Italian Economic Journal
https://doi.org/10.1007/s40797-026-00394-1

RESEARCH PAPER

The Archer in the Mist: The Acquiring-a-Company game and Weak Learning Direction Theory

Antonio Cosma[1] · **Luca Panaccione**[2]

Received: 31 March 2025 / Accepted: 7 May 2026
© The Author(s) 2026

Abstract
We analyse the bidding behaviour of buyers in a modified Acquiring-a-Company game allowing for strategic interaction and extensive learning in light of the decisions of the sellers. The benchmark solution, based on risk-neutral players, does not organise the data well. However, we show that the buyers' choices are consistent with a suitable version of Learning Direction Theory, which we name Weak Learning Direction Theory.

Keywords Acquiring-a-Company · Winner's Curse · Learning Direction Theory

JEL Classification C78 · C91 · D82 · D91

1 Introduction

In the Acquiring-a-Company (hereafter AaC) game, one party, the buyer, makes a bid to buy an item, the company, from another party, the seller, knowing only the distribution of firm values and the percentage difference between the buyer's and the seller's valuations. The seller, instead, is aware of the actual value of the company and decides whether to accept or reject the bid. In the case of acceptance, the buyer earns the difference between the buyer's valuation of the company and the bid, while the seller earns the difference between the bid and the seller's valuation. In case of rejection, both earn nil. Since it is assumed that the seller's valuation is lower than the buyer's, there are always ex-post gains from trade.

✉ Luca Panaccione
luca.panaccione@uniroma1.it

Antonio Cosma
antonio.cosma@unibg.it

1 Department of Management, Università degli Studi di Bergamo, Via dei Caniana, 2, 24127 Bergamo, Italy

2 Department of Economics and Law, Sapienza University of Rome, Via del Castro Laurenziano 9, 00161 Rome, Italy

Published online: 11 June 2026 Springer

The benchmark solution, based on the maximisation of the seller's and buyer's (expected) earnings, implies that the optimal bid should induce trade whenever the seller's valuation of the company is low enough compared to the buyer's, while the buyer should refrain from trading when the seller's valuation is relatively high. In this case, as in the seminal analysis of George (1970), asymmetric information about the value of the sale item prevents trades that would benefit both parties.

However, since Bazerman and Samuelson (1983) there is robust experimental evidence that buyer participants adopt bidding strategies inconsistent with the above benchmark solution. The sub-optimal choices usually persist even when extensive learning via repeated play and information feedback is allowed, leading to the well-known *winner's curse* phenomenon. Experimental evidence of sub-optimal bidding strategies has been found both in experiments with *robot* sellers, i.e., with computers implementing a (monotonic) minimum acceptance-threshold rule,[1] and in experiments with *human* seller participants, whose focus is mainly the strategic interaction between participants in opposite roles.[2]

The failure of the benchmark solution to satisfactorily account for the experimental evidence has called for alternative interpretations of the behaviour of buyer participants.[3] Among these alternative strategies, Learning Direction Theory (henceforth LDT) has proven able to account for the observed sequences of individual bids in an experimental AaC game with robot sellers (see Selten et al. 2005).[4] In this paper, we contribute to the analysis of LDT in an AaC game experiment with human, instead of robot, sellers.

To better frame our contribution, we recall that LDT is a qualitative learning theory inspired by ex-post rationality: "One looks at what might have been better last time and adjusts the decision in this direction" (Selten et al. 2005, p. 6).[5] Formally, LDT differs from reinforcement learning in that it is based, in a repeated game, on comparisons of counterfactual payoffs corresponding to choices not selected in the previous round.

[1] See, e.g., Ball et al. (1991), Selten et al. (2005), Grosskopf et al. (2007), Bereby-Meyer and Grosskopf (2008), Charness and Levin (2009), Dennis et al. (2012), Fudenberg and Peysakhovich (2016) and Martínez-Marquina et al. (2019).

[2] See Di Cagno et al. (2016, 2017), Galliera (2018), Angelovski et al. (2020) and Di Cagno et al. (2024) for experiments in which the privately informed participant sends (cheap-talk) value messages to the other participant, or Güth et al. (2019) for experiments with compulsory disclosure of information.

[3] The persistent sub-optimal behaviour in the AaC game is often imputed to the failure of buyers to reason in terms of the conditional expectation of the company value and to anticipate the selective acceptance of sellers (see, e.g., Ball et al. 1991). Some experiments with robot sellers aim at testing whether and how buyers discover the benchmark bidding strategy (see, e.g., Ball et al. 1991; Grosskopf et al. 2007; Bereby-Meyer and Grosskopf 2008, and also Foreman and Keith 1996). Other experiments investigate whether different equilibrium concepts (see, e.g., Charness and Levin 2009; Fudenberg and Peysakhovich 2016) or different learning theories (see, e.g., Dennis et al. 2012) can organise the data more successfully.

[4] Selten and Stoecker (1986) first introduced LDT to analyse a prisoner's dilemma experiment. It has subsequently been used in several experimental settings; see Selten (2004) for an early review of studies using LDT. Recent applications include Neugebauer and Selten (2006), Neugebauer and Selten (2019) and Alberti et al. (2021).

[5] The archer recalled in our title is often used as an illustrative example of LDT: "Consider an archer who tries to hit the trunk of a tree. If the arrow misses the tree on the left side, then the archer will tend to aim more to the right, and in the case of a miss to the right the aim will be more to the left" (Selten et al. 2005, p. 5).

In a setting where exactly the same decision problem is repeated (a “groundhog day” environment), this ex-post perspective can be interpreted literally: one can imagine replaying the very same situation and choosing the action that would have yielded a higher payoff. In the Acquiring-a-Company task, however, the “target” (the optimal bid) moves randomly across rounds because the firm value changes. As Selten et al. (2005) point out, this makes the Samuelson–Bazerman task a challenge for learning direction theory, since the bid that would have been better in the previous period need not be optimal in the current one. Nevertheless, their data show that subjects tend to adjust their bids in the direction that ex-post rationality would recommend for the previous round, i.e. behaviour is shaped *as if* ex-post rationality were still a useful guide even though the environment is stochastic. From now on, when we refer to ex-post rationality, we mean it in this sense.

