Journal of Economic Behavior & Organization 119 (2015) 139–152

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Does feedback really matter in one-shot first-price auctions?

Peter Katuščák[a,b], Fabio Michelucci[a,*], Miroslav Zajíček[b]

[a] Center for Economic Research and Graduate Education – Economics Institute (CERGE-EI), Politických vězňů 7, 110 00 Prague, Czech Republic[1]

[b] Department of Institutional, Environmental and Experimental Economics, Faculty of Economics, University of Economics in Prague, Náměstí Winstona Churchilla 4, 130 67 Praha 3, Czech Republic

ARTICLE INFO

Article history:
Received 17 September 2014
Received in revised form 29 July 2015
Accepted 3 August 2015
Available online 10 August 2015

JEL classification:
C91
C92
D44

Keywords:
Auctions
Feedback
Regret
Market design

ABSTRACT

Does the *type* of posterior feedback affect how people decide in one-shot environments? We revisit this question in first-price auction markets. We consider three feedback types: minimal (only knowing whether winning or not), loser (also knowing the winning bid) and winner (knowing the second highest bid if winning). Filiz-Ozbay and Ozbay (2007) find that loser as opposed to minimal or winner feedback increases bids. We use three novel protocols and additionally replicate theirs. Using a sample of 624 subjects, we find that bidders' *ex ante* knowledge of posterior feedback type has no systematic effect on the average bid/value ratios.

1. Introduction

Consider a one-time decision problem under risk or uncertainty. Upon choosing one of the alternatives, a decision-maker (DM) is typically informed about payoff outcome of the chosen alternative. She may or may not receive posterior feedback about payoff outcomes of the unchosen alternative(s). (Subjective) Expected Utility Theory predicts that knowing *ex ante* whether such additional information will or will not be available *ex post* has no impact on the decision. However, several studies have proposed that availability of such information may indeed play a role through anticipated regret (Bell, 1982; Loomes and Sudgen, 1982; Engelbrecht-Wiggans, 1989).

This possibility is interesting for mechanism design. If the presence of such feedback can affect people's behavior in the direction preferred by the mechanism designer, selection of the right kind of feedback becomes an important aspect of the design.

In this paper, we investigate whether the type of posterior feedback affects bidding in one-shot first-price sealed-bid auction with private values (FPA). At the end of the auction, each bidder learns whether she has won or not (*minimal feedback*). The auctioneer may, however, give bidders additional feedback and announce this fact before the bidding starts. For example, he may publicly announce he will disclose the winning bid to all the bidders after the auction (*loser feedback*).

[1] CERGE-EI is a joint workplace of Charles University in Prague and the Economics Institute of the Academy of Sciences of the Czech Republic.

* Corresponding author. Tel.: +420 224005117
E-mail addresses: peter.katuscak@cerge-ei.cz (P. Katuščák), fabiomichelucci@gmail.com (F. Michelucci), miroslav.zajicek@vse.cz (M. Zajíček).

http://dx.doi.org/10.1016/j.jebo.2015.08.002

Or, alternatively, he may publicly announce he will disclose the second highest bid to the winner after the auction (*winner feedback*). These are the three feedback types we focus on in this paper.

There is some existing literature on this topic. Filiz-Ozbay and Ozbay (2007) (referred to as FO hereafter) find that, in one-shot bidding, bidders bid more aggressively (higher bid/value ratios) under loser feedback as opposed to minimal or winner feedback. In particular, FO find that the average bid/value ratio in their four-bidder FPA increases from .79 (.77) under minimal (winner) feedback to .87 under loser feedback. To assess how remarkable this increase is, notice that, under the assumption of uniformly distributed values and the theoretical prediction of risk-neutral Nash equilibrium, it is roughly equivalent to the effect of increasing the number of bidders from 4 to 6. However, while increasing competition by attracting new bidders is costly (and sometimes infeasible), changing the type of feedback is easy and costless.[2]

FO connect their findings to anticipated regret. They argue that loser as opposed to minimal feedback makes bidders worried about potential posterior regret of losing and learning that they could have profitably won by bidding more than the winning bid. To reduce the likelihood of such event, they bid more. There is also an argument for why winner as opposed to minimal feedback can make bidders bid less. This is because bidders might be worried about potential posterior regret of winning and learning that they could have won with a much smaller bid. To reduce the likelihood of such event, they bid less. This latter hypothesis is not supported by FO data, though.

Given the one-shot nature of the auction, this interpretation requires that bidders must anticipate regret even though they (most likely) do not have any recent experience of it in an auction of a similar kind.[3] However, this assumption does not square well with findings of several studies that investigate repeated FPAs (Isaac and Walker, 1985; Ockenfels and Selten, 2005; Neugebauer and Selten, 2006; Neugebauer and Perote, 2008; Engelbrecht-Wiggans and Katok, 2008). These studies generally find that loser (winner) feedback leads to higher (lower) bid/value ratios compared to minimal feedback in later rounds of repeated bidding.[4] However, most of these studies find no effect of feedback in the first bidding round (Ockenfels and Selten, 2005; Neugebauer and Selten, 2006; Neugebauer and Perote, 2008). These contradictory findings have recently sparked a debate about this important market design issue (see, for instance, Neugebauer and Perote (2008)[5] and Kagel and Levin (2011)[6]). In this paper, we contribute to this debate by comprehensively testing for the presence of feedback effects in one-shot bidding environments using four different protocols involving one computerized opponent, one and three human opponents, and a replication of the protocol used by FO.

The first protocol we analyze, denoted HC, identifies feedback effects under the simplest possible non-strategic environment. In this setting, a human bidder faces a computerized opponent that draws its bid from a known uniform distribution. The second protocol, denoted 2H, is an auction with two *ex ante* symmetric human bidders. This is a natural extension of HC to the simplest possible strategic environment. In this setting, the effect of feedback operates not only through a direct preference channel, but also through an indirect channel of beliefs about how feedback might affect the bidding strategy of the opponent (plus all the higher order beliefs). The comparison with HC highlights the impact of these indirect effects. In both of these two protocols, we use all three feedback types: minimal, loser and winner.

We find that the type of feedback has no significant impact on the average bid/value ratio in either of the two protocols. Regarding the effect of winner vs. minimal feedback, this finding is consistent with the finding of FO. However, regarding the effect of loser vs. minimal feedback, our finding contradicts their conclusions.

Since FO use four human bidders, our findings from the two-bidder auctions (HC and 2H) might suggest that the effects of feedback might be sensitive to the level of bidding competition.[7] To examine this possibility, we introduce the third protocol, denoted 4H. This is an auction with four human bidders that otherwise uses the same procedure as 2H. In light of

[2] Although setting a positive reserve price is also a costless way of increasing expected revenue (Myerson, 1981; Riley and Samuelson, 1981), its effect may be limited if bidder competition is already strong, and it exposes the auctioneer to the risk of retaining the object. Moreover, setting an optimal reserve price requires fine knowledge of distributions of bidders' values and preferences, which might not be known in many applications.

[3] In addition, this interpretation requires that bidder preferences change with feedback type. In the absence of this assumption, one can use the Law of Iterated Expectations to show that feedback has no effect on equilibrium bids. Please refer to the working paper version of this study Katuščák et al. (2013) for more details.

[4] Engelbrecht-Wiggans and Katok (2008) is the only study that utilizes winner feedback. The absolute size of the effect vis-á-vis minimal feedback is smaller than the one for loser feedback.

[5] Neugebauer and Perote (2008) write in footnote 14: "The possibility of anticipating regret has been proposed by Ozbay and Filiz (in press). They report that subjects change their bid according to the anticipated feedback in a one-shot, contingent bid first price auction. However, this pattern seems difficult to reproduce in the first round of the repeated setting. Our results as well as the data reported in Neugebauer and Selten (2006) and Dufwenberg and Gneezy (2002) rather suggest that the first round bids are the same across feedback treatments." In addition, see also evidence from Fig. 1 in Ockenfels and Selten (2005). On the other hand, Fig. 1 in Isaac and Walker (1985) and Figs. 1 and 2 in Engelbrecht-Wiggans and Katok (2008) suggest that loser feedback generates the highest bid/value ratios even in the first round (Isaac and Walker (1985) compare loser with complete feedback, i.e., revealing all bids, and Engelbrecht-Wiggans and Katok (2008) compare minimal, loser, winner and full feedback, i.e., revealing the highest bid of the opponents). In neither of the two studies do the authors report any formal test results for the first round. Further, in Engelbrecht-Wiggans and Katok (2008), loser feedback includes explicit information on money left on the table and Isaac and Walker (1985) include explicit information about the identity of the bidders. Both differences hinder comparability of these two studies to our study and the other cited studies.

