Do Anomalies Disappear in Repeated Markets?

Graham Loomes[*], Chris Starmer[#] and Robert Sugden[*]

***University of East Anglia**

 # University of Nottingham

Introduction

There is a large literature demonstrating the existence of 'anomalies' in individual choice behaviour (see Camerer 1995, Starmer 2000). Taken at face value, such anomalies pose a major challenge to choice theorists, applied economists and policy analysts. But much of this evidence comes from experiments involving one-off decisions in non-market settings, and some economists have questioned its economic significance by raising doubts about whether these anomalies will arise or persist in real market environments. For example, Binmore (1994, 1999) argues that anomalous behaviour is economically significant only if it survives in an environment in which individuals repeatedly face the same decision problems, receive feedback on the outcomes of their decisions, and have adequate incentives.

In fact, there is some evidence that specific anomalies may decay in repeated experimental markets. Two studies showing this result are of particular interest in the context of this paper: Shogren et al (2001) report erosion of the gap between willingness to pay (WTP) and willingness to accept (WTA) when values of goods are elicited in repeated auctions; and Cox and Grether (1996) report decay of preference reversal when lottery values are elicited in repeated auctions. These studies have two interesting findings in common. First, each reports the *persistence* of the particular anomaly even in an environment with repeated decisions, feedback and incentives. Second, each reports erosion of the anomaly in an environment with an added ingredient, namely, the *presence of a market mechanism*. However, other experiments using apparently similar repeated market designs have found that anomalies persist: for example, preference reversal decays in only one of the two experimental markets investigated by Cox and Grether, and Knetsch et al (2001) find persistence of the WTP–WTA disparity.

These findings pose an important research question: what is the mechanism by which repeated markets influence behaviour? This study is motivated by that question. We consider three alternative hypotheses about how stated values evolve in repeated markets and we present the results of an experiment designed to discriminate between them.

1. Three Hypotheses

The refining hypothesis

The refining hypothesis is that market experience helps individuals to make decisions that more accurately reflect their preferences. If preferences satisfy standard consistency requirements and anomalies result from errors, the refining hypothesis predicts a tendency for anomalies to become less frequent as market experience accumulates. The refining hypothesis does not specify the mechanism, or mechanisms, which promote error reduction. It is simply an empirical conjecture.

As such, we take it to be in the spirit of Plott's (1996) 'discovered preference hypothesis'. Plott suggests that rationality is a 'process of discovery': when individuals face unfamiliar tasks, their behaviour can be influenced by various biases, but with incentives and practice, they arrive at 'considered choices' that reflect stable underlying preferences (p. 248). Much of the evidence cited by Plott concerns the remarkable tendency for various experimental market institutions to converge on the equilibrium predictions of standard economic models. Plott conjectures that this convergence is at least partly due to subjects' discovering how best to satisfy their preferences.

While the mechanisms of discovery are not fully specified by Plott, neither are they totally mysterious. Like Plott and Binmore, we find it plausible to suppose that repetition, feedback and incentives might each play a role. Repetition allows subjects to become more familiar with decision tasks and the objects of choice; feedback allows subjects to experience the consequences of particular choices; incentives provide a general motivation to attend carefully to tasks. If such factors tend to reduce error propensity, their operation could provide part of a more fleshed out version of the refining hypothesis.

Notice, however, that none of these factors is unique to a market context: it is possible to construct decision environments with repetition, feedback and incentives, but no market. On the other hand, the evidence motivating our study suggests that some anomalies persist in an environment with each of these three features present, yet decay with the addition of a market mechanism. As we have stated it, the refining hypothesis does not give any special significance to the market mechanism per se. The other two hypotheses do.

The market discipline hypothesis (MDH)

Like the refining hypothesis, MDH assumes that agents have stable underlying preferences, and that they may commit errors when attempting to act on those preferences within a market institution. However, MDH distinguishes between two types of error: those which, ex post, are costly to the agent once the market outcome is known, and those which are not. The hypothesis is that agents adjust their behaviour to correct errors if and only if those errors have proved costly.

