0.85	0.94	0.67	3.00***
AAD	0.48	0.49	0.50	0***
Round 15				
Quadratic score ($)	2.15	2.47	2.30	3.50***
Linear score ($)	0.72	0.94*	0.82	3.00***
AAD	0.50	0.46*	0.49	0***
Markets with 6 players				
Round 1				
Quadratic score ($)	2.30	2.47*	2.33	3.50***
Linear score ($)	0.78	0.91*	0.67	3.00***
AAD	0.45	0.44*	0.50	0***
Round 15				
Quadratic score ($)	2.35	2.22	2.47**	3.20***
Linear score ($)	0.80	0.66†	0.96*	2.40***
AAD	0.40	0.43	0.39*	0v
Markets with 8 players				
Round 1				
Quadratic score ($)	2.35	2.41**	2.34	2.82***
Linear score ($)	0.82	0.84	0.67†	1.65***
AAD	0.32	0.29**	0.36†	0***

Table 14 continued

	Subjective beliefs	BNE beliefs	Empirical beliefs	Perfect foresight
Round 15				
Quadratic score ($)	2.33	2.40	2.39	2.87***
Linear score ($)	0.81	0.83	0.84	1.74***
AAD	0.32	0.28	0.29**	0***

As a benchmark, the accuracy of perfect foresight is the one that would be achieved by correctly forecasting the actual behavior of other market participants

One-side Wilcoxon signed rank test against no difference in accuracy across pairs of beliefs. Pairs are defined as subjective beliefs vs. BNE beliefs, empirical beliefs, or perfect foresight

† † †, ††, † Denotes $p<0.001$, $p<0.01$, $p<0.05$ when the alternative hypothesis is that the accuracy of subjective beliefs is higher

***, **, * Denotes $p<0.001$, $p<0.01$, $p<0.05$ when the alternative hypothesis is that the accuracy of subjective beliefs is lower

Two-side Wilcoxon signed rank test rejects the null hypothesis of no difference in the accuracy of subjective beliefs between round 1 and 15 only in the case of market size 6 and accuracy measured by *AAD*. For market size 4 (quadratic rule $p = 0.4545$, linear rule $p = 0.3877$, *AAD* $p = 0.4240$), for market size 6 (quadratic rule $p = 0.8145$, linear rule $p = 0.4807$, *AAD* $p = 0.0075$), for market size 8 (quadratic rule $p = 1.0000$, linear rule $p = 0.7201$, *AAD* $p = 0.7201$)

Table 15 Calibration: empirical relative frequency of bids (offers) falling in each interval, by subjective beliefs. Subjective beliefs are represented by the vector of probabilities assigned to each interval

Subjective beliefs	No.	Choice					
		$[0,2]	$[2.01,4]	$[4.01,6]	$[6.01,8]	$[8.01,10]	> $10
Bids							
(0,0,1,0,0,0)	20	0.27	0.16	0.30	0.24	0.03	0.00
(0,0.2,0.3,0.3,0.2,0)	12	0.21	0.24	0.27	0.11	0.17	0.00
(0,0.25,0.25,0.25,0.25,0)	15	0.35	0.09	0.37	0.14	0.05	0.00
(0,0.25,0.5,0.25,0,0)	12	0.35	0.15	0.26	0.13	0.10	0.00
(0.1,0.2,0.2,0.2,0.2,0.1)	15	0.16	0.17	0.35	0.25	0.07	0.00
(0.1,0.2,0.25,0.25,0.2,0)	15	0.33	0.13	0.35	0.10	0.08	0.00
(0.1,0.3,0.3,0.2,0.1,0)	11	0.28	0.27	0.32	0.11	0.02	0.00
(0.16,0.17,0.17,0.17,0.17,0.16)	12	0.20	0.20	0.20	0.26	0.13	0.00
(0.2,0.2,0.2,0.2,0.2,0)	122	0.25	0.18	0.23	0.23	0.11	0.00
(0.25,0.25,0.25,0.25,0,0)	36	0.22	0.27	0.24	0.13	0.14	0.00
(0.3,0.3,0.3,0.1,0,0)	14	0.15	0.21	0.26	0.24	0.12	0.02
(0.5,0.5,0,0,0,0)	15	0.27	0.20	0.30	0.11	0.08	0.03
Offers							
(0,0,1,0,0,0)	15	0.21	0.22	0.25	0.19	0.13	0.00
(0,0.2,0.3,0.3,0.2,0)	14	0.11	0.14	0.38	0.22	0.15	0.00
(0,0.25,0.25,0.25,0.25,0)	22	0.10	0.13	0.23	0.29	0.25	0.00
(0,0.25,0.5,0.25,0,0)	18	0.17	0.16	0.15	0.18	0.34	0.00
(0.1,0.2,0.2,0.2,0.2,0.1)	14	0.19	0.12	0.23	0.26	0.20	0.00