[6] Kagel and Levin (2011) note: "To sum up: FO's introduction of regret theory to explain bidding above the RNNE in FPSB auctions is quite innovative. However, the statistical significance of their results are suspect, and Engelbrecht-Wiggans and Katok (2008) and Neugebauer and Selten (2006) fail to replicate their results."

[7] Even though it is not *a priori* clear why market size should interact with feedback, it is worth to carry out this robustness check. In fact, within a repeated bidding environment, Neugebauer and Selten (2006) find that the effect of feedback on bidding is more pronounced in auctions against a higher number of computerized opponents.

our motivation, we implement only two feedback types: minimal and loser.[8] Again, we find that the type of feedback has no significant impact on the average bid/value ratio. As a result, the effect of loser vs. minimal feedback on bidding does not appear to be sensitive to the level of bidding competition, at least not in the range we investigate. A similar null result is reported by Ratan (2013) for the case of one human bidder competing against three computerized opponents programmed to play the risk-neutral Nash equilibrium bidding strategy.

This result is hard to reconcile with the findings of FO since 4H is similar to their protocol. However, it is not identical. It is therefore possible that the effect of feedback on bidding may be sensitive to fine details of experimental implementation. To investigate this possibility, we introduce the fourth protocol, denoted 4HR. This protocol also uses four human bidders and replicates the experimental procedure used by FO. As in 4H, we implement only two feedback treatments: minimal and loser. Again, we find no effect of feedback on the average bid/value ratio. Hence, our previous findings do not appear to be sensitive to details of the utilized experimental procedure.

The outcome of 4HR makes the discrepancy between our empirical findings and the findings of FO even more striking. What can account for the difference? Despite our replication of the procedure of FO in 4HR, there are a few remaining differences in the implementation. First, we use a different subject pool. Our subject sample is drawn from undergraduate and graduate students from universities in Prague, the majority of whom have an economics or business major, whereas FO utilize undergraduate students from New York University whose major composition is not known to us. Second, unlike FO, in each session we use an equal number of men and women to make sure that our results are unaffected by a potential interaction between feedback and gender.[9] Third, we use arguably higher stakes than FO do (see Section 2.4 for further details). Fourth, we use a much larger sample size. Our findings on the effect of loser vs. minimal feedback are based on 144 subjects in HC, 144 subjects in 2H, 96 subjects in 4H, and 96 subjects in 4HR. In comparison, the finding of FO is based on 64 subjects, of which 28 are in the minimal and 36 are in the loser feedback treatment. Note that, given the number of subjects we use in 4H and 4HR, if the effect of loser vs. minimal feedback identified by FO were real and if their data provides a good description of the idiosyncratic variance in bidding behavior, a two-tailed test would have a power of over 0.95 against the null hypothesis of no effect. If we instead use the variance implied by our data, the power of the test ranges from 0.73 to 0.95, depending on the protocol and the outcome measure.[10] The gender control, the higher stakes and the larger sample size all contribute to reducing noise in a given subject population. Hence, our failure to replicate the finding of FO points to subject pool difference as a possible explanation. To shed light on this hypothesis, we carry out a separate analysis of bids of undergraduate vs. more advanced students and of economics or business major students vs. students with other majors. In none of these cases we find any systematic effects of feedback on bidding.

Overall, our results provide evidence that the type of feedback does not have any systematic effect on the average bid/value ratio in one-shot first-price auctions.[11] Our results therefore do not support the notion that the average aggressiveness of bidding in one-shot auctions can be manipulated by the announcement of a specific posterior feedback type, at least not by switching among the three types of feedback we consider. Reflecting on the existence of feedback effects in repeated bidding, we also separately investigate feedback effects in the subsample of subjects with some field experience with auctions and do not find any significant effects either. As an additional result, we find that the type of feedback has a weak effect on the heterogeneity of slopes of bidding strategies among subjects, with this heterogeneity being the smallest under loser feedback. If confirmed by future research, such an effect may be of interest for the impact of feedback on auction efficiency.

The rest of the paper is structured as follows. Section 2 presents details of our experimental design. Section 3 presents our results, and Section 4 concludes. An online appendix contains a representative set of instructions (4H, both printed and on screen instructions) and a demographic questionnaire we used at the end of the experiment.

2. Experimental design

The first protocol, denoted HC, is an auction in which a human bidder plays against one computerized opponent. The aim of this protocol is to face the human bidder with a simple tradeoff between the probability of winning and the surplus conditional on winning. To achieve this aim, we draw the computerized bidder's bid from the uniform distribution on [0, 100] (U[0, 100]). This removes any strategic uncertainty on the part of the human bidder: a bid from the range of [0, 100] is equal to the percentage probability of winning, whereas the difference between the value and the bid is equal to the size

[8] Recall that FO do not find any statistically significant difference between providing winning and minimal feedback. This is further confirmed by our results in HC and 2H. Therefore we do not implement the winning feedback treatment in this protocol.

[9] We asked FO about the gender composition of their sample, but they informed us they had not recorded this information. For more details on how we perform gender balancing, see Section 2.

[10] See Section 3.1 for details.

[11] In the working paper version of this study Katuščák et al. (2013), we extend our investigation also to auction revenue and efficiency, and obtain analogous conclusions.

of the payoff conditional on winning.[12,13] We somewhat deviate from the simplicity of the environment when choosing human bidders' values. Although it would be simplest to draw a single value from $U[0, 100]$, we instead elicit bids by the strategy method based on six random value draws from $U[0, 100]$ presented sequentially (see Section 2.2 for more details). Given the one-shot nature of the auction, we use this procedure to obtain more information about human bidders' bidding strategies.[14] From now on, we will refer to the set of six values and their ordering as a "value list." We implement three feedback types within this protocol: minimal, loser and winner.

The second protocol, denoted 2H, is an auction with two human bidders. The procedure is analogous to HC except that the computerized opponent is replaced by an *ex ante* symmetric human opponent. This protocol is the simplest possible extension of HC to an environment with uncertainty about the bidding strategy used by the opponent. It allows feedback to operate not only directly through preference shifts, but also indirectly through a change in beliefs about the bidding strategy of the opponent. As in HC, we implement three feedback treatments within this protocol: minimal, loser and winner.

Our results from HC and 2H reveal that loser, as opposed to minimal, feedback has no systematic effect on the average bid/value ratio. This is a different finding than the one obtained by FO, who find that loser feedback increases bid/value ratios. This difference in findings could be, among other things, due to a different number of bidders (HC and 2H use two, FO use four) or due to other differences in the experimental protocols. These differences motivate our next two protocols, in which we implement only two feedback treatments: minimal and loser.

The third protocol, denoted 4H, is a computerized auction with four human bidders that otherwise uses the same procedure as 2H. The comparison of 4H and 2H investigates the possibility that effects of feedback might be sensitive to the level of bidding competition.[15]

The fourth protocol, denoted 4HR, is a paper and pencil auction with four human bidders that exactly replicates the procedure used by FO. In this protocol, each bidder's value is drawn from $U[0, 100]$, but this time bidding is implemented by the strategy method with ten randomly generated values presented simultaneously (see Section 2.2 for more details). The comparison of 4HR and 4H investigates the possibility that the impact of feedback may be sensitive to procedural details of the utilized protocol. In particular, in comparison to 4HR (which replicates FO), our previous protocols use longer printed instructions and combine them with on-screen instructions, including a practice round of bidding. The aim of such approach is to maximize subjects' understanding of the auction and the computer interface used to input the bids.[16] However, one could argue that this makes it cognitively more demanding to understand the instructions, which in turn crowds out the attention that subjects pay to potential effects of feedback. Also, the previous protocols insert a questionnaire at the end of the printed instructions that verifies subjects' understanding of the auction, while this step is omitted in 4HR. Again, the aim of the questionnaire is to make sure that subjects understand the auction environment. However, one could argue that the questionnaire, by including questions on the type of feedback received after the auction, makes the possibility of receiving loser feedback salient even in the minimal feedback treatment, and this reduces the effective size of the treatment effect. Implementing 4HR addresses all these potential criticisms by replicating the exact protocol used by FO.

