

Does Anticipated Regret Really Matter? Revisiting the Role of Feedback in Auction Bidding*

Peter Katuščák[†] Fabio Michelucci[‡] Miroslav Zajíček[§]

April 20, 2013

Abstract

Does the type of post-auction feedback affect bidding behavior in the first price auction? Filiz-Ozbay and Ozbay (2007) find that such manipulation can increase bids in a one-shot auction. They explain their finding by anticipated regret combined with the assumption that feedback directly affects salience of regret relative to material payoff. We revisit this important market design issue using four different auction protocols and a large sample of subjects. We do not find any systematic effect of feedback on the average bid/value ratio. This evidence contradicts a presence of anticipated regret or its manipulability by feedback in one-shot auctions.

Keywords: auctions, bidding, feedback, regret

JEL Classification: C91, C92, D03, D44

*We would like to thank Avner Shaked, Rani Spiegler and Jakub Steiner for helpful discussions on the design of the experiment. Special thanks go to Tomáš Miklánek for his excellent job with programming and organizing the experiment. We would also like to thank Lei Ba, Oleksandra Pokhylko and, especially, Michal Bauer for comments on the manuscript. This study was supported by the Grant Agency of the Czech Republic (GACR 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 (UEP). All opinions expressed are those of the authors and have not been endorsed by GACR, GDN, CERGE-EI or LEE/UEP.

[†]CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Academy of Sciences of the Czech Republic, Politických vězňů 7, 111 21 Prague, Czech Republic, Peter.Katuscak@cerge-ei.cz.

[‡]CERGE-EI, Fabio.Michelucci@cerge-ei.cz (corresponding author).

[§]University of Economics in Prague, Winston Churchill Sq. 4, 130 67 Praha 3, Czech Republic, miroslav.zajicek@vse.cz.

1 Introduction

Consider a first-price sealed-bid auction with private values (FPA). At the end of the auction, each bidder learns whether she won or not. This is labeled as *minimal feedback*. The auctioneer may, however, give bidders additional feedback and announce this fact before the bidding starts. For example, he may publicly announce the winning bid (*loser feedback*). Or, alternatively, he may inform the winner about the second highest bid (*winner feedback*). The auctioneer might employ these (or other) alternative feedback types if they help him to achieve his objectives, such as expected revenue or efficiency maximization. However, auction theory based on standard preferences predicts that, in an equilibrium of a single-round auction, an *ex ante* known form of post-auction feedback, in excess of the minimal feedback, has no impact on bidding, and hence on expected revenue or efficiency.¹ Contrary to this prediction, Filiz-Ozbay and Ozbay (2007) (referred to as FO in what follows) find that bidders bid more aggressively (higher bid/value ratios) on average under loser as opposed to minimal or winner feedback.

FO interpret this finding using the theory of anticipated regret originally developed by Bell (1982) and Loomes and Sudgen (1982), and first applied to auctions by Engelbrecht-Wiggans (1989). This theory introduces one or two additively separable welfare-lowering regret components into the utility function besides the usual concern for material payoff. *Ex post*, a bidder experiences regret if, in light of the available information, a different bid would have made him/her materially better off. In case of losing and receiving loser feedback, this would be a bid in excess of the winning bid but below one's value for the object (if such bid is feasible). In case of winning and receiving winner feedback, this would be a lower bid that would still be higher than the second-highest bid. The regret of foregone material payoff (if any) in the former case is called *loser regret*, whereas in the latter case it is called *winner regret*. Both types of regret are assumed to be (weakly) convex in foregone payoffs. Because such regret disutilities are anticipated *ex ante*, a bidder chooses his/her bid so as to optimally trade-off material payoff and regret considerations in expectation. For example, to reduce expected loser (winner) regret, a bidder would bid more (less) than in the absence of such regret consideration. A bidder may place different weights on the two types of regret, which includes a possibility that only one of the two regret types is present in the utility function.²

However, such modification of theory is not sufficient to generate the prediction that loser/winner feedback makes bidders bid more/less aggressively. With loser (winner) feedback, the amount of regret that is realized *ex post* is rationally expected *ex ante*. Under min-

¹See, for example, Krishna (2002).

²Turocy and Watson (2012) provide experimental evidence consistent with the hypothesis that loser regret is more important than winner regret. This is because the FPA auction, as typically framed in the laboratory, faces bidders with only two possible outcomes: winning the auction or walking away empty-handed (with the exception of a show-up fee). Indeed, when not winning the auction does not automatically imply a zero payoff (because of an outside option to buy at a fixed price), the authors show that bidders bid less aggressively.

imal feedback, it is not exactly realized even *ex post*. Rather, it is expected, conditional on winning or losing the auction. Hence, *ex ante*, one has expectations over expectations under minimal feedback. However, by the Law of Iterated Expectations, the ultimate *ex ante* anticipated loser or winner regret is the same in either case for a given bid. So in order to explain why feedback affects bidding, FO additionally assume that the form of feedback *directly* manipulates the relative weights placed on the two types of regret, as opposed to the material payoff, in the utility function.³ In particular, they assume that the weight a bidder places on loser (winner) regret increases under loser (winner) feedback in comparison to minimal feedback. The intuition is that loser (winner) feedback makes loser (winner) regret more salient relative to the material payoff (and the other type of regret) at the time of determining the bid. We call this the *salience assumption*. Note that feedback-manipulated salience of loser and winner regret does not need to be the same. In particular, the finding of FO that only loser (as opposed to minimal) feedback is effective in changing bids can be explained by restricting the salience assumption only to loser regret (or, alternatively, by the absence of winner regret considerations in the utility function altogether).

The finding of FO and its theoretical interpretation are of great relevance for auction design, and they also raise new questions. If a simple and nearly costless manipulation of posterior feedback can generate higher bids and more revenue, auctioneers should adopt it widely in practical applications. Moreover, one can further inquire about what type of feedback would be the most effective in stimulating anticipated regret and hence maximizing the expected auction revenue. Finally, an analogous manipulation of feedback might work even in market institutions other than auctions, and this possibility calls for further empirical evidence.

However, before applying the anticipated regret theory with the salience assumption (henceforth “the theory”) to a wide variety of applications, we think it is necessary to further examine its validity in a simple decision-theoretic environment. Let us highlight the reasons why we think the theory may not be as robust as it might first seem. First, even if some bidders may be regret-prone *ex post*, they might not correctly anticipate such feelings *ex ante* in a one-shot auction. It might take a few previous regret experiences in the same auction environment for bidders to learn about their posterior regret feelings and hence to start anticipating them in subsequent auctions. This intuition is supported by results of studies that manipulate feedback in a repeated bidding environment (Ockenfels and Selten 2005, Neugebauer and Selten 2006, Engelbrecht-Wiggans and Katok 2008, Neugebauer and Perote 2008). These studies identify a significant positive effect of loser and, less robustly, a negative effect of winner feedback on bids in later auction rounds, whereas there are no such effects in the first auction round.⁴ Ockenfels and Selten (2005) and Neugebauer and Selten (2006) develop a di-

³Engelbrecht-Wiggans (1989) informally suggests the same assumption.

⁴The earliest study we are aware of on the effect of feedback in a repeated first-price auction is Isaac and Walker (1985). However, the authors do not present a separate analysis of bidding in the first round.

rection learning theory, called *impulse balance theory*, to explain these findings. Indeed, one way to interpret this theory is that bidders learn to anticipate their regret feelings from regret experiences in the previous rounds of bidding and then apply these “learned” preferences in the following rounds.