Selten et al. (2005) apply LDT to the choices of buyers in an AaC experiment with a robot seller. Since the buyer’s profit equals the difference between the value of the firm and the bid in the case of acceptance and is nil in the case of rejection, two experience conditions are relevant. In the *over-payment* condition, the accepted bid is larger than the minimum required for acceptance and profits would have been greater by bidding less. In the *lost opportunity* condition, the bid is smaller than the minimum required for acceptance and the buyer would have gained by bidding more. On this basis, they predict that bid revisions, when they occur, should more frequently than by chance go in the direction indicated by ex-post rationality, and they show that this prediction fits how buyer participants adjust their decisions in light of past outcomes.

The main goal of our note is to test, on experimental data, whether LDT is still successful in organising behaviour in a more challenging informational environment than in Selten et al. (2005). Specifically, our experimental protocol:

1. allows for explicit strategic interactions, since sellers are human, and not robot, participants;
2. provides the buyer participants with indirect feedback on the realised value of the company via the payoff only in case of acceptance; and
3. changes repeatedly the percentage difference between the buyer’s and the seller’s valuation and, accordingly, the benchmark bid.

These features of the bargaining protocol imply that the cognitive burden of buyer participants is rather different from that in a protocol with robot sellers: in this latter case, they face an individual decision task under stochastic uncertainty, since they do not know the value of the company when bidding; in our protocol, they additionally face strategic uncertainty about the behaviour of the sellers. Sellers may, for example, adopt a monotonic response strategy with a rather idiosyncratic acceptance threshold, or they may not follow a monotonic response strategy at all. More importantly, the acceptance threshold of seller participants is not known to buyers.

In Selten et al. (2005), the past realisation of the firm value serves as a reference point for the buyer when revising the bid. In our setting, the buyer cannot observe the corresponding reference point, namely the seller’s acceptance threshold, but only a noisy version of it through the seller’s accept/reject decision. For these reasons, we refer to our version of LDT as Weak Learning Direction Theory (WLDT). Consequently, in order to define testable predictions, we need to adjust the description of the experience conditions on which bid revisions are based.

Our contribution is threefold: (i) we adapt LDT to this informational environment (WLDT) and discuss the testable directional predictions for bid revisions; (ii) we implement permutation-based tests suited to the panel structure of bid revisions; and (iii) we show that, despite strategic interaction and noisier feedback, buyers' bid revisions remain strongly aligned with WLDT's directional predictions.

The paper is organised as follows: In Sect. 2.1, we introduce the AaC game and characterise the benchmark solution. In Sect. 2.2 we provide details of the experimental implementation. Section 3.1 presents the testable assumptions, while Sect. 3.2 outlines the testing procedure and reports the results.

2 The Model and the Experimental Protocol

The data we analyse in Sect. 3.2 are a subset of a larger dataset collected for a broader experimental research on the AaC game.[6] In what follows, we set out the essential features of the game and experimental protocol, paying special attention to the features most relevant to our analysis.

2.1 The Benchmark Analysis

We start by describing the AaC game when the buyer and seller maximise (expected) profits. Then, given that our experimental implementation of the AaC game involves human sellers, we will incorporate an explicit behavioural dimension by extending the benchmark analysis to allow for inequity-averse preferences (Fehr and Schmidt 1999).[7] In the first case, the (random) value v of the firm is drawn from a uniform distribution on [0, 1], and the seller's valuation of the firm is qv, with $0 < q < 1$. The seller knows the realised value v of the firm. Although the buyer does not know the realised value v of the firm, they are aware of the discount factor q. The buyer offers the seller a price of p to acquire the firm. If the seller accepts the offer, the buyer earns $v - p$, and the seller earns $p - qv$; otherwise, they both earn nil (Fig. 1 represents the AaC game in its basic form and does include some additional features introduced in the experimental implementation and described at the end of Sect. 2.2). In the benchmark analysis of the game, acceptance by the seller is expected only when $p \geqslant qv$ (i.e., when $v \leqslant p/q$). Since $p > q$ is a dominated choice, a natural restriction on the price the buyer can offer is $0 \leqslant p \leqslant q$. A risk-neutral buyer chooses a bid p^* to maximize expected payoff. Let

$$p^*(q) \in \underset{0 \leq p \leq q}{\operatorname{argmax}} \int_0^{p/q} (v - p)\, dv. \tag{1}$$

[6] The main results of this research project are presented in Di Cagno et al. (2016, 2017) and Galliera (2018).

[7] We thank an anonymous referee for suggesting this extension.

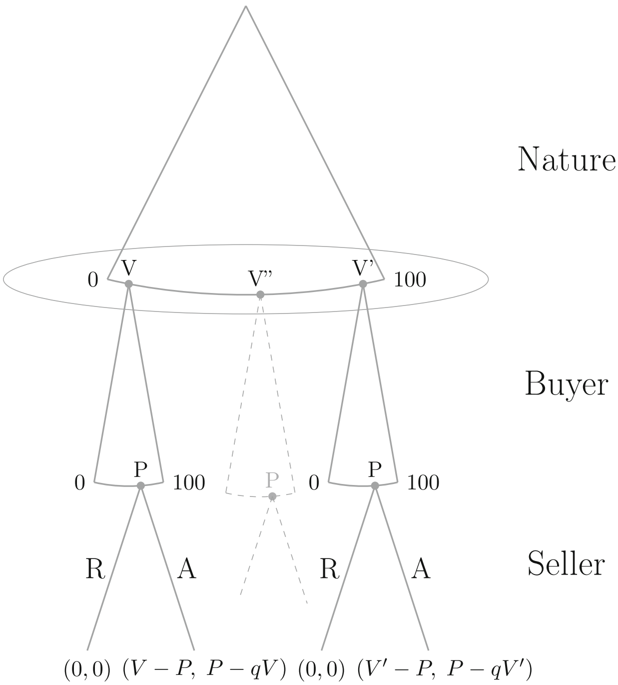

Fig. 1 The AaC game tree. R: *Reject*, A: *Accept*

Evaluating the integral yields:

$$\int_0^{p/q} (v - p)\, dv = \left[\frac{v^2}{2} - pv\right]_0^{p/q} = \frac{1}{2}\frac{p^2}{q^2} - \frac{p^2}{q} = \frac{1}{q}\left(\frac{1}{2q} - 1\right)p^2. \tag{2}$$

The objective function is a quadratic function that strictly increases (decreases) with p when $q < 0.5$ ($q > 0.5$), and it does not vary with p when $q = 0.5$. Therefore, the optimal bid is $p^* = q$ when $q < 0.5$, $p^* = 0$ when $q > 0.5$, and $p^* \in [0, 0.5]$ when $q = 0.5$.