In each session in all four protocols, we use an equal number of men and women.[17] For a given gender, we generate a set of value lists, which is then used across all three feedback types. This is done in order to preclude any omitted variable bias or sampling noise driven by a potential interaction between feedback type and gender or feedback type and values assigned to subjects.

Finally, we also make the size of stakes comparable within each pair of protocols featuring a given number of bidders (two and four). For example, in HC, a subject faces tougher competition than in 2H. To correct for this, we double the exchange rate in HC relative to 2H so that the expected risk-neutral Nash equilibrium payoff of a human bidder is the same in both protocols.[18] Also, the exchange rate used in 4H is the same as the one in 4HR (see Section 2.4 for more details).

[12] Some studies, such as Engelbrecht-Wiggans and Katok (2008), specify the distribution of the computerized bidder's bids indirectly by specifying their value distribution and having them play a symmetric risk-neutral Nash equilibrium strategy. We did not opt for such design since our objective is to make the bid distribution of the computerized opponent as transparent as possible.

[13] Note that, given that the human bidder faces only one as opposed to multiple computerized opponents, she does not need to think about the highest order statistic of the opponents' bids. The question of how feedback about different distributions of bids affects bidding behavior is analyzed in Jehiel et al. (2015).

[14] A variant of this procedure is also adopted by FO.

[15] See Footnote 7 for further comments.

[16] Note that since we focus on a one-shot environment, in the practice round subjects are not matched with an opponent (or a group of opponents) and thus they are not informed how their bid relates to those of other subjects in the practice round. What they are told is the type of feedback they would receive depending on the chosen bid and the hypothetical realization of the highest bid of the opponents.

[17] We recruited men and women separately by sending separate invitation emails and creating two artificial gender-specific sessions for signing up. Potential subjects were not aware of this gender-segregated recruitment since the emails did not indicate any gender preference and they could not observe who else received the email and who signed up for the session. When running the actual session, we made sure to allow an equal number of men and women into the lab, whereas the remaining potential subjects were paid a show-up fee and encouraged to sign up for another session.

[18] We acknowledge that experimental subjects often "overbid" the risk-neutral Nash prediction, implying it would be better to correct for such "deviation" when comparing expected auction payoffs. However, in the absence of a generally accepted theory of "overbidding," we follow a simple approach based on the risk-neutral Nash equilibrium.

2.1. Instructions

In HC, 2H and 4H, subjects were provided with a set of printed instructions (see the online appendix) at the beginning of the experiment. The instructions informed them that they would go through 4 stages (instructions, decision stage, demographic questionnaire, feedback), and explained the auction setting they would face. The decision stage consisted of bidding, eliciting beliefs and eliciting emotions. The subjects were initially told they could earn experimental currency units (ECUs) by winning the auction, and no further details were given at that time as to whether there would be additional opportunities to earn ECUs. In particular, there was no mention of the upcoming beliefs and emotions elicitation. The subjects were also informed about the exchange rate between ECUs and Czech crowns (CZK). At the end of the printed instructions, we asked the subjects to respond to several quiz questions to check their understanding of the instructions. We then checked each subject's answers and any incorrect answers were corrected and an explanation was provided to the subject.[19] Before the actual bidding, subjects had an opportunity to practice submitting a bid in a practice round. The subjects then proceeded to bidding for a real payoff. In 4HR, we followed the procedure of FO. This involved a set of printed instructions with a shorter description of the auction and without any quiz questions. In all protocols, information about the post-auction feedback was presented on a separate page of the instructions, minimizing the probability that subjects would omit reading it.

2.2. Value assignment and bidding

In HC, the subjects were told, in non-technical terms, that each of them was competing against a computerized bidder's bid drawn from $U[0, 100]$. This was framed as the opponent's bid being a number between 0 and 100, including non-integers, with each number being equally likely to be drawn. In 2H and 4H, subjects were randomly and anonymously matched into bidding groups of two and four bidders, respectively. The subjects also knew that all the other subjects in the session, and hence all the potential bidding opponents, received the very same instructions and faced the very same decision environment. In particular, they were explicitly told, using the same framing as above, that the payoff-relevant value(s) of their opponent(s) was (were) drawn from $U[0, 100]$.

In all protocols except for 4HR, we elicited bids by the strategy method in which the subjects bid for six potential values. The six values were drawn from the intervals $[0, 100/6]$, $(100/6, 200/6]$, . . ., $(500/6, 100]$, respectively. We chose this method of value generation so as to have even coverage of the entire value support for each subject. The six values were presented sequentially on separate screens in random order. A subject entered his/her bid on the given screen before proceeding to the next one. After the subjects submitted all six bids, one value-bid pair was chosen at random, each with an equal probability of 1/6, to be payoff-relevant. This means that each subjects' payoff-relevant value was effectively drawn from $U[0, 100]$. Each subject was then informed which pair had been randomly selected for him/her. In 4HR, like FO, we use a strategy method with ten values presented simultaneously on a sheet of paper.[20,21]

In HC, we generated 36 value lists for men and another 36 value lists for women.[22] These two sets of value lists were then used in three sessions of 24 subjects, once for each of the three feedback types (9 sessions altogether). We repeated an analogous procedure in 2H. In 4H, we generated 24 value lists. This set of value lists was then used in each of two sessions of 24 subjects for each feedback type (minimal and loser). Moreover, between the two sessions of the same feedback type, we switched the value lists across genders. In 4HR, by the original protocol of FO, we used only four generated value lists, with each value drawn independently from $U[0, 100]$. Each group of four bidders received these four value lists, with all the values from a given list presented simultaneously on a sheet of paper. Hence, by construction, the set of value lists was identical across treatments. We further strengthened this protocol by switching the value lists across genders between two sessions of 24 subjects of the same feedback type. This way we control for any potential interaction of feedback type and particular values assigned to individual subjects on one hand and gender on the other hand. This reduces noise in the estimates of treatment effects.

In HC, 2H and 4H, subjects were not allowed to overbid their values. In 4HR, following the original protocol of FO, we allowed any non-negative bid.

2.3. Other stages

In HC, 2H and 4H, following the announcement of the payoff-relevant value and bid, but before announcing auction feedback and payoffs, we collected additional data on subject beliefs, emotions and demographics. Instructions for these stages were presented on-screen.

[19] Incorrect answers were infrequent, suggesting good understanding of the instructions.

[20] In the other three protocols, we use six values to reduce the cognitive demand placed on the subjects. We use the sequential presentation because we believe it is more appropriate for focusing subject attention on one bidding situation at a time.

[21] We are not aware of any study that would compare the effect of feedback under the strategy method and under the direct-response method involving one-shot, one-value bidding. More generally, Brandts and Charness (2011) survey the literature on strategy vs. direct-response method for various experiments. They conclude that, in the surveyed studies, if an effect is identified by the strategy method, it is also identified by the direct-response method.

[22] Although the value lists were not identical across the genders by design, the overall empirical distribution of values ended up being almost identical across genders.

Table 1
Numbers of subjects and exchange rates by experimental protocol and treatment.