Second, it is not apparent why, for example, a provision of loser as opposed to minimal feedback makes a bidder more concerned about loser regret. The intuition for this is based on the iterated expectation argument we discussed earlier. Indeed, receiving loser feedback might result in a more painful regret experience from a narrow-margin loss. But it might also result in relieving the bidder of regret concerns in case of a big-margin loss. It is therefore not clear why bidders would place a larger weight on loser regret relative to material payoff under loser feedback. An analogous argument can also be made about winner vs. minimal feedback. Moreover, we consider the salience assumption to be less than fully satisfactory from the point of view of theoretical modeling since the theory is vacuous as to how one should be guided to think about the impact on anticipated regret of alternative feedback types.

Our initial experimental plan includes two designs. The first design, denoted HC, aims to test the theory 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. This design eliminates any strategic uncertainty about the bidding strategy used by the opponent, particularly how this strategy might be affected by the type of feedback. We would argue that this is the cleanest environment and the natural benchmark to test the theory. Since the theory is based on feedback-driven preference shifts, we should observe that, in comparison to minimal feedback, loser feedback makes bidders bid more aggressively, whereas the opposite (perhaps less strongly in light of the findings of FO) is true of winner feedback.

The second design, denoted 2H, is an auction with two *ex ante* symmetric human bidders. This is a natural extension of HC to the simplest possible environment with strategic uncertainty. In this setting, the theory predicts that the effect of feedback operates not only through the direct preference channel, but also through an indirect channel of beliefs about how feedback affects the bidding strategy of the opponent (plus all the higher order beliefs). The comparison with HC can highlight the impact of these indirect effects. To assess if any eventual difference can be attributed to difference in these beliefs between the two treatments, we also elicit subject beliefs about the opponent’s bid (see Subsection 2.3 for further details). The theory predicts that the symmetric equilibrium bidding strategy is affected in the same direction as in HC (see FO, remarks 1 and 3). In both of these two designs, 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 designs. This indicates that either bidders are not affected by winner and loser regret, or the salience assumption does not apply. Regarding the effect of winner vs. minimal feedback, this finding is consistent with the finding and interpretation of FO.

However, regarding the effect of loser vs. minimal feedback, our finding and the message for the theory we draw from it contradicts their conclusions.

Since FO use four human bidders, our findings from the two-bidder auctions (HC and 2H) might suggest that either the presence of loser regret or the effectiveness of the salience assumption might be sensitive to the level of bidding competition. In our third design, denoted 4H, we test this hypothesis by conducting an auction with four human bidders that otherwise uses the same procedure as 2H. In light of our motivation, we implement only two feedback types: minimal and loser. 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 and the implications for the theory we have drawn from our first two designs do not appear to be sensitive to the level of bidding competition, at least not in the range we investigate.

Vis-à-vis the findings in FO, this result presents a puzzle since 4H is very similar to the design used by FO. 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 such possibility, we conduct the fourth design, denoted 4HR, which 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 and their message for the theory do not appear to be sensitive to details of the utilized experimental procedure.

The outcome of this latter robustness check makes the discrepancy between our empirical findings and the findings of FO even more striking. What can account for the difference? Despite going all the way to replicating 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 economics or business major, whereas FO utilize undergraduate students from New York University (of major distribution we are not aware of). Second, unlike FO, in each session we use an equal number of men and women and make feedback orthogonal to the exact assignment of bidder values within each gender. This is done in order to preclude any omitted variable bias or sampling noise driven by a potential interaction between feedback type and gender.⁵ Third, we use arguably higher stakes than FO do (see Subsection 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. The gender control, the higher stakes and the larger sample size all contribute to reducing noise in our estimates in comparison to the estimates of FO. Hence our failure to replicate their finding is unlikely to be driven by sampling noise. This leaves the subject pool difference as a possible explanation.

⁵We have asked FO about the gender composition of their dataset, but they did not record this information.

To shed light on this hypothesis, we carry out a separate analysis of bids of undergraduate and more advanced students, and also of economics or business major students and students with other majors. In neither case we find any systematic effects of feedback on bidding. Of course, there are other (unobservable) differences between the two subject pools that might account for the difference. One way or another, if there is an effect of feedback on bidding, it appears to be dependent on place and characteristics of bidders.⁶

Overall, our results suggest that the type of feedback does not have any systematic effect on the average bid/value ratio in one-shot FPA auctions. Moreover, we extend our investigation into auction revenue and efficiency, and obtain analogous conclusions. Our results therefore do not support the notion that the average aggressiveness of bidding can be manipulated by posterior feedback provision, at least not by switching among the three types of feedback we consider. This conclusion, when looked at through the lens of the anticipated regret theory with the salience assumption, implies that either anticipated regret is not an important component of bidder decision utility⁷ in one-shot auctions or the salience assumption does not apply. Following FO, we also collect anticipated emotions data in all of our designs. On the one hand, we find that loser as opposed to minimal feedback makes bidders more worried about regret when losing by a narrow margin, but it also makes them less worried about regret when losing by a big margin, as suggested by our remarks regarding the potential weakness of the salience assumption. On the other hand, however, intensities of various anticipated emotions are not robustly correlated with bidding behavior in a cross-section of subjects. This may be due to regret not being a part of decision utility, but it may also be due to the emotion intensity measures being noisy and hard to compare across subjects. The results based on the emotions data are therefore inconclusive as to which part of the theory is more likely to be inapplicable.

The rest of the paper is structured as follows. Section 2 presents details of our experimental design. Section 3 presents our results. Finally, Section 4 concludes. The Appendix contains a representative set of instructions.

2 Experimental Design

The first design, denoted HC, is a computerized auction with two bidders, one human and one computerized. A bidder's *value* is drawn from the uniform distribution on $[0, 100]$ ($U[0, 100]$). The *bid* of the computerized opponent is drawn from $U[0, 100]$ as well. We implement three feedback types within this design: minimal, loser and winner. This design abstracts from any strategic uncertainty and hence provides the cleanest environment for testing the theory. Several features of the design are aimed at facing the subjects with the simplest possible decision

⁶We plan to investigate subject heterogeneity in response to feedback in more detail in a companion paper.

⁷“Decision utility” is an objective that a decision-maker uses to choose her action. In case the decision-maker does not fully appreciate all welfare consequences of her decision at the time of making it, this might differ from “experienced utility”, see Kahneman, Wakker and Sarin (1997).

environment. We draw the computerized bidder's bid from $U[0, 100]$ as opposed to some other distribution because this setting exposes human bidders to a simple trade-off between payoff and the probability of winning. In particular, an own bid from the range of $[0, 100]$ is equal to the percentage probability of winning, whereas the difference between the own value and the own bid is equal to the size of the payoff conditional on winning. We use one computerized opponent as opposed to multiple such opponents because the former design does not require subjects to think about the maximum order statistic of the opponents' bids.⁸ We somewhat deviate from 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 Subsection 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.

The second design, denoted 2H, is a computerized 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 design is the simplest possible extension of HC to the environment with strategic uncertainty. Under the theory, 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 design: minimal, loser and winner.

The next two designs aim to bridge the gap between the two initial designs and the four-human-bidder design utilized by FO. They are motivated by the difference in results regarding the effect of loser vs. minimal feedback between HC and 2H and the four bidders' design of FO. For this reason, we implement only two feedback treatments within this design: minimal and loser. The third design, 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 either the presence of loser regret or the effectiveness of the salience assumption might be sensitive to the level of bidding competition.⁹

The fourth design, denoted 4HR, is a paper and pencil auction with four human bidders that exactly replicates the procedure used by FO. Again, 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 Subsection 2.2 for more details).

