Bonn Econ Discussion Papers

Discussion Paper 10/2001

Learning Direction Theory and the Winners Curse

by

Reinhard Selten, Klaus Abbink, and Ricarda Cox

January 2001

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Learning Direction Theory and the Winner's Curse

by REINHARD SELTEN[*], KLAUS ABBINK[**], and RICARDA COX[*]

[*] University of Bonn
[**] Institut d'Anàlisi Econòmica CSIC and Universitat Pompeu Fabra, Barcelona

January 2001

Abstract

We report an experiment on a decision task by SAMUELSON and BAZERMAN (1985). Subjects submit a bid for an item with an unknown value. A winner's curse phenomenon arises when subjects bid too high and make losses. Learning direction theory can account for this. However, other influences on behaviour can also be identified. We introduce *impulse balance theory* to make quantitative predictions on the basis of learning direction theory. We also look at *monotonic ladder processes*. It is shown that for this kind of Markov chains the impulse balance point is connected to the mode of the stationary distribution.

Keywords

Experimental economics, learning, individual decision making

JEL Classification Codes

C91, D81, D83

Acknowledgements

Financial support by the *Deutsche Forschungsgemeinschaft* through the *Sonderforschungsbereich 303*, the European Union through the TMR research network ENDEAR (FMRX-CT98-0238), and the *Land Nordrhein-Westfalen* is gratefully acknowledged.

Correspondence Address

Laboratorium für experimentelle Wirtschaftsforschung
Uni Bonn
Adenauerallee 24-42
53113 Bonn, Germany
phone 49-228-73-9192
fax 49-228-73-9193
e-mail selten@lab.econ1.uni-bonn.de
klaus.abbink@econ.upf.es

1. Learning Direction Theory

Learning direction theory (SELTEN and STOECKER 1986, SELTEN and BUCHTA 1999) is a qualitative theory about learning in repetitive decision tasks. The theory is quite simple and can best be introduced with the help of an example. Consider an archer who tries to hit the trunk of a tree. If the arrow misses the tree on the left side, then the archer will tend to aim more to the right, and in the case of a miss to the right the aim will be more to the left.

The example may seem to be trivial, but it is important to realise that the behaviour of the archer is based on a qualitative causal picture of the world. The direction of a change of the aim is seen as resulting in a corresponding change of the direction of the flight of the arrow. If the archer had to find his aim by looking into a mirror image, the causal relationships could be reversed.

Another feature of the situation is the fact that the archer sees whether the arrow misses to the left or to the right. This feedback permits the use of the qualitative causal picture of the environment in which the decision is made. Learning direction theory does not only require a qualitative causal picture of the world, but also the right kind of feedback which makes it possible to make use of it.

The way in which the decision is based on experience may be described as “ex-post rationality”. One looks at what might have been better last time and adjusts the decision in this direction.

Learning direction theory is applicable to repetitive decision tasks in an environment which can be described as follows.

> ***Learning situation:*** A parameter p_t has to be chosen in periods 1...T. After each period t the decision maker receives feedback which permits causal inference on what might have been better in this period.

The decision is guided by the following principle:

> ***Ex-post rationality:*** p_t # p_{t-1}, if $p < p_t$, but not $p > p_t$ might have been the better parameter choice last period. Similarly $p_t \$ p_{t-1}$, if $p > p_t$, but not $p < p_t$ might have been the better parameter choice last period.

Consider the example of a sealed-bid auction. Let b be the bid, v be the value of the object to the bidder, and p be the price. After an auction, the bidder may find himself in one of three experience situations:

1. Success: The bidder receives the object.	$b = p$	b possibly too high.
2. Lost opportunity:	$b < p < v$	b too low.
3. Outpriced bid:	$p > v$	

In the first experience condition, the bidder might have obtained the object for a lower bid, and a higher bid would have been less advantageous. Therefore, in this condition, the bid was possibly too high, and according to ex-post rationality one can expect a tendency to a lower bid next time, if the bid is changed at all. In the case of a lost opportunity, the bidder could have profitably obtained the object by overbidding *p*. According to ex-post rationality, one can expect a tendency to an increase of the bid if it is changed at all.

Learning direction theory is not meant to be a complete explanation of adaptive behaviour. Therefore, it does not propose the hypothesis that ex-post rationality prevails in all cases. Sometimes other influences might move the decision in the opposite direction. However, it is assumed that ex-post rationality is the strongest influence. These considerations can be summarised as follows.

> ***Prediction:*** More frequently than randomly expected, parameter changes, if they occur, are in the direction indicated by ex-post rationality.