Figure 2 conveys the intuition behind the two terms entering the expected profit. The surface $\frac{p^2}{q}$ of the grey rectangle represents the costs expected by the buyer for a given offer p if the bid is accepted. The hatched triangle represents the expected gains to the buyer, $\frac{1}{2}\frac{p^2}{q^2}$, in case of a successful bid. The grey-and-hatched region represents the fraction of expected gains and costs that offset each other. The net expected profit to the buyer is then given by the difference between the hatched-only and the grey-only triangles. The surfaces of the two triangles are equal if $q = 1/2$. Furthermore, expected profits are positive if $q < 1/2$ and negative if $q > 1/2$.

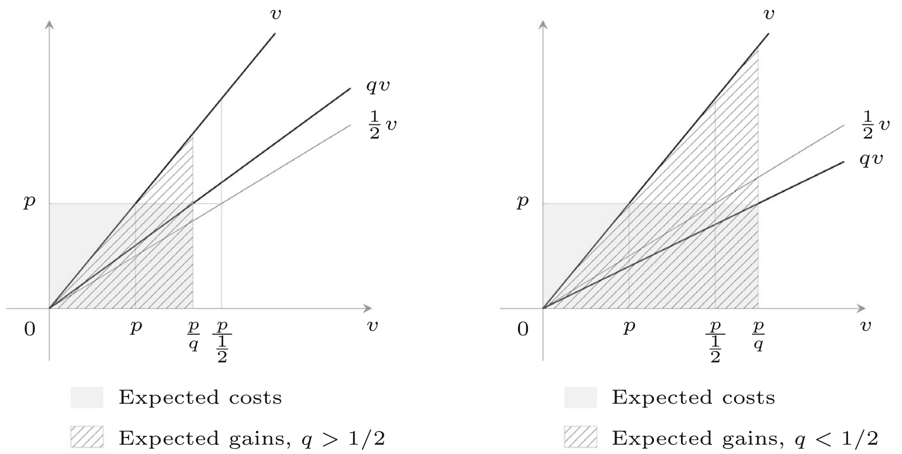

Fig. 2 Expected Gains and Losses for a given value p of the buyer's bid

We now extend the above analysis by including inequity-averse preferences. These preferences penalise payoff differences, both when they are advantageous and when they are disadvantageous to the decision-maker (buyer or seller). Because our primary interest is bid revisions rather than optimal bidding levels, we do not pursue a complete analysis with simultaneous inequity aversion on both sides, so as not to divert from the paper's main focus. Nevertheless, for completeness, we study two one-sided extensions in which either the buyer or the seller is inequity averse (but not both).[8] In both cases, the optimal bidding rule remains qualitatively similar to the baseline: the buyer chooses either 0 or q, with a cutoff q^* that depends on the inequity parameters (governing aversion to advantageous and disadvantageous inequality). As the inequity parameters vanish, q^* converges to the baseline cutoff $1/2$. Intuitively, inequity concerns change the effective attractiveness of trade by introducing a psychological cost of unequal payoffs, which shifts the cutoff relative to the benchmark without inequity aversion.

2.2 The Experimental Protocol

In the experimental implementation from which our data are gathered, participants repeatedly play the (one-shot) AaC game. Specifically, the experiment features four stages: the first stage consists of a single incentivised (trial) round of the AaC game; the second stage consists of thirty rounds of the AaC game, with one round randomly selected for payment; the third stage consists of twelve additional rounds of the AaC game with a different incentive scheme than in stage two; finally, the fourth stage features a risk-preference elicitation (individual) task.

In our analysis we only use data from the second stage, since its setting is the closest to Selten et al. (2005). In this stage, participants are randomly assigned to the constant (across rounds) role of either seller or buyer, and interact with a randomly

[8] Details are provided in the Online Appendix. See Eva and Hoppe (2013) for an analysis in which both participants are inequity-averse.

selected (in each round) participant in the opposite role. The discretized, and equally likely, values of the firm, v, and of the discount factor, q, selected in each round, are $\{5, 10, 15, \ldots, 95\}$ and $\{10\%, 20\%, 30\%, \ldots, 90\%\}$, respectively. The gains from play are communicated at the end of each round; one of these rounds is randomly selected as payment for this phase at the end of the experiment. Based on their role, sellers are informed of both v and q, while buyers are informed only of q. The buyer proposes a price p, which cannot exceed 100. After receiving the price proposal, the seller accepts or refuses it. In case of acceptance, the firm is sold at the offered price; the buyer earns $v - p$ and the seller earns $p - qv$; these payoffs are communicated to both players. In case of refusal, they both earn nil and receive no further feedback.

The experimental protocol has additional features that we do not integrate into our analysis, namely: (i) the buyer and seller receive differential (across four treatments) information on the gender and the field of study of the counterpart,[9] (ii) before receiving the offer, the seller, who is aware of the company value v, sends to the buyer a (cheap-talk) value message, $\hat{v}$.[10] We do not explicitly model these additional features. Our approach follows Selten et al.'s own positioning of LDT: it is not meant to provide "…a complete explanation of adaptive behaviour …" and explicitly allows for "other influences" (Selten et al. 2005, p. 6) that may affect choices, while assuming that the ex-post rationality motive is the strongest influence. In our setting, the additional design features naturally fall into these "other influences". Our empirical objective is therefore to assess whether, despite them, bid revisions remain predominantly aligned with LDT's directional predictions.

The experiment was run in the laboratory of Max Planck Institute of Economics in Jena in 12 sessions with 376 participants (11 sessions with 32 participants each and one session with 24). The participants were recruited among the undergraduate population of Jena University using Orsee (Greiner 2015). The experiment was fully computerized using z-Tree (Fischbacher 2007) and fully incentivised: in addition to the payment for the second phase, payments for the experiment include the gains from the three other phases and the initial endowment, which was assigned to avoid bankruptcy issues.[11]

9 Specifically, in treatment U (Unknown) and G (awareness of Gender constellation) participants are aware and unaware, respectively, of the gender of the paired participant; in treatment OC (Other Confound) they are informed about the field of study (Economics versus Non-Economics) instead and, finally, in treatment E (Embedded gender constellation) they are informed about both the gender and field of study.

10 The value message, which can be true ($\hat{v} = v$) or not ($\hat{v} \neq v$), does not affect the benchmark solution of the game and is sent to the buyer after the seller learns the values of v and q.

