European Economic Review ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

European Economic Review

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

Market experience eliminates some anomalies—and creates new ones

Jacinto Braga [a], Steven J. Humphrey [b,*], Chris Starmer [b]

[a] Department of Economics and International Development, University of Bath, Bath BA2 7AY, UK
[b] School of Economics, University of Nottingham, Nottingham NG7 2RD, UK

ARTICLE INFO

Article history:
Received 11 December 2006
Accepted 9 June 2008

JEL classification:
C91
D81
D83

Keywords:
Preference reversal
Vickrey auction
Refining
Loss experience
Price following

ABSTRACT

We report two experiments which investigate whether experience of decision making in repeated markets purges behaviour of preference reversals. We investigate two behavioural mechanisms that may be shaping bids in repeated auctions: A tendency to adjust bids towards previously observed market prices, and a tendency to reduce bids following bad market outcomes. We find little support for the former but strong support for the latter. Also, whilst 'just enough' market exposure eliminates the typical preference reversal phenomenon, continued exposure fosters the mirror image anomaly. Therefore, although market experience shapes behaviour, in our experiments, it does not generally promote consistency with standard preference theory.

Experimental research spanning over 50 years has documented a wide range of behavioural phenomena running contrary to core economic theories of preference such as expected utility theory and Hicksian consumer theory (Camerer, 1995; Starmer, 2000). At face value, these 'anomalies' challenge the descriptive validity of those theories and a broader set of economic models constructed using them. Some suggest that these anomalies prompt questions about the welfare evaluations of market outcomes and corresponding policy judgements (see the symposium in Environmental and Resource Economics, 2005, pp. 1–181).

Emerging evidence, however, suggests that some preference anomalies subside with experience in particular types of experimental market. The strongest evidence of this relates to the disparity

* Corresponding author. Tel.: +44 115 951 5472; fax: +44 115 951 4159.
E-mail address: steve.humphrey@nottingham.ac.uk (S.J. Humphrey).

0014-2921/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.euroecorev.2008.06.005

between willingness to accept (WTA) and willingness to pay (WTP). Studies consistently reveal that WTA exceeds WTP by an amount that is hard to rationalise with standard preference theory (Bateman et al., 1997). However, the disparity tends to close when valuations are elicited in experiments featuring various kinds of repeated market (Coursey et al., 1987; List and Shogren, 1999; Shogren et al., 1994, 2001; Loomes et al., 2003; Plott and Zeiler, 2005). Recent work by List (2003) provides some complementary field evidence for a naturally occurring market.

Related, though more sparse, findings have been reported for the preference reversal phenomenon. Preference reversal (PR for short) is a *systematic* inconsistency in the preference ranking over a pair of alternatives when elicited using different, but theoretically equivalent, methods (usually choice and monetary valuation). PR is a particularly troubling anomaly for economics because it challenges not only particular classes of preference theory, but also the more fundamental assumption that human choices can be adequately modelled in terms of stable preferences (Grether and Plott, 1979). It has also proved to be robust to a wide range of experimental controls (Seidl, 2002; Cubitt et al., 2004). However, paralleling the evidence related to the WTP/WTA disparity, Cox and Grether (1996) observe that PR decays when valuations are elicited in a repeated Vickrey (1961) auction. Although earlier studies have shown that it is possible to erode PR by explicitly punishing inconsistent preferences (Chu and Chu, 1990), the results of Cox and Grether represent an important landmark in providing the first evidence that PR might decay simply as a consequence of participation in something resembling a *naturally occurring* market.

Such evidence has prompted some to conjecture that the experience of participation in markets might have some general tendency to erode preference anomalies. Plott's (1996) *discovered preference hypothesis* is an explicit conjecture of this form and, more recently, Loomes et al. (2003) discuss the closely related *refining hypothesis*. As we interpret them, these hypotheses share the following core assumptions: (i) Agents have stable underlying preferences which (ii) are 'anomaly-free' (i.e., expected utility risk preferences; Hicksian consumer preferences); (iii) preferences stated in the context of particular elicitation mechanisms may not always coincide with underlying preferences; (iv) market experience, by giving agents incentives and feedback, tends to promote greater consistency between stated and underlying preferences. We will use the label 'refining' for any theory satisfying (i)–(iv) above. If markets have general refining tendencies, that would have significant implications for the development and use of economic theory: it could, for instance, vindicate the use of standard preference theories for modelling behaviour in real markets, despite the laboratory anomalies.

The existing evidence, however, reveals that while some forms of market experience may drive out anomalies, others do not. For instance, Knetsch et al. (2001) found that while the WTA/WTP disparity disappeared in a second-price Vickrey auction (henceforth SPA), it persisted in a strategically equivalent, second-to-last price Vickrey auction. Similarly, while Cox and Grether found that PR subsided in the SPA, it persisted in the Becker, DeGroot, Marschak (1964) variant of the Vickrey mechanism (henceforth BDM). This evidence of mechanism dependence shows that it is not market experience per se that drives out anomalies. So, to understand the dynamics of stated preference, it is necessary to explore which particular aspects of market experience promote changes in stated preferences, and the extent to which they foster consistency with standard preference theory.

To this end, we have looked for causal factors which might explain the observed dynamics of stated preference in existing studies. In order to establish purchase on the problem, in this paper, we confine attention to a single anomaly (studies of PR) and a particular class of market mechanism (variants of repeated Vickrey auctions). Focussing in this way, we have identified two behavioural mechanisms which might contribute to explaining the existing data: *Price following* and *loss experience*. Price following is a hypothesised tendency for participants in repeated markets to adjust their 'bids' towards previously observed market prices.[1] Loss experience involves subjects adjusting their bids following an aversive outcome in the previous market period.

[1] In this paper, we deal only with auctions where the auctioneer is buying from the subjects, and throughout 'bid' means the minimum selling price stated by a subject.

We report two new experiments that test for the operation of these mechanisms. We find only weak support for price following, but strong support for the operation of loss experience. In the second study, we exploit PR as a vehicle for investigating whether loss experience leads to anomaly-free stated preferences. Our results suggest that it does not. With 'just enough' exposure to loss experience, our markets engender the disappearance of the standard asymmetric pattern of PR, but with longer exposure a new anomaly emerges which is the mirror image of typical PR. In what follows, Section 1 introduces our hypotheses; Sections 2 and 3 report the experiments and Section 4 concludes.

1. Background—formulating hypotheses for the case of preference reversal

In the classic PR setup, subjects choose between paired bets and place separate monetary values on them. One of the bets, the '*P*-bet', offers a relatively large probability of a moderate positive outcome, while the other, the '*$*-bet', offers a relatively small chance of a larger win (for compactness, we sometimes refer to these bets as simply '*P*' and '*$*', respectively). Experiments implementing variants of this basic design reliably reveal systematically different preference rankings of the two gambles:[2] Many subjects choose *P* and value *$* more highly (a 'standard reversal') but very few do the opposite, that is, choose *$* and value *P* more highly (a 'non-standard reversal'). Interest in PR stems largely from the fact that observed inconsistencies tend to be patterned in this highly predictable way.[3]

While there is no generally agreed theory of what causes PR, a common interpretation (see Tversky et al., 1990) is that it arises as a consequence of biases in the valuation of bets and, in particular, a tendency towards relative 'overvaluation' of *$*-bets.[4] Overvaluation might stem from the anchoring of valuation on the positive outcome and insufficiently adjusting from the anchor (Lichtenstein and Slovic, 1971). Alternatively, and in the context of experiments that employ incentive compatible revelation mechanisms such as the BDM or the SPA, subjects might deliberately overstate reservation prices believing (erroneously) that it is in their interests to do so (Bateman et al., 1997). If bets are overvalued, *$*-bets may be relatively more prone to it than *P*-bets which typically offer a near certain payoff. Given that assumption, overvaluation might explain the typical asymmetric pattern of PR. On this interpretation, PR arises as a consequence of biased value elicitation procedures. Other theories, however, suggest that PR is a property of properly measured, but non-standard preferences (Loomes and Sugden, 1983; Sugden, 2003). Whatever the correct explanation for PR, it has proved to be a highly replicable phenomenon. Given this, the work of Cox and Grether is important because their findings suggested that PR may be less robust in those environments that are of central concern to economics, that is, market environments.

1.1. Cox and Grether's experiment

Prior to Cox and Grether's work, most PR studies involved one-shot decisions. In contrast, Cox and Grether used repeated market valuation tasks and provided subjects with immediate feedback on the outcome of their decisions: In each of ten valuation rounds (a sequence of 5 in a row each for *P* and *$*), subjects were endowed with a lottery and given the opportunity to sell it. Those who sold received a sure sum of money (the price); those who did not sell played out their lotteries and either won or

[2] Early examples include Lichtenstein and Slovic (1971, 1973) and Grether and Plott (1979) but the literature is large and an extensive review is available in Seidl (2002).