We consider the implications of MDH for an auction in which each of n agents submits a bid to buy one unit each of a good; the supply of the good is $k < n$ units. (A corresponding analysis applies for selling). We assume that, for each agent j, the good has a *value* v_j , indexed so that $v_1 \geq v_2 \geq ... \geq v_n$. There is *market-clearing equilibrium* in the Walrasian sense if the price p is in the interval $v_k \geq p \geq v_{k+1}$. The market institution is a generalisation of a Vickrey auction: the k agents submitting the highest bids buy at a price equal to the (k + 1)th highest bid. (The commonly used second price auction institution is the special case where k = 1.) Notice that if all agents bid their values, the price is equal to v_{k+1} and there is market-clearing equilibrium. It is well known that, for every participant in such an auction, it is a weakly dominant strategy to bid their value. However, since MDH permits errors, it does not predict that bids are always equal to values.

Now suppose that the price is *not* market-clearing. First, suppose $p > v_k$. Then there must be at least one agent j who, as a result of bidding more than his value, buys at a price greater than that value. This is a costly error. Alternatively, suppose $p < v_{k+1}$. Then there must be at least one agent j who, as a result of bidding less than her value, misses an opportunity to buy at a price below that value. This too is a costly error. We take MDH to imply that, in a repeated auction, an agent who makes a costly error in one round reacts by reducing the discrepancy between bid and value in the following round (if there is one). A simple version of this hypothesis is the partial adjustment model $b_{j,r+1} = v_j + \theta (v_j - b_{j,r})$, with $0 \leq \theta < 1$, where $b_{j,r}$ is the bid of agent j in round r and where that bid was a costly error. In this model, price converges to market-clearing equilibrium.

However, MDH does *not* imply that *all* agents' bids converge to their values. There can be a market-clearing equilibrium in which no costly errors are made, but some non-marginal bids remain 'incorrect'. (Any profile of bids which involves no costly errors is a Nash equilibrium.) For example, there may be individuals who understate their values, but have no incentive to adjust because the market price is either above their value or below their bid. So, if MDH is true, even indefinitely repeated auction mechanisms cannot be assumed to reveal the underlying preferences of *all* bidders. What *is* revealed, given sufficient repetition, is the market-clearing price, which reflects the preferences of *marginal* traders.

The shaping hypothesis

The shaping hypothesis is that, in repeated market environments, there is a tendency for agents to adjust their bids towards the price observed in the previous market period. In the context of our auction institution, a simple version of this hypothesis would be the partial adjustment model $b_{j,r+1} = p_r + \theta (p_r - b_{j,r})$, with $0 \leq \theta < 1$, where p_r is the price in round r.

The implications of shaping for the value-revealing properties of markets will differ between environments. For example, an observed price can contain information about others' willingness to pay. This may be relevant to an agent's own value of a good if he buys it with an expectation of selling it later, or if he is uncertain about the utility he will derive from consuming it and believes this utility to be correlated with the utilities of other market participants. In such cases, a tendency for bids to adjust towards observed market prices would be consistent with the refining hypothesis: agents would be using market experience to help them to 'discover' their underlying preferences.

However, if agents' values are not correlated and if there is no expectation of resale, adjusting bids toward an observed market price is inconsistent with either the refining or the market discipline hypotheses. Nevertheless, it is possible that people use this adjustment rule as a heuristic even in cases in which it leads to divergences between bids and values. That they do so is the shaping hypothesis.

2. Explaining Existing Data

The elimination of anomalies reported by Shogren et al and by Cox and Grether is consistent with each of the three hypotheses just discussed.

Consider first Shogren et al. They report the typical disparity between WTP and WTA in the early rounds of a second price auction. But with repetition, WTP rises and WTA falls; after a few rounds, *average* WTP and WTA converge. This observation is consistent with both the refining hypothesis and MDH, if each is supplemented by two additional assumptions. The first is that underlying preferences have the conventional property that WTP ≈ WTA. The second is that inexperienced subjects are liable to exhibit *strategic bias:* that is, to bid low when buying and to ask high when selling. On the refining hypothesis, we should expect such a bias to be eliminated through market experience. If MDH is true, subjects who experience costly errors will adjust their bids towards their values. Such adjustments will tend to reduce the understatement of values by buyers (thus increasing average WTP) and to reduce the overstatement of values by sellers (decreasing average WTA).

But these data are also consistent with shaping. When subjects are buying in a second price auction, the market price is the second *highest* WTP; when they are selling, the market price is the second *lowest* WTA. So, when there are more than three bidders in these auctions (as was the case in Shogren et al's experiment), shaping will generate a downward trend in WTA and an upward trend in WTP.