Protocol	Description	Feedback treatment			Total	Exch. rate (CZK/ECU)
		Minimal	Loser	Winner		
HC	1 human against computer	72	72	72	216	20
2H	2 human bidders	72	72	72	216	10
4H	4 human bidders	48	48		96	25
4HR	4 human bidders (FO repl.)	48	48		96	25
Total		240	240	144	624	

First, subjects were told about an additional opportunity to earn ECUs through reporting their beliefs about the highest bid of their opponent(s). We elicited unconditional beliefs, beliefs conditional on winning and on losing, and also beliefs about the probability of winning/losing. The elicitation was incentivized by quadratic scoring combined with the strategy method in that we paid for only one randomly chosen belief report in order to minimize the possibility of hedging.[23] Second, we collected data on various experienced emotions. Third, we administered a demographic questionnaire in which we collected information about age, country of origin, number of siblings, academic major, the highest achieved academic degree, self-reported risk-tolerance (on a scale of 1–7) and previous experience with online and offline auctions (note that, by the protocol of the sampling procedure, we already knew each subject's gender). In addition, we also collected information on menstrual cycle from female subjects. Of these additional data, we use only information on academic major, the highest achieved academic degree and previous field experience with auctions in this paper.[24]

Finally, the subjects were presented with feedback about the auction outcome (winning vs. losing and, depending on feedback type, further information about the highest bid of the opponent(s)) and their payoffs from the auction and from the belief elicitation procedure.[25]

In 4HR, we followed a slightly different procedure to exactly replicate the protocol of FO.[26]

2.4. Logistics and subjects

Table 1 presents the number of subjects in all protocol-treatment combinations. Altogether, we have data on 624 subjects, of which 216 are in HC, 216 in 2H, 96 in 4H and 96 in 4HR. Across all protocols, 240 subjects are in the minimal feedback treatment, another 240 in the loser feedback treatment and 144 in the winner feedback treatment. The data come from 26 experimental sessions of 24 subjects each.[27] In order to control for potential interactions of feedback type with gender, each session utilized 12 male and 12 female subjects. All subjects in a given session participated in the same protocol and treatment. All the sessions were conducted in the Laboratory of Experimental Economics (LEE) at the University of Economics in Prague (VŠE)). For the three original protocols, HC, 2H and 4H, we used a computerized interface programmed in z-Tree Fischbacher (2007), while 4HR was conducted by paper and pencil. All the sessions, with the exception of 4HR, were conducted in English (this was known to subjects at the time of recruitment).[28] In case of 4HR, we distributed to subjects the original instructions in English taken from FO as well as their translation into Czech.[29]

The subjects were recruited using the Online Recruitment System for Economic Experiments (Greiner, 2015) among students from the University of Economics in Prague and various other universities in Prague. Of all subjects, 49 percent do not hold any degree, 42 percent hold a bachelor's degree, 8 percent hold a master's degree and 1 percent hold a post-graduate degree. Regarding the field of study, 4 percent have a mathematics or statistics major, 9 percent have a science, engineering or medicine major, 70 percent have an economics or business major, 6 percent have a social science major other than economics or business, and 10 percent have a humanities or some other major. Almost 97 percent of our subjects are between 18 and 27 years old, with the remainder being older (up to 39). Also, 43 percent of subjects claim to have had a previous experience with online auctions only, 4 percent with offline auctions only and 5 percent claim experience with both types.

[23] See Blanco et al. (2010) for a discussion of the problems inherent to hedging when eliciting beliefs in experiments.

[24] The belief and emotions data were collected in order to shed light on theories that may underlie feedback effects on bidding in case such effects would exist. The menstrual cycle data was collected in order to continue a line of research started by Chen et al. (2013) and Pearson and Schipper (2013). These studies document that women's menstrual cycle has an impact on their bidding behavior in FPAs. We plan to explore these as well as the remaining demographic data in a companion paper.

[25] In 4H only, following the auction feedback, we re-asked the emotions questions. See our working paper Katuščák et al. (2013) for details.

[26] As for the emotions elicitation in this protocol, see our working paper Katuščák et al. (2013) for details.

[27] There is only one exception to this pattern. Due to an unusually low number of subjects who showed up, we ran one of the HC sessions with 20 subjects and we ran the following HC session with 28 subjects, making up for the four missing subjects in the previous session. Since HC is a protocol based on individual decision-making, without any interaction with other subjects, we believe that this shift does not affect the data for the involved subjects.

[28] The student pool on which we draw consists mainly of Czech and Slovak students with a good command of English. This is in part due to a substantial fraction of university curricula being taught in English in many local universities. This allows routine running of experiments in English.

[29] In 4HR, we wanted to have a further control for language since we did not use our own instructions and thus could not make them as detailed as in the other three protocols.

Table 2
Slopes *(and their standard errors)* of the average bidding functions.

Protocol	Feedback treatment			Treatment differences		
	Minimal (M)	Loser (L)	Winner (W)	L–M	W–M	L–W
HC	0.706 *(0.016)*	0.691 *(0.010)*	0.711 *(0.016)*	−0.015 *(0.019)*	0.005 *(0.021)*	−0.020 *(0.019)*
2H	0.688 *(0.016)*	0.678 *(0.014)*	0.690 *(0.015)*	−0.010 *(0.019)*	0.002 *(0.023)*	−0.011 *(0.021)*
4H	0.756 *(0.022)*	0.785 *(0.014)*		0.029 *(0.026)*		
4HR	0.807 *(0.028)*	0.812 *(0.022)*		0.005 *(0.033)*		

The subjects were paid in cash in Czech crowns (CZK) at the end of their session. Table 1 presents exchange rates used in the four protocols.[30] Under these exchange rates, the risk-neutral Nash equilibrium (RNNE) expected payoff is calibrated to be 167 CZK in HC and 2H and 125 CZK in 4H and 4HR.[31] The actual payoffs from the auction were consistent with overbidding compared to the RNNE and amounted to 102 CZK and 93 CZK in HC and 2H, respectively, and 88 CZK and 35 CZK in 4H and 4HR, respectively. The sessions in HC, 2H and 4H lasted approximately 90 min with an average earning of 380 CZK, of which 150 CZK was the show-up fee.[32] The sessions in 4HR lasted only about 50 min since they were implemented by paper and pencil and we did not collect any beliefs data. These sessions recorded an average earning of 290 CZK, of which 250 CZK was the show-up fee.[33]

For the purpose of comparison with FO, note that the cash value of 1 ECU in our experiment is always at least 0.5 USD (the exchange rate FO used). In fact, in protocols with four bidders, including the replication protocol 4HR, each ECU is worth about 1.3 USD. Although the purchasing power of nominally equivalent amounts changes over time and space, we would argue that the stakes in our study are significantly larger than those used by FO.

3. Results

3.1. All subjects

Fig. 1 presents a scatterplot of bids against values by protocol and feedback treatment, in each case also plotting an OLS estimate of the average linear bidding function with zero intercept.[34] The plots of the bidding function estimates give a clear overall picture: feedback has little effect on the average bidding function. This observation is confirmed by Table 2. The table presents estimates of the slope of the average bidding function by protocol and feedback type and their differences by feedback within protocol. All standard errors are adjusted for clustering at subject level. Moreover, when computing the standard errors for slope differences, we first difference bids within pairs of subjects, one in each of the two treatments, and both of the same gender, facing the same value list. We then compute the clustered standard error from the regression of the paired differences on the respective values, imposing a zero intercept. This method of computing the standard errors increases efficiency in case a particular ordering of values, gender, or their combination, has a systematic effect on bidding.[35] In HC, the estimated slope is almost identical across the three feedback treatments, varying between 0.691 and 0.711. The situation is similar in 2H, with the slope varying between 0.678 and 0.69. Likewise, in 4HR, the slope varies between 0.807 and 0.812 across the two feedback types. There is a somewhat more sizeable difference between the two slopes in 4H only.

[30] The average currency exchange rate over the duration of the experiment was approximately 19 CZK to 1 USD and 25 CZK to 1 EUR.

[31] In HC, conditional on a value of v, the optimal bid is $v/2$. Conditional on winning, which happens with probability $v/200$, the realized surplus is $v/2$. Hence the expected payoff is $v^2/400$. The expected value of this expression is 8.33 ECU. Given the exchange rate of 20 CZK/ECU, this comes out to be about 167 CZK. An analogous computation applies in 2H (with $v/2$ now being the RNNE bid), except that the probability of winning conditional on v is $v/100$. This is exactly compensated for by halving the exchange rate relative to HC, resulting in the same expected payoff. In 4H and 4HR, given v, the RNNE bid is $0.75v$. Conditional on winning, which happens with probability $(v/100)^3$, the realized surplus is $v/4$. Hence the expected payoff is $v^4/(4 \times 100^3)$. The expected value of this expression is 5 ECU. Given the exchange rate of 25 CZK/ECU, this comes out to be 125 CZK.

[32] The overall subject earnings are larger than the sum of the auction payoff and the show-up fee because they also include the payoff for the belief elicitation. Moreover, the resulting raw payoff is rounded up to the nearest multiple of 10 CZK. The rounding up applies to 4HR as well.