[3] The existence of some inconsistency between an individuals' reported rankings in choice and valuation tasks, by itself, is not especially surprising. People might, for instance, make a mistake in one or more task leading to some inconsistent rankings. The fact that inconsistencies are concentrated on standard reversals, however, renders their interpretation as 'mistakes' less convincing.

[4] To fix the notion of 'overvaluation' suppose that for any lottery L, an individual has a well-defined certainty equivalent C^L which we define behaviourally: Given a direct choice between L and a sure amount of money M, the agent strictly prefers L if $M < C^L$ and does not strictly prefer L if $M \geqslant C^L$. An agent overvalues a lottery L when their reserve in a selling task is greater than C^L.

lost. In testing for PR, Cox and Grether compared these market-elicited valuations with the ordering of *P* and *\$* revealed in a choice task at the end of the experiment. This involved a one-off choice between the two lotteries—as has been standard in the PR literature—and no market mechanism was involved in this task.

Our interest centres on differences in the incidence of PR comparing across two different mechanisms used to incentivise the valuation tasks.[5] For one group of subjects, valuations were elicited in a SPA. The SPA is a special case of the generic *n*th price Vickrey auction which, in the case of selling, works as follows: N agents are each endowed with a single unit of a good to sell; each agent submits a (sealed) bid and the $n-1$ agents submitting the lowest bids sell at a uniform price equal to the *n*th lowest bid. For each participant in such an auction, it is a weakly dominant strategy for them to bid their true value. The SPA is the special case of this mechanism where $n=2$. For another group of subjects, valuations were elicited using the BDM mechanism. In the BDM, a subject is endowed with a good and then states a minimum selling price (r). The experimenter draws a random bid (b) and if $b>r$, the individual sells and receives b; if $b\leqslant r$ the individual does not sell. It is well-known that an expected utility maximiser should report a value of r equal to their true value.

In both treatments, initial valuations of *P* and *\$* were systematically inconsistent with the choice, displaying the typical preponderance of standard over non-standard reversals. But with repetition, behaviour in the two treatments diverged. Comparing the final valuations with the choice, the typical, asymmetric, reversal pattern was still present in the BDM treatment but had disappeared in the SPA. In the latter, inconsistencies were still frequent, but the two types of reversal occurred with approximately equal frequency.[6] Since the BDM can be understood as an SPA in which a subject competes against one other bidder who bids at random, it is intriguing that the two institutions should produce different results. We now present two hypotheses which might account for this difference between these widely used incentive mechanisms.

1.2. Two mechanisms

The *price following hypothesis* is an empirical conjecture that subjects participating in repeated experimental auctions have some tendency to adjust their bids towards previously observed market prices. Such a tendency has been invoked as a possible account of behaviour in some studies concerned with the WTA–WTP disparity (Knetsch et al., 2001; Loomes et al., 2003) and is suggested by a wide range of evidence from psychological studies (for example, Simonson and Drolet, 2004). Price following would be sufficient to explain the erosion of PR in Cox and Grether's SPA. There were five participants in each of the SPA markets and since the price was the second lowest bid, the majority of bids would have been above it, so price following would have tended to decrease average bids. There is good reason to expect that such a trend would be more pronounced for \$-bids than for *P*-bets because the two *P*-bets used in these studies offered almost sure wins (with probabilities of 97% and 81%), so their valuations would have been highly concentrated just below the value of the lottery prize. Consequently, the difference between the second lowest and other bids would have tended to be small, leaving relatively little scope for adjustment towards the price. By contrast, bids for \$-bets tend to be more dispersed, leaving considerably more scope for price following to effect downward adjustments. A relatively large reduction of the \$-bids relative to the *P*-bids could then have made standard PR less frequent and non-standard PR more frequent, leading to the approximately equal frequencies of the two types of reversal that Cox and Grether observed by the final (fifth) round of their SPA treatment.

There are two reasons why price following might have eroded PR in the SPA but not in the BDM. One is that the BDM price is known to be just a random number which varies erratically from round to round. Consequently, it may have been a less salient cue than the price in a SPA. The other reason

[5] In addition to the two mechanisms that we discuss, Cox and Grether also elicited valuations using an English clock auction. However, since our concern focuses on what properties of mechanisms may cause PR to decay, this treatment is of less interest to us because it produced no evidence of PR to begin with.

[6] This is based on data from a subset of participants who faced real financial incentives.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

requires the plausible assumption that most subjects will keep their bids below the highest outcome of the lottery. To erode PR, some subjects have to bring their *$*-bids below their *P*-bids. Cox and Grether's BDM price was uniformly distributed between *$* 0 and 9.99, and thus would have tended to be too high to draw *$*-bids below the highest outcomes of the *P*-bets (which were *$* 4.00 and 2.00).

As a second candidate explanation of the difference between the dynamics of these two mechanisms, we propose the *loss experience hypothesis*. This hypothesis assumes a tendency for subjects to reduce their stated valuations of lotteries, in an incentive compatible mechanism, following the experience of *losing*.[7] There is more than one reason why individuals might be prone to a loss experience effect. One follows from the 'overvaluation' interpretation of PR: If subjects have a tendency to overvalue bets, it is conceivable that losses might prompt them to reconsider their previously stated valuations.[8] Another follows from the psychology literature on affect which suggests losses will generate negative feelings which might, in turn, reduce a subject's perceived value of a lottery (for a survey of some relevant findings see Isen, 1999).

If subjects were prone to a loss experience effect, it would tend to reduce standard PR and to increase non-standard PR, thus eroding the typical asymmetry, in any environment where individuals repeatedly face opportunities to trade *P*- and *$*-bets and receive round-by-round feedback on lottery outcomes. The reason is that bids for *$*-bets will tend to fall more than those for *P*-bets simply because subjects face a much higher probability of losing with the *$*-bets. The loss experience hypothesis also leads us to expect stronger erosion of the typical reversal asymmetry in the SPA than in the BDM treatment of Cox and Grether's study. This is because, in Cox and Grether's experiment, subjects were relatively more likely to play their bets in the SPA treatment.[9] Thus, relative to the SPA, the BDM mechanism would have led to fewer opportunities for loss experience to operate.

In what follows, we present two experiments which test the hypotheses just discussed. Before describing the experiments, however, we make two methodological observations. The first is that we have formulated our hypotheses as simple empirical tendencies rather than as formal models of bid adjustment. This reflects an underlying scientific strategy. The development of formal models would require assumptions that we do not yet have an evidential basis for. For example, there are different ways in which one might formalise the loss experience hypothesis: Agents might have precise underlying preferences and loss experience might prompt adjustment of bids towards them; alternatively, agents might have imprecise preferences where losses, and possibly gains too, lead to (perhaps bounded) changes in an agent's perceived value of a bet. Given these different possible formulations, our strategy is to begin by exploring simple but *testable* empirical conjectures that presuppose minimal theoretical structure. We see this as a natural first step towards developing more sophisticated, but empirically grounded, models of bid dynamics. The second observation is that our two hypotheses stand in quite different relation to the refining hypothesis. A loss experience effect might account for some refining tendencies of markets if it prompted the correction of errors in earlier decisions (though, a loss experience effect in the context of a model of imprecise preferences need not necessarily be refining). By contrast, price following cannot be a refining mechanism because its operation would cause bids to become *endogenous* to the market (i.e., a function of other agents' bids and features of the market such as how the market clearing price is selected). Consequently, bids influenced by price following could not generally be expected to gravitate towards underlying preferences that are independent of the market experience. This important conceptual distinction provides an added motive for investigating which, if either, of these mechanisms is at work.

[7] In this paper, 'losing' means receiving the non-positive outcome of the lottery, which is zero in our experiments but negative in many PR experiments, such as Cox and Grether's.

[8] On this account, we could interpret the loss experience effect as either prompting the agent to reconsider the merits of deliberate mis-statements of value, or prompting them to realise that they had, inadvertently, mis-stated.

[9] In the SPA, four out of five subjects in each auction (i.e., 80%) would fail to sell and would therefore play their bets. In the BDM group, subjects would play the gamble only if their bid was above the random BDM price. As Cox and Grether's BDM offer was uniformly distributed between 0 and $ 9.99, it would take a bid of $ 8.00 for the ex-ante probability of playing the bet to reach 80%. A bid of $ 8.00 would be rather high for the *$*-bets used in Cox and Grether's experiment: {$ 16, 31%; −$ 1.50, 69%} and {$ 9, 19%; −$ 0.50, 81%}.