Table 15 continued

Subjective beliefs	No.	Choice					
		\$[0,2]	\$[2.01,4]	\$[4.01,6]	\$[6.01,8]	\$[8.01,10]	> \$10
(0.1,0.3,0.3,0.2,0.1,0)	18	0.25	0.25	0.22	0.14	0.08	0.06
(0.16,0.17,0.17,0.17,0.17,0.16)	11	0.08	0.11	0.29	0.22	0.30	0.00
(0.2,0.2,0.2,0.2,0.2,0)	123	0.21	0.21	0.23	0.15	0.20	0.01
(0.2,0.25,0.25,0.2,0.1,0)	16	0.14	0.31	0.36	0.10	0.09	0.00
(0.25,0.25,0.25,0.25,0,0)	23	0.22	0.23	0.31	0.09	0.15	0.00
(0.3,0.3,0.3,0.1,0,0)	12	0.33	0.25	0.13	0.19	0.10	0.00
(0.5,0.5,0,0,0,0)	13	0.51	0.13	0.11	0.10	0.14	0.00

Among all stated vectors, only those with more than ten observations are displayed

Appendix 2: The double auction model

This section provides an overview of the double auction model by Rustichini et al. (1994).

The market is populated with n buyers and n sellers. Each buyer has an endowment of \$10 and each seller has one unit of an indivisible good. Each buyer can purchase at most a single unit of the good from any seller, and each seller could sell a single unit of the good to any buyer. Each buyer has a private value and each seller has a private cost for the good. Private values and costs are independently drawn from the uniform distribution over $[0, 10]$. A subject's private value or cost is her own private information, and the process by which private values and costs are drawn is common knowledge among subjects.

All buyers and sellers simultaneously submit their bids or offers. Sorting all bids and offers in increasing order as $\psi_{(1)} \leq \psi_{(2)} \leq \cdots \leq \psi_{(2n)}$, the market price is set at the midpoint between the middle two figures, at $p = 0.5\psi_{(n)} + 0.5\psi_{(n+1)}$.[61] Buyers who submit a bid larger than or equal to p and sellers who submit an offer smaller than or equal to p trade. If trading at p, a buyer with private value v earns a payoff equal to $v - p$, and a seller with private cost c earns a payoff equal to $p - c$. Participants who do not trade earn zero.[62]

Lets consider a buyer i with private value v. According to the auction rules, the relationship between her bid b and the bids and offers of the other $2n - 1$ participants determines what her payoff will be, depending on whether she trades and, if so, at what price. Sorting all other $2n - 1$ bids and offers in increasing order as $\zeta_{(1)} \leq \zeta_{(2)} \leq \cdots \leq \zeta_{(2n-1)}$, buyer i trades if $b > \zeta_{(n)}$. If $\zeta_{(n)} < b < \zeta_{(n+1)}$, the price is $p = 0.5\zeta_{(n)} + 0.5b$. If $b > \zeta_{(n+1)}$, the price is $p = 0.5\zeta_{(n)} + 0.5\zeta_{(n+1)}$. Buyer i

[61] This procedure is equivalent to determining the interval over which the aggregated demand and supply schedules cross, and setting the market price at its midpoint, as the example in Fig. 2 shows.

[62] If excess demand or excess supply arises, priority is given to sellers whose offers are smallest and to buyers whose bids are largest. A fair lottery then determines who trades among the remaining subjects on the long side of the market.

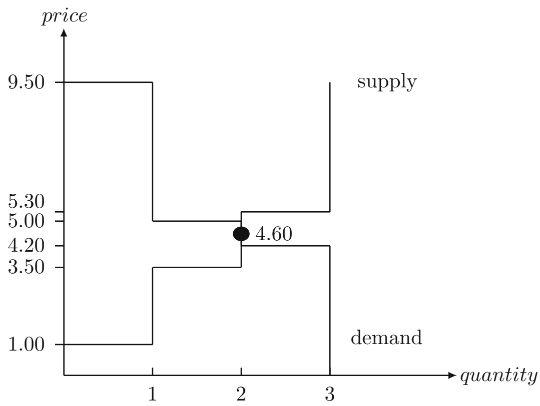

Fig. 2 Example: a market with three buyers and three sellers. Note: Buyers bid at 9.50, 5.00 and 4.20 and sellers offer at 1.00, 3.50 and 5.30. Bids and offers imply a demand and supply curve, respectively. The crossing of the demand and supply curves determines the interval [4.20, 5.00], from which a market-clearing price can be selected. The market-clearing price is selected at the midpoint of the interval, $p = 4.60$

chooses bid $b \geq 0$ without knowing the other $2n$-1 bids and offers but holding probabilistic beliefs about them.[63]