11 The instructions of the experiment, reproduced from Di Cagno et al. (2017, Appendix 2) are reported in the Appendix.

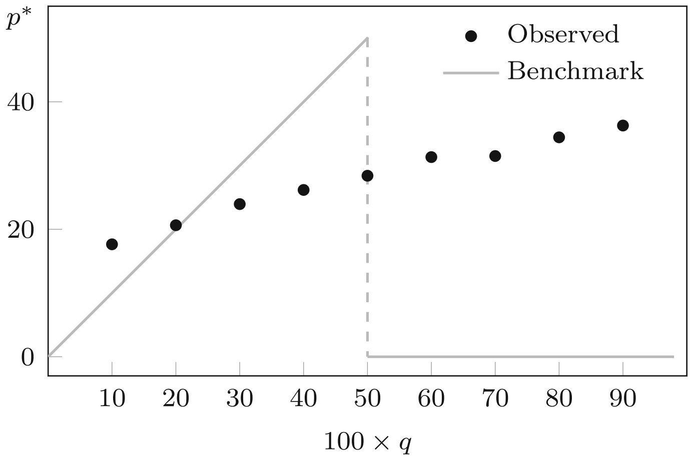

Fig. 3 Average and benchmark bid for each level of q

3 Validation of WLDT

3.1 Predictions of WLDT

It is not surprising that participants' behaviour deviates from the benchmark solution, even with the opportunity to learn through several rounds of play. Learning the optimal bid—already notoriously difficult when sellers are robots (see the discussion in the Introduction)—becomes even more complex due to the strategic interaction with human sellers and the fact that q changes randomly in each round. This is clearly shown in Fig. 3, which depicts how the average bid for a given q departs from the benchmark solution p^*. Average bids appear to increase monotonically, remaining below the benchmark for most of the $q < 0.5$ region and above it for $q > 0.5$. Moreover, because incorporating inequity aversion leaves (see Sect. 2.1) the benchmark bidding rule qualitatively unchanged (it remains a corner solution with a shifted cutoff), the observed deviations remain substantial also relative to the inequity-adjusted benchmark.

To define predictions of bid revisions, we assume that at the start of each round, buyers re-evaluate their previous choices in the same spirit as in LDT but with modifications dictated by the different information environment. Predictions are in line with classic LDT, but justifications differ.

In the case of acceptance, the buyer learns v but not the minimum threshold below which the seller would have rejected the bid. The buyer faces two conflicting tensions. On one hand, having secured the firm, he may feel a sense of possible overpayment, which pushes him to decrease his bid, feeling sufficiently optimistic that the seller will still accept and that he will achieve a higher profit. On the other hand, he may fear not locking in his current profit, as lowering his bid could result in an offer that falls below the seller's acceptance threshold. The extent to which one of these forces prevails likely depends on the buyer's subjective assessment of their relative strength.

Table 1 Descriptive statistics on bid revisions

Number of buyers/sellers, N	188×				
Number of rounds, T	30 =				
Number of Observations	5640				
Seller's decision at t, d_{it}	*Buyer's bid revision at t+1,* Δp_{it+1}				
	Negative	Zero	Positive	Correct	Not correct
Rejection	703	436	1163	1599	703
%	31	19	50	69	31
Acceptance	1629	346	1175	1975	1175
%	52	11	37	63	37
Total	2332	782	2338	3574	1878
%	43	14	43	66	34

Depending on which of these forces dominates, we predict that the buyer either does not change his bid or decreases it.

Prediction 1: After acceptance in round t, the buyer will not increase the bid in round $t + 1$.

In the case of rejection, the only information available to the buyer is that the seller's acceptance threshold is higher than the previous bid. The buyer is weighing two conflicting pressures. On one hand, he experiences the condition of lost opportunity, as only by increasing the bid does he stand a chance of acquiring the firm. On the other hand, risk aversion discourages him from raising the bid, since he has no clear target value, and an excessive increase—an overshoot—may result in overpayment or even a loss. Depending on which of these forces dominates, we predict that the buyer either does not change his bid or increases it.

Prediction 2: After rejection in round t, the buyer will not decrease the bid in round $t + 1$.

3.2 Testing WLDT

We test our WLDT predictions regarding bid revisions using data from the experiment detailed in Sect. 2. To align closely with the setup in Selten et al. (2005), we focus exclusively on observations from the second stage of the experiment. This yields a dataset of 5640 pairs of buyer bids and seller acceptance decisions. Each observation contains p_{it}, the bid of the buyer i in round t, with $i = 1, 2, \ldots N$ ($N = 188$), and with $t = 1, 2, \ldots T$ ($T = 30$), as well as $d_{it} \in \{Acceptance, Rejection\}$, the decision of the seller paired with buyer i in round t. Table 1 provides descriptive statistics, including bid revisions $\Delta p_{it+1} = p_{it+1} - p_{it}$. Regression analysis, as observed in Charness and Levin (2009) and Fudenberg and Peysakhovich (2016) among others, shows that rejections are typically followed by higher bids ($\Delta p_{it+1} > 0$), and acceptances by lower ones ($\Delta p_{it+1} < 0$). However, our focus is not on the average bid levels given the sellers' decision, but on the sequential pattern of acceptance/rejection and the sign of bid revision in the following round. A negative regression coefficient of Δp_{it} on $d_{i,t-1}$ might suggest compliance with WLDT. Yet, the following example shows that it may not always be the case. By coding $d_{it} = 1$ in case of acceptance of the bid by the seller, a negative regression coefficient means that the expected bid revision given acceptance is negative. Consider now a sequence of observations in which three-fourths of acceptances lead to positive bid revisions of 10 and one-fourth to large negative revisions of − 34. The expected bid revision given acceptance is − 1, but the sequential pattern does not agree with WLDT.

The first three columns of Table 1 summarise bid revisions: negative, zero and positive. Unconditionally, negative and positive revisions occur equally (43%), with 14% unchanged. Conditioning on seller decisions in the previous round reveal distinct patterns. Following rejections, positive revisions dominate (1163 vs. 703 negatives, 436 unchanged). The opposite pattern appears after acceptances: negative bid revisions prevail (1629 vs. 1175, 346 unchanged).[12] These patterns are further confirmed if we aggregate no-revisions with positive ones as *Correct* bid revisions (i.e., in agreement with Predictions 1 and 2), after a rejection, and no-revisions with negative ones as *Correct* after an acceptance, as in the last two columns of Table 1. Overall, these empirical findings support WLDT.

In sum, our testable hypothesis is:

Hypothesis: Buyers adjust their current bids consistent with WLDT more often than random chance would suggest.