2. Experiment A—testing the price following hypothesis

The primary objective of Experiment A was to test the price following hypothesis. A subsidiary objective (discussed in Section 2.3) was to begin exploration of the loss experience hypothesis. The design involved a comparison of behaviour across two treatments. One of them (which we label SPA) was set up to be a close replication of the Cox and Grether SPA treatment. In our SPA, following Cox and Grether, we observed individual reservation values for a pair of *P* and *$*-bets in a repeated second price Vickrey auction plus a one-off choice between *P* and *$* elicited at the end. This design allows examination of how PR changes as bids evolve in a repeated market against the baseline of a single (non-market) choice. To the extent that we wish to assess how PR is affected by 'market experience', one might query our decision to adopt a choice task which features neither a market nor repetition. Several factors motivate this decision. The first flows from our desire to replicate Cox and Grether's SPA treatment as part of Experiment A.[10] Notwithstanding that motivation, we contend that a straightforward pairwise choice task is the simplest and most transparent way of eliciting an agent's ranking of a pair of gambles. As such, it provides a natural benchmark against which to compare rankings elicited in other, inevitably more complex, ways (e.g., market valuations). Of course, it is possible that choices may be influenced by experience too, perhaps as a consequence of experiencing gamble outcomes.[11] We think this is an interesting possibility and one which Experiment B is designed, in part, to explore.

In SPA, we implemented a standard second price auction buying just one lottery from the individual with the lowest bid at a price equal to the second lowest bid. If price following operates in this market, it should reduce *$*-bids thereby eroding the typical asymmetric PR pattern. This would be in line with what Cox and Grether observed in their SPA condition. But if price following were the *primary* cause of PR erosion in the SPA, we would expect to see the reverse effect in markets where the price is near the top of the distribution of bids. To test this implication, in our second condition, subjects face an identical decision environment except that the auction rule is a second-to-last-price auction (henceforth StLPA). This is another variant of the Vickrey mechanism, but instead of buying one unit at the second lowest price, we buy the lotteries from all but the two highest bidders at a uniform price equal to the second *highest* bid (i.e., second-to-last lowest bid). In this market, on the price following hypothesis, *$*-bids should *increase*, standard reversals should become *more* frequent and the standard asymmetry between the two types of reversal should become *stronger*. Hence, if we are able to replicate Cox and Grether's findings in the SPA, comparison with the StLPA will test the hypothesis that price following is the primary cause of bid adjustment.

2.1. Procedures

The SPA and StLPA conditions were identical except for the number of lotteries bought and the price we paid. Table 1 outlines the experimental design and indicates the sample sizes. We used two pairs of lotteries, a high-value pair and a low-value pair (see the first column of Table 1). Each subject dealt with one pair only. Subjects in groups of either seven or eight were each given a lottery and then submitted a sealed bid to sell it. Bids above the largest outcome of the bet were not allowed. Once all bids had been submitted, subjects learned the market price and whether they had sold their bet. Those who failed to sell played them out and immediately learned the outcome. After each round, all lotteries were resolved via a single public draw of a numbered chip from a bag.[12] There were five consecutive auctions for one bet, then five for the other bet. The order of the bets was

[10] Although their design features other treatments beyond the SPA, in every one of these, Cox and Grether test for PR by comparing a pair of valuations with a non-market choice. In one of these treatments, they did implement what might be called 'market-choice' tasks: these involved decisions in an English clock auction which were then used to generate valuations for comparison with behaviour in the standard non-market choice task. It may, however, be interesting for future research to explore the effects of comparing market valuations with choices also elicited through some market mechanism.

[11] There is little or anything that we think a subject could learn about the task itself.

[12] This procedure has the feature that subjects who had sold their bets knew what their payoff would have been had they kept them.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

Table 1
Outline and number of subjects (and auction groups) of Experiment A

Pairs of *P*- and $-bets[a]	SPA—second price auction		StLPA—second-to-last price auction		Total
	7 bidders	8 bidders	7 bidders	8 bidders	
High value: (£8, 97%), (£32, 31%)[a]	21 (3)	24 (3)	21 (3)	24 (3)	90 (12)
Low value: (£5, 81%), (£18, 19%)[a]	21 (3)	24 (3)	21 (3)	24 (3)	90 (12)
Total	42 (6)	48 (6)	42 (6)	48 (6)	180 (24)

[a] Bets offer the stated amount of money with the stated probability and zero otherwise.

randomised. In the final task, subjects made a straight choice between the *P*- and $-bets. Subjects had the details of all their previous tasks permanently displayed on their computer screens; for each auction round, they could see the lottery, their bid, the market price, whether they had sold or played the lottery, and their outcome (earnings) in the task.

We designed Experiment A with the intention of making our SPA condition closely replicate that implemented by Cox and Grether. There are, however, two exceptions to this. One difference is that, at the level of the individual gamble, we increased payoffs roughly threefold.[13] The second difference is that our design employed the random lottery incentive procedure and paid subjects according to the outcome of only one of their tasks (selected at random at the end of the experiment). In contrast, in Cox and Grether's experiment (for the subjects facing real incentives), all tasks counted for real. The random lottery incentive procedure has the advantage of avoiding potential portfolio and income effects that could arise when all tasks are for real[14], and, with a fixed budget, it allows larger incentives at the level of the individual task. There is, however, a literature which has questioned the validity of the procedure. Specifically, Holt (1986) has shown that if preferences do not obey the independence axiom of expected utility theory, then the random lottery incentive procedure may be a biased elicitation mechanism which could, under certain assumptions, give rise to spurious observation of PR. While Holt's argument correctly identified a logical possibility, subsequent experimental research has discounted Holt's explanation of PR (Cubitt et al., 2004). Moreover, extensive direct tests of the mechanism have failed to produce any evidence that the mechanism leads to biases in elicited preferences in practice (Starmer and Sugden 1991; Cubitt et al., 1998). Hence, like many other researchers in the field (e.g., Hey and Lee, 2005a, b), we conclude that the random lottery incentive procedure can be robustly defended as a useful practical tool of experimental research.

The experiment was conducted in the CeDEx Laboratory at the University of Nottingham. Subjects were recruited by email to pre-registered volunteers from across the student population. The treatments implemented in each session were determined randomly. Instructions explaining the general structure of the experiment were read aloud from a script. As the two types of auction could take place simultaneously in a session, the relevant auction was explained by means of a computerised practice session.[15] The task for which each subject would be paid was determined randomly for each subject.[16]

[13] With one exception, our lotteries can be obtained from theirs by replacing the small loss with a zero outcome, substituting pounds for dollars, doubling the nominal amounts, and rounding the probabilities from thirty sixths to percentage points. The exception is the *P*-bet of the low-value pair, where the above transformation would yield a positive outcome of £4, which we thought may be too small to engage subjects.

[14] There is some evidence of an income effect in Cox and Grether's results (though it is significant only at 10%).

[15] The practice session involved three mock auctions with each subject competing with computer-simulated bidders. Subjects could play the mock auctions as many times as they wished, but had to play each of them at least once so that the qualitatively distinct outcomes of the auctions were illustrated (i.e., selling the lottery, not selling the lottery, and randomized tie-breaking by the computer). The presentation of the practice auctions corresponded with that of the proper tasks but with the addition of a superimposed instructions window, which subjects could scroll up and down at will. The instructions are available from the authors on request.

[16] Each subject randomly selected an envelope at the start of each session which they were told contained the number of their 'real' task. They were not allowed to open the envelope until the end of the experiment.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

Table 2
Frequency of responses in Experiment A and mean difference between *P*- and *$*-bids

	Choice	SPA—second price auction (highest bid)			StLPA—second-to-last price auction (highest bid)		
		P	*$*	Equal	*P*	*$*	Equal
Round 1	*P*	7 (0.9)	**39** (8.7)	5	9 (3.6)	**41** (7.3)	0
	$	**4** (3.4)	34 (9.3)	1	**5** (3.8)	35 (9.2)	0
				$p = 0.000$			$p = 0.000$
Round 5	*P*	35 (2.6)	**15** (5.6)	1	17 (2.5)	**30** (4.9)	3
	$	**16** (2.7)	21 (5.9)	2	**7** (2.7)	32 (7.1)	1
				$p = 0.640$			$p = 0.000$

Note: The table reports frequencies of responses with reversals in bold. Values within parentheses are mean differences between *P*- and *$*- bids. *p*-Values result from a test (binomial distribution) of the null hypothesis that standard and non-standard reversals are equally likely against the alternative hypothesis that standard reversals are more likely.