Assuming buyer i behaves as a risk-neutral subjective expected utility maximizer, given her beliefs about the realizations of $\zeta_{(n)}$ and $\zeta_{(n+1)}$, represented by the joint density $f_{(n),(n+1)}$, she chooses bid b in order to maximize her expected payoff:

$$\begin{aligned} \pi(v,b) = & \int_b^{10} \int_0^b (v - [0.5s + 0.5b]) f_{(n),(n+1)}(s,t) ds dt \\ & + \int_0^b \int_0^t (v - [0.5s + 0.5t]) f_{(n),(n+1)}(s,t) ds dt, \end{aligned} \tag{1}$$

where $\zeta_{(n)}$ and $\zeta_{(n+1)}$ are defined over the set $[0, 10]$, the first term is the expected payoff if $\zeta_{(n)} < b < \zeta_{(n+1)}$ and the second term is the expected payoff if $b > \zeta_{(n+1)}$. The inner integral integrates over all possible values of $\zeta_{(n)}$ and the outer one over all possible values of $\zeta_{(n+1)}$.

Analogously, the beliefs of seller i, who has private cost c and submits offer a, can be represented by $f_{(n-1),(n)}$, the joint density of $\zeta_{(n-1)}$ and $\zeta_{(n)}$, and her expected payoff can be written as

$$\begin{aligned} \pi(c,a) = & \int_a^{10} \int_0^a (0.5a + 0.5t - c) f_{(n-1),(n)}(s,t) ds dt \\ & + \int_s^{10} \int_a^{10} (0.5s + 0.5t - c) f_{(n-1),(n)}(s,t) ds dt. \end{aligned} \tag{2}$$

where $\zeta_{(n-1)}$ and $\zeta_{(n)}$ are defined over the set $[0, 10]$, the first term is the expected

[63] $\psi_{(n+1)}$ is the $n+1$-th order statistic among the $2n$ choices submitted by all agents. $\zeta_{(n)}$ is the n-th order statistic among the $2n$-1 choices submitted by the other agents. In terms of all $2n$ choices (including one's own), a winning bid b must be at least equal to $\psi_{(n+1)}$. In terms of the other $2n$-1 choices (exclusing one's own), a winning bid b must be at least equal to $\zeta_{(n)}$.

payoff if $\zeta_{(n-1)} < a < \zeta_{(n)}$ and the second term is the expected payoff if $a < \zeta_{(n-1)}$. The inner integral integrates over all possible values of $\zeta_{(n-1)}$ and the outer one over all possible values of $\zeta_{(n)}$.

Summing up, a buyer's beliefs are represented by joint density $f_{(n),(n+1)}$ and a seller's beliefs are represented by joint density $f_{(n-1),(n)}$. Denoting the density functions of other buyers' bids and other sellers' offers with g_b and g_s respectively, and their cumulative density functions with G_b and G_s respectively, $f_{(n),(n+1)}$ and $f_{(n-1),(n)}$ can be written as function of the beliefs about other buyers' bids (g_b and G_b) and the beliefs about other sellers' offers (g_s and G_s).

To show it, consider a general formulation with m buyers and n sellers. Denote $k = m$. Denote with $f_{(k),(k+1)}$ the joint density of the mth and $(m+1)$th order statistics of the $m + n - 1$ bids and offers. Then:

$$\begin{aligned}
f_{(k),(k+1)}(x, y) = {} & n(n-1)g_s(x)g_s(y) \sum_{\substack{i+j=k-1\\ 0 \le i \le m-1\\ 0 \le j \le n-2}} \binom{m-1}{i}\binom{n-2}{j} G_b(x)^i G_s(x)^j \\
& \times (1 - G_b(y))^{m-1-i}(1 - G_s(y))^{n-2-j} + +n(m-1)g_s(x)g_b(y) \\
& \times \sum_{\substack{i+j=k-1\\ 0 \le i \le m-2\\ 0 \le j \le n-1}} \binom{m-2}{i}\binom{n-1}{j} G_b(x)^i G_s(x)^j (1 - G_b(y))^{m-2-i}(1 - G_s(y))^{n-1-j} \\
& + (m-1)n g_b(x)g_s(y) \sum_{\substack{i+j=k-1\\ 0 \le i \le m-2\\ 0 \le j \le n-1}} \binom{m-2}{i}\binom{n-1}{j} G_b(x)^i G_s(x)^j (1 - G_b(y))^{m-2-i} \\
& \times (1 - G_s(y))^{n-1-j} + (m-1)(m-2)g_b(x)g_b(y) \sum_{\substack{i+j=k-1\\ 0 \le i \le m-3\\ 0 \le j \le n}} \binom{m-3}{i} \\
& \times \binom{n}{j} G_b(x)^i G_s(x)^j \\
& \times (1 - G_b(y))^{m-3-i}(1 - G_s(y))^{n-j}
\end{aligned}$$