[33] For a comparison, an hourly wage that students can earn in research assistant or manual jobs typically ranges from 50 to 100 CZK. As a result, even accounting for the time needed to get to the lab, the rewards are at least competitive with the typical outside options available to our subjects.

[34] For the sake of making the four plots comparable, we have removed one bid of 120 from the plot under 4HR and minimal feedback. However, this bid is accounted for in the estimate of the respective average bidding line. Overall, in 4HR, we observe 45 overbids out of 960 bids. Of these, 20 came from two subjects who overbid for all possible values. Overall, 9 subjects out of 96 overbid for at least one value.

[35] In HC, 2H and 4H, each subject of a given gender within a treatment faces a unique value list, so the pairing of subjects for the purpose of differencing is unique. In 4HR, there are only 4 value lists and 8 combinations of gender and value lists. Hence, the analogous pairing is non-unique. In this case we construct 1,000 random pairings of 48 subjects in the two treatments, always pairing within gender and a particular value list. Each time, we compute the robust variance (square of the standard error) from the regression of the paired differences on the respective values, imposing a zero intercept. We then average these variances across the 1,000 pairings. The standard error presented in Table 2 is the square root of this average.

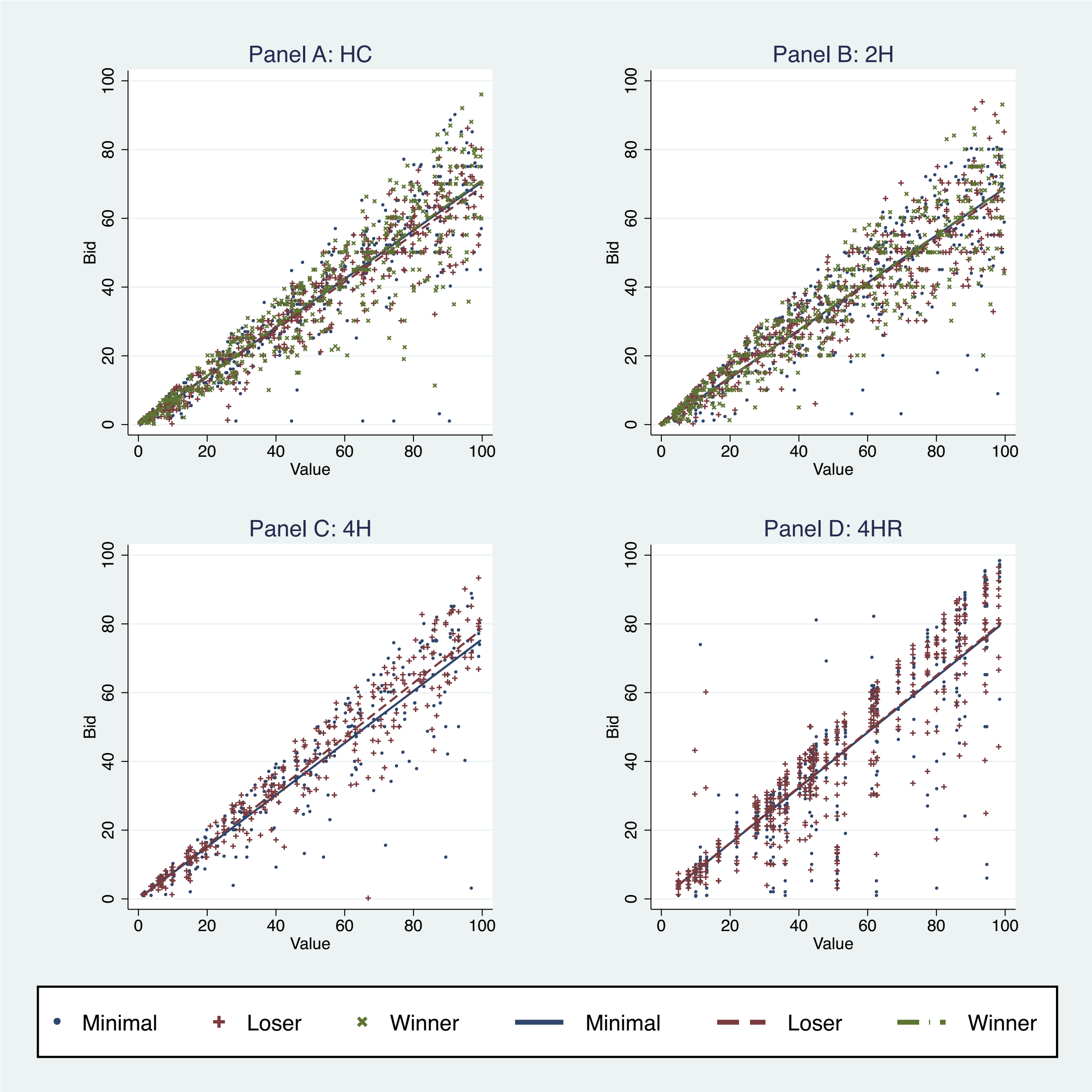

Fig. 1. Scatterplots of bids and average bidding functions.

Under minimal feedback, it is 0.756, whereas under loser feedback it is 0.785. However, neither this treatment difference (p-value 0.268), nor any other one, is statistically significant.[36]

Using the identical approach, the slope of the average bidding function in the FO data is 0.790 (with the standard error of 0.029) in case of minimal feedback, 0.883 (0.011) in case of loser feedback and 0.768 (0.028) in case of winner feedback. There is a statistically significant difference of 0.0931 (0.0305, p-value of 0.003) between the slopes under loser and minimal feedback, but the difference of −0.0217 (0.0399, p-value of 0.589) between the slopes under winner and minimal feedback is statistically insignificant.[37]

[36] All statistical tests performed in this paper are two-tailed.

[37] We computed these results using the FO data with standard errors clustered at subject level. The slope coefficients coincide with the ones reported by FO in Footnote 8.

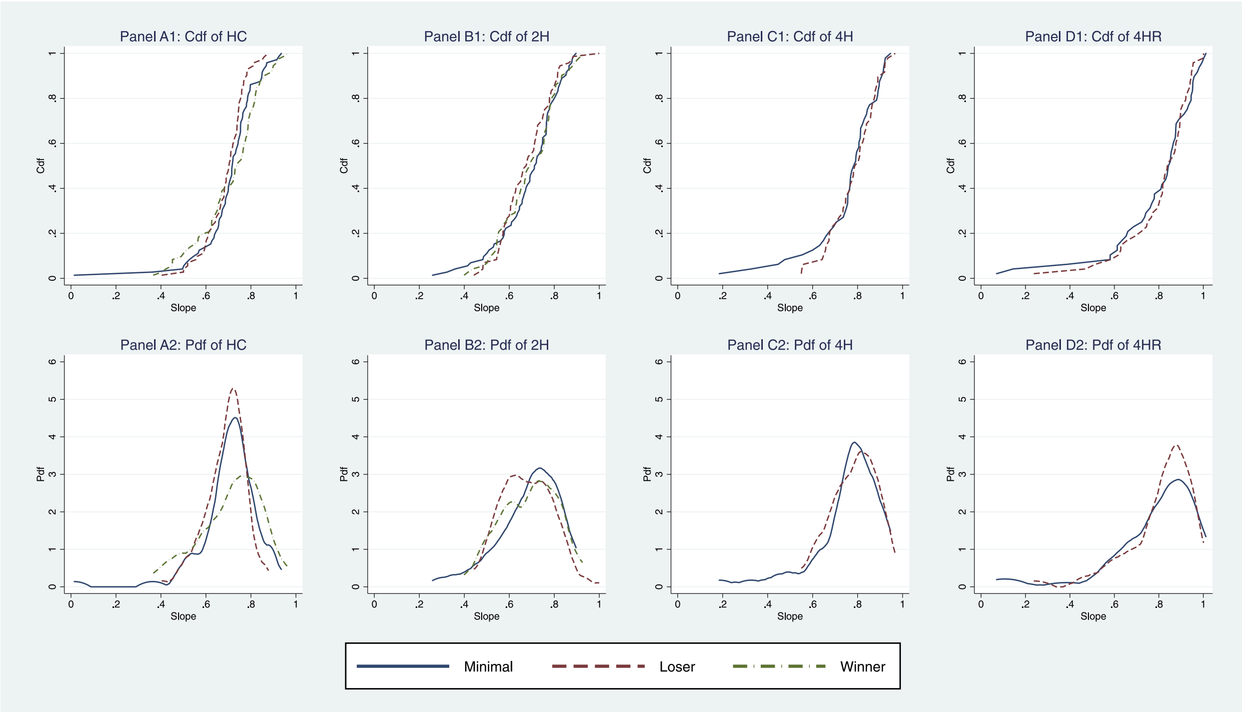

Fig. 2. Distributions of individual bidding function slopes.