2.2. *Results for Experiment A*

Sessions lasted approximately 50 minutes and average earnings were £8.60 per subject. Out of 180 participants, 100 were male. Table 2 provides a summary of the incidence of PR, based on either first or final valuations, for the two auction types. The results are aggregated across pairs of lotteries and group sizes. The table reports the frequencies of the possible response types, where a 'response' is a combination of two preference rankings: One revealed by a subject's choice between *P* and *$*; the other implied by their bids for *P* and *$*. A response displays a standard reversal if *P* is chosen and *$* is valued more highly, a non-standard reversal if *$* is chosen and *P* is valued more highly or a consistent preference when one bet is chosen and valued at least highly as the other. Since subjects had to choose one of the lotteries, subjects bidding equal values for them can be interpreted as being indifferent. These responses are identified in a separate column ('equal') and are not counted in either category of reversal. Frequencies of reversals are shown in bold face. We conduct statistical tests for asymmetry between the two types of reversal based on the null hypothesis that standard and non-standard reversals are equally likely.[17]

The first question we address is whether behaviour in our SPA condition replicates the findings of Cox and Grether. The answer is 'yes'. In the first round of the SPA, we observe the typical asymmetric pattern of PR. The 39 standard reversals account for almost half of all responses whereas non-standard reversals are rare (only 4). The null hypothesis that the two types of reversal are equally likely is confidently rejected ($p = 0.000$). By round five, standard reversals have fallen to less than half their initial rate and the frequency of non-standard reversals has quadrupled. Just as Cox and Grether observed by their fifth round, these changes render the frequencies of the two types of reversal approximately equal (15 standard and 16 non-standard). This leads to our first main finding: The experience of participating in a second price Vickrey auction with the opportunity to sell *P*- and *$*-bets has a replicable tendency to reduce standard reversals and increase non-standard reversals.

The behaviour observed in the SPA is consistent with all three of the hypotheses we have discussed: Refining, loss experience, and price following. We now turn to examine behaviour in the StLPA. In this context, price following has the distinctive implication that average bids should *rise*.

[17] This test allows for the possibility that PR of either type might occur as a consequence of some non-systematic components of decisions generating 'error' or 'noise' in stated choices or valuations. Whatever the sources of those errors, the null hypothesis presumes that the two types of PR are equally likely to occur as a consequence of error. This assumption is simple and has some appeal given the apparent symmetry of the two PR categories. However, in the absence of a generally agreed theory of decision error, it is to some extent ad hoc. A recent literature has emerged around modelling decision errors and/or preference imprecision (see Loomes, 2005; Hey 2005). While the majority of the most widely discussed theories would be consistent with the null hypothesis tested here, there is one exception that we are aware of discussed in Butler and Loomes (2007).

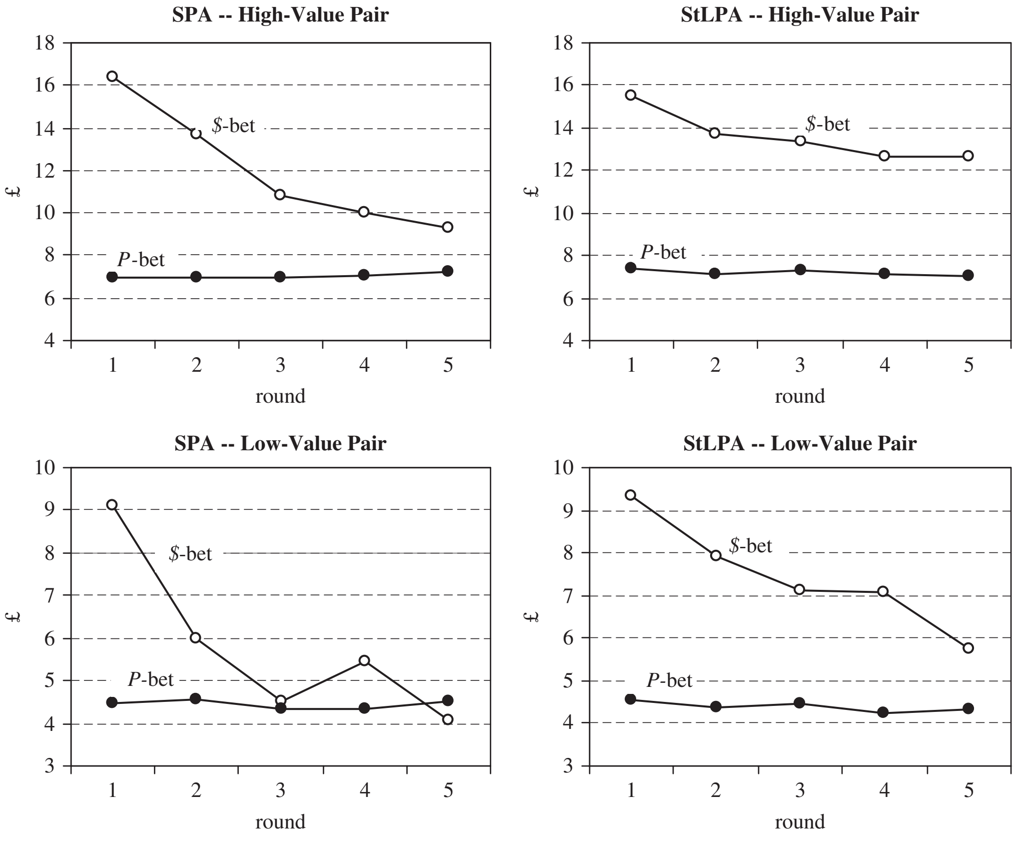

Fig. 1. Mean bids in Experiment A.

For reasons set out already, we expect the effects of price following to be relatively more pronounced in $-bids, thereby increasing the incidence of standard reversals and making the typical asymmetry *stronger*. Fig. 1 shows the evolution of the average bid for each auction type and each bet. In all cases, the initial mean $ valuation is considerably higher than the corresponding mean for *P*.[18] Across auction rounds, there is relatively little variation in the mean *P*-bids. By contrast, clear *downward* trends are apparent in the mean $-bids for *both* auction types. The decline in $-bids for the StLPA auction is inconsistent with price following being the primary cause of bid dynamics. This leads to our second main finding: We reject the price following hypothesis as the primary cause of bid evolution.

It is apparent from Fig. 1 that $ valuations decline relatively slowly in the StLPA and, for each pair of bets, the final round difference between the mean valuations of the *P*- and $-bets is considerably larger in the StLPA. This difference naturally influences the incidence of PR. Whereas the rapid decrease of the $-bids in SPA has eliminated systematic PR by the fifth round, in the StLPA the asymmetry decreases but remains significant ($p<0.000$). The differences between the markets are statistically significant as we demonstrate by testing the hypothesis that two observed response distributions (relative frequencies of consistent *P*, consistent $, standard reversal and non-standard reversal) are random samples from the same probability distribution. Based on a chi-squared

[18] There is some tendency for *P* choosers to make lower $-bids than $ choosers. Even so, in round 1 of both auctions, *P* choosers bid an average of £8.8 for (£18, 19%) and £16 for (£32, 31%), far more than they bid for the paired *P*-bets.

goodness-of-fit test, the change between rounds 1 and 5 is not significant in the StLPA ($p = 0.15$) but is in the SPA ($p < 0.0001$). Accordingly, the difference between the two auctions is not significant in the first round ($p = 0.95$) but is in the fifth ($p = 0.0009$).[19] These results lead to our third main finding: Different, but strategically equivalent, auction types differ in their propensity to reduce the preference reversal anomaly.

2.3. Results: Subjects' reactions to previous outcomes

Given the downward trajectory of average \$-bids in the StLPA, price following cannot be the predominant mechanism involved in bid adjustment. But it is possible that price following acts alongside other factors, at least one of which exerts an opposing influence strong enough to outweigh it. One candidate factor is loss experience, which would create a downward pull in both markets. Our data allow us to test for the separate impacts of loss experience and price following and to investigate the form these effects take, if they operate. For instance, we can test if subjects respond differentially to directly experienced losses versus losses they observe others to experience on bets that they *might* have faced had they not sold them. We may also examine whether subjects respond differently to the experiences of winning and losing. To this end, we compute the mean bid change across all subjects and auction rounds ($\Delta\text{bid}_t = \text{bid}_t - \text{bid}_{t-1}$) associated with four events: Playing a lottery and losing (*loss*), playing and winning (*win*), selling and observing a losing number being drawn (*obLoss*) or selling and observing a winning number (*obWin*). This is equivalent to regressing the bid change on a set of dummies:

$$\Delta\text{bid}_t = b_{\text{loss}} \times \text{loss}_{t-1} + b_{\text{win}} \times \text{win}_{t-1} + b_{\text{obLoss}} \times \mathit{obLoss}_{t-1} + b_{\text{obWin}} \times \mathit{obWin}_{t-1}.$$

Since each subject participated in five auctions for each bet, for each subject/bet we have four observations on Δbid_t, ($t = 2, 3, 4, 5$). Because bids made by the same subject across multiple auction rounds or by different subjects trading in the same group are not independent observations, we employ Huber–White sandwich estimates of variance to take account of within-cluster correlation.