In each session in all four designs, we use an equal number of men and women and make feedback orthogonal to the exact assignment of bidder values within each gender. This is done

⁸Some 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.

⁹For example, 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.

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.

Another important consideration is the size of stakes across the four designs given the different levels of competition. Clearly, in HC, 4H and 4HR, a subject faces a tougher competition than in 2H. Under the same exchange rate, human bidders would therefore typically face different real payoff incentives in different designs. To correct for this, we adjust the exchange rate so that the expected risk-neutral Nash equilibrium payoff of a human bidder is similar in all four designs (see Subsection 2.4 for more details).

2.1 Instructions

In HC, 2H and 4H, the subjects were provided with a set of printed instructions (see the 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 explaining the auction setting they would face. The decision stage consisted of bidding, beliefs elicitation and emotions elicitation. 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.¹⁰ 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 designs, information about the post-auction feedback was presented on a separate page of the instructions, minimizing the probability that subjects would omit reading it.¹¹

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

¹⁰Incorrect answers were infrequent, suggesting a good understanding of the instructions.

¹¹The Appendix presents a representative set of instructions.

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 of their opponent(s) was drawn from $U[0, 100]$.

In all designs 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 an even coverage of the entire value support for each subject. The six values were presented sequentially on separate screens in a 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, following FO, we use a strategy method with ten values presented simultaneously on a sheet of paper.¹²

In 2H, we generated 36 sets of values, together with their ordering, for men and another such set for women.¹³ This pattern was then used in three sessions of 24 subjects, once for each of the three feedback types. As a result, the distribution of values and the order of their presentation was identical across treatments within each gender. This way we control for any potential interaction of feedback type with the pattern of value assignment (overall and also within each gender) and also potentially reduce noise in the estimates of treatment effects. We repeated an analogous procedure in HC. In 4H, we generated 24 sets of six values, together with their ordering. This pattern was then used in each of two sessions of 24 subjects for each of the two feedback types (minimal and loser). Moreover, between the two sessions of a given feedback type, we switched the pattern of values across the two genders. This way, feedback type is orthogonal to the pattern of value assignment within each gender as in HC and 2H, but, moreover, the pattern is identical for men and women both within and across treatments. In 4HR, by the original design of FO, there are 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, this pattern of values is orthogonal to treatment. We further strengthened this design by switching the pattern of values across the two genders in the two sessions of 24 subjects we ran for each feedback treatment. As a result, the pattern of values is orthogonal to treatment within each gender and, moreover, the pattern is also identical for men and women.

¹²In the other three designs, 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.

¹³Although value pattern was not replicated across genders by design, the overall empirical distribution of values ended up being almost identical across the two genders.

In HC, 2H and 4H, subjects were not allowed to overbid their value. In 4HR, following the original design 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.

First, subjects were told about an additional opportunity to earn ECUs through reporting their beliefs about the highest bid of the opponents. We elicited unconditional beliefs, expected beliefs conditional on winning and on losing, and also beliefs about 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.^{14,15}

Second, we asked subjects to rate their expected emotions conditional on various possible realizations of feedback about the auction outcome they might receive in the feedback stage (non-incentivized). This elicitation method is very similar to the one used by FO, who adopt it from Zeelenberg and Pieters (2004). The subjects rated the emotions on a scale of 1 to 9, with 1 meaning “not at all” and 9 meaning “very much.” We elicited intensity of the following emotions: anger, disappointment, envy, happiness, irritation, regret, rejoice, relief and sadness. Compared to FO, we additionally elicited the intensity of disappointment and rejoice, but did not elicit the intensity of elation.¹⁶ Under minimal feedback, each subject was asked to rate his/her emotions conditional on winning and conditional on losing. Under loser feedback, each subject was asked to rate his/her emotions conditional on winning, conditional on losing to a “slightly high” winning bid and conditional on losing to a “quite a bit higher” winning bid. Under winner feedback, each subject was asked to rate his/her emotions conditional on winning with the second highest bid being “slightly lower” than the own bid, conditional on winning with the second highest bid being “quite a bit lower” than the own bid and conditional on losing. We implemented the conditioning on the winning/losing margin so as to be able to distinguish between cases when precise feedback may stimulate regret and cases when it may dampen it in comparison to minimal feedback. In order to reduce noise in these measures, we provided the subjects with a list of definitions of these emotions, which we present in Appendix B.¹⁷

¹⁴See Blanco, Engelmann, Koch and Normann (2010) for a discussion of the problems inherent to hedging in experiments.

¹⁵We do not use beliefs data in this paper. However, we did perform tests to see whether feedback had any effect on beliefs within each design, but did not find any significant differences. We plan to explore this data in a companion paper.

¹⁶We judge elation to be sufficiently indistinguishable from happiness.

¹⁷In 4H only, after the auction feedback, we re-asked the emotions questions, this time unconditionally in a

Table 1: Numbers of Subjects and Exchange Rates across Experimental Designs and Treatments

Design	Feedback Treatment			Total	Exchange Rate (CZK/ECU)
	Minimal	Loser	Winner		
HC	72	72	72	216	20
2H	72	72	72	216	10
4H	48	48		96	25
4HR	48	48		96	25
Total	240	240	144	624	

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 to 7) and previous experience with online and offline auctions (note that, by the design of the sampling procedure, we already knew each subject’s gender). In addition, we also collected information on menstrual cycle from female subjects.¹⁸

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 opponents) and their payoffs from the auction and from the belief elicitation procedure.

In 4HR, we followed a slightly different procedure that replicated the design of FO. First, we did not elicit any beliefs. Second, given the motivation behind this treatment, we elicited emotions only conditional on losing. Third, when eliciting emotions under loser feedback, we did not condition on the size of the losing margin. Also, as discussed before, we elicited the intensity of elation but did not elicit the intensity of disappointment and rejoice.

2.4 Logistics and Subject Pool

Table 1 presents the number of subjects in all design-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 designs, 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

“hot state.” We did this in order to gauge how well subjects’ anticipation of the intensity of various emotions matches the intensity of actually experienced emotions. We find that the latter are in general smaller than the former. At the individual level, the two intensities are significantly positively correlated with the exception of happiness in case of losing and irritation in case of winning (both correlations insignificantly different from zero). For example, the correlation for regret is 0.71 when losing and receiving minimal feedback, 0.63 when losing and learning that one lost by a margin of at least 10 ECUs and 0.84 when winning (and not knowing by what margin). Only 4 subjects in 4H lost by a margin of less than 10 ECUs, and many correlations are not significant in this small sample.