Note that if the effect of loser vs. minimal feedback identified by FO (0.0305) were real, then, given our sample size in 4H and 4HR, we would be very likely to reject the null hypothesis of no effect in favor of a positive effect in a two-tailed test. From the *ex ante* perspective, relying on the standard error of 0.0305 based on the FO sample of 64 subjects, our sample size of 96 subjects implies an expected standard error of 0.0249, and hence a rejection probability of approximately 0.96.[38] Analogously, from the *ex post* perspective, relying on our standard errors of 0.0256 in 4H and 0.0329 in 4HR (see the column "L–M" in Table 2, where more rounded values are presented), the rejection probability would be 0.95 and 0.80, respectively.

Although informative, the average bidding functions hide individual heterogeneity in reaction to feedback. We therefore go a step further and estimate the slope of the bidding function for each individual subject using OLS. We assume that this function is linear and has a zero intercept. With v_{ij} denoting the values and b_{ij} denoting the corresponding bids of subject i, with $j \in \{1, 2, \ldots, 6\}$ indexing the order in which the individual values are presented, the estimate of the slope for subject i is given by

$$\widehat{slope}_i = \frac{\sum_{j=1}^{6} v_{ij} b_{ij}}{\sum_{j=1}^{6} v_{ij}^2} = \sum_{j=1}^{6} \left(\frac{v_{ij}^2}{\sum_{k=1}^{6} v_{ik}^2} \right) \frac{b_{ij}}{v_{ij}}. \tag{1}$$

That is, the estimated slope turns out to be a square-value-weighted average of the six (or ten, in 4HR) individual bid/value ratios. We then compare distributions of these slopes across different feedback types within protocol. Fig. 2 plots cumulative distribution functions (top row) and kernel estimates of the respective densities (bottom row) of the empirical distributions of the slopes by protocol and feedback treatment. The figure reveals the same overall picture: feedback has little impact on the mean of the slope distribution.

To test this null hypothesis, Table 3 presents estimates of means and treatment effects on means of these distributions together with their standard errors obtained analogously to Table 2 (except that no clustering is necessary in this case). The means are very close to the means of the average bidding function slopes presented in Table 2. Looking at the estimated treatment effects and their standard errors reveals that feedback type has no significant effect on means of the slope distributions in any of the four protocols. This confirms the visual observation drawn from Fig. 2.

[38] Since our sample size is 50 percent larger than the one used by FO (96 vs. 64), the expected standard error on the treatment effect is $0.0305/\sqrt{1.5} \doteq 0.0249$ with our sample size. In turn, with 96 data clusters and two degrees of freedom used to compute the slope coefficients, the two-tailed critical value for the rejection of the null hypothesis in favor of a positive effect is given by $t_{94,0.025} \times 0.0249 \doteq 0.0494$. The probability of such rejection if the FO effect were real would be $1 - T_{94}[(0.0494 - 0.0931)/0.0249] \doteq 0.959$, where $T_{94}(\cdot)$ is the cumulative distribution function of Student's t-distribution with 94 degrees of freedom.

Table 3
Means *(and their standard errors)* of the bidding function slope distributions.

Protocol	Feedback treatment			Treatment differences		
	Minimal (M)	Loser (L)	Winner (W)	L–M	W–M	L–W
HC	0.706	0.690	0.711	−0.015	0.005	−0.020
	(0.016)	*(0.010)*	*(0.016)*	*(0.019)*	*(0.020)*	*(0.019)*
2H	0.688	0.677	0.689	−0.011	0.001	−0.012
	(0.016)	*(0.013)*	*(0.015)*	*(0.019)*	*(0.023)*	*(0.021)*
4H	0.757	0.783		0.026		
	(0.022)	*(0.015)*		*(0.026)*		
4HR	0.792	0.811		0.019		
	(0.028)	*(0.022)*		*(0.033)*		

Table 4
Interquantile ranges *(and their standard errors)* of the bidding function slope distributions.

Prot.	Statistic	Feedback treatment			Treatment differences		
		Min. (M)	Loser (L)	Winner (W)	L–M	W–M	L–W
HC	Q_{75}–Q_{25}	0.111	0.102	0.178	−0.009	0.067**	−0.076**
		(0.017)	*(0.019)*	*(0.025)*	*(0.025)*	*(0.030)*	*(0.031)*
	Q_{90}–Q_{10}	0.284	0.194	0.362	−0.090	0.079	−0.168***
		(0.046)	*(0.030)*	*(0.048)*	*(0.058)*	*(0.069)*	*(0.053)*
2H	Q_{75}–Q_{25}	0.154	0.169	0.187	0.015	0.033	−0.018
		(0.029)	*(0.020)*	*(0.029)*	*(0.036)*	*(0.040)*	*(0.036)*
	Q_{90}–Q_{10}	0.351	0.265	0.304	−0.086	−0.047	−0.039
		(0.042)	*(0.027)*	*(0.033)*	*(0.051)*	*(0.050)*	*(0.043)*
4H	Q_{75}–Q_{25}	0.108	0.144		0.037		
		(0.040)	*(0.027)*		*(0.050)*		
	Q_{90}–Q_{10}	0.361	0.270		−0.090		
		(0.090)	*(0.042)*		*(0.104)*		
4HR	Q_{75}–Q_{25}	0.191	0.133		−0.058		
		(0.047)	*(0.036)*		*(0.059)*		
	Q_{90}–Q_{10}	0.392	0.328		−0.064		
		(0.113)	*(0.060)*		*(0.126)*		

Note: In the tests for treatment differences, results are significant at: **5 percent level, ***1 percent level.

The fact that feedback type has no significant effect on the average bidding behavior, as represented by the means of the slope distributions, is the main conclusion of our analysis. However, in principle, feedback type might affect other moments of the slope distribution. The variance is of particular interest. To see why, consider the implications of increased variance, given the mean, for auction efficiency. The higher the variance is, the higher the likelihood of the object not being allocated to the highest-value bidder is. As a result, expected auction efficiency (the winner's value as a fraction of the maximum value) is reduced.

A glance at Fig. 2 provides a visual evaluation of the null hypothesis that variance of the slope distributions does not change with feedback type. In particular, it appears that the distribution of slopes is more concentrated under loser feedback than under minimal feedback in HC and 4HR, whereas the latter is more concentrated than the distribution under winner feedback in HC. The difference is most profound between loser and winner feedback in HC. On the other hand, with the possible exception of the lower tail under minimal feedback in 4H, there does not appear to be any significant difference in how concentrated the various distributions under 2H and 4H are.

To formally test for these effects, Table 4 presents estimates of two interquantile ranges (IRs) of the slope distributions and treatment effects on these IRs, together with their standard errors. The first IR is the difference between the 75th and the 25th percentile (Q_{75}–Q_{25}). The second IR is the difference between the 90th and the 10th percentile (Q_{90}–Q_{10}). The standard errors are obtained by bootstrapping with 1000 replications, with the individual bootstrap draws clustered at the level of value lists and gender. That is, if a subject with a particular value list and gender is drawn under one treatment, the corresponding subjects are also drawn for the other (two) treatment(s). Moreover, all bootstrap draws are stratified at gender level, meaning that each bootstrap draw contains the same number of men and women within each feedback

Table 5
Pairwise tests for the equality of bidding function slope distributions.

Protocol	Test	p-Values of treatment differences		
		Loser–minimal	Winner–minimal	Loser–winner
HC	Wilcoxon	0.142	0.595	0.064
	K–S	0.213	0.146	0.001
2H	Wilcoxon	0.228	0.823	0.487
	K–S	0.419	0.947	0.304
4H	Wilcoxon	0.747		
	K–S	0.993		
4HR	Wilcoxon	0.918		
	K–S	0.934		

treatment.[39] The clustering and stratification are implemented in order to minimize the amount of noise contained in the individual bootstrap realizations of the statistics of interest.