On the price following hypothesis, we expect those who have sold lotteries to increase their bids in the next round (i.e., $b_{\text{obLoss}} > 0$ and $b_{\text{obWin}} > 0$) and those who have kept lotteries to reduce them (i.e., $b_{\text{Loss}} < 0$ and $b_{\text{Win}} < 0$). A *direct* loss experience effect occurs if $b_{\text{loss}} < 0$ and $b_{\text{loss}} < b_{\text{win}}$. The second part of this condition implies that we interpret a negative response to losses as evidence of direct loss experience only if bids fall *more* following a loss than they do following a win. Analogously, there is an *indirect loss experience effect* if merely observing a losing lottery is enough to drive bids downwards. The hypothesis that there is indirect loss experience would lead us to expect $b_{\text{obLoss}} < 0$ and $b_{\text{obLoss}} < b_{\text{obWin}}$. Our analysis is summarised in Table 3 where we report separate regressions modelling changes in bids for *P* and *\$*. For each bet, we report three regressions: One based on data from the SPA condition only; one based on data from the StLPA condition only; and one based on pooled data from the two conditions.

The patterning of signs on the coefficients is suggestive of a price following effect: Coefficients on the dummies for *loss* and *win* are consistently negative indicating that bids tend to fall after subjects have played the lottery (i.e., they bid above the price) regardless of whether they won or lost; coefficients on *obLoss* and *obWin* are consistently positive indicating that the sellers in a given round tend to subsequently increase their bid regardless of whether they see others win or lose on that bet. However, while the majority of these coefficients are significant, not all are, and the coefficient of *obWin* is never significant at the 5% level. The support for price following is, therefore, only modest.

There is much clearer support for *direct* loss experience. The coefficient b_{loss} is always negative and significant (highly so for \$-bets). The secondary condition ($b_{\text{loss}} < b_{\text{win}}$) also holds in all regressions indicating that subjects who failed to sell and played a lottery tended to reduce bids by more after a loss than after a win. This difference is statistically significant and large in the case of

[19] Cases of equal bids were not frequent enough to form a category for this test and were split equally between the *P* and *\$* columns of Table 2. Splitting them differently does not change the results.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

Table 3
Subjects' mean reaction to previous outcome—Experiment A

Sample	*P*-bets pooled		P-bets in SPA		*P*-bets in StLPA		$-bets pooled		$-bets in SPA		$-bets in StLPA	
Observations[a]	720		360		360		720		360		360	
Clusters	24		12		12		24		12		12	
p (all $b=0$)	0.001		0.005		0.000		0.000		0.002		0.000	
R^2	0.022		0.128		0.022		0.127		0.114		0.277	
Δbid_t	b	p-Value	b	p-Value	b	p-Value	b	p-Value	b	p-Value	b	p-Value
$loss_{t-1}$	−0.26	0.003	−0.22	0.024	−0.46	0.021	−3.07	0.000	−2.44	0.000	−5.36	0.000
win_{t-1}	−0.12	0.017	−0.06	0.037	−0.28	0.087	−1.63	0.019	−0.80	0.324	−3.40	0.007
$obLoss_{t-1}$	0.36	0.025	0.84	0.053	0.22	0.129	1.00	0.001	1.61	0.046	0.88	0.011
$obWin_{t-1}$	0.11	0.188	0.86	0.054	0.01	0.819	1.03	0.203	8.57	0.052	0.10	0.798
Test for secondary condition of loss experience hypothesis[b] (p-values)												
H_0: $b_{loss}=b_{win}$, H_a: $b_{loss}<b_{win}$	0.065		0.055		0.226		0.036		0.045		0.077	

[a] There were 180 subjects.
[b] We do not report a corresponding test for the indirect effect since the primary condition is not satisfied.

$-bets: Subjects reduced their $-bids by £3.07 after they lost but only by £1.63 after they won.[20] There is no support for indirect loss experience since b_{obLoss} is consistently positive (hence, we do not test the secondary condition).

3. Experiment B

The results of Experiment A encouraged us to further explore the role of loss experience. In Experiment B, we test the loss experience hypothesis directly by comparing two treatments: In one, we expect loss experience to operate, while the other is designed to 'turn off' any loss experience effect. A second main objective of Experiment B was to test whether the loss experience effect, if it operates, is a *refining mechanism*: That is, does it tend to suppress anomalous behaviour relative to standard preference theory? We again use PR to explore this issue. While the data from our (and Cox and Grether's) SPA market would be consistent with the claim that asymmetric PR is eliminated in the context of a repeated SPA that conclusion warrants further investigation. Re-inspection of Fig. 1 suggests that, in Experiment A, $-bids may not have completely stabilised by the final auction round. This raises an interesting question: What would happen to $-bids if there were further rounds?. If $-bids continued to fall with longer experience of the auction market, that could lead to a new type of systematic anomaly whereby non-standard reversals outnumber standard reversals. If loss experience produced that outcome, we should not conclude that it is a refining mechanism.[21] To shed light on the longer run impact of loss experience, in Experiment B, we run new auctions with double the number of rounds.

A third main objective was to explore the effects of experience on choice behaviour. In the market based studies which show PR to decay, the evidence is of rankings revealed in repeated market valuations becoming more consistent with a one-off choice, and in all of these studies, the choice ranking was administered at the end of the experiment following multiple market periods. This suggests another interesting question: Is the choice ranking also influenced by experience?. If the answer were 'yes', then it may be that the effects of experience work (at least partly) through subjects learning about aspects of risk (e.g., getting a better 'feel' for probability, or experiencing how bad it feels to lose). On the other hand, if choices were largely unaffected by experience but valuations are, that would suggest that subjects are learning how to make 'better' valuations in the context of a particular mechanism (e.g., learning to avoid overvaluation). To explore this question, Experiment B featured three choice tasks, with the first choice elicited prior to any rounds of market experience. Hence, a novel feature of Experiment B is that it produced data on both 'experienced' and 'inexperienced' *choice* behaviour.

3.1. Procedures

In Experiment B, each subject faced 23 tasks. The first task was a straight choice between a *P*-bet and a $-bet; followed by 10 auction rounds in which they repeatedly valued one of these two bets; then a second choice between the same *P*/$ pair; then another 10 auction rounds for the other bet of the pair; finally, there was a third choice between the *P*/$ pair. For all subjects, the *P*-bet gave a 97% chance of winning £8 (otherwise nothing) while the $-bet gave a 31% chance of winning a prize of just over £30—(£32, 31%; £0, 69%) in some trading groups, (£32, 16%; £31, 15%, £0, 69%) in others.[22]

Apart from where specific experimental objectives dictated otherwise, we retained design features of Experiment A (e.g., recruitment procedures, the random lottery incentive scheme and so on). We focus here on the distinctive aspects of Experiment B. One such feature is that valuations were elicited via a *random-price Vickrey auction*. In each round, after all subjects had submitted their

[20] The relevant test is reported in the last row of Table 3.

[21] Relative to our earlier characterisation of refining theories in terms of four underlying assumptions, such evidence would require relaxation of one or both of assumptions (i) and (ii).

[22] This variation was introduced to test a hypotheses unrelated to the present discussion. The manipulation actually made very little difference and we do not discuss it further in this report.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

bids, the number of units that the experimenters would buy (n) was determined by a (public) random draw. The n lowest bidders then sold their lotteries at a uniform price equal to the (n+1)th lowest bid. The value of n was determined, round-by-round, by drawing from a deck comprised of cards running from the ace (one) up to the number of bidders minus two. Hence, the price varied between the second, and second-to-last-lowest bid. This auction mechanism was chosen with a view to neutralising price following: To the extent that participants had a tendency to price-follow, in a random price environment, that effect should have no *systematic* impact on bids.

To facilitate a test of the loss experience hypothesis, the design featured two treatments which differed according to whether or not subjects received round-by-round feedback on lottery outcomes. In the *Feedback* treatment, any subject who held a lottery at the end of a choice or an auction task would immediately play it. As in Experiment A, lotteries were resolved by one subject drawing a single number that determined the outcome of all lotteries being played in that task; and, as the draw was public, even subjects who had sold their lotteries received feedback on whether they would have won or lost, had they kept the lottery. In the *No-Feedback* treatment, no lotteries were played out until the end of the experiment. In this condition, a subject who held a lottery at the end of a particular task knew that they would only discover its outcome if it turned out to be the one selected to be real, for them, under the random lottery incentives. If the decline of $-bids and erosion of PR observed in SPA (Experiment A) was driven by loss experience, we should expect $-bids to fall and standard PR to decline across auction rounds of the Feedback treatment. But, in the No-Feedback treatment, there can be no round-by-round loss experience effect on $-bids, hence, standard PR would be expected to persist despite repeated market experience.

3.2. *Results for Experiment B*

One hundred and eighty-two subjects participated in the experiment (98 males) in 24 trading groups (of seven or eight subjects each), 12 for each of the feedback conditions. This resulted in respective sample sizes of 94 and 88 for the Feedback and No-Feedback treatments. Sessions lasted about an hour and average winnings were £8.04. Table 4 reports the frequencies of response patterns.[23] The left hand side of the table reports results for the Feedback treatment; the right hand side reports corresponding results for the No-Feedback treatment. The first row of the table reports responses, based on rankings from the first choice task and first-time valuations for each of the bets.[24] The final row of the table reports the analogous results based on the final choice and tenth valuation for each bet. The other three rows report intermediate cases.