¹⁸In this paper, we use information only on academic major and the highest achieved academic degree. We plan to explore the additional data in a companion paper. Chen, Katuščák and Ozdenoren (2013) document the impact that women’s menstrual cycle has on bidding behavior.

experimental sessions of 24 subjects.¹⁹ In order to control for potential interactions of feedback type with gender, each session utilizes 12 male and 12 female subjects.²⁰ All subjects in a given session participated in the same design and treatment. All the sessions were conducted at the Laboratory of Experimental Economics (LEE) at the University of Economics in Prague. For the three original designs, HC, 2H and 4H, we used a computerized interface programmed in Z-Tree (Fischbacher 2007), whereas 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). In case of 4HR, we distributed to subjects the original instructions in English taken from FO as well as their translation into Czech.²¹

The subjects were recruited using the Online Recruitment System for Economic Experiments (Greiner 2004) among students from the University of Economics 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 a previous experience with online auctions, 4 percent with offline auctions and 5 percent claim both types of experiences.^{22,23}

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 designs.²⁴ Under these exchange rates, the risk-neutral Nash equilibrium expected payoff is calibrated to be 167 CZK in HC and 2H and 156 CZK in 4H and 4HR. The sessions in HC, 2H and 4H lasted approximately 90 minutes with an average earning of 380 CZK, of which 150 CZK was the show-up fee. The sessions in 4HR lasted only about 50 minutes since they were implemented by pen and pencil and we did not collect any belief data. These sessions recorded an average earning of 290 CZK, of which 250

¹⁹There 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 design based on individual decision-making, without any interaction with other subjects, we believe that this shift does not affect the observations for the involved subjects.

²⁰We plan to investigate gender differences in reaction to feedback in a companion paper.

²¹The student pool on which we draw consists mainly of Czech and Slovak students (see below) with a good command of English. However, in 4HR we wanted to have a further control for language since we did not use our own instructions and thus could not design them to be as detailed as in the other three designs.

²²We also collected data on the number of siblings, risk attitude self-assessment and menstrual cycle information for female subjects. We plan to analyze these variables in a companion paper. We do not have any demographic information for one subject.

²³There is some evidence that subject experience and education level can affect behavior, see for instance, List (2003).

²⁴The average exchange rate over the duration of the experiment was approximately 19 CZK to 1 USD and 25 CZK to 1 EUR.

CZK was the show-up fee.

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 they used). In fact, in designs with four bidders, including the replication design 4HR, each ECU is worth about 1.3 USD. Although the purchasing power of nominally equivalent amounts changes over time and space, taking into account the experimental exchange rates and the level of competition in various designs, we would argue that we face our subjects with non-negligibly larger stakes than FO do.

3 Results

In this section, we present our results. Subsection 3.1 presents the analysis of treatment effects on bidding. Subsection 3.2 presents the analysis of treatment effects on auction revenue and efficiency. Subsection 3.3 complements our bidding analysis with the analysis of anticipated emotions.

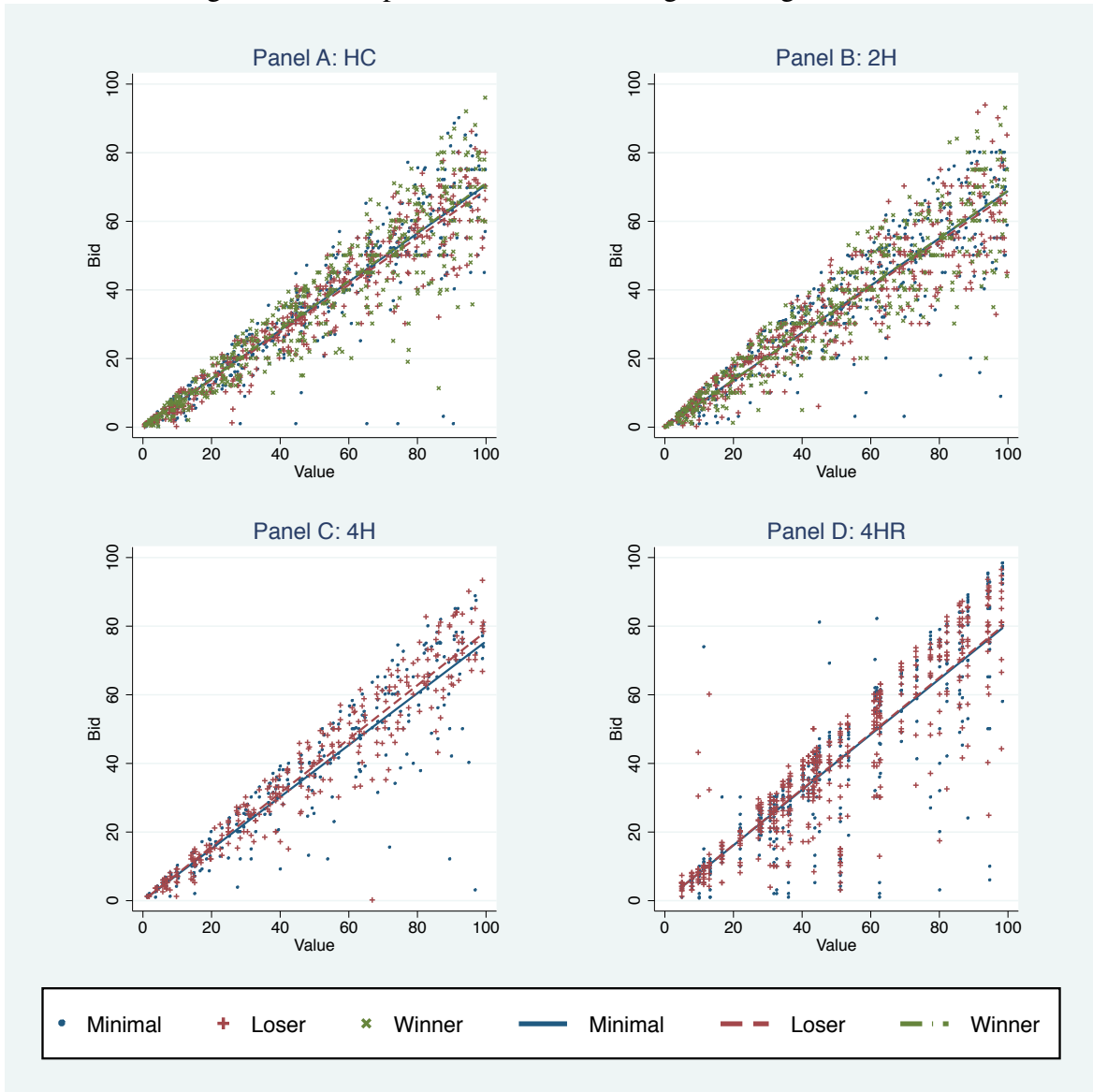
3.1 Bidding

Figure 1 presents a scatterplot of bids against values by design and feedback treatment, in each case also plotting an OLS estimate of the average linear bidding function with zero intercept.²⁵ The plot gives 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 design and feedback type and their differences by feedback within design. 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 set and ordering of values, and then compute the clustered standard error of the average difference. Such way of computing the standard errors increases efficiency in case a particular set and/or ordering of values, or gender, or their combination, has a systematic effect on bidding.²⁶ In HC, the estimated slope is almost identical across the three feedback treatments, varying in a narrow band from 0.691 to 0.711. The situation is similar in 2H, with the slope varying in the band from 0.678 to 0.69. Likewise, in 4HR, the slope varies

²⁵For 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 960 bids. Out 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.

²⁶In HC, 2H and 4H, each subject of a given gender within a treatment faces a unique set and ordering of values, 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 each treatment, always pairing within gender and a particular value list, each time computing the variance for the estimate of the average difference. We then average these variances across 1,000 pairings. The standard error presented in Table 2 is the square root of this average.

Figure 1: Scatterplots of Bids and Average Bidding Functions



between 0.807 and 0.812 across the two feedback types. Only in 4H there is a somewhat more sizeable difference between the two slopes. Under minimal feedback, it is 0.756, whereas under loser feedback it is 0.785. However, neither this treatment difference, nor any other one, is statistically significant.