The Cox and Grether data on preference reversal are also consistent with all three hypotheses. In the classic preference reversal experiment, subjects confront two bets – a *$ bet* offering a small chance of a relatively large prize, and a *P bet* offering a larger chance of a smaller prize. Subjects make straight choices between the two bets, and report a WTA valuation for each of them. The widely-observed anomaly is that many subjects choose the P bet while at the same time valuing the $-bet more highly. Cox and Grether elicit asks (i.e. bids to sell) for P and $ bets in repeated second price auctions; after five asks have been elicited for each of the gambles, subjects make the straight choice. While the usual preference reversal pattern is observed when comparing first round bids with choices, there is no systematic pattern of reversals in the fifth round. This change occurs because asks for $ bets fall markedly across rounds.

Since this is a fall in WTA valuations, it is compatible with all three hypotheses for exactly the same reasons as are Shogren et al's findings. On all three hypotheses, of course, a declining trend of WTA valuations is also predicted for P bets. But because P bets offer a high probability of winning, initial valuations of a P bet tend to be concentrated in the narrow interval of 'reasonable' or 'credible' values bounded by the bet's expected value and the value of its prize. Thus, strategic bias is likely to induce much more overstatement of valuations for $ bets than for P bets. For the same reason, initial asks are much more dispersed for $ bets than for P bets. Hence, shaping is likely to have a much stronger impact on the distribution of valuations for $ bets.

3. A new experimental design

Our experiment was designed to discriminate between the three hypotheses, and to test for the existence and persistence of the disparity between WTP and WTA in a repeated market context.

A key feature common to both the refining hypothesis and MDH is that market experience tends to produce convergence of behaviour on underlying preferences. It follows that the direction, and ultimately the termination, of the convergence process should be independent of the particular prices generated along the way. In our experiment, we examine whether individuals' valuations of a good respond systematically to price variability even in an experimental environment where the source of variability is controlled such that it contains no information relevant for the agent's own value. If we were to find evidence of subjects responding to price information in such an environment, that would be evidence consistent with shaping but inconsistent with either the refining hypothesis or the MDH.

In the experiment, subjects took part in a series of repeated auctions. All of the auctions were *median price auctions* which operated as follows. Each auction involved an odd number of participants, and a computer program elicited a bid from each subject equal to the price at which they were *just* not willing to trade. The *median* bid was then computed and announced as the market price. The program implemented all trades consistent with subjects' bids at the market price. So, in buying auctions, only subjects with bids *above* the market price bought; in selling auctions, only subjects with asks *below* the market price sold. The mechanism was sealed bid, but subjects learned the market price, and whether they had bought or sold, immediately at the end of each auction round. The median price institution is a form of Vickrey auction; it provides incentives for truthful revelation.

The most common test for the payment disparity in auction markets compares comparing *average* buying and selling bids, round by round, and most of the data has been generated from second price auctions[i]. While the comparison of average bids across a series of second price auction is an appropriate test for erosion of the payment disparity under the refining hypothesis, it is an inappropriate test given MDH. As we noted above, one implication of MDH is that repeated auction mechanisms cannot be assumed to reveal the underlying preferences of *all* bidders: only marginal bids are reliable signals of value. Indeed, evidence from experiments using induced values shows that deviations of bids from values are not uncommon among non-marginal traders (Miller and Plott, 1985; Franciosi et al 1993). If we are entitled to use only marginal bids as data, we must ensure that marginal bids are comparable. The buying and selling prices resulting from second price auctions will generally be non-comparable because they reveal, respectively, the second highest WTP and the second lowest WTA, which are directly comparable only in the case where there are exactly three traders.

One response to this was provided by Shogren et al who conducted repeated random nth price auctions[ii]. Since in this mechanism every trader is potentially marginal, the comparison of averages is legitimate. We follow a different route in using median price auctions. This affords us two different tests. We can test the refining hypothesis by looking for convergence of average bids over time. And we can test MDH by comparing the marginal bids of buyers and sellers over time. This is a legitimate comparison in the median price institution since in each case marginal bids provide a value measure at the middle of the value distribution.

In the first stage of the experiment, subjects bought or sold *vouchers* which had predetermined redemption values: that is, a subject holding a voucher at the end of the experiment was entitled to claim a fixed sum of money for sure. The experiment involved two vouchers with different redemption values: VL (the 'low value' voucher, with a redemption value of £2.25) and VH (the 'high value' voucher, worth £4.25). Subjects took part in two repeated voucher auctions (one buying, one selling) under the conditions set out in Table 1.