We first examine the incidence of PR comparing ‘inexperienced’ decisions (i.e., first time choice and valuations); ‘moderately’ experienced decisions (using second choice and 5th round valuations) and final decisions (third choice and 10th valuations). Inexperienced decisions display strong PR of the typical form in both feedback conditions, with standard reversals occurring much more frequently than non-standard ones: 35 against 5 (Feedback); 25 against 3 (No-Feedback). Subsequently, behaviour under the two conditions diverges sharply. In the No-Feedback treatment, the typical PR asymmetry persists: Based on final decisions, there are 22 standard and only 4 non-standard reversals. In the Feedback treatment, moderately experienced decisions (second group of rows of Table 4) show a roughly symmetric reversal pattern. This resembles the pattern that obtained with five repetitions of the SPA in Experiment A. However, it is apparent that behaviour in the Feedback treatment had not stabilised by round 5 and, when final choices are compared with bids from auctions towards the end of each series, non-standard reversals outnumber the standard type. This asymmetry becomes significant towards the end of the auction series ($p = 0.026$ and $p = 0.068$ in rounds 9 and 10, respectively). This leads to our first main finding based on Experiment B: As predicted by the loss experience hypothesis, lottery feedback reduces standard reversals and increases non-standard ones; this process appears not to terminate in there being equal numbers of the two types of reversal; with sufficient lottery feedback, a new form of systematic PR becomes apparent in which non-standard reversals predominate.

[23] We aggregate results across subjects who faced the two slightly different version of the $-bet.

[24] Note that the first bid for a type of bet was either the second or thirteenth task for a subject.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

Table 4
Frequency of responses in Experiment B and mean difference between *P*- and $-bids

	Choice	Feedback (highest bid)			No-Feedback (highest bid)		
		P	$	Equal	*P*	$	Equal
Round 1 bids and first choice	*P*	19 (2.8)	**35** (8.5)	6	19 (2.8)	**25** (6.7)	4
	$	**5** (2.9)	29 (7.8)	0	**3** (1.1)	36 (7.3)	1
				$p = 0.000$			$p = 0.000$
Round 5 bids and mid choice	*P*	42 (3.1)	**16** (4.1)	1	25 (3.0)	**21** (2.4)	7
	$	**13** (2.1)	20 (4.3)	2	**8** (2.8)	26 (4.2)	1
				$p = 0.356$			$p = 0.012$
Round 8 bids and last choice	*P*	46 (3.1)	**10** (3.8)	6	32 (2.5)	**18** (2.4)	6
	$	**17** (3.3)	13 (2.3)	2	**9** (3.0)	22 (3.6)	1
				$p = 0.939$			$p = 0.061$
				($p = 0.124$)			
Round 9 bids and last choice	*P*	49 (3.3)	**8** (3.3)	5	31 (2.6)	**20** (3.5)	5
	$	**19** (3.1)	12 (2.9)	1	**7** (3.9)	24 (4.7)	1
				$p = 0.990$			$p = 0.010$
				($p = 0.026$)			
Round 10 bids and last choice	*P*	48 (3.6)	**10** (3.2)	4	28 (2.6)	**22** (2.0)	6
	$	**19** (3.4)	8 (8.2)	5	**4** (4.0)	27 (5.1)	1
				$p = 0.969$			$p = 0.000$
				($p = 0.068$)			

Note: The table reports frequencies of responses with reversals in bold. Values within parentheses are mean differences between *P*- and $- bids. *p*-Values result from a test (binomial distribution) of the null hypothesis that standard and non-standard reversals are equally likely against the alternative hypothesis that standard reversals are more likely. *p*-Values in parentheses test the same null against the alternative that non-standard reversals are more likely.

Because Experiment B observes both choices and valuations with different levels of experience, we can examine the extent to which the evolution of PR, as summarised in Table 4, is driven by changes operating through choice and/or valuation behaviour. While there is considerable switching of individual responses across repeated choice tasks (41 out of 94 subjects in the Feedback treatment and 29 out of 88 subjects in the No-Feedback treatment changed their choice at least once), this variation appears to be random.[25] Individual switching therefore cancels out in the aggregate: There are no significant differences in the split between *P* and $ choices when comparing first, middle and last choices within a treatment; and also no significant differences when comparing first, middle or last choices across treatments.[26] Even so, it would still be possible for variation in choice behaviour to be impacting on the pattern of PR. For instance, if those who have committed reversals are more likely to switch choices than those who have not, that could result in a change in reversal asymmetry without any change in the aggregate split of preference between *P* and $.

With this in mind, and in order to examine whether changes in the incidence of PR might be partly due to variation in choice behaviour, we undertook parallel analysis to that reported in Table 4 separately for each of the three choices. This analysis (which we do not report, but is available from the authors upon request) reveals that the evolution of responses in Table 4 is not attributable to changes in choices. In the No-Feedback treatment, reversal patterns would have become marginally less asymmetric if subjects had stuck to their first or second choices, but significant asymmetry would have remained. In the Feedback treatment, the response pattern obtained with any auction round is

[25] The presence of a significant stochastic component in choice, commensurate with that observed here, is a typical finding of experimental research involving repeated choices among lotteries (Hey, 2005; Loomes, 2005).

[26] It is possible, in principle, that increasing the number of choice repetitions could lead to systematic trends in behaviour. Our experience with valuation tasks, however, suggests that the impact of experience is typically strongest in early repetitions and that trends in valuation behaviour are clearly visible over three rounds.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005

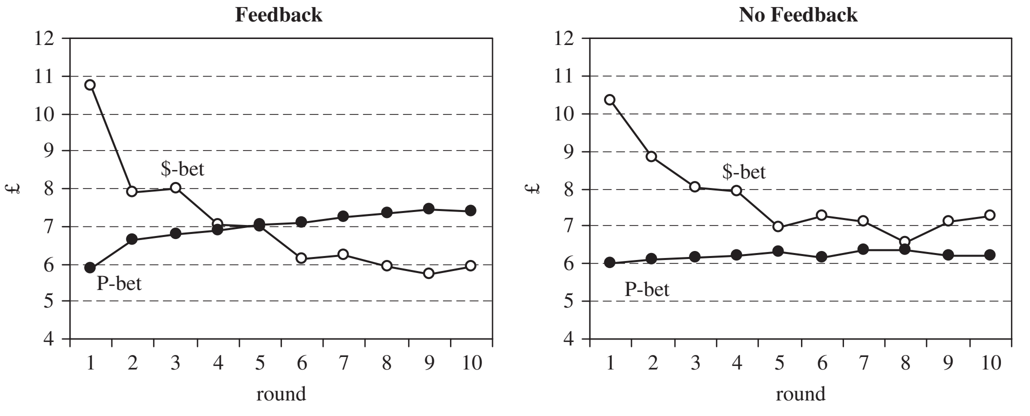

Fig. 2. Mean bids in Experiment B.

qualitatively the same (and insignificantly different) regardless of whether the first, second or last choice is used. It may be particularly worth noting that the non-standard asymmetry observed in final decisions for the Feedback treatment appears just as clearly if final valuations are compared with inexperienced choices. Using the first choice with final valuations, we observe 17 non-standard and 7 standard reversals as compared with 19 standard and 10 non-standard reversals using the final choice with final valuations (i.e., last row Table 4). This is reassuring because it confirms that one of our most striking findings is not attributable to variation in choice behaviour. So while there is clearly variation with respect to which particular individuals appear in particular response categories, changes in the aggregate pattern of reversals are not attributable to choice behaviour. This leads to our second main finding in Experiment B: While there is variation in choice behaviour at the individual level in both feedback conditions, it appears to have little or no effect on the aggregate pattern of reversals.

Given that experience appears to have little or no systematic impact on choice behaviour, in the remainder of this section, we focus on the analysis of bidding. Fig. 2 plots mean bids for *P*- and *$*-bets, across auction rounds, with separate panels for the two feedback treatments. In the No-Feedback treatment, mean bids for *P* are essentially stable, and while *$* values fall rapidly in early rounds, they level out at a mean value consistently above the average *P*-bid. Two qualitative differences are apparent in the corresponding plots for the feedback condition: *$*-bids fall further and *P*-bids show a tendency to rise. The net effect of these two changes is that, by round 6, mean *P*-bids are above those for *$*. This crossover is consistent with the falling incidence of standard PR and rising incidence of non-standard PR. It is intriguing that feedback should systematically reverse the mean ranking of *P*- and *$*-bets revealed in market valuations while having no systematic influence in choices. This is consistent with the view that choice and valuation tasks invoke different mental processes. It is also striking that lottery feedback is, apparently, causing bids for both *$*- and *P*-bets to converge to stable but different levels when compared with the markets without lottery feedback: For a given bet, the difference between the mean bids of the two treatments is not significant in round 1 but tends towards significance in later rounds (for *$*-bids, $p = 0.051$ in the ninth round, $p = 0.068$ in the tenth round; for *P*-bids, $p = 0.03$ in the ninth round, $p = 0.037$ in the tenth round; in both cases, we used the one-tail test). This leads to our third main finding: Average bids for lotteries converge to different levels in repeated auctions depending on whether or not there is feedback on lottery outcomes.