The table shows that feedback type has no significant effect on either of the IRs, with the exception of HC. In that protocol, the distribution of slopes is more dispersed under winner feedback in comparison to loser feedback (the two IR differences are significant at 5 and 1 percent level, respectively). Also, at the level of the interquartile range only, the distribution of slopes is more dispersed under winner feedback in comparison to minimal feedback. This formal evidence confirms some of our visual observations on the comparison of distribution dispersions drawn from Fig. 2.

Table 5 presents p-values of two types of tests of equality of slope distributions between pairs of feedback treatments within each protocol. We use Mann–Whitney ranksum test and Kolmogorov–Smirnov test. None of the pairs of distributions are statistically significantly different at conventional levels with the exception of loser and winner feedback in HC (respective p-values of 0.064 and 0.001). These findings are consistent with those in Fig. 2 and Tables 3 and 4. In particular, the spread of the slope distribution rather than its mean seems to be responsible for the difference of the slope distributions under loser and winner feedback in HC.

In contrast to our approach, FO compute the slope of each individual subject's bidding function as the simple average of the ten recorded bid/value ratios. They find that the average slope under loser feedback is significantly higher than under minimal feedback, but that there is no significant difference between the average slopes in the minimal and winner feedback treatments. To verify that the difference between their results and ours is not driven by different methodology, we apply our methodology to their data and obtain results that are qualitatively equivalent to their reported findings. In particular, the average slope of the bidding function is 0.797 (with the standard error of 0.027) under minimal feedback, 0.883 (0.010) under loser feedback, and 0.774 (0.029) under winner feedback. The difference between the average slope under loser and minimal feedback is 0.0858 (0.0287, p-value of 0.004) and it is statistically significant, whereas the difference between the average slope under winner and minimal feedback is −0.0231 (0.0391, p-value of 0.557) and it is not statistically significant.[40]

Note that if the effect of loser vs. minimal feedback identified from the FO data (0.0858) were real, then, given our sample size in 4H and 4HR, we would be quite likely to reject the null hypothesis of no effect in favor of a positive effect in a two-tailed test. From the *ex ante* perspective, relying on the standard error on the estimate based on the FO data (0.0287), with our sample size of 96, the expected rejection probability would be approximately 0.95. Analogously, from the *ex post* perspective, relying on our standard errors of 0.0263 in 4H and 0.0331 in 4HR (see the column "L–M" in Table 3, where more rounded values are presented), the rejection probability would be 0.90 and 0.73, respectively.[41]

FO also compare the slope distributions under different feedback treatments. They report that the distribution of the slopes under loser feedback first-order stochastically dominates the distribution under minimal feedback, but there is no significant difference in the distributions between the minimal and winner feedback treatments. They do not, however, report results of any formal statistical tests to support these conclusions. Using their data and our construction of the slopes, we test for any distribution differences using Mann–Whitney and Kolmogorov–Smirnov tests. We indeed find that the distributions under minimal and winner feedback are not significantly different, whereas the distributions under minimal and loser feedback are (p-values of 0.004 and 0.015, respectively).

To summarize, our results, as opposed to those of FO, document that feedback type has little or no effect on the mean of bid/value ratios. This conclusion is robust to all four protocols that we employ. In addition, the type of feedback has a weak effect on the heterogeneity of slopes of bidding strategies, with this heterogeneity being smallest under loser feedback. Regarding heterogeneity, FO make a similar observation in their paper.

[39] In HC, 2H and 4H, each subject of a given gender within a treatment faces a unique value list, so the process of clustering is straightforward. In 4HR, there are only 4 value lists and 8 combinations of gender and value lists. In each bootstrap draw we therefore first randomize the order of subjects within each combination of value-list and gender and then proceed as in the previous three protocols.

[40] We computed these results using the FO data, with robust standard errors.

[41] The methodology of power calculations is analogous to the one presented in Footnote 38.

3.2. Experienced subjects

In the previous section, we found that feedback had little impact on the average bid/value ratios. From the point of view of auction design, this conclusion suggests that feedback is irrelevant for bidding and revenue in one-shot FPA auctions. However, as we discussed in the introduction, several studies (Isaac and Walker, 1985; Ockenfels and Selten, 2005; Neugebauer and Selten, 2006; Neugebauer and Perote, 2008; Engelbrecht-Wiggans and Katok, 2008) document that feedback does matter in later rounds of repeated bidding, suggesting that previous experience with the same type of auction and its feedback might change the way bidders react to it. To the extent that bidders in many one-shot real-world auctions may have experience from previous auctions of a similar kind, our suggested conclusion may not be robust to the level of bidders' experience.[42]

We address this concern by investigating the robustness of our null findings in Tables 2 and 3 to the self-reported level of subject field experience with auctions. Our post-experiment questionnaire includes questions on whether a subject has ever been bidding in an online or in an offline auction. Based on these two questions, we split our subjects into experienced ones (answering "yes" to at least one of the two questions) and inexperienced ones (the rest). Of all 624 subjects, 333 are classified as experienced, 290 as inexperienced and we are not able to classify one subject due to not answering either of the two questions.[43] If previous, but non-immediate, experience matters for the effect of feedback, we might be able to observe some effects of feedback for our experienced subjects.

To test this hypothesis, we rerun the tests presented in Tables 2 and 3 for experienced subjects only.[44] All the qualitative findings are unchanged: feedback type has no significant impact on the average bid/value ratios.[45] Admittedly, our indicator of experience includes *any* previous field experience with auctions, some of which might not be similar to FPA (for example, many bidding mechanisms utilize the English auction). Also, we do not know how recent such experience may be. As a result, the test is arguably not very powerful.

3.3. Other robustness checks

One could raise an objection that bidding data for low values is noisier than that for high values in that the probability of winning is relatively small and hence subjects think less carefully about how much to bid. However, as we have already pointed out, the estimated slopes of the individual bidding functions are square-value-weighted averages of the individual bid/value ratios (see Eq. (1)). As a result, these slope estimates are already significantly weighted toward bid/value ratios for higher values. Nevertheless, in order to check the robustness of the null baseline findings to excluding low values, we repeat all the tests presented in Tables 2 and when the data is restricted to only the upper half of values, then the two highest values, and finally to only the highest value for each bidder and the corresponding bid(s). Qualitatively, we obtain the same results as those presented in Section 3.1.[46]

In comparison to FO, we use a different subject pool. Our subject sample is drawn from students from several universities in Prague, whereas FO utilize students from New York University. Our sample consists of both undergraduate and more advanced students in approximately equal proportions, whereas FO sample only undergraduate students. Also, about two thirds of our subjects have economics or business major (most of them being students of the University of Economics in Prague). We are not aware of the academic major distribution in the FO sample,[47] but if they had drawn a representative sample from the population of NYU undergraduate students, we would expect that their share of economics or business majors would be much smaller than ours. Although we are not able to evaluate the effect of different geographical location on the results, we are able to evaluate the effects of both student seniority and academic major, since we inquire about both in our demographic questionnaire. We rerun all the tests in Tables 2 and 3 separately for undergraduate and more advanced students, but do not find any significant effects of feedback on average bid/value ratios in either group.[48] We also separately analyze bidding behavior of subjects with economics or business major and other subjects. Again, we do not find any robust feedback effects in either group.[49,50] We therefore conclude that if the difference between our results and the results of FO is driven by subject pool differences, it does not appear to be driven by student seniority or academic major.

---

[42] There is some evidence from other domains that subject experience can affect behavior. See, for instance, List (2003).

[43] We have also considered approximating auction experience with the number of times a subject has participated in an auction experiment in the Laboratory of Experimental Economics before. However, only one such experiment was run over the previous several years. Only 18 men and 10 women from our overall sample have participated in that experiment. This provides an insufficient degree of variation for identifying the effect of this type of experience on the reaction to feedback.

[44] Because of a non-balanced presence of experienced subjects in different sessions and treatments, we resort to using a simple regression-based standard error adjustment for clustering (in case of bids) and heteroscedasticity (in case of slopes).

[45] Detailed results are available from the authors upon request.

[46] Detailed results are available from the authors upon request.

[47] FO informed us in a private conversation that they did not collect such information.

[48] If anything, the undergraduate students tend to bid more under winner as opposed to loser feedback, but only in HC.