To test this, we evaluate how large the observed proportion $\hat{\pi}$ of correct revisions must be to reject the null hypothesis, which assumes no relationship between seller's decisions and bid revisions (this null is hereafter referred to as the "randomness assumption"). We use two permutation-based tests, which require minimal assumptions about the data generation process (Charles et al. 2023). The first strategy is completely distribution free under the randomness assumption, in the sense that it relies only on the permutation distribution obtained by reassigning seller decisions while preserving their empirical distribution. It answers the question: "Under randomness, how likely is it that we observe a proportion of bid revisions in accordance with WLDT as large as $\hat{\pi}$ or larger?" The second strategy, used for robustness check and to facilitate comparison, replicates the procedure by Selten et al. (2005) and addresses the question: "How likely is it to observe a $\hat{\pi}$ this much larger than the expected value of π under randomness?". Test statistics and *p-values* of the two testing strategies are reported in Table 2.

In our first approach, we approximate directly the null distribution $f_{\pi,0}$ of π, the proportion of WLDT-consistent revisions under randomness. We compare the observed $\hat{\pi} = 0.656$ (see the first row of Table 2) to this null distribution, calculating a *p-value* as $\text{Prob}_{f_{\pi,0}}(\pi \geq \hat{\pi})$. To this end, if Δp_{it+1} are realisations of the random variable ΔP_{it+1} and d_{it} of D_{it}, we have:

$$H_0 : \Delta P_{it+1} \perp\!\!\!\perp D_{it}, \qquad \text{against} \qquad H_1 : \Delta P_{it+1} \not\perp\!\!\!\perp D_{it}. \tag{3}$$

[12] Please note that the sum for the *Total* row in the table is 5452 (rather than 5640) because we compare the bid change at $t + 1$ with the decision at t, so we lose 188 observations, one for each buyer).

Table 2 Test outputs

Distribution-free	$\hat{\pi}$	se($\hat{\pi}$)	$\overline{\pi}$	$\hat{\pi} - \overline{\pi}$	*p-value*
	0.656	0.008	0.572	0.084	0.0001
t-test	$\overline{s}$	sd(s)	se($\overline{s}$)	$\hat{t}$	*p-value*
	0.085	0.091	0.007	12.9	$< 10^{-4}$

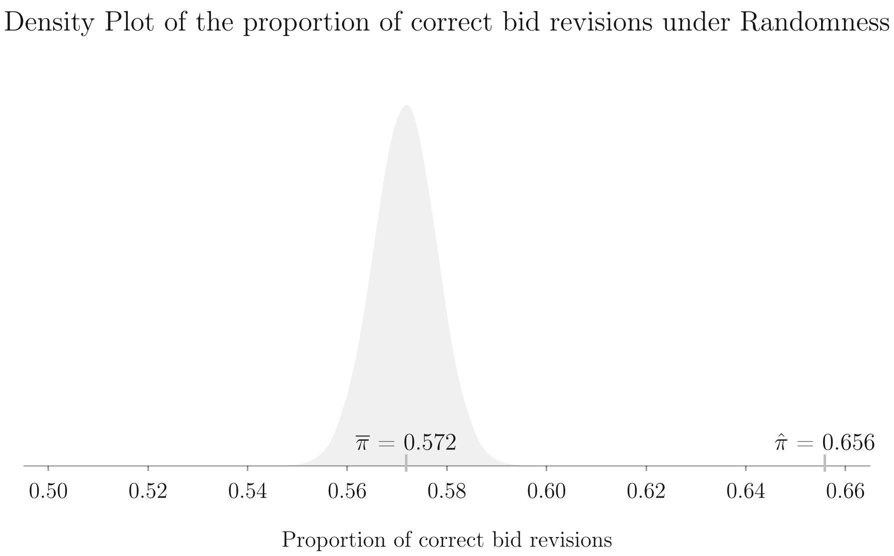

Fig. 4 Density plot of correct proportion of bid revision under randomness, with highlighted the observed value $\hat{\pi} = 0.656$

Additionally, ΔP_{it+1} is also independent of D_{jt} since a buyer is unaware of the seller's decision in another buyer/seller pair. We can therefore stack the bid-change sequences for the N buyers in a vector ΔP of dimension $(T-1) \times N$ and the lagged sellers decision sequences in a vector D of the same dimension. To generate an approximation $\hat{f}_{\pi,0}$ of $f_{\pi,0}$, we permute D while keeping ΔP fixed. With $B = 10{,}000$ permutations, Fig. 4 shows the resulting smoothed null distribution. The *p-value* is computed from the permutation distribution as

$$\textit{p-value} = \frac{1}{B+1} \sum_{b=0}^{B} \mathbf{1}\{\pi^{(b)} \geq \hat{\pi}\},$$

where $\mathbf{1}\{\cdot\}$ is the indicator function (equal to 1 if the condition holds and 0 otherwise), and where $\pi^{(b)}, b = 1, \ldots, B$, denotes the statistic obtained after the b-th permutation, with $\pi^{(0)} = \hat{\pi}$.[13] The test statistic and the corresponding *p-value* are reported in the

[13] Including the observed statistic $\hat{\pi}$ among the $B + 1$ values ensures that the permutation *p-value* is never reported as zero Phipson and Smyth (2010)