Denote $k = m - 1$. Denote with $f_{(k),(k+1)}$ the joint density of the $(m-1)$th and mth order statistics of the m+n-1 bids and offers. Then:

$$f_{(k),(k+1)}(x,y) = (n-1)(n-2)g_s(x)g_s(y) \sum_{\substack{i+j=k-1\\0\le i\le m\\0\le j\le n-3}} \binom{m}{i}\binom{n-3}{j} G_b(x)^i G_s(x)^j$$
$$\times\ (1-G_b(y))^{m-i}(1-G_s(y))^{n-3-j} + (n-1)mg_s(x)g_b(y)$$
$$\times \sum_{\substack{i+j=k-1\\0\le i\le m-1\\0\le j\le n-2}} \binom{m-1}{i}\binom{n-2}{j} G_b(x)^i G_s(x)^j (1-G_b(y))^{m-1-i}(1-G_s(y))^{n-2-j}$$
$$+\ m(n-1)g_b(x)g_s(y) \sum_{\substack{i+j=k-1\\0\le i\le m-1\\0\le j\le n-2}} \binom{m-1}{i}\binom{n-2}{j} G_b(x)^i G_s(x)^j (1-G_b(y))^{m-1-i}$$
$$\times\ (1-G_s(y))^{n-2-j} + m(m-1)g_b(x)g_b(y) \sum_{\substack{i+j=k-1\\0\le i\le m-2\\0\le j\le n-1}} \binom{m-2}{i}$$
$$\times \binom{n-1}{j} G_b(x)^i G_s(x)^j$$
$$\times\ (1-G_b(y))^{m-2-i}(1-G_s(y))^{n-1-j}$$

The formulas for the marginal densities of order statistics are as follows. Denote $k = m$.

For the buyer:

$$f_{(k)}(x) = ng_s(x) \sum_{\substack{i+j=k-1\\0\le i\le m-1\\0\le j\le n-1}} \binom{m-1}{i}\binom{n-1}{j} G_b(x)^i G_s(x)^j (1-G_b(x))^{m-1-i}$$
$$(1-G_s(x))^{n-1-j}$$
$$+\ (m-1)g_b(x) \sum_{\substack{i+j=k-1\\0\le i\le m-2\\0\le j\le n}} \binom{m-2}{i}\binom{n}{j} G_b(x)^i G_s(x)^j (1-G_b(x))^{m-2-i}$$
$$(1-G_s(x))^{n-j}$$

For the seller:

$$f_{(k)}(x) = (n-1)g_s(x) \sum_{\substack{i+j=k-1 \\ 0 \le i \le m \\ 0 \le j \le n-2}} \binom{m}{i}\binom{n-2}{j} G_b(x)^i G_s(x)^j (1-G_b(x))^{m-i}$$
$$\times (1-G_s(x))^{n-2-j}$$
$$+ m g_b(x) \sum_{\substack{i+j=k-1 \\ 0 \le i \le m-1 \\ 0 \le j \le n-1}} \binom{m-1}{i}\binom{n-1}{j} G_b(x)^i G_s(x)^j$$
$$\times (1-G_b(x))^{m-1-i}(1-G_s(x))^{n-1-j}$$

Appendix 3: Parametric fitting of beliefs

Depending on (a) the number of intervals where positive probability is placed, and (b) whether the intervals, where positive probability is placed, are adjacent to each other, I fit:

(1) *a triangular distribution* when positive probability is placed (i) over a single interval or (ii) over two adjacent intervals,
(2) *the union of two triangular distributions* when positive probability is placed (i) over two non-adjacent intervals or (ii) over three intervals of which two but not all three are adjacent (iii) over four intervals, consisting of two disjoint pairs of adjacent intervals,
(3) *the union of three triangular distributions* when positive probability is placed over three interval none of which is adjacent to any other,
(4) *a unimodal beta distribution* when positive probability is placed over three or more intervals all adjacent to each other,
(5) *the union of a beta distribution and a triangular distribution* when positive probability is placed over more than three intervals, of which at least three, but not all, are adjacent to each other.

Here below I describe how the fitting is performed in each case. Figure 3 shows an example for each case, the most common being the case in which beliefs are fitted with a unimodal beta distribution (approximately 90 % of observations), followed by the case in which beliefs are fitted with a triangular distribution (approximately 7 % of observations).

Triangular distribution

If positive probability is assigned to only one interval, $[l, r]$, then I assume that the support of the subjective distribution is $[l, r]$ and I fit a triangular distribution over it.