Using the identical approach, FO estimate the slope of the average bidding function to be 0.790 in case of minimal feedback, 0.883 in case of loser feedback and 0.768 in case of winner feedback. There is a statistically significant difference in these slopes between loser and minimal feedback (p-value of 0.003), but not between winner and minimal feedback.

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 by OLS. We assume that this function is linear and has a zero-

Table 2: Slopes (*and Standard Errors*) of the Average Bidding Functions

Design	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)		

intercept. With v_{ij} denoting the values and b_{ij} denoting the corresponding bids of subject i , with j (or k) $\in \{1, 2, \dots, 6\}$ indexing the order in which the individual values are presented, the estimate of slope for subject i is given by

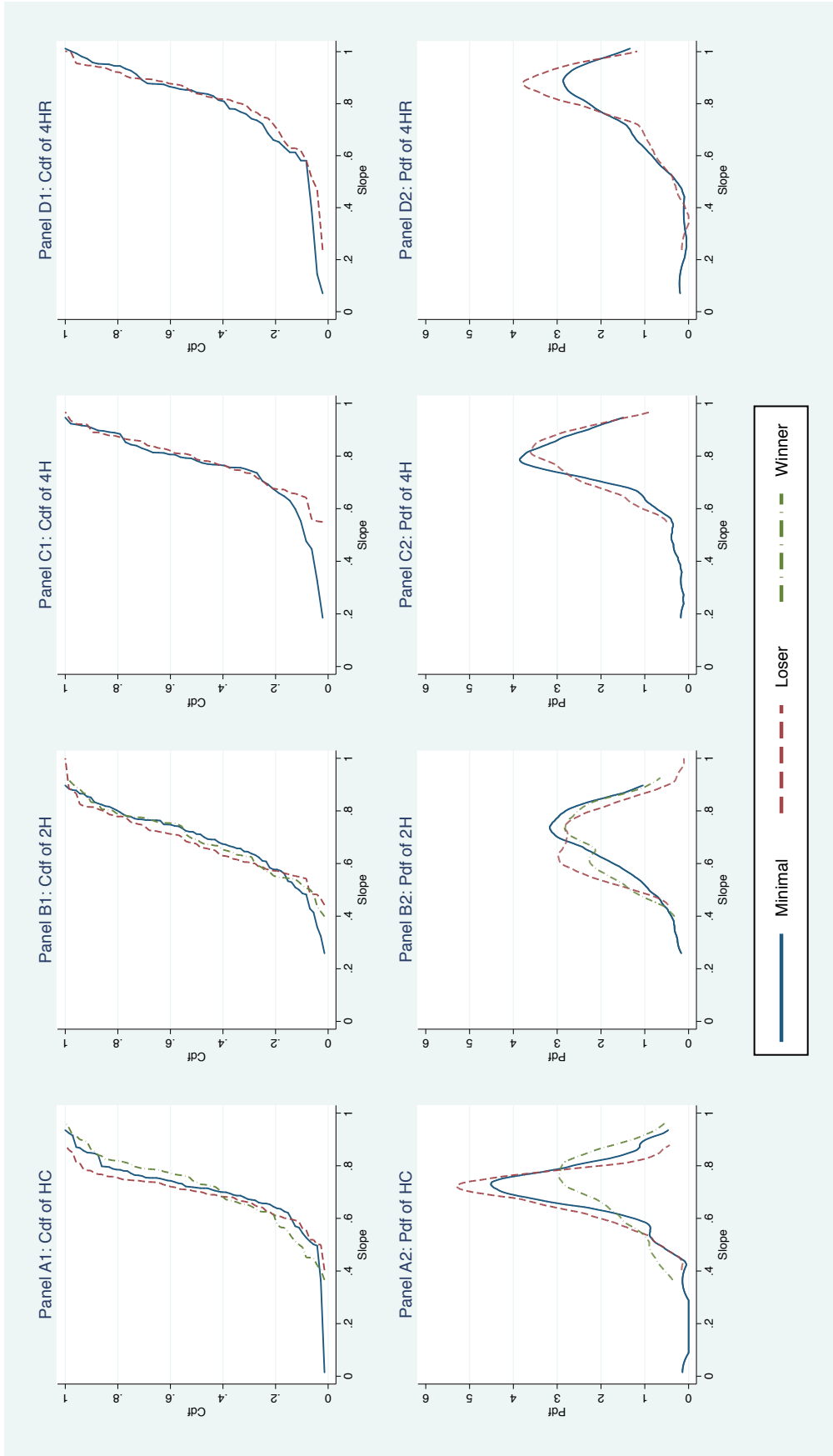
$$\widehat{slope}_i = \frac{\sum_{j=1}^6 v_{ij} b_{ij}}{\sum_{j=1}^6 v_{ij}^2} = \sum_{j=1}^6 \left(\frac{v_{ij}^2}{\sum_{k=1}^6 v_{ik}^2} \right) \frac{b_{ij}}{v_{ij}}. \quad (1)$$

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

On the other hand, the figure also reveals that feedback might have some impact on individual heterogeneity in bidding function slopes. 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 a 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.

Table 3 presents estimates of means and two interquartile ranges (IRs) of these distributions, 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 table also presents estimates of treatment differences in means and IRs, their standard errors and statistical significance. The standard errors on mean differences are obtained analogously to Table 2, except that no clustering is necessary in this case. The

Figure 2: Distributions of Individual Bidding Function Slopes



standard errors on IRs and their differences are obtained by bootstrapping with 1,000 replications. In case of the standard errors on differences, we cluster the bootstrap draws at the level of sets of values and their ordering. That is, if a subject with a particular combination of values and their ordering (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 treatment.²⁷ 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 means of these distributions are very close to the means of the average bidding function slopes presented in Table 2. Looking at the estimated differences and their standard errors reveals that feedback type has no significant effect on means of slope distributions in any of the four designs. This confirms the first casual observation drawn from Figure 2. Regarding the dispersion of the distributions, feedback type has no significant effect on either of the IRs with the exception of HC. In that design, 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 casual observations on the comparison of distribution dispersions drawn from Figure 2.

In contrast to our approach, FO compute the slope of each individual subject 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 and our results 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 (standard deviation/standard error for the average) of the individual bid/value ratios is 0.798(0.142/0.027) under minimal feedback, 0.883(0.063/0.010) under loser feedback and 0.774(0.162/0.029) under winner feedback. The difference between the former two is statistically significant (p-value of 0.004), whereas the difference between the latter two is not.

Table 4 presents p-values of two types of tests of equality of slope distributions between pairs of feedback treatments within each design. 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

²⁷In HC, 2H and 4H, each subject of a given gender within a treatment faces a unique set and ordering of values, 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 designs.