top part ("Distribution free") of Table 2. In our data, no permuted value exceeds $\hat{\pi}$, so $p = 1/(B+1)$; for $B = 10{,}000$ this yields $p = 1/10{,}001 \approx 0.0001$. We therefore reject the null. In Table 2 we also report the permutation mean $\overline{\pi}$ (computed over the B permutations), the effect size $\hat{\pi} - \overline{\pi}$, and the standard error of $\hat{\pi}$, denoted $\mathrm{se}(\hat{\pi})$. Because the WLDT-compliance indicator may be correlated across rounds within a buyer, we compute $\mathrm{se}(\hat{\pi})$ via a buyer-cluster bootstrap with $B^{(\mathrm{boot})} = 10{,}000$ resamples, drawing entire buyer histories (all rounds for a buyer) with replacement from the original data and recomputing $\hat{\pi}$ in each resample.[14] A key difference between our setup and that of Selten et al. (2005) is that in their design the parameter governing the wedge between buyer and seller valuations is fixed, whereas in our experiment q is redrawn each round. We therefore examine whether compliance with WLDT depends on the realised value of q. To do so, we rerun the same procedure as in the baseline analysis (in particular, using $B = 10{,}000$ permutations and $B^{(\mathrm{boot})} = 10{,}000$ bootstrap resamples) conditioning on q_t, the value of q in the round in which the buyer posts P_t. For each value q, an observation pairs the seller's decision in round $t-1$ with the buyer's bid revision $\Delta P_t = P_t - P_{t-1}$, computed between consecutive rounds as in the baseline analysis. By construction, P_{t-1} was typically chosen under a different realisation of q. Our objective is thus not to track within-q adjustments, but to assess whether the directional response implied by WLDT is weaker for some values of q or, conversely, whether certain ranges of q are associated with revisions that systematically run counter to WLDT predictions. We also report, as a reference, the horizontal line at the value $\hat{\pi}$ obtained in the unconditional analysis. Overall, there are no major differences across values of q. All observed $\hat{\pi}^{(q)}$ lie more than two standard errors above the corresponding expected values under randomness, $\overline{\pi}^{(q)}$, so we strongly reject the null under the randomness assumption (the *p-values* are 0.0001 for all q except for $q = 0.7$, where it is 0.0002). There is no sharp change in buyers' bid-revision behaviour at $q = 0.5$, as might have been suggested by the benchmark solution. We only observe a slight decline in $\hat{\pi}^{(q)}$ at $q = 0.9$, which is the only value for which the error bars do not touch the unconditional benchmark $\hat{\pi} = 0.656$. A possible explanation is the upward trend in bid levels as a function of q shown in Fig. 3. Even though our focus is on bid revisions rather than bid levels, a higher q reduces the wedge between buyer and seller valuations. Buyers who are primarily motivated to secure the firm (rather than to maximise expected profits as in the benchmark solution) may therefore keep bids high when q is large. Such inertia in bidding could in turn lead to slightly lower WLDT compliance.

As a second testing strategy, following Selten et al. (2005), we compare $\hat{\pi}$ with $\mathsf{E}(\pi)$ under $f_{\pi,0}$, using a buyer-level analysis. For each buyer i, we calculate r_i, the proportion of WLDT-consistent revisions, and $\bar{r}_i^B$, the average proportion over $B = 10{,}000$ permutations of D_{it}. The spread $s_i = r_i - \bar{r}_i^B$, with $\mu = \mathsf{E}_{f_{\pi,0}}(s_i)$, tests:

[14] Treating $\Delta P = 0$ as compliant with WLDT has a sizable mechanical effect on the reported compliance rate because unchanged bids are frequent in our data. We therefore run a conservative stress test that codes $\Delta P = 0$ as non-compliant. In this stress test, because a sizable fraction of choices are zero changes and are automatically counted as non-compliant, the compliance rate is lower both in the observed data and in each permutation of the test. We still reject the null with *p-value* $\approx 10^{-4}$: conditional on moving, the direction of bid revisions is markedly more aligned with the WLDT prediction than would be expected under the randomness benchmark. We thank an anonymous referee for pointing this out.

$$H_0 : \mu = 0, \qquad \text{against} \qquad H_1 : \mu > 0. \tag{4}$$

A *t-test* rejects H_0 at all reasonable significance levels (*p-value* $< 10^{-4}$).[15] Table 2 reports $\bar{s}$, its standard error (computed as $\mathrm{sd}(s)/\sqrt{188}$, where $\mathrm{sd}(s)$ is the sample standard deviation of $\{s_i\}_{i=1}^{188}$), the resulting t-statistic, and the *p-value*.

We conclude this section by providing some insights on the difference between the two testing strategies. The main difficulty in both protocols is that we do not know which is the proportion of correct bid revisions under randomness. The second protocol overcomes this issue by computing an expected value of π at a buyer level, and exploits the central limit theorem to test whether the average across buyers of the difference between observed and expected π is larger than 0. The logic of the first protocol is instead closely related to that of Fisher's exact test (Guido et al. 2015): rather than estimating unknown expectations, we condition on observed outcomes, the sequence of observed ΔP and the number of *Acceptances* and *Rejections*, and assess whether the observed association between bid revisions and seller's decisions is stronger than what would be expected by chance. This approach allows us to derive *p-values* without relying on a first step estimation of unknown quantities.

4 Conclusions

In order to apply LDT in an experimental AaC game with explicit strategic interaction between buyers and sellers, limited feedback and randomly changing percentage difference in the company valuation, the notion of ex-post optimality shall be adapted to take into account subjective evaluations about which bid would have been better in the previous interaction. We propose predictions based on rather natural evaluations and we show that even in this modified informational environment, buyers' bid revisions are strongly aligned with WLDT's directional predictions (relative to the appropriate no-association benchmark). Therefore, our analysis confirms and extends the validity of the predicting power of WLDT in experimental settings.

These findings suggest that WLDT could be successfully applied in other cognitively challenging environments in which counterfactual reasoning allows for directional learning. A possible application is in consumer choice under price uncertainty. In a repeated bundle-choice environment with stochastic prices and immediate partial feedback on past outcomes, one could test whether participants select bundles consistent with ex-ante optimisation (i.e., minimising expected expenditure) or instead adapt round-by-round in the direction suggested by (W)LDT.

Appendix

Introduction

Welcome to our experiment!

[15] A *Wilcoxon signed-rank* test also rejects H_0, with the size of the *p-value* comparable to that in the *t-test*. Given the robustness of the result and magnitude of the *p-values*, we are not concerned about symmetry or the normality of the distribution of s_i.

During this experiment, you will be asked to make several decisions and so will the other participants. Please read the instructions carefully. Your decisions, as well as the decisions of the other participants, will determine your earnings according to some rules, which will be shortly explained later. In addition to your earnings from your decisions over the course of the experiment, you will receive a participation fee of 10 euro. Besides this amount, you can earn more euro. However, there is also a possibility of losing part of the participation fee, as it will be explained in the next section of these instructions. *But do not worry: you will never be asked to pay with your own money, as your losses during the tasks will be covered by the participation fee.* The participation fee and any additional amount of money you will earn during the experiment will be paid individually immediately at the end of the experiment; no other participant will know how much you earned. All monetary amounts in the experiment will be computed in ECU (Experimental Currency Units). At the end of the experiment, all earned in ECUs will be converted into euro using the following exchange rate:

30 ECU = 1 euro

You will be making your decisions by clicking on appropriate buttons on the screen. All the participants are reading the same instructions and taking part in this experiment for the first time, as you are.

Please note that hereafter any form of communication between the participants is strictly prohibited. If you violate this rule, you will be excluded from the experiment with no payment. If you have any questions, please raise your hand. The experimenter will come to you and answer your questions individually.