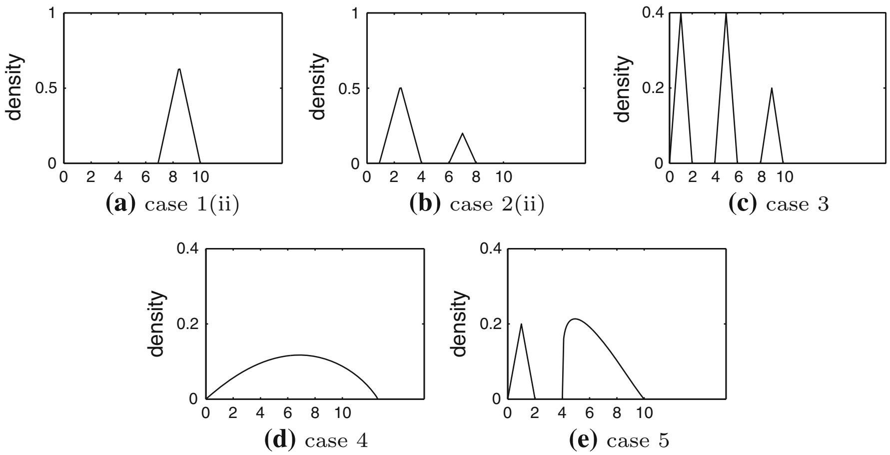

Fig. 3 A selection of the fitting methods 1–5

The fitted isosceles triangle has base $r - l$ and height $\frac{2}{r-l}$. If positive probability is assigned to two adjacent intervals and if equal probability is assigned to each interval, then I assume that the support of the subjective distribution is the union of the two intervals and the fitted isosceles triangle has base 4 and height $\frac{1}{2}$.

If positive probability is assigned to two adjacent intervals and a higher probability is assigned to one interval than to the other, then I assume that the subjective distribution has the shape of an isosceles triangle and that its support contains entirely the interval that was assigned a higher probability and partly the other interval. If the subject assigns probability α and $1 - \alpha$ to the intervals $[y, y + 2]$ and $(y + 2, y + 4]$, respectively, where $\alpha < 0.5$, then, the fitted isosceles triangle has base with endpoints $y + 2 - t$ and $y + 4$ and height $h = \frac{2}{t+2}$, with $t = \frac{2\sqrt{\frac{\alpha}{2}}}{1-\sqrt{\frac{\alpha}{2}}}$.[64]

Union of two triangular distributions

If positive probability is assigned to two non-adjacent intervals, then I assume that the support of the subjective distribution is the union of the two intervals and I fit a triangular distribution over each interval. For example, probability α is assigned to the interval $[l_1, r_1]$ and probability $1 - \alpha$ to the interval $[l_2, r_2]$, then I assume that the support of the distribution is the union of $[l_1, r_1]$ and $[l_2, r_2]$. The isosceles triangle fitted over $[l_1, r_1]$ has base $r_1 - l_1$ and height $\frac{2\alpha}{r_1-l_1}$. The isosceles triangle fitted over $[l_2, r_3]$ has base $r_2 - l_2$ and height $\frac{2(1-\alpha)}{r_2-l_2}$.

If positive probability is assigned to three intervals of which two but not all three are adjacent, then I assume that the support of the subjective distribution is the union

[64] Analogously, if the subject assigns probability α and $1 - \alpha$ to the intervals $[y, y + 2]$ and $(y + 2, y + 4]$ respectively, where $\alpha > 0.5$, then I let $t = \frac{2\sqrt{\frac{1-\alpha}{2}}}{1-\sqrt{\frac{1-\alpha}{2}}}$. Then the subjective probability density function takes the form of a triangle with a base with endpoints y and $y + 2 + t$ and with a height $h = \frac{2}{t+2}$.

of the intervals and I fit two triangular distributions, one over the two adjacent intervals and another other the non-adjacent interval. For example, suppose that the intervals $[l_1, r_1]$, $[l_2, r_2]$ and $[l_3, r_3]$ are assigned probability α, β and $1-\alpha-\beta$, respectively, and that $[l_1, r_1]$ and $[l_2, r_2]$ are adjacent to each other (with $l_2 > l_1$), while $[l_3, r_3]$ is not adjacent to any of the other intervals. Then one triangle is fitted over the union of $[l_1, r_1]$ are $[l_2, r_2]$ and one triangle is fitted over $[l_3, r_3]$, following the procedures already described for fitting one triangular distribution.[65]

If positive probability is assigned to four intervals, consisting of two disjoint pairs of adjacent intervals, then I also fit two triangular distributions. Each triangular distribution has support over a pair of adjacent intervals and the fitting is done following the procedure already described for fitting one triangular distribution.

Union of three triangular distributions

If positive probability is assigned to three intervals none of which is adjacent to any other, then I assume that the support of the subjective distribution is the union of the intervals and I fit a triangular distribution over each intervals, following the procedure already described for fitting one triangular distribution.