Table 1	*Majority (group I)*	*Minority (group II)*
Buy	VH (treatment T9)	VL (treatment T10)
Sell	VL (treatment T11)	VH (treatment T12)

In each series of auctions, the n participating subjects were randomly subdivided into *majority* and *minority* subgroups, which we label I and II respectively. The majority group contained (n+1)/2 subjects, the *minority* group, (n – 1)/2. In each buying auction, the majority bid for VH while the minority bid for VL.[iii] In each selling auction, the majority had the opportunity to sell VL while the minority had the opportunity to sell VH. Because each subject's private value for their voucher is its redemption value, we can identify a *correct* bid for each subject, and therefore identify the market price for each auction predicted by standard theory.

These voucher auctions serve two purposes in our design. First, they provide some indication of whether subjects understand the experimental environment. If subjects understand the environment, and if they always follow their dominant strategies without error, bids should always be equal to redemption values. That, of course, is a very strong assumption; as we have noted, there is substantial experimental evidence of deviations between bids and private values. However, we wished to see whether our own experimental procedure was broadly comparable with those that had been used by other researchers.

Second, if strategic bias (bidding low when buying, asking high when selling) explains the existence of the payment disparity, it is possible that it may be observed, at least in early rounds, in bids for vouchers (i.e. $WTP_{VL} < WTA_{VL}$ and/or $WTP_{VH} < WTA_{VH}$, where WTP_{VL} denotes WTP for lottery VL, and so on). On the refining hypotheses, we would expect any such disparities between bids and redemption values to be completely eliminated through market experience. In contrast, if MDH is true, we may observe some differences between bids and redemption values for non-marginal bidders even after repeated market feedback; but we should expect marginal bidders to learn to bid correctly.

In the second stage of the experiment, subjects bought and sold *lotteries*. In order to test for a shaping effect, we sought to construct an environment in which agents are not sure of the value to them of the goods being traded and where different groups of agents, bidding for identical goods in repeated markets, experience systematically different price feedback. Given that we required the market feedback to be a price genuinely produced by a freely operating market mechanism (and not fabricated by us as experimenters) the problem was how to engineer systematically different price feedback between the groups. Our solution was as follows.

We constructed median price auctions in which subjects bid to buy or sell binary lotteries of the form (x, p) where p is the probability of winning a money prize x (and with probability 1 – p, the lottery gives zero). The attractiveness of binary lotteries can be controlled in objective dimensions (prizes and chances of winning). Since such lotteries are not often bought and sold in markets outside the laboratory, our subjects would not have observed typical market prices for them, nor have previously refined values for them.

Two specific lotteries were involved: a ‘high probability’ lottery, H = (£12, 0.8) and a ‘low probability’ lottery L = (£12, 0.2). We elicited bids and asks for both lotteries. As in the voucher auctions, the participants in each auctions were divided into a majority subgroup I and an minority subgroup II. A crucial feature of our design was that within each market one subgroup was bidding for L while the other was bidding for H. Subjects were not told that some auction participants were bidding for different lotteries *in the same auction*, and we did not expect them to deduce this fact. Through this manipulation, we generated observations in eight different treatments, T1-T8, as set out in Table 2. Each row of the Table describes the conditions faced by two different subgroups in a single auction. So, for example, the first row describes a buying auction in which the majority subgroup I bid for H (treatment T1), while the minority subgroup II bid for L (Treatment T2). Given that the market price is the median bid, and given the large difference in expected value of the bets, we hoped that this manipulation would tend to produce a higher market price in auctions in which the majority bid for H than in those in which the majority bid for L.

Table 2	*Majority (I)*	*Minority (II)*
Buy	H (treatment T1)	L (treatment T2)
Sell	H (treatment T3)	L (treatment T4)
Buy	L (treatment T5)	H (treatment T6)
Sell	L (treatment T7)	H (treatment T8)

Comparisons between bids in the eight treatments allow us to conduct a number of tests. We can test for a WTP-WTA disparity by comparing buying and selling bids for a given lottery, while controlling for shaping by holding constant the lottery the majority are bidding for (e.g. by comparing T1 with T3 or T5 with T7). We can also examine whether the extent of any such disparity changes across rounds. For the refining hypothesis, the relevant test is to compare average bids; for MDH, it is to compare median bids. Our tests for shaping involve comparisons such as that between T2 and T5. Both of these conditions produce data on WTP_L; but in T5 the majority of subjects are bidding for L, while in T2 the majority are bidding for H (in T2. Although we have good reason to expect higher market prices in T2, that difference in price contains no information that is relevant to agents valuing L. So, if either the refining hypothesis or MDH is correct, although bids for L in these two conditions may change through rounds of a repeated auction, we have no reason to expect systematic differences in the trajectory of bids. If, on the other hand, bids for L are systematically higher in T2, that is evidence of shaping. We can conduct similar tests for shaping using WTA_L, WTP_H and WTA_H.