Description of the experiment

This experiment is fully computerized. This experiment consists of the following **four phases, each composed by a different number of rounds**: Phase I of 1 round, Phase II of 30 rounds, Phase III of 12 rounds, and Phase IV of 10 rounds. After completing Phase I, you will proceed to Phase II; after completing Phase II, you will proceed to Phase III; after completing Phase III you will proceed to Phase IV. You can earn money in each phase of the experiment.

At the beginning and at the end of the Experiment, you are asked to reply to a short questionnaire.

At the beginning of the Experiment, each participant is randomly assigned one of two possible roles. Half the participants will be assigned the role of **Buyer**; the other half will be assigned the role of **Seller**. You will remain in the same role you have been assigned throughout the experiment.

In each of Phase I, II and III and in each of their rounds you will be matched with a different participant randomly assigned to you. In Phase IV you will decide individually and independently of your role.

Description of the task: Phase I

In Phase I selling of a firm between a Seller, who owns the firm, and Buyer can take place. You will be told if you are Buyer or Seller, and will be matched with one of the other participant in the other role. For example, if you are selected as Buyer, then you will be randomly and anonymously matched with another participant who is a Seller.

The computer will randomly select the value of the firm among the following values: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90 and 95 (all the values are equally likely). This value will be communicated only to the Seller. The Buyer will not learn the value of the firm selected randomly by the computer.

The Seller's evaluation of the firm is proportional to the value of the firm selected by the computer. This proportion will be randomly selected by the computer and can only take one of the following values: 10, 20, 30, 40, 50, 60, 70, 80 or 90% (all the values are equally likely). The Seller's evaluation is the value of the firm multiplied by the selected proportion. The proportion will be communicated to both, Buyer and Seller, whereas the value of the firm will be known only to the Seller. *Do not worry: the software will provide the information on the decision screen, depending on your role, Seller or Buyer.*

As an example, suppose that the computer selected a value of the firm equal to 90 and a proportion of 50%, so that the Seller's evaluation of the firm will be 45, corresponding to 50% of 90. *In this case, the Seller will find on the screen of the computer that the value of the firm is 90, the proportion is 50% and that the Seller's evaluation is 45; the Buyer will find on the screen only the proportion of 50%.* Another example: suppose that the computer selected a value of the firm equal to 90 and a proportion of 80%. In this case, the Seller's evaluation will be equal to 72, corresponding to 80% of 90. *In this case, the Seller will find on the screen of the computer that the value of the firm is 90, the proportion is 80 percent and that the Seller's evaluation is 72; the Buyer will find on the screen only the proportion of 80%.*

The Seller sends a value message to the Buyer about the value of the firm, which can be either true or false. Therefore, the value message is not necessarily equal to the firm value nor to the Seller's evaluation of the firm. The message consists of an integer value between 0 and 100.

After having received the message, the Buyer makes a take-it-or-leave-it offer to the Seller by proposing a price, an integer number between 0 and 100. When making this offer, the Buyer just knows the value message and by which proportion of the value the Seller evaluates the firm.

After having received the price offer of the Buyer, the Seller decides whether to accept it or not. If she accepts, the firm will be sold for the offered price to the Buyer. If she does not accept, no trade takes place. After the Seller has decided, the payoffs of Buyer and of Seller are calculated and individually communicated at the end of Phase I. These payoffs are calculated as explained below, and they are paid to all participants at the end of the experiment.

Calculation of the payoff in Phase I

The payoff of the unique round in Phase I does not depend on the value message and is calculated as follows:

If the Seller has accepted the offered price, the payoffs are:

- The Buyer earns the difference between the value of the firm and the accepted price
- The Seller earns the difference between the accepted price and the Seller's evaluation of the firm

An example: suppose that the firm value is equal to 45 and that the proportion of the firm value is 80%, so that the Seller's evaluation of the firm is 36. Suppose the Buyer offers a price equal to 40 and that the Seller accepts it. In this case, the Buyer earns $45 - 40 = 5$, and the Seller earns $40 - 36 = 4$.

Another example: suppose that the firm value is equal to 45 and that the proportion of the firm value is 80%, so that the Seller's evaluation of the firm is 36. Suppose the Buyer offers a price equal to 55 and that the Seller accepts it. In this case, the Buyer earns $45 - 55 = -10$, and the Seller earns $55 - 36 = 19$.

If the Seller does not accept the Buyer's offer, the payoffs are 0 for both Seller and Buyer.

Description of the task: Phase II

In Phase II, you will face for 30 rounds the same situation as in Phase I. As in the previous Phase, in each of the rounds you will be matched with a different participant randomly assigned to you.

The same instructions as in Phase I apply to Phase II, also the calculation of the payoffs.

The payment from this Phase will consist of the payoff of **one of the 30 rounds randomly selected**. For example, if round number five is selected, your payment for Phase II will be the payoff you earned in that round.

Calculation of the payoff in each round in Phase II

The payoff of each round in Phase II does not depend on the value message and is calculated as follows:

If the Seller has accepted the offered price, the payoffs are:

- The Buyer earns the difference between the value of the firm and the accepted price
- The Seller earns the difference between the accepted price and the Seller's evaluation of the firm

An example: suppose that the firm value is equal to 45 and that the proportion of the firm value is 80%, so that the Seller's evaluation of the firm is 36. Suppose the Buyer offers a price equal to 40 and that the Seller accepts it. In this case, the Buyer earns $45 - 40 = 5$, and the Seller earns $40 - 36 = 4$.

Another example: suppose that the firm value is equal to 45 and that the proportion of the firm value is 80%, so that the Seller's evaluation of the firm is 36. Suppose the

Buyer offers a price equal to 55 and that the Seller accepts it. In this case, the Buyer earns $45 - 55 = 10$, and the Seller earns $55 - 36 = 19$.

If the Seller does not accept the Buyer's offer, the payoffs are 0 for both Seller and Buyer.

Description of the task: Phase III

In Phase III, you will face for 12 rounds the same situation as in Phase I. As in the previous Phase, in each of the rounds you will be matched with a different participant randomly assigned to you.

The same instructions as in Phase I apply to Phase III.

At the beginning of the Phase, you will be asked if you prefer to be paid on the basis of the payoff of **one of the 12 rounds randomly selected** *or* on the basis of **the average payoff of the 12 rounds**. On the basis of your choice, the computer will calculate your payoff for this Phase.