While the between treatment difference in mean *$*-bids is an implication of the loss experience hypothesis, the upward trend in *P*-bet valuations in the feedback condition is not. Loss experience, although rare for *P*-bets, would lead to a decline in *P*-bids. This leads to our fourth main finding for Experiment B: Feedback on lottery outcomes appears to promote reductions in *$*-bids and increases in *P*-bids. While the former is predicted by the loss experience hypothesis, the latter is not.

Although the effects of outcome feedback on the dynamics of bids are clear in our data, we cannot be completely sure that these effects are monotonic. The observed terminal values of the mean

P- and \$-bids could, for example, reflect over-reactions to feedback and, on that assumption, further repetitions might moderate overreactions generating increases in \$-bids, decreases in *P*-bids and a reduction in the final asymmetry observed between the PR categories. This is a logical possibility which, in principle, could be explored via further research although, on the basis of our data, we do not think it is very likely. While Fig. 1 provides reason to believe that \$-bids had not stabilised after the five auction rounds of Experiment A (which is why we doubled this number in Experiment B), the last five rounds of the feedback condition for Experiment B (illustrated in the left hand panel of Fig. 2) show little evidence of any remaining trend.

3.3. Results: Subjects' mean reaction to previous outcome

To further explore what factors may be influencing bids, we analysed subjects' reactions to various auction/lottery outcomes, using the full dataset for Experiment B. As in Experiment A, we estimate subjects' mean reactions to various events. For the Feedback treatment, the events are, as before, keeping and losing (*loss*); keeping and winning (*win*); selling and observing a win (*obWin*); selling and observing a loss (*obLoss*). In the No-Feedback treatment, there are two events: Ending the auction by selling the bet (*sell*) or by keeping the bet (*keep*). Hence, our procedure is equivalent to estimating the model:

$$\Delta\text{bid}_t = b_{\text{loss}} + \text{loss}_{t-1} + b_{\text{win}} \times win_{t-1} + b_{\text{obLoss}} \times obLoss_{t-1} + b_{\text{obWin}} \times obWin_{t-1} + b_{\text{keep}} \times keep_{t-1} + b_{\text{Sell}} \times sell_{t-1}.$$

We estimate separate equations for changes in *P*- and \$-bids based on 1638 observations each (there were 182 subjects each generating nine observations on Δbid_t, $t = 2, \ldots, 10$, per bet). As before, we apply the Huber–White sandwich estimate of variance to allow for within-cluster correlation. Results are reported in Table 5. The results for the No-Feedback treatment reveal a simple pattern which applies to both *P*- and *\$-bets*: On average, those who had sold increased their bids next round, while those who had not sold reduced them. These effects are all significant and it is noteworthy that this pattern should emerge in the random price environment.

Cross-group comparisons allow us to assess the effects of feedback on bids. This differs for *P*- and \$-bets and we first examine the \$-bets. Table 5 shows that those who kept \$-bets and won behaved similarly to those who kept but received no feedback: There are mean reductions of, respectively, 1.10 and 1.07 and the between treatment test reported toward the bottom of the page confirms the eye-ball impression of no significant difference between b_{win} and b_{keep} ($p = 0.937$). By comparison, those who lost on \$-bets reduced their subsequent bids by an average of 1.72. This coefficient is significant ($p = 0.000$) and significantly more negative than the response for either winners or keepers. The relevant statistics are reported under the heading 'secondary tests for direct loss experience' and the effects are weakly significant within treatment ($p = 0.081$) and strongly so for the cross-treatment comparison ($p = 0.014$). We interpret these results as demonstrating a clear *direct* loss experience effect in \$-bets. Those who sold \$-bets and learned that they would have lost behave similarly to those who receive no feedback at all (i.e., $b_{\text{sell}} = b_{\text{obLoss}}$ is not rejected, $p = 0.774$). Hence, as in Experiment A, there is no support for the existence of an *indirect* loss experience effect (since b_{obLoss} is positive and insignificant, we do not report tests of secondary conditions). In contrast, *observing that you would have won* leads to a significantly bigger increase in \$-bids compared with those who sell but receive no lottery feedback (b_{obWin} is significantly larger than b_{sell}, $p = 0.059$).

We summarise the influence of feedback on \$-bidding in the following way: Feedback matters, but only when it is '*bad news*'. Subjects who receive 'positive' feedback (i.e., by keeping and winning, or by selling and seeing that they would have lost) behave in qualitatively similar ways to subjects who receive no feedback. Negative feedback, on the other hand, reinforces tendencies that already operate in the absence of feedback: Those who keep and lose subsequently reduce their bids acting as if they are more willing to sell next round; those who see that they would have won had they not sold tend to increase their subsequent bids, acting as if they are now less willing to sell.

We now examine the impact of feedback on *P*-bids. Among sellers of *P*-bets, feedback makes no significant difference to average bid adjustment (i.e., we cannot reject the hypothesis that

Table 5
Subjects' mean reaction to previous outcome—Experiment B

P						*$*					
Observations			1638			Observations					1638
Clusters			24			Clusters					24
F (all $b = 0$)			0.000			F (all $b = 0$)					0.000
R^2			0.069			R^2					0.076
Feedback			*No-feedback*			*Feedback*			*No-feedback*		
	b	p-Value		b	p-Value		b	p-Value		b	p-Value
loss_{t-1}	−0.20	0.091	keep_{t-1}	−0.21	0.009	loss_{t-1}	−1.72	0.000	keep_{t-1}	−1.07	0.000
win_{t-1}	−0.02	0.605				win_{t-1}	−1.10	0.013			
obLoss_{t-1}	0.00	1.000	sell_{t-1}	0.38	0.000	obLoss_{t-1}	0.50	0.138	sell_{t-1}	0.61	0.008
obWin_{t-1}	0.44	0.000				obWin_{t-1}	1.38	0.002			

Secondary test for direct loss experience (p-values, one-tailed)[a]

$b_{\text{loss}} = b_{\text{win}}$ vs. $b_{\text{loss}} < b_{\text{win}}$	0.072	$b_{\text{loss}} = b_{\text{win}}$ vs. $b_{\text{loss}} < b_{\text{win}}$	0.081
$b_{\text{loss}} = b_{\text{keep}}$ vs. $b_{\text{loss}} < b_{\text{keep}}$	0.523	$b_{\text{loss}} = b_{\text{keep}}$ vs. $b_{\text{loss}} < b_{\text{keep}}$	0.014

Other between treatment tests (p-values, two-tailed)

Keepers of *P*		Sellers of *P*		Keepers of *$*		Sellers of *$*	
$b_{\text{win}} = b_{\text{keep}}$	0.023	$b_{\text{obLoss}} = b_{\text{sell}}$	0.768	$b_{\text{win}} = b_{\text{keep}}$	0.937	$b_{\text{obLoss}} = b_{\text{sell}}$	0.774
		$b_{\text{obWin}} = b_{\text{sell}}$	0.684			$b_{\text{obWin}} = b_{\text{sell}}$	0.059

[a] Since b_{obLoss} is consistently positive, we do not report secondary tests for indirect loss experience.

$b_{\text{obloss}} = b_{\text{sell}}$ ($p = 0.768$) or that $b_{\text{obwin}} = b_{\text{sell}}$ ($p = 0.684$)). For subjects who keep and play bets, however, feedback does affect mean adjustment. In contrast to $-bids, bad news does not significantly affect *P*-bids: Subjects who play and lose significantly reduce their bids ($b_{\text{loss}} = -0.20$, $p = 0.091$), but by only as much as those who keep their lotteries without playing them ($b_{\text{keep}} = -0.21$, $p = 0.009$) and we cannot reject the hypothesis that $b_{\text{loss}} = b_{\text{keep}}$ ($p = 0.523$). Hence, we find no evidence that direct loss experience reinforces the tendency of keepers to reduce their bids. By contrast, bids for *P*-bets are affected by *good news*: The evidence for this is that subjects who win on *P*-bets tend not to change their bids. Thus, in the case of *P*-bets, the experience of winning appears to counteract the tendency of keepers to reduce their bids. (The difference is statistically significant i.e., we reject $b_{\text{win}} = b_{\text{keep}}$, $p = 0.023$).