4. Experimental procedures

The data reported in this paper come from an experiment conducted at the University of East Anglia. At the start of every experimental session, each subject was assigned to a group of size n (n = 5 or 7), which then took part in a separate sequence of auctions with group membership unchanged throughout. A total of 175 subjects[iv] took part[v]. Of these, 104 subjects were in the majority subgroup (I), the remaining 74 were in the minority subgroup (II). Throughout the experiment, subjects sat at individual computers with partitions between each person. In buying auctions, each subject was endowed with an amount of cash (which was always more than the lottery prize or redemption value). In selling auctions, they were each endowed with either a voucher or a lottery. In each round, the computer program elicited bids/asks from each trader, identified the median value and fed that information back, at the same time telling each trader whether or not she had bought/sold at that price. Each auction was repeated six times in sequence.

Before responding to any of the tasks, subjects took part in ‘practice’ auctions. During the practice auctions, we explained the procedures[vi] and allowed subjects to become familiar with the auction rules and elicitation mechanism. They first practised in two voucher auctions (one buying and one selling), and then responded to the two experimental voucher tasks for real money. They then practised in two auctions for lotteries (again, one buying and one selling) before proceeding to the lottery tasks involving real monetary payoffs.

In each auction round, the program elicited a ‘just not willing to trade’ value from each subject. We designed our elicitation procedure to be as simple and as transparent as possible and we developed and refined it through pilot experiments. To illustrate how it worked, consider a subject in an auction with an opportunity to buy. The computer program asked a series of questions of the form: ‘Would you be willing to pay £x?’ By asking a series of such questions, adjusting the value of x according to previous answers, it was possible to approximate the subject’s maximum WTP. At the end of the sequence, the computer summarised the implications of the subject’s answers as follows:

‘You have said you are willing to pay x but you are not willing to play x + e.’

where x was the largest amount they had said they would pay and x+e was the smallest amount they had said they would not pay. The computer then asked the subject to confirm that they were happy with this statement. If the subject confirmed, x+e was recorded as the subject’s ‘just not willing to trade’ value; if the subject did not confirm, the elicitation procedure recommenced. The computer recorded bids on a 26 point scale with £0.5 intervals[vii]. We report results as values on this scale.

Random lottery incentives were used to motivate subjects: they knew from the outset that one round of one auction would be selected, at the end, to be played out for real. If the selected round was a selling round and the subject had sold their voucher or lottery during that round, they were paid whatever had been the market price. If they had kept their endowment, they were paid the redemption value if it was a voucher or, if it was a lottery, the lottery was played out and the subject received the outcome. If the round selected was a buying round, in addition to the payoff from any lottery or voucher that may have been purchased, subjects were paid any remaining cash endowment (net of the purchase price if they had bought).

5. Results: voucher auctions

Since the vouchers have fixed redemption values, a lack of coherence in subjects’ valuations of vouchers might be interpreted as a signal of misunderstanding by subjects. There are, however, very clear patterns in the data which we take as reassurance on this score. The correct bid is very common – it is made by between 44% and 73% of subjects, depending on the task and round. We are inclined to take the general proximity of bids to values as reassurance that subjects were responding coherently and sensibly within our auction mechanism.

However, as Table 3 shows, there is underbidding by buyers of the high value voucher and overbidding by sellers of the low value vouchers. (In the tables of results, 'Buy L/ Maj H' refers to those subjects who bid to buy voucher L in auctions in which the majority of particpants were bidding to buy H, and so on.) These deviations from redemption values are each significant at 1% and persist through the 6 auction rounds. There is no evidence of much learning among those whose bids deviate from their values; and thus there is little support for the refining hypothesis.