Table 3: Means and Interquartile Ranges (*and their Standard Errors*) of the Bidding Function Slope Distributions

Des.	Statistic	Feedback Treatment			Treatment Differences		
		Min. (M)	Loser (L)	Winner (W)	L - M	W - M	L - W
HC	Mean	0.706 (0.016)	0.690 (0.010)	0.711 (0.016)	-0.015 (0.019)	0.005 (0.020)	-0.020 (0.019)
	$Q_{75} - Q_{25}$	0.111 (0.017)	0.102 (0.019)	0.178 (0.025)	-0.009 (0.025)	0.067** (0.030)	-0.076** (0.031)
	$Q_{90} - Q_{10}$	0.284 (0.046)	0.194 (0.030)	0.362 (0.048)	-0.090 (0.058)	0.079 (0.069)	-0.168*** (0.053)
2H	Mean	0.688 (0.016)	0.677 (0.013)	0.689 (0.015)	-0.011 (0.019)	0.001 (0.023)	-0.012 (0.021)
	$Q_{75} - Q_{25}$	0.154 (0.029)	0.169 (0.020)	0.187 (0.029)	0.015 (0.036)	0.033 (0.040)	-0.018 (0.036)
	$Q_{90} - Q_{10}$	0.351 (0.042)	0.265 (0.027)	0.304 (0.033)	-0.086 (0.051)	-0.047 (0.050)	-0.039 (0.043)
4H	Mean	0.757 (0.022)	0.783 (0.015)		0.026 (0.026)		
	$Q_{75} - Q_{25}$	0.108 (0.040)	0.144 (0.027)		0.037 (0.050)		
	$Q_{90} - Q_{10}$	0.361 (0.090)	0.270 (0.042)		-0.090 (0.104)		
4HR	Mean	0.792 (0.028)	0.811 (0.022)		0.019 (0.033)		
	$Q_{75} - Q_{25}$	0.191 (0.047)	0.133 (0.036)		-0.058 (0.059)		
	$Q_{90} - Q_{10}$	0.392 (0.113)	0.328 (0.060)		-0.064 (0.126)		

Notes:

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

p-values of 0.064 and 0.001). These findings are consistent with those in Figure 2 and Table 3. 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.

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

Table 4: Pairwise Tests for the Equality of Bidding Function Slope Distributions

Design	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		

(p-values of 0.004 and 0.008, respectively).

One could raise an objection that bidding data for low values is more noisy than that for high values in that the probability of winning is relatively small and hence the subjects think less carefully about how much to bid. However, as we already pointed out, the estimated slopes of the individual bidding functions are square-value-weighted averages of the individual bid/value ratios (see equation 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 our test results to excluding low values, we repeat all the tests presented in Tables 3 and 4 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 almost the same results as those presented in Tables 3 and 4, with a few small exceptions. In HC, the difference in the interquartile range between winner and minimal feedback is no longer significant in any of the robustness checks and the difference in the interquartile range between winner and loser feedback is now only weakly significant (p-value of 0.081) and only for the upper half of values.²⁸

To summarize, the results document that feedback type has little effect on the mean of bid/value ratios. This conclusion is robust to all four designs 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. FO make a similar observation in their paper.

In comparison to FO, we use a different subject pool. Our subject sample is drawn from students from 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

²⁸Detailed results are available from the authors upon request.

of Economics in Prague). We are not aware of the major distribution in the FO sample,²⁹ 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 (undergraduate vs. graduate or post-graduate) and academic major, since we measure both in our data. We carry out a separate analysis of bidding for subjects with economics or business major and other subjects. We do not find any robust feedback effects in either group.³⁰ We also separately analyze bids of undergraduate and more advanced students, but again do not find any significant effect of feedback in either group.³¹ We therefore conclude that if the difference between our results and the results of FO is driven by subject pool differences, it is not driven by student seniority or academic major.

Even though our results suggest there are no feedback effects on the mean of the bid/value ratio, we take seriously the 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). To investigate such possibilities, we turn to analysis of auction revenue and efficiency in the next subsection.

3.2 Auction Revenue and Efficiency

Table 5 presents our estimates of average auction revenue and efficiency by design and feedback treatment, as well as their treatment differences, together with standard errors for all the estimates. In HC, given a particular bid b_{ij} , $j \in \{1, \dots, 6\}$ of subject $i \in \{1, \dots, 72\}$ in a given feedback treatment, we first compute the expected revenue conditional on this human bid. Since the bid of the computer is drawn from $U[0, 100]$, this expectation is given by

$$\frac{b_{ij}}{100} b_{ij} + \frac{100 - b_{ij}}{100} \frac{b_{ij} + 100}{2} = \frac{10,000 + b_{ij}^2}{200}. \quad (2)$$

²⁹FO informed us in a private conversation that they did not collect such information.

³⁰If anything, the non-economics/business students tend to bid less under loser vs. minimal feedback, but only in HC.

³¹If anything, the undergraduate students tend to bid more under winner as opposed to loser feedback, but only in HC.

Table 5: Average Auction Revenue and Efficiency (*and their Standard Errors*)

Design	Statistic	Feedback Treatment			Treatment Differences		
		Minimal (M)	Loser (L)	Winner (W)	L - M	W - M	L - W
HC	Av. revenue	58.78 (0.32)	58.24 (0.26)	58.90 (0.38)	-0.54 (0.36)	0.12 (0.44)	-0.66 (0.43)
	Efficiency						
2H	Av. revenue	47.23 (0.94)	46.19 (0.96)	47.08 (1.01)	-1.04 (1.21)	-0.15 (1.37)	-0.89 (1.44)
	Efficiency	97.53 (0.48)	98.31 (0.15)	98.09 (0.21)	0.779 (0.50)	0.559 (0.52)	0.220 (0.23)
4H	Av. revenue	63.20 (1.30)	64.58 (1.20)		1.38 (1.37)		
	Efficiency	96.73 (0.69)	98.25 (0.23)		1.52** (0.76)		
4HR	Av. revenue	68.11 (2.40)	66.74 (1.61)		-1.37 (1.88)		
	Efficiency	95.88 (1.10)	96.83 (0.74)		0.95 (1.27)		

Notes:

¹ Efficiency figures are stated in percentage points.

² In the tests for treatment differences, significant at: ** 5 percent level.

To obtain the average revenue, we then average this expectation over all six bids of all 72 subjects in a given feedback treatment. Because computerized bidders do not have an assigned value, we do not define efficiency in HC.

In 2H, 4H and 4HR, we do not rely just on the realized bidding groups and value/bid realizations in these groups that we used to determine subject payoffs, but instead consider all possible auction realizations given by different bidding groups and value/bid realizations. In each such auction realization, the revenue is equal to the highest bid. The average revenue is then computed as an average of realized revenue across all possible auction realizations for the given feedback treatment. Average efficiency is computed as the average realized value of the winner divided by the average maximum value across all possible auction realizations for the given feedback treatment.

In 2H, there are 72 subjects bidding in groups of two in each feedback treatment and each subject bids for six potential value realizations. Hence there are altogether $\binom{72}{2} \times 6^2 = 92,016$ possible auction realizations for each feedback type. In 4H, there are 48 subjects bidding in groups of four in each feedback treatment, each bidding for six potential value realizations. This gives $\binom{48}{4} \times 6^4 = 252,175,680$ possible auction realizations for each feedback type. In 4HR, there are 48 subjects bidding in groups of four in each feedback treatment, each bidding for ten potential value realizations. This gives $\binom{48}{4} \times 10^4 = 1,945,800,000$ possible auction

realizations for each feedback type.

The standard errors presented in Table 5 are computed by repeating the procedure described above based on 1,000 bootstrap samples clustered at subject level and stratified by gender. Moreover, for standard errors on treatment differences, we cluster the bootstrap draws across different feedback types at the level of sets of values and their ordering. That is, if a subject of a particular gender and with a particular combination of values and their ordering is drawn under one treatment, the corresponding subject(s) are also drawn for the other (two) treatment(s). 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.³²

The average revenue varies in a narrow band of 58.24 to 58.90 in HC, 46.19 to 47.23 in 2H, 63.20 to 64.58 in 4H and 66.74 to 68.11 in 4HR. None of the treatment differences are statistically significant in any of the four designs. The estimates of average efficiency are overall close to 100 percent, varying in a narrow band from 95.88 to 98.31 across all designs and feedback types. None of the treatment differences in 2H and 4HR are statistically significant. However, in 4H, the average efficiency is higher under loser feedback in comparison to minimal feedback (p-value of 0.049). As discussed in the previous subsection, this is probably a consequence of distribution of slopes being less dispersed in the tails under loser feedback.