Unimodal beta distribution

If positive probability is assigned to three or more intervals all adjacent to each other, then I fit a generalized unimodal Beta distribution over the intervals. The cumulative distribution function for a unimodal Beta distribution evaluated at x is denoted $Beta(x, \alpha, \beta, l, r)$, where α and β are shape parameters and l and r are location parameters determining the support for the distribution over the range $[l, r]$. If a subject does not assign positive probability to the right tail interval $[10.01, \infty)$, then the lower bound l of the support for the fitted Beta distribution will coincide with the left endpoint of the leftmost interval with positive probability and the upper bound r of the support will coincide with the right endpoint of the rightmost interval with positive probability and, therefore, the parameters l and r will be fixed. Thus, fitting the data with a Beta distribution requires solving the problem $\min_{\alpha, \beta} \sum_{j=1}^{6} [Beta(r_j, \alpha, \beta, l, r) - G(r_j)]^2$, where $G(r_j)$ is the sum of the subjective probabilities assigned up to the interval with right endpoint r_j, inclusive.

If instead a subject assigns positive probability to the upper unbounded interval [\$10.01,∞), I let the location parameter r be a free parameter in the minimization of the least squares problem. I restrict r to lie within the most extreme value recorded in the data, which is 18.50. Thus, the problem becomes $\min_{\alpha, \beta, r < 18.50}$

[65] Therefore, if $\alpha < \beta$ then the first triangle has base with endpoints $l_1 + 2 - t$ and $l_1 + 4$ and height $h = \frac{2(\alpha+\beta)}{t+2}$ with $t = \frac{2\sqrt{\frac{\alpha}{2(\alpha+\beta)}}}{1-\sqrt{\frac{\alpha}{2(\alpha+\beta)}}}$. If $\alpha > \beta$, then the first triangle has base with endpoints l_1 and $l_1 + 2 + t$ and height $h = \frac{2(\alpha+\beta)}{t+2}$ with $t = \frac{2\sqrt{\frac{\beta}{2(\alpha+\beta)}}}{1-\sqrt{\frac{\beta}{2(\alpha+\beta)}}}$. The second triangle has base with endpoints l_3 and r_3 and height $\frac{2(1-\alpha-\beta)}{r_3-l_3}$.

$\sum_{j=1}^{6}[Beta(r_j, \alpha, \beta, l, r) - G(r_j)]^2$.

Union of a Beta distribution and a triangular distributions

If positive probability is assigned to more than three intervals, of which at least three but not all are adjacent to each other, then I fit a unimodal beta distribution over the three or more adjacent intervals and a triangular distribution over the remaining one or two intervals. I follow the procedures already described for fitting a triangular distribution and a unimodal Beta distribution.

Goodness of fit is assessed by the average absolute deviation between the fitted and the elicited beliefs. In most cases the average absolute deviation is below 0.01, and in all cases below 0.06.

Appendix 4: Descriptive Results

Descriptive results about bidding behavior

In this section I illustrate how the observed bidding choices compare with the bidding behavior prescribed by the risk-neutral BNE model, as presented by Rustichini et al. (1994).[66] Figure 4 illustrates the approximate risk-neutral bidding functions predicted by the model.

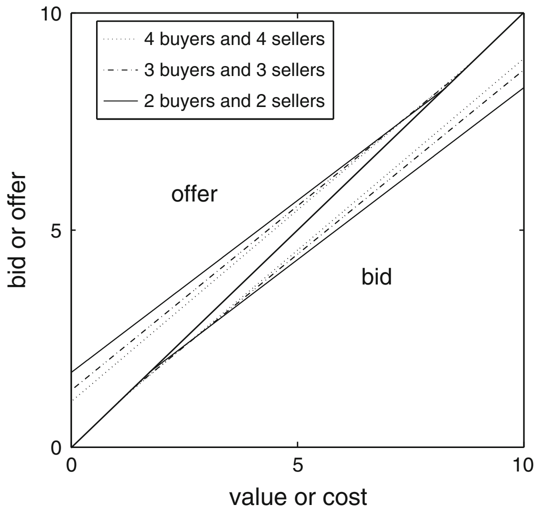

Fig. 4 Approximate risk-neutral BNE bid and offer functions

[66] Rustichini et al. (1994) show that there exists a family of asymmetric smooth equilibria, which can be computed numerically. To simplify the analysis, I consider approximate symmetric bid and offer functions. The approximation should not affect the analysis since the family of asymmetric smooth equilibria is contained in a small neighborhood of the symmetric bid and offer functions. Also Cason and Friedman (1997) use approximate symmetric bid and offer functions in their experimental implementation of the Rustichini et al. (1994) double auction.

Table 16 Underrevelation of private information and deviation of the BNE best response from the observed choice

		(i) All (%)	(ii) Bid $\leq$ value or offer $\geq$ cost (%)	(iii) Bid > value or offer < cost (%)
Underrevelation of private information				
$VUR(v,b) = (v-b)/v$	Median	3***	7***	−34***
	[Q1,Q3]	[0,20]	[0,23]	[−203, −13]
$CUR(c,a) = (a-c)/(9.99-c)$	Median	1***	4***	−36***
	[Q1,Q3]	[0,13]	[0, 20]	[−141, −10]
Deviation of BNE best response from choice				
$D(b_{BNE},b) = (b_{BNE}-b)/b_{BNE}$	Median	−4***	0	−46***
	[Q1,Q3]	[−12, 10]	[−10, 14]	[−208, −28]
$D(a_{BNE},a) = (a_{BNE}-a)/a_{BNE}$	Median	6***	2***	42***
	[Q1,Q3]	[−2, 32]	[−4, 12]	[25, 59]
Obs.		990	802	188