Table 3: testing for deviations in aggregate in voucher auctions
(H_o: mean bid = correct bid; H_1: mean bid ≠ correct bid)

	Round	*Mean bid*	*Correct bid*	*Z*	*Sig*	*Bias observed*
Buy H/ Maj H	1	8.17	9	-4.38	1%	underbidding
	6	7.74	9	-6.52	1%	underbidding
Buy L/ Maj H	1	4.92	5	-0.53	no	none
	6	5.03	5	0.173	no	none
Sell L/ Maj L	1	5.67	5	4.27	1%	overbidding
	6	5.58	5	4.19	1%	overbidding
Sell H/ Maj L	1	8.96	9	-0.293	no	none
	6	8.76	9	-1.58	no	none

To investigate whether voucher bids are consistent with MDH, we need to consider the marginal traders. If MDH is true, underbidding by *marginal* traders in buying auctions, and overbidding by *marginal* traders in selling auctions, should tend to disappear as market experience accumulates. In our data, however, the practice of understating in buying auctions and overstating in selling auctions persists even at the margin, as shown by Table 4. Although median WTA bids show *some* tendency to move towards the level standard theory predicts, it is still the case that by round 6 more than half of all market prices are at least one point on the scale higher than the relevant redemption value. For buying auctions there is even less evidence of market discipline moving prices towards the predicted level: even in round 6, more than half are at least two points lower than they should be.

Table 4: Testing for deviations at the margin in voucher auctions MEDIAN BIDS						
ROUND	*1*	*2*	*3*	*4*	*5*	*6*
WTP Majority: H (value = 9)	7	7	7	6	7	7
WTA Majority: L (value = 5)	7	7	7	6	6	6

6. Results: lottery auctions

First we consider whether, as the refining hypothesis predicts, WTP-WTA disparities are eroded by market experience. Here the relevant tests involve comparisons of mean bids. The mean bidding data for lottery auctions are summarised in Figures 1 and 2. Figure 1 presents data for the low value lottery L; Figure 2 presents the same data for H. Each data point records the average bid for a particular treatment in a given auction and round. Buying (WTP) bids are represented as solid diamonds, while selling (WTA) bids are unfilled circles. The darker lines connect data-points for groups were the majority of subjects were bidding for H. It is apparent that time profiles of average bids for L and H are quite different: the most obvious feature of the data for H (the very strong downward time trend) is not present in the data for L. Our main interest here, however, is in whether there is a payment disparity (with WTP < WTA), and if there is, whether it persists across rounds.

Figures 1 and 2 suggest the presence of a payment disparity for both High and Low lotteries in round 1, though inspection of the graphs suggests a tendency for this to decline over time. This general impression is confirmed by a series of hypothesis tests reported in Table 5. In each case the null hypothesis of no disparity is compared against a one-tailed alternative. Each row of the table presents buying and selling means from for the relevant condition with standard deviations in parentheses. We also report the sample size (n), the difference between means (WTA – WTP) and the standardised measure of the difference (z). Where there is evidence of a significant payment disparity, we report the level of significance at which the null is rejected (a 'no' in the final column means that we cannot reject the null at 10%).

Table 5: Testing for a WTP-WTA disparity in lottery auctions
MEAN BIDS
(H_o:WTA=WTP) (H_1: WTA > WTP)

	WTP(sd)	*WTA (sd)*	*n*	*WTA-WTP*	*z*	*signif*
Low:maj L						
round1	7.03 (4.11)	8.72 (4.65)	104	1.69	2.78	1%
round6	7.26 (3.81)	8.10 (4.11)	104	0.84	1.53	10%
High:maj L						
round1	14.15 (5.08)	16.89 (3.56)	71	2.74	3.72	1%
round6	12.06 (4.63)	14.24 (5.09)	71	2.18	2.67	1%
Low: maj H						
round1	7.42 (3.61)	9.24 (4.17)	71	1.82	2.78	1%
round6	8.41 (3.86)	9.41 (3.59)	71	1	1.6	10%
High: maj H						
round1	14.12 (4.98)	15.06 (4.92)	104	0.94	1.37	10%
round6	12.48 (4.55)	13.29 (4.8)	104	0.81	1.25	no

In round 1 we find a significant payment disparity in all conditions (significant at 1% in three out of four cases). There is a tendency for this to decline across rounds but, after six rounds, there remains a significant disparity in one case at 1% (and 10% in two other cases). These results do not give strong support for the refining hypothesis.