Overall, the type of feedback has little effect on the average revenue or efficiency. Even though it does have a significant impact on average efficiency in 4H, the effect is quantitatively small. As a result, in terms of these two ultimate auctioneer objectives, we mirror our conclusion of little feedback effect from the previous subsection.

3.3 Bidders' Emotions

The results presented so far contradict the predictions of the anticipated regret theory combined with the salience assumption. However, it is not clear which of the following two components of the theory is contradicted by the results. The first possibility is that regret is present in the utility function, but feedback is ineffective in manipulating its salience. The second possibility is that regret is not in the decision utility function at all, either because the bidders are not prone to regret or because they fail to anticipate it. The data on the intensity of anticipated emotions could help addressing this question, so we turn to their analysis in this subsection.

Table 6 reports mean expected intensities of various emotions and their standard errors.³³ We also perform t-tests for the equality of means under loser and minimal feedback and winner and minimal feedback treatments.³⁴ Focusing on loser regret, in HC, 2H and 4H, we first

³²The same details as in footnote 27 apply here as well.

³³Calculation of these means ignores individual scale subjectivity and the ordinal rather than cardinal nature of the emotion measures. We acknowledge this limitation but carry on with the analysis in order to see how emotions are affected by feedback.

³⁴Note that, apart from 4HR, the averages are computed across HC, 2H and 4H, with the exception of the

Table 6: Means of Anticipated Intensities of Emotions (and their Standard Errors)

Emotion	Conditional on Losing (and losing margin)						Conditional on Winning (and winning margin)					
	Feedback Treatment (HC, 2H, 4H)			Feedback Treat. (4HR)			Feedback Treatment (HC, 2H, 4H)			Feedback Treatment (HC, 2H, 4H)		
	Minimal	Loser, small	Loser, big	Winner ²	Minimal	Loser	Minimal	Loser	Minimal	Loser	Winner, small ²	Winner, big ²
Anger	4.08 (0.18)	5.48*** (0.19)	4.12 (0.18)	4.38 (0.20)	3.27 (0.30)	3.23 (0.30)	1.50 (0.09)	1.28* (0.07)	1.51 (0.11)	1.53 (0.11)	1.92 (0.14)	2.72*** (0.18)
Disappointment	6.05 (0.16)	6.53** (0.16)	5.22*** (0.17)	6.44 (0.19)			2.10 (0.14)	1.73** (0.12)	1.84 (0.15)	6.58 (0.21)	7.34*** (0.21)	3.23*** (0.21)
Elation					1.57 (0.17)	1.58 (0.17)						
Envy	4.34 (0.18)	4.53 (0.20)	3.83** (0.17)	4.04 (0.19)	3.75 (0.33)	3.79 (0.32)	1.70 (0.11)	1.53 (0.09)	1.53 (0.11)	1.53 (0.11)	1.92 (0.14)	1.92 (0.14)
Happiness	1.53 (0.08)	1.40 (0.08)	1.74* (0.09)	1.51 (0.11)	1.38 (0.11)	1.50 (0.16)	6.71 (0.18)	7.34*** (0.15)	6.58 (0.21)	6.58 (0.21)	5.56*** (0.22)	5.56*** (0.22)
Irritation	4.21 (0.17)	5.37*** (0.20)	4.31 (0.16)	4.61 (0.19)	2.94 (0.28)	3.54 (0.33)	2.12 (0.13)	1.90 (0.12)	1.77* (0.13)	1.77* (0.13)	2.94*** (0.19)	2.94*** (0.19)
Regret	4.53 (0.17)	5.49*** (0.18)	3.95** (0.17)	4.76 (0.21)	4.02 (0.27)	4.08 (0.38)	2.05 (0.13)	1.87 (0.12)	2.15 (0.17)	2.15 (0.17)	3.75*** (0.22)	3.75*** (0.22)
Relief	2.46 (0.13)	2.15* (0.11)	2.54 (0.14)	2.49 (0.17)	1.52 (0.18)	1.81 (0.20)	4.99 (0.20)	4.95 (0.19)	5.81*** (0.23)	4.44* (0.23)	4.44* (0.22)	4.44* (0.22)
Rejoice	2.06 (0.11)	2.38* (0.14)	2.29 (0.12)	2.18 (0.14)			4.48 (0.19)	4.62 (0.18)	5.08** (0.23)	5.08** (0.23)	4.16 (0.20)	4.16 (0.20)
Sadness	4.70 (0.17)	5.10 (0.21)	4.32 (0.20)	4.89 (0.20)	3.15 (0.26)	3.40 (0.34)	1.67 (0.11)	1.25*** (0.05)	1.46 (0.13)	1.46 (0.13)	2.93*** (0.24)	2.93*** (0.24)

Notes:

¹ All emotions are measured on a scale of 1 (“not at all”) to 9 (“very much”).

² These averages are computed for observations in HC and 2H only.

³ Significantly different from mean intensity under minimal feedback at: * 10 percent level, ** 5 percent level, *** 1 percent level.

⁴ For the equality of means tests between minimal and winner feedback, all the test results are robust to using only observations from HC and 2H, with the exception of: (1) irritation if winning by a small margin under winner feedback is lower than irritation when winning under minimal feedback at the 5 percent level; (2) relief if winning by a small margin under winner feedback is higher than relief when winning under minimal feedback at the 10 percent level; (3) relief if winning by a big margin under winner feedback is smaller than relief when winning under minimal feedback at the 5 percent level; (4) rejoice if winning by a small margin under winner feedback is no longer statistically significantly different from rejoice when winning under minimal feedback.

observe that the averages are significantly different from 1, the minimum possible level of regret. We also observe that the average under winner feedback is insignificantly different from the average under minimal feedback. Similarly, in 4HR, the average under loser feedback is insignificantly different from the average under minimal feedback. However, when, in HC, 2H and 4H, the expectations are conditioned on the size of the losing margin, we do find significant differences. In particular, the average under loser feedback, when expecting to lose by a small margin (5.49), is significantly higher than the average under minimal feedback (4.53). On the other hand, the average under loser feedback, when expecting to lose by a big margin (3.95), is significantly lower than the average under minimal feedback. Also, the expected loser regret when losing by a small margin is statistically significantly higher than when losing by a big margin (p -value < 0.01). These observations suggest that loser regret considerations might indeed be present in bidders' utilities at the time of the elicitation procedure. Also, as the theory would suggest, loser regret is decreasing with the size of the losing margin. However, judging from 4HR, average loser regret considerations do not seem to be affected by feedback type.

Turning to winner regret, also in this case the averages are different from 1, but the average expected intensity of winner is uniformly lower than the intensity of loser regret. The average under winner feedback, when winning by a big margin (3.75), is significantly higher than the average under minimal feedback (2.05). This time, there is no offsetting difference of the opposite sign when winning by a small margin (2.15). Also, the expected winner regret when winning by a small margin is statistically significantly lower than when winning by a big margin (p -value < 0.01). Moreover, the average winner regret under loser feedback (1.87) is insignificantly different from the average winner regret under minimal feedback. These observations suggest that winner regret considerations might also be present in bidders' utilities at the time of the elicitation procedure, although to a smaller degree than considerations loser regret. Also, as the theory would suggest, winner regret is increasing with the size of the winning margin. The data is also consistent with the possibility that winner feedback increases the average intensity of anticipated winner regret in comparison to minimal feedback.