Median, 1st quartile (Q1), and 3rd quartile (Q3). Wilcoxon signed-rank test of the null hypothesis that the median equals zero

*** Denotes significantly different from zero at 0.1 %; ** at 1 %; * at 5 %

Since the double auction is a strategic environment with private information, strategic misrepresentation of private information is a key feature within the BNE model. As a measure of how much subjects reveal of the private information they hold, I use the value underrevelation ratio and the cost underrevelation ratio, as defined by Cason and Friedman (1997). The value underrevelation ratio $VUR(v,b)$ is the fraction of the buyer's private value, v, discounted in the chosen bid, b, and the cost underrevelation ratio $CUR(c,a)$ is the cost mark-up relative to the highest possible cost. Thus, $VUR(v,b) = (v-b)/v$ and $CUR(c,a) = (a-c)/(9.99-c)$.[67] A positive ratio corresponds to underrevelation and a negative ratio corresponds to overrevelation.[68]

The upper panel of Table 16 reports the median value and cost underrevelation ratios. Results are reported for (i) the entire sample, and separately for (ii) the subsample in which underrevelation occurs and (iii) the subsample in which

[67] Recall that the highest possible private cost in the experiment is \$9.99.

[68] Table 16 reveals that 19 % of the observations (188 out of 990) consist of either a buyer bidding above value or a seller offering below cost. While a percentage of 19 % is surprisingly high, several qualifications needs to be made. Differences exist across subjects: some subjects incur in overrevelation more often than others. Among participants, 70 % of them overreveal their private information in 3 or fewer of the 15 total rounds. Most importantly, the expected monetary losses generated by overrevelation are small. I compute, based on the elicited subjective beliefs, the expected payoffs generated by the actual bidding choices above value or below cost and I compare them with the expected payoffs generated by bidding at value or at cost. The percentage of observations, for which the expected monetary loss generated by overrevelation is greater than or equal to 10 cents (50 cents) is equal to 10 % (5 %).

Table 17 Market performance

	2 Buyers, 2 sellers	3 Buyers, 3 sellers	4 Buyers, 4 sellers	All markets
Number of trades				
0	20 %			7 %
1	67 %	53 %	27 %	48 %
2	13 %	44 %	55 %	37 %
3		2 %	18 %	7 %
Mean profits	$0.94	$0.78	$0.91	$0.88
Trading efficiency	84.8	66.4	79.8	78
Cason and Friedman (1997)			87.3	
Within CE				
Price interval	68.3 %	37.8 %	41.7 %	50 %
Cason and Friedman (1997)			45 %	

overrevelation occurs. Within subsample (ii) the median $VUR(v, b)$ and $CUR(c, a)$ are 7 and 4 %, respectively.[69]

The lower panel of Table 16 reports the magnitude of the percent deviation of the BNE best response from the observed choice, defined as $D(b_{BNE}, b) = (b_{BNE} - b)/b_{BNE}$ for buyers and $D(a_{BNE}, a) = (a_{BNE} - a)/a_{BNE}$ for sellers. Within subsample (ii), the median deviation is approximately 0 for buyers and 2 % for sellers. [70]

Descriptive results about market performance

Table 17 reports a description of market performance across the experimental auctions. The number of trades taking place in each round ranges between zero and three, with one or two trades per round being the most common outcome. Over all observations, including those when no trade occurs, mean profits are $0.88. In a competitive equilibrium (CE) buyers and sellers submit respectively their private values and private costs without engaging in any strategic misrepresentation of their private information. Across all experimental auctions, trading efficiency, defined as

[69] I compare $VUR(v, b)$ and $CUR(c, a)$, which are computed with respect to the observed bids and offers b and a, with the underrevelation ratios computed with respect to the bids and offers prescribed by the BNE strategies b_{BNE} and a_{BNE}, defined as $VUR(v, b_{BNE}) = (v - b_{BNE})/v$ and $CUR(c, a_{BNE}) = (a_{BNE} - c)/(9.99 - c)$. The medians of $VUR(v, b_{BNE})$ and $CUR(c, a_{BNE})$ are equal and approximately 10 %, indicating that experiment participants reveal more private information than they would according to the BNE.

[70] Cason and Friedman (1997) report qualitatively comparable results, in spite of several differences in the experimental design. First, buyers choose a bid above what the BNE would prescribe and sellers choose an offer below what BNE would prescribe. Second, subjects reveal more of their private information than what they would according to BNE strategies: for inexperienced subjects the median value underrevelation ratio is 6.5 % and the median cost underrevelation ratio is 9.1 % (for experienced subjects the ratios are 2.5 and 2.4 %, respectively).

the percentage of the gains from exchange realized by traders in comparison with the CE, is 78 % and prices are within the CE price interval in 50 % of the rounds.[71]

References

Armantier, O., & Treich, N. (2009). Subjective probabilities in games: An application to the overbidding puzzle. *International Economic Review*, *50*(4), 1079–1102.