[49] If anything, the non-economics/non-business students tend to bid less under loser feedback in comparison to minimal feedback, but only in HC.

[50] Because of non-balanced presence of the two respective sub-groups in different sessions and treatments, we resort to using a simple regression-based adjustment for clustering (in case of bids) and heteroscedasticity (in case of slopes) when computing the standard errors.

Even though our results suggest there are no feedback effects on the mean of the bid/value ratio, there is a possibility that the type of feedback may affect bidding in ways that do not easily manifest themselves in the average bid/value ratio, but may instead affect other features of bidding behavior. For example, our results show that, in HC, winner feedback leads to a larger heterogeneity of slopes of the individual bidding functions in comparison to loser feedback, especially in the tails of the respective distributions. Such differences may have implications for average revenue or efficiency, the objectives that an auctioneer ultimately cares about. For example, a larger dispersion might increase auction revenue by shifting the distribution of the highest bid to the right. On the other hand, a smaller dispersion might increase auction efficiency by increasing the probability that the auction allocates the object to (one of) the highest value bidder(s). We investigate such possibilities in our working paper Katuščák et al. (2013) and do not find any significant feedback effects with the exception of 4H, where efficiency is higher under loser feedback in comparison to minimal feedback, but the effect is quantitatively small. As a result, in terms of these two ultimate auctioneer objectives, we mirror our conclusion of little feedback effect on bidding.

4. Conclusion

This paper presents new evidence on the important market design issue of whether an auctioneer in first-price auction can induce bidders to bid more aggressively by controlling the type of feedback that bidders receive after the auction. Several studies (Isaac and Walker, 1985; Ockenfels and Selten, 2005; Neugebauer and Selten, 2006; Neugebauer and Perote, 2008; Engelbrecht-Wiggans and Katok, 2008) suggest that this is possible in a repeated bidding environment. However, since many real-world auctions are non-repeated, it is interesting to answer the question specifically within the one-shot environment.

The studies mentioned above do not provide a clear evidence for the first round of bidding. So far the only direct evidence on the effect of feedback on bidding in one-shot auctions is due to Filiz-Ozbay and Ozbay (2007) (FO). They find that bidders bid more relative to their values under loser as opposed to minimal (or winner) feedback. We revisit this finding with a more articulated multi-dimensional design that employs a larger sample size. In comparison to minimal feedback, we do not find any systematic effect of loser or winner feedback on the average bid/value ratio in any of our protocols, including the protocol that replicates the protocol used by FO.

In comparison to FO, we control for gender composition of the sample, use larger stakes and use a much larger sample size. Therefore, the sampling noise contained in our estimates should be smaller than the one contained in theirs in a given subject population. Also, larger stakes should imply a stronger external validity of our results. Although the difference in the result on the effect of loser vs. minimal feedback between our study and FO could be due to the subject pool difference, the tests that we run regarding known differences between the two subject pools (student seniority and academic major) do not reveal any significant differences in subjects' bidding behavior. Likewise, our conclusions are robust to using only subjects with some field experience with auctions.

We conclude that in one-shot auctions with two and, in case of minimal and loser feedback, four bidders, using loser or winner as opposed to minimal feedback has no systematic effect on the average bid/value ratio. From the point of view of auction design, our results imply that such feedback manipulation is unlikely to be useful in increasing revenue in one-shot first-price auctions. Whether other auction environments, in connection with loser or winner or, potentially, alternative feedback types, are capable of systematically affecting the bid/value ratios in one-shot environments is an open question for future research.

Acknowledgements

Special thanks go to Tomáš Miklánek for his excellent job with programming and organizing the experiment. We would also like to thank Emel Filiz-Ozbay for sharing relevant information regarding the design in Filiz-Ozbay and Ozbay (2007), as well as for comments on our manuscript. We are grateful to our editor, associate editor and several anonymous referees for comments and suggestions that improved the content and clarity of presentation. We also thank Michal Bauer, Avner Shaked, Rani Spiegler and Jakub Steiner for helpful discussions and comments. Finally, we thank seminar and conference participants at the University of Arkansas, Bratislava Economic Meeting 2012, the 1st ESEI Market Design Workshop in Prague, University of Siena, University of Alicante, 2013 Florence Workshop on Behavioral and Experimental Economics, CERGE-EI, M-BEES 2013, ESA World Meeting 2013 and the 8th Nordic Conference on Behavioral and Experimental Economics. This study was supported by the Grant Agency of the Czech Republic (GAČR 402/11/1726) and by the CERGE-EI Foundation under a program of the Global Development Network (GDN 11/014), and resources of the Laboratory of Experimental Economics (LEE) at the University of Economics in Prague (VŠE). All opinions expressed are those of the authors and have not been endorsed by GAČR, GDN, CERGE-EI or LEE/VŠE.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2015.08.002.

References

Bell, D.E., 1982. Regret in decision making under uncertainty. Oper. Res. 30 (5), 961–981.
Blanco, M., Engelmann, D., Koch, A.K., Normann, H.T., 2010. Belief elicitation in experiments: is there a hedging problem? Exp. Econ. 13, 412–438.
Brandts, J., Charness, G., 2011. The strategy versus the direct-response method: a first survey of experimental comparisons. Exp. Econ. 14 (3), 375–398.
Chen, Y., Katuščák, P., Ozdenoren, E., 2013. Why can't a woman bid more like a man? Games Econ. Behav. 77 (January (1)), 181–213.
Dufwenberg, M., Gneezy, U., 2002. Information disclosure in auctions: an experiment. J. Econ. Behav. Organ. 48, 431–444.
Engelbrecht-Wiggans, R., 1989. The effect of regret on optimal bidding in auctions. Manag. Sci. 35 (6), 685–692.
Engelbrecht-Wiggans, R., Katok, E., 2008. Regret and feedback information in first-price sealed-bid auctions. Manag. Sci. 54 (April (4)), 808–819.
Filiz-Ozbay, E., Ozbay, E.Y., 2007. Auction with anticipated regret: theory and experiments. Am. Econ. Rev. 54 (April (4)), 1407–1418.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1 (July (1)), 114–125.
Isaac, M.R., Walker, J.M., 1985. Information and conspiracy in sealed bid auctions. J. Econ. Behav. Org. 6, 139–159.
Jehiel, P., Katuščák, P., Michelucci, F., 2015. How to Boost Revenues in First-price Auctions? The Magic of Disclosing Only Winning Bids from Past Auctions. CERGE-EI, mimeo.
Kagel, J.H., Levin, D., 2015. Auctions: a survey of experimental research. In: Kagel, J., Roth, A. (Eds.), The Handbook of Experimental Economics, vol. 2, forthcoming.
Katuščák, P., Michelucci, F., Zajíček, M., 2013. Does Anticipated Regret Really Matter? Revisiting the Role of Feedback in Auction Bidding. CERGE-EI Working Paper No. 487, Available from: http://www.cerge-ei.cz/pdf/wp/Wp487.pdf
List, J.A., 2003. Does market experience eliminate market anomalies? Q. J. Econ. 118 (1), 41–71.
Loomes, G., Sudgen, R., 1982. Regret theory: an alternative information theory of rational choice under uncertainty. Econ. J. 92 (368), 805–824.
Myerson, R.B., 1981. Optimal auction design. Math. Oper. Res. 6 (February (1)), 58–73.
Neugebauer, T., Perote, J., 2008. Bidding 'as if' risk neutral in experimental first price auctions without information feedback. Exp. Econ. 11, 190–202.
Neugebauer, T., Selten, R., 2006. Individual behavior of first-price auctions: the importance of information feedback in computerized experimental markets. Games Econ. Behav. 54, 183–204.
Ockenfels, A., Selten, R., 2005. Impulse balance equilibrium and feedback in first price auctions. Games Econ. Behav. 51 (1), 155–170.
Pearson, M., Schipper, B.C., 2013. Menstrual cycle and competitive bidding. Games Econ. Behav. 78 (2), 1–20.
Ratan, A., 2013. Is Decision Regret relevant for Auctions? A Reconsideration of Anticipated Regret in First Price Auctions. Monash University, mimeo.
Riley, J.G., Samuelson, W.F., 1981. Optimal auctions. Am. Econ. Rev. 71 (June (3)), 381–392.