Calculation of the payoff in each round in Phase III

The payoff of each round in Phase II does not depend on the value message and is calculated as follows:

If the Seller has accepted the offered price, the payoffs are:

- The Buyer earns the difference between the value of the firm and the accepted price
- The Seller earns the difference between the accepted price and the Seller's evaluation of the firm

An example: suppose that the firm value is equal to 45 and that the proportion of the firm value is 80%, so that the Seller's evaluation of the firm is 36. Suppose the Buyer offers a price equal to 40 and that the Seller accepts it. In this case, the Buyer earns $45 - 40 = 5$, and the Seller earns $40 - 36 = 4$.

Another example: suppose that the firm value is equal to 45 and that the proportion of the firm value is 80%, so that the Seller's evaluation of the firm is 36. Suppose the Buyer offers a price equal to 55 and that the Seller accepts it. In this case, the Buyer earns $45 - 55 = 10$, and the Seller earns $55 - 36 = 19$.

If the Seller does not accept the Buyer's offer, the payoffs are 0 for both Seller and Buyer.

Description of the task: Phase IV

Phase IV consists of 10 rounds; during this Phase, you will not interact with other participants. During this phase, you are asked to choose between pairs of lotteries. In particular, in each round for each lottery pair you have to assess which one you would prefer to play.

At the end of the experiment, one round will be randomly selected for payment, and the computer will play on your screen the lottery that you have preferred in this round. The payment of Phase IV is given by the result of this lottery.

Your final payment

Your final payment will be displayed on the screen at the end of the experiment. It is determined as the sum of:

- Payoff from the unique round in Phase I (in euro)

- Payoff from one randomly selected round in Phase II (in euro)
- Payoff from EITHER one randomly selected round OR an average payment between 12 rounds from Phase III (in euro)
- Payoff from one randomly selected round in Phase IV (in euro)
- Participation fee.

Acknowledgements We thank Daniela Di Cagno and Werner Güth for useful comments.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

Data Availability Computations are carried out using R Core Team (2021) codes are available at https://github.com/ollimca/WLDT.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Akerlof GA (1970) The market for lemons: quality uncertainty and the market mechanism. Q J Econ 84(3):488–500

Alberti F, Cartwright A, Cartwright E (2021) Predicting efficiency in threshold public good games: a learning direction theory approach. Working papers in economics and finance 2021-01, University of Portsmouth, Portsmouth Business School

Angelovski A, Di Cagno D, Güth W, Marazzi F (2020) Telling the other what one knows? Strategic lying in a modified acquiring-a-company experiment with two-sided private information. Theor Decis 88(1):97–119

Ball SB, Bazerman MH, Carroll JS (1991) An evaluation of learning in the bilateral winner's curse. Organ Behav Hum Decis Process 48(1):1–22

Bazerman MH, Samuelson WF (1983) The winner's curse: an empirical investigation. In: Tietz R (ed) Aspiration levels in bargaining and economic decision making: Proceedings of the third conference on experimental economics, Winzenhohl, Germany. Springer-Verlag, pp 186–200

Bereby-Meyer Y, Grosskopf B (2008) Overcoming the winner's curse: an adaptive learning perspective. J Behav Decis Mak 21(1):15–27

Charness G, Levin D (2009) The origin of the winner's curse: a laboratory study. Am Econ J Microecon 1(1):207–36

Di Cagno D, Galliera A, Güth W, Pace N, Panaccione L (2016) Make-up and suspicion in bargaining with cheap talk: an experiment controlling for gender and gender constellation. Theor Decis 80(3):463–471

Di Cagno D, Galliera A, Güth W, Pace N, Panaccione L (2017) Experience and gender effects in acquisition experiment with value messages. Small Bus Econ 48(1):71–97

Di Cagno D, Güth W, Lohse T, Marazzi F, Spadoni L (2024) Who cares when value (mis)reporting may be found out? An acquiring-a-company experiment with value messages and information leaks. J Behav Exp Econ 108:102151

Dittrich DAV, Güth W, Kocher MG, Pezanis-Christou P (2012) Loss aversion and learning to bid. Economica 79(314):226–257
Eva I, Hoppe PW (2013) Schmitz Contracting under incomplete information and social preferences: an experimental study. Rev Econ Stud 80(4):1516–1544
Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114(3):817–868
Fischbacher U (2007) z-tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10(2):171–178
Foreman P, Murnighan JK (1996) Learning to avoid the winner's curse. Organ Behav Hum Decis Process 67(2):170–180
Fudenberg D, Peysakhovich A (2016) Recency, records, and recaps: learning and nonequilibrium behavior in a simple decision problem. ACM Trans Econ Comput (TEAC) 4(4):1–18
Galliera A (2018) Self-selecting random or cumulative pay? A bargaining experiment. J Behav Exp Econ 72:106–120
Greiner B (2015) Subject pool recruitment procedures: organizing experiments with orsee. J Econ Sci Assoc 1(1):114–125
Grosskopf B, Bereby-Meyer Y, Bazerman M (2007) On the robustness of the winner's curse phenomenon. Theor Decis 63(4):389–418
Imbens GW, Rubin DB (2015) Causal inference for statistics, social, and biomedical sciences: an introduction. See chapter 5: randomization inference and exact tests. Cambridge University Press, New York
Güth W, Pull K, Stadler M, Zaby AK (2019) Compulsory disclosure of private information: theoretical and experimental results for the acquiring-a-company game. J Inst Theor Econ 175(3):502–523
Holt CA, Sullivan SP (2023) Permutation tests for experimental data. Exp Econ 26(4):775–812
Martínez-Marquina A, Niederle M, Vespa E (2019) Failures in contingent reasoning: the role of uncertainty. Am Econ Rev 109(10):3437–74
Neugebauer T, Selten R (2006) Individual behavior of first-price auctions: the importance of information feedback in computerized experimental markets. Games Econ Behav 54(1):183–204
Neugebauer T, Selten R (2019) Experimental stock market dynamics: excess bids, directional learning, and adaptive style-investing in a call-auction with multiple multi-period lived assets. J Econ Behav Org 157:209–224
Phipson and Smyth (2010) Permutation P-values should never be zero: calculating exact p-values when permutations are randomly drawn, https://doi.org/10.2202/1544-6115.1585
R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Selten R (2004) Learning direction theory and impulse balance equilibrium. In: Daniel F, Alessandra C (eds) Economics lab: an intensive course in experimental economics, Routledge advances in experimental and computable economics, chapter 12. Routledge, Taylor & Francis, London, pp 133–140
Selten R, Stoecker R (1986) End behavior in sequences of finite prisoner's dilemma supergames a learning theory approach. J Econ Behav Organ 7(1):47–70
Selten R, Abbink K, Cox R (2005) Learning direction theory and the winner's curse. Exp Econ 8(1):5–20

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.