Next, we consider whether, as MDH predicts, payment disparities *in marginal bids* diminish with experience. Since our design uses median price auctions, comparisons of median bids provide the appropriate tests. The relevant data are presented in Table 6. The first two rows of the table report, respectively, the medians of buying and selling bids when the majority bid for H. Each row pools data for two groups: each of these values is the median of a distribution which contains both bids for L and bids for H, but since the majority in each auction is bidding for H, it is legitimate to interpret the median as a bid for H.

Table 6: Testing for a WTP-WTA disparity in lottery auctions MEDIAN BIDS						
ROUND	*1*	*2*	*3*	*4*	*5*	*6*
Majority: H (so median bid usually for H)						
WTP (T1+T2)	12	11	11	11	11	11
WTA (T3+T4)	12	12	12	12	12	11
Majority: L (so median bid usually for L)						
WTP (T5+T6)	9	9	9	9	9	9
WTA (T7+T8)	12	11	10	10	9	9

In round 1, there is evidence of a payment disparity, but only in auctions where the majority bid for L. By round 6, however, there is no evidence of a payment disparity at all. By contrast with what we found in the voucher auctions, these results are consistent with the MDH.

Finally, we test for shaping effects by comparing bids made for a given lottery (either H or L) in two conditions: one where the majority of subjects are bidding for H, the other where the majority are bidding for L. The shaping hypothesis provides no reason to expect any significant differences between these treatments in round 1, but for subsequent rounds it gives us reason to expect higher bids in groups where the majority of subjects bid for H.

Table 7: Testing for shaping in lottery auctions **(H_o: bid/MajH = bid/ Maj L) (H_1: bid/ Maj H > bid/ Maj L)**						
		Maj H (I)	*Maj L (II)*	*difference (H-L)*	*z*	*sig*
Buy High	*round 1*	14.12 (4.98)	14.15 (5.08)	-0.03	-0.04	no
	round 6	12.48 (4.55)	12.06 (4.63)	0.42	0.59	no
Buy Low	*round 1*	7.42 (3.61)	7.03 (4.11)	0.39	0.66	no
	round 6	8.41 (3.86)	7.26 (3.81)	1.15	1.95	5%

Sell Low	*round 1*	9.24 (4.17)	8.72 (4.65)	0.52	0.77	no
	round 6	9.41 (3.59)	8.10 (4.11)	1.31	2.23	5%
Sell High	*round 1*	15.06 (4.92)	16.89 (3.56)	-1.83	-2.85	no
	round 6	13.29 (4.08)	14.24 (5.09)	-0.95	-1.24	no

Table 7 presents a series of tests for shaping effects with separate tests for round 1 and round 6. In each case, the null hypotheses that average bids are equal in the two conditions is tested against the one tail alternative that bids are higher when the majority bid for H.

Looking first at round 1 data, we cannot reject the null (no shaping) in any of the four cases. In round 6, however, we do find significant differences between subgroups in bids for L. Values revealed for both buying and selling are higher for the subgroups where the majority were valuing H. These differences are in the direction predicted by the shaping hypothesis and they are significant at the 5% level. Moreover, the fact that the statistically significant effects occur with the L lotteries rather than with the H lotteries is reminiscent of Cox and Grether's finding that it was the more highly dispersed $ bet values that changed most in the context of Vickrey auctions rather than the more tightly bunched P-bet values: our L lotteries have the characteristics of $-bets (low probability of winning, and low expected value relative to high payoff) whereas our H lotteries more closely resemble P-bets. We conjecture that this may be at least part of the reason why shaping effects were more pronounced for L than for H. These data cannot be explained either by the refining hypothesis or by MDH.

7. Conclusions

Recent experimental research has led some to claim that experience of a market mechanism may eliminate certain well-attested violations of the standard model of preference. As yet, the evidence for this claim is relatively modest and there has been little attempt to set out what role the market – distinct from the role of repetition and incentives – might play in such observations.

We have presented a new experiment which had two main objectives. First, we sought to provide fresh evidence on the robustness of the WTP-WTA disparity in a repeated market environment. Our tests, using the median price auction mechanism, provide a new and, relative to most existing studies, more neutral framework in which to test for such a disparity. We found significant evidence of the disparity; and while there is evidence that it tends to decay across rounds, we do not find that it is eliminated within our design.