Armantier, O., & Treich, N. (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review*, *62*, 17–40.

Bellemare, C., Sebald, A., & Strobel, M. (2010). Measuring the willingness to pay to avoid guilt: Estimation using equilibrium and stated belief models. *Journal of Applied Econometrics*. https://ideas.repec.org/p/iza/izadps/dp4803.html.

Bellemare, C., Sebald, A., & Strobel, M. (2011). Measuring the willingness to pay to avoid guilt: Estimation using equilibrium and stated belief models. *Journal of Applied Econometrics*, *26*, 437–453.

Bellemare, C., Bissonnette, L., & Kröger, S. (2012). Flexible approximation of subjective expectations using probability questions. *Journal of Business and Economic Statistics*, *30*(1), 213–245.

Blanco, M., Engelmann, D., Normann, H., & Koch, A. (2010). Belief elicitation in experiments: Is there a hedging problem? *Experimental Economics*, *13*(4), 412–438.

Camerer, C., Ho, T., Chong, J., Weigelt, K. (2002). Strategic teaching and equilibrium models of repeated trust and entry games. Mimeo: CalTech.

Camerer, C., & Hua Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67*(4), 827–874.

Cason, T., & Friedman, D. (1997). Price formation in single call markets. *Econometrica*, *65*(2), 311–345.

Cheung, Y. W., & Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, *19*(1), 46–76.

Costa-Gomes, M., & Weizsäcker, G. (2008). Stated beliefs and play in normal-form games. *Review of Economic Studies*, *75*(3), 729–762.

Danz, D. N., Fehr, D., & Kübler, D. (2012). Information and beliefs in a repeated normal-form game. *Experimental Economics*, *15*(4), 622–640.

Engelberg, J., Manski, C., & Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics, 27*(1), 30–41.

Erev, I., & Roth, A. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review, 88*(4), 848–881.

Fudenberg, D., & Levine, D. (1998). Learning in games. *European Economic Review*, *42*(3), 631–639.

Haruvy, E., Lahav, Y., & Noussair, C. (2007). Traders' expectations in asset markets: Experimental evidence. *The American Economic Review*, *97*(5), 1901–1920.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290*(5500), 2261–2262.

Hommes, C., Sonnemans, J., Tuinstra, J., & Velden, H. V. D. (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies*, *18*(3), 955.

Hyndman, K., Özbay, E., Schotter, A., & Ehrblatt, W. Z. (2012). Convergence: An experimental study of teaching and learning in repeated games. *Journal of the European Economic Association*, *10*(3), 573–604.

Kirchkamp, O., & Reiß, J. P. (2011). Out-of-equilibrium bids in first-price auctions: Wrong expectations or wrong bids. *The Economic Journal*, *121*(557), 1361–1397.

Manski, C. (2002). Identification of decision rules in experiments on simple games of proposal and response. *European Economic Review*, *46*(4–5), 880–891.

Manski, C. (2004). Measuring expectations. *Econometrica*, *72*(5), 1329–1376.

Manski, C., & Neri, C. (2013). First-and second-order subjective expectations in strategic decision-making: Experimental evidence. *Games and Economic Behavior*, *81*(C), 232–254.

[71] A comparison with Cason and Friedman (1997) is possible for the markets with eight subjects. They find a higher trading efficiency and a slightly higher fraction of prices within the CE price interval. Recall that in their experiment there are more trading rounds (30 instead of 15).

Nyarko, Y., & Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica*, *70*(3), 971–1005.

Palfrey, T., & Wang, S. (2009). On eliciting beliefs in strategic games. *Journal of Economic Behavior & Organization*, *71*(2), 98–109.

Rey-Biel, P. (2009). Equilibrium play and best response to (stated) beliefs in normal form games. *Games and Economic Behavior*, *65*(2), 572–585.

Rustichini, A., Satterthwaite, M., & Williams, S. (1994). Convergence to efficiency in a simple market with incomplete information. *Econometrica*, *62*(5), 1041–1063.

Rutström, E., & Wilcox, N. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, *67*(2), 616–632.

Seidenfeld, T. (1985). Calibration, coherence, and scoring rules. *Philosophy of Science*, *52*, 274–294.

Smith, V., Suchanek, G., & Williams, A. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, *56*(5), 1119–1151.

Sonnemans, J., Hommes, C., Tuinstra, J., & Velden, H. V. D. (2004). The instability of a heterogeneous cobweb economy: A strategy experiment on expectation formation. *Journal of Economic Behavior & Organization*, *54*(4), 453–481.