It is also interesting to observe that regret does not appear to be the most relevant emotion at play. Indeed, conditional on losing, disappointment ranks highest in the scale of reported emotions, while, conditional on winning, happiness ranks the highest. Moreover, disappointment displays a qualitatively very similar pattern to regret. Although this might be an artifact of subjects confusing disappointment with regret, it suggests that focusing on regret alone might be too narrow an approach to capture emotions driven by feedback.

Recall, however, that the emotions questions are asked after bidding and after the payoff-relevant bid-value pair is disclosed to each subject (but before receiving feedback on the auc-

winner feedback treatment, for which these averages are computed only for HC and 2H. See footnotes 2 and 4 below the table.

tion). As a result, we cannot evaluate to what extent these emotions are taken into account by the subjects at the time of deciding on their bids. Indeed, when we investigate whether the emotion reports correlate with bidding behavior, particularly with the slopes of the individual bidding functions, we find that none of the correlations are robustly significant across treatments. This could be due to emotions playing only a minor role relative to material payoff in bidders' utility functions, or due to these emotions not being well-anticipated *ex ante*. However, given that the emotions data is based on non-incentivized self-reports and the responses might not be comparable across subjects, the lack of correlation could also be due to noise in the data. We are not able to distinguish between these possibilities. Given this observation, we do not feel confident in drawing any conclusions regarding which part of the theory is inapplicable.

4 Conclusion

This paper presents new evidence on the important market design issue of whether an auctioneer in a one-shot first-price auction can make bidders bid more aggressively, and hence improve revenue, by controlling the type of feedback that bidders receive after the auction. At the same time, it provides a comprehensive test of the anticipated regret theory in combination with the assumption that feedback type directly manipulates the salience of winner and loser regret in the utility function (salience assumption). Despite the importance of the problem from theoretical as well as practical point of view, it has received only a limited attention in the auction literature up until now (Filiz-Ozbay and Ozbay 2007) (FO).

Compared to FO, we use 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 designs, including the design that exactly replicates the design used by FO. In regard to the effect of winner vs. minimal feedback, our finding replicates the finding of FO. On the other hand, the finding on loser vs. minimal feedback is in a stark contrast to the finding by FO that loser feedback significantly increases the average bid/value ratio in an auction with four human bidders. We also find that the type of feedback has no systematic effect on average revenue and little effect on average efficiency of the auction.

Since, in comparison with FO, we control for gender balance, use arguably larger stakes and use a much larger sample size, we believe that sampling noise contained in our estimates is no larger than the one contained in the estimates of FO. Also, larger stakes should imply a stronger external validity of our results. We speculate that the difference in the result on the effect of loser vs. minimal feedback between our study and FO may be due to the subject pool difference. However, it appears unrelated to student seniority or academic major. In the absence of further information on the sample of subjects used by FO, we are not able to

evaluate this hypothesis any further.

We conclude that in one-shot auctions, at least those 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 or revenue, and little effect on efficiency. From the point of view of auction design, these results imply that such feedback manipulation is unlikely to be useful in increasing revenue in first-price auctions. Furthermore, loser feedback does not appear to be a good explanation of overbidding relative to the risk-neutral Nash equilibrium. This conclusion follows not only from the null effect of feedback, but also from a great amount of "overbidding" under minimal feedback in two-bidder auctions.

Our empirical results also contradict predictions of the anticipated regret theory augmented by the salience assumption. From the point of view of the theory, this implies that anticipated regret is not an important component of bidder decision utility or that the type of feedback has little effectiveness in manipulating the salience of regret. The former could be due to bidders either not being regret-prone or not anticipating their posterior regret. We are not able to discriminate between these two explanations. Whether alternative auction environments, in connection with loser or winner or, potentially, alternative feedback types, are capable to systematically affect the bid/value ratios in one-shot environments is an open question for future research.

References

- Bell, David E.**, “Regret in Decision Making under Uncertainty,” *Operations Research*, 1982, 30 (5), 961–981.
- Blanco, Mariana, Dirk Engelmann, Alexander K. Koch, and Hans Theo Normann**, “Belief elicitation in experiments: is there a hedging problem?,” *Experimental Economics*, 2010, 13, 412–438.
- Chen, Yan, Peter Katuščák, and Emre Ozdenoren**, “Why Can’t a Woman Bid More Like a Man?,” *Games and Economic Behavior*, January 2013, 77 (1), 181–213.
- Engelbrecht-Wiggans, Richard**, “The Effect of Regret on Optimal Bidding in Auctions,” *Management Science*, 1989, 35 (6), 685–692.
- ___ and **Elena Katok**, “Regret and Feedback Information in First-Price Sealed-Bid Auctions,” *Management Science*, April 2008, 54 (4), 808–819.
- Filiz-Ozbay, Emel and Erkut Y. Ozbay**, “Auction with Anticipated Regret: Theory and Experiments,” *American Economic Review*, April 2007, 54 (4), 1407–1418.
- Fischbacher, Urs**, “z-Tree: Zurich Toolbox for Ready-made Economic Experiments,” *Experimental Economics*, 2007, 10 (2), 171–178.
- Greiner, Ben**, “An Online Recruitment System for Economic Experiments,” *Kurt Kremer, Volker Macho (Eds.); Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen : Ges. f§r Wiss. Datenverarbeitung*, 2004, pp. 79–93.
- Isaac, Mark R. and James M. Walker**, “Information and Conspiracy in Sealed Bid Auctions,” *Journal of Economic Behavior and Organization*, 1985, 6, 139–159.
- Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin**, “Back to Bentham? Explorations of Experienced Utility,” *The Quarterly Journal of Economics*, 1997, 112 (2), 375–406.
- Krishna, Vijay**, *Auction Theory*, first ed., Academic Press, 2002.
- List, John A.**, “Does Market Experience Eliminate Market Anomalies?,” *Quarterly Journal of Economics*, 2003, 118 (1), 41–71.
- Loomes, Graham and Robert Sudgen**, “Regret Theory: An Alternative Information Theory of Rational Choice under Uncertainty,” *The Economic Journal*, 1982, 92 (368), 805–824.
- Neugebauer, Tibor and Javier Perote**, “Bidding Ôas if Ô risk neutral in experimental first price auctions without information feedback,” *Experimental Economics*, 2008, 11, 190–202.
- ___ and **Reinhard Selten**, “Individual behavior of first-price auctions: The importance of information feedback in computerized experimental markets,” *Games And Economic Behavior*, 2006, 54, 183–204.
- Ockenfels, Axel and Reinhard Selten**, “Impulse Balance Equilibrium and Feedback in First Price Auctions,” *Games And Economic Behavior*, 2005, 51 (1), 155–170.
- Turocy, Theodore L. and Elizabeth Watson**, “Reservation Values and Regret in Laboratory

First-Price Auctions: Context and Bidding Behavior,” *Southern Economic Journal*, 2012, 78 (4), 1163–1180.

Zeelenberg, Marcel and Rik Pieters, “Consequences of regret aversion in real life: The case of the Dutch postcode lottery,” *Organizational Behavior and Human Decision Processes*, 2004, 93, 155–168.