Our second main objective was to examine alternative hypotheses about the effect market experience may have on stated preferences. We set out three hypotheses. The main finding of our experiment in this regard is that we have observed evidence consistent with the shaping hypothesis. This evidence is not consistent with any other hypothesis that we are aware of. The discovery of a shaping effect has potentially far-reaching theoretical consequences. For example, claims concerning the efficiency of competitive markets typically assume that preferences are independent of market activity. If it were the case that values are 'contaminated' by price feedback through market participation, that would warrant serious reconsideration of the foundations of standard economic theory. An illustrative worry at a more practical level would be that the advice we offer to practitioners of contingent valuation would differ depending on which of our hypotheses is more powerful. If markets discipline and/or refine stated values, then values revealed in a repeated market context will be more reliable indicators of individuals' preferences than values observed in 'one-shot' contexts. But if the shaping hypothesis is correct, the reverse may be true. While our experiment takes only one tentative step in the direction of understanding the influence of market participation on stated values, we suggest that the issues raised warrant further investigation.

References

Binmore, Ken. (1994) Playing Fair, MIT Press.

___________ (1999) “Why Experiment in Economics?”, Economic Journal, 109, pp. F16-24.

Camerer, Colin F. (1995) “Individual Decision Making,” in Handbook of Experimental Economics. J. Kagel and A. E. Roth, eds. Princeton: Princeton U. Press.

James C. Cox and David M. Grether (1996) “The Preference Reversal Phenomenon: Response Mode, Markets and Incentives.” Econ. Theory, 7, pp. 381–405.

Franciosi, R, R.M. Isaac, D. Pingry and S Reynolds (1993) “An Experimental Investigation of the Hahn-Noll revenue neutral auction for emissions licenses”, Journal of Environ. Econ. Management, 24, pp. 679-90.

Knetsch, Jack., Fang. F. Tang and Richard Thaler (2001) “The Endowment Effect And Repeated Market Trials: Is The Vickrey Auction Demand Revealing?” Experimental Economics, 4, pp. 257-000.

Miller, G and Charles Plott. (1985) “Revenue Generating Properties of Sealed-Bid Auctions: An Experimental Analysis of One-Price and Discriminative Auctions”. pp. 159-182 in Smith, Vernon (Ed.), Research in Experimental Economics, Vol. 3, JAI Press, Inc. Greenwich CT.

Plott, Charles, R. (1996) “Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis,” in *Rational Foundations of Economic Behavior*. K. Arrow, E. Colombatto, M. Perleman, and C. Schmidt, eds. London: Macmillan and NY: St. Martin’s, pp. 225–50.

Shogren, Jason, F., Sungwon Cho, Cannon Koo, John List, Changwon Park, Pablo Polo and Robert Wilhelmi. (2001) “Auction Mechanisms and the Measurement of WTP and WTA”, Resource and Energy Economics, 23, 97-109.

Starmer, Chris. (2000) "Developments In Non-Expected Utility Theory: The Hunt For A Descriptive Theory Of Choice Under Risk", Journal of Economic Literature, XXXVIII, 332-382, 2000.

[i] Exceptions are Knetsch et al, and Shogren et al 2001 who respectively use, ninth and random nth price auctions in addition to the second price institution.

[ii] In this mechanisms, bids are rank ordered, and the nth is chosen (at random after all bids have been submitted to be the price). In buying (selling) auctions, n-1 subjects with bids above (below) the price buy (sell).

[iii] We told subjects that the face value of vouchers was liable to vary between subjects in any particular auction. We did this so that subjects would then not be surprised by market prices which were different from the value of their own lottery.

[iv] This excludes data for one additional group of subjects that was lost due to a computer failure in one session.

[v] In all the subjects took part in 10 repeated auctions though in this paper we report data from only 6. The other 4 auctions and two additional choice tasks will be reported elsewhere. The

additional auctions were, from the point of view of the subjects, lottery auctions; the choice tasks, though different, were always the final tasks in the experiment and so subjects' exposure to these could not have affected the data reported here.

[vi] Verbal instructions were read from a script, a copy of which can be downloaded from www.***

[vii] The computer recorded a value (v) of 1 for subjects with a 'just not willing to trade' value of £0.01; v = 2 for subjects just not willing to trade at 50p; v = 3 for subjects just not willing to trade at £1; and so on in 50p intervals up to v = 26. Notice that an adjustment is needed to compare buying and selling values on this scale. For example, a subject who states v = 3 for 'just not willing to *sell*' has reported a value between 3 and 4. A subject who states a v = 4 for 'just not willing to *buy*' also has a value between 3 and 4. We adopt the convention of reporting WTA = v and WTP = v - 1 throughout so the data we present are pre-adjusted to allow immediate comparison of buy and sell bids.