Learning direction theory is qualitative. It does not specify probabilities and sizes of adjustments. However, the principles of learning direction theory can be incorporated into quantitative theories. In the last part of this paper, we shall introduce the *impulse balance theory*, which makes an attempt at the description of long-run consequences of direction learning.

It should be emphasised that learning direction theory is very different from the usual models of reinforcement learning (BUSH and MOSTELLER 1955, ROTH and EREV 1995). In these theories, the degree of reinforcement of an action depends on the size of the payoff obtained by it. In learning direction theory the level of the payoff obtained for last period's decision does not matter at all. What is important is the additional payoff which might have been gained by other actions. It is not experienced payoffs alone, but rather the comparison of experienced payoffs with hypothetical payoffs which guides the decision maker. The counterfactual causal reasoning about the past is a crucial feature of learning direction theory.

There are at least twelve studies in which learning direction theory has been successfully applied. Table 1 lists them and shortly characterises the area of research to which they belong. It can be seen that learning direction theory is applicable to a wide variety of contexts.

2. Winner's Curse

The winner's curse phenomenon was first observed in the context of oil field auctions by CAPEN, CLAPP, and CAMPBELL (1971). They pointed out that on the average, those firms who obtained oil fields as auction winners made a loss. The explanation for this is that every bidder bases its bid on a value estimate suggested by geological studies. The higher this estimate, the higher will be the bid. Consequently, the object is likely to be obtained by the bidder with the highest over-estimate. The bidders do not sufficiently take this into account and therefore make a loss in the case that they get the object. Of course, in game-theoretic equilibrium, the right corrections are made, and no losses appear.

Table 1. Applications of learning direction theory in the literature

Selten and Stoecker (1986)	Prisoners' dilemma end effect
Mitzkewitz and Nagel (1993)	Ultimatum game
Kuon (1994)	Bargaining
Ryll (1995)	Bargaining
Nagel (1996)	Beauty Contest game
Cason and Friedman (1997)	Markets
Berninghaus and Ehrhart (1998)	Co-ordination
Sadrieh (1998)	Markets
Kagel and Levin (1999)	Auctions
Selten and Buchta (1999)	Auctions
Grosskopf, Erev, and Yechiam (2000)	Individual decision making
this study	Winner's Curse

The winner's curse has been experimentally observed in common value auctions (for an overview see Kagel 1995). For this study, the experimental work of Samuelson and Bazerman (1985) is of special significance because our experimental set-up is very similar to theirs. They experimented with a situation which was first theoretically investigated by Akerlof (1970). The experimental set-up of Samuelson and Bazerman is as follows.

A sells a firm to B. The value of the firm for A is v, where v is a random variable with a uniform distribution over the interval $0 \# v \# 100$. The value of the firm for B is $1.5v$. The information about v is different for both players. A knows the realisation of v, but B knows the distribution of v only. B has to name a price x, and A accepts this offer for $x \$ v$, and rejects it for $x < v$. In experiments B is represented by a subject, but A is simulated by a computer, and B is informed about the behaviour of A. The situation is repeated over many periods, and after each period B is informed about the realisation of v regardless of whether his offer was accepted or not.

In the experiments subjects are often guided by the expected values $E(v) = 50$ for A, and $1.5E(v) = 75$ for B. In choosing the price x B may think that he has to give a little more than 50 to A in order to induce him to sell, and that this can be done because B's expected value is 75. However, this reasoning about absolute expectations is wrong. What matters are not the absolute expectations, but the conditional expectations

$$E(v * x \# v) = ½ x.$$

The offer x is accepted only in the case $x \# v$, and the mean value of x in this case is ½x. Therefore, B can expect

$$1.5\, E(v * x \# v) = ¾ x$$

as the mean value of x for B. Since B has to pay the price x, this has the consequence that B makes a loss of ¼ x for every $x > 0$. The only way to avoid losses is the bid $x = 0$.

BALL, BAZERMAN, and CARROLL (1991) conduct an experiment in which 37 subjects perform this decision task 20 times. The authors observe average bids which are clearly higher than the optimal bid of $x = 0$. Moreover, no tendency of the average bid towards the optimum can be detected.

In the literature the reason for the winner's curse phenomenon in the SAMUELSON-BAZERMAN task is seen in the wrong orientation at absolute values which has been discussed above. It is plausible that this is the correct explanation for the first period, but one should expect that over time, the error is corrected by learning. Why does learning fail to produce a tendency towards the optimum? Learning direction theory provides an explanation for this.

Suppose that B chooses $x = 50$. Then, B's offer will be rejected with probability ½ and accepted with probability ½. Ex-post rationality indicates an increase of the bid in the case of rejection, and a decrease of the bid in the case of acceptance. Of course, the impulses in both directions might differ in strength, but somewhere in the middle of