One interpretation of why the effects of feedback differ between *P*- and $-bets is 'probability distortion'. There is evidence that individuals choose among risks as if, relative to expected utility theory, they 'overweight' small probabilities and 'underweight' larger ones (e.g., see Prelec, 1998). A model with that property would lead to 'overvaluation' of $-bets and 'undervaluation' of *P*-bets. Some theorists interpret these tendencies as biases which sophisticated decision makers would, or should, learn to avoid (e.g., Bleichrodt et al., 2001). Consistent with this, one might conjecture that continual winning on *P*-bets tends to erode undervaluation; while continual losing on $-bets tends to erode overvaluation. This conjecture may be worthy of future investigation.

4. Discussion and conclusion

There is voluminous evidence that choice behaviour deviates in predictable ways from basic assumptions of standard preference theory. Set against this, accumulating evidence indicates that some anomalies in stated preference decay when preferences are elicited in repeated markets. Close reading of this evidence, however, suggests that it is not market experience per se that has these effects, but some aspects of experience which are associated with some market types and not others. The new studies reported here explore two research questions prompted by this evidence. The first is: Can we identify specific mechanisms that promote changes in stated preference in some markets?. If we can, a second question arises: Can we determine whether those mechanisms are 'refining'. That is, do any of these mechanisms have some general tendency to promote greater consistency between behaviour and standard preference theory? We suggest that the pursuit of these questions is an important endeavour for those seeking to understand the significance of preference anomalies for behaviour in real markets and the welfare evaluation of their outcomes.

In this paper, we have focussed on understanding the dynamics of risk preference in repeated markets. We proposed two behavioural mechanisms which might organise features of the prior evidence: The price following hypothesis and the loss experience hypothesis. We then presented the results of two experiments intended to test them. While our results provide only modest support for the price following hypothesis, we find strong evidence consistent with the loss experience hypothesis. Our experiments produce particularly clear evidence of a *direct* loss experience effect. This is revealed as a tendency for subjects to reduce stated valuations for a lottery following the experience of holding it and losing. Since we do not find similarly strong (*indirect*) effects when subjects simply observe others losing on bets, this direct loss experience effect cannot be straightforwardly understood as learning about probabilities. Our results also suggest that loss experience is not the only mechanism at work and an interesting but tentative finding is that experiencing positive lottery outcomes can promote higher lottery valuations.

We believe that our experiments are the first to identify the effects of loss experience in experimental markets and we conjecture that its operation may have important implications for the understanding of naturally occurring markets. For example, the presence of a corresponding behavioural effect in financial markets might contribute to an explanation of phenomena such as stock market crashes and, in particular, the initiation of financial crises such as those in the mid 1990s, which were characterised by the selling of stocks following bad market outcomes, despite market fundamentals remaining strong. Loss experience might also have implications for insurance markets. For example, if individuals become more risk averse as they experience losses, this raises a question about whether there may be some tendency for under-insurance against low probability losses.

Perhaps our most surprising findings relate to the refining properties of market experience. Experiment B examined the extent to which Vickrey auctions erode perhaps the most infamous challenge to choice theory, the preference reversal phenomenon. A main finding is that in the absence of lottery feedback, PR exists and persists in the market; and while feedback has a definite effect, with enough feedback, the mirror image anomaly emerges in which non-standard reversals predominate. So our results do not show that repeated market experience eliminates PR: We find that it eliminates standard PR but only to replace it with a non-standard variety. Of course our findings are based on behaviour in a particular environment involving 10 repetitions of a random-price Vickrey auction and it may be worthwhile for future research to examine the robustness and generalisability of our results. Natural extensions to our work might investigate other variants of the Vickrey auction and larger amounts of repeated market experience. Nevertheless, our findings add to the evidence against the general presumption that market experience is refining and provide further motivation for deepening our understanding of the mechanisms that promote changes in stated preference.

Acknowledgements

We are grateful to an associate editor and two anonymous referees for helpful comments on an earlier version of this paper. We have also benefited from many comments and suggestions from numerous participants at conferences and seminars where we have presented this work.

References

Bateman, I., Munro, A., Rhodes, B., Starmer, C., Sugden, R., 1997. A test of the theory of reference dependent preferences. Quarterly Journal of Economics 112, 479–505.
Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behavioral Science 9, 226–232.
Bleichrodt, H., Pinto, J.L., Wakker, P.P., 2001. Making descriptive use of prospect theory to improve the prescriptive use of expected utility. Management Science 47, 1498–1514.
Butler, D.J., Loomes, G., 2007. Imprecision as an account of the preference reversal phenomenon. American Economic Review 97, 277–297.
Camerer, C., 1995. Individual decision making. In: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 587–703.
Chu, Y.-P., Chu, R.-L., 1990. The subsidence of preference reversals in simplified and marketlike experimental settings: A note. American Economic Review 80, 902–911.
Coursey, D., Hovis, J., Schultz, W., 1987. The disparity between willingness to accept and willingness to pay measures of value. Quarterly Journal of Economics 102, 679–690.
Cox, J.C., Grether, D.M., 1996. The preference reversal phenomenon: Response mode, markets and incentives. Economic Theory 7, 381–405.
Cubitt, R., Starmer, C., Sudgen, R., 1998. On the validity of the random lottery incentive system. Experimental Economics 1, 115–131.
Cubitt, R.C., Munro, A., Starmer, C., 2004. Testing explanations of preference reversal. Economic Journal 114, 709–726.
Grether, D.M., Plott, C.R., 1979. Economic theory of choice and the preference reversal phenomenon. American Economic Review 69, 623–638.
Hey, J.D., 2005. Why we should not be silent about noise. Experimental Economics 8, 325–345.
Hey, J.D., Lee, J., 2005a. Do subjects remember the past? Applied Economics 37, 9–18.
Hey, J.D., Lee, J., 2005b. Do subjects separate (or are they sophisticated)? Experimental Economics 8, 233–265.
Holt, C., 1986. Preference reversals and the independence axiom. American Economic Review 76, 508–515.
Isen, A., 1999. Positive affect. In: Dalgleish, T., Power, M. (Eds.), Handbook of Cognition and Emotion. Wiley, New York, pp. 521–539.
Knetsch, J., Tang, F.-F., Thaler, R., 2001. The endowment effect and repeated market trials: Is the Vickrey auction demand revealing? Experimental Economics 4, 257–269.
Lichtenstein, S., Slovic, P., 1971. Reversals of preference between bids and choices in gambling decisions. Journal of Experimental Psychology 89, 46–55.
Lichtenstein, S., Slovic, P., 1973. Response-induced reversals of preference in gambling: An extended replication in Las Vegas. Journal of Experimental Psychology 101, 16–20.
List, J., 2003. Does market experience eliminate market anomalies? Quarterly Journal of Economics 118, 41–71.
List, J., Shogren, J., 1999. Price information and bidding behavior in repeated second-price auctions. American Journal of Agricultural Economics 81, 942–949.
Loomes, G., 2005. Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. Experimental Economics 8, 301–323.
Loomes, G., Starmer, C., Sugden, R., 2003. Do anomalies disappear in repeated markets? The Economic Journal 113, 153–166.
Plott, C., 1996. Rational individual behavior in markets and social choice processes: The discovered preference hypothesis. In: Arrow, K., Colombatto, E., Perlman, M., Schmidt, C. (Eds.), The Rational Foundations of Economic Behavior. St Martin's Press, London: Macmillan and New York, pp. 220–250.
Plott, C.R., Zeiler, K., 2005. The willingness to pay–willingness to accept gap, the 'endowment effect', subject misconceptions, and experimental procedures for eliciting valuations. The American Economic Review 95, 530–545.
Prelec, D., 1998. The probability weighting function. Econometrica 66, 497–527.
Seidl, C., 2002. Preference reversal. Journal of Economic Surveys 6, 621–655.
Shogren, J., Shin, S., Hayes, D., Kliebenstein, J., 1994. Resolving differences in willingness to pay and willingness to accept. American Economic Review 84, 255–270.
Shogren, J., Cho, S., Koo, C., List, J., Park, C., Polo, P., Wilhelmi, R., 2001. Auction mechanisms and the measurement of WTP and WTA. Resource and Energy Economics 23, 97–109.
Simonson, I., Drolet, A., 2004. Anchoring effects on consumer's willingness-to-pay and willingness-to-accept. Journal of Consumer Research 31, 681–690.
Starmer, C., 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. Journal of Economic Literature 38, 332–382.
Starmer, C., Sugden, R., 1991. Does the random-lottery incentive system elicit true preferences? An experimental investigation. American Economic Review 81, 971–978.
Sugden, R., 2003. Reference-dependent subjective expected utility. Journal of Economic Theory 111, 172–191.
Tversky, A., Slovic, P., Kahneman, D., 1990. The causes of preference reversal. American Economic Review 80, 206–217.
Vickrey, W., 1961. Counterspeculation, auctions, and competitive sealed tenders. Journal of Finance 16, 8–37.

Please cite this article as: Braga, J., et al., Market experience eliminates some anomalies—and creates new ones. European Economic Review (2008), doi:10.1016/j.euroecorev.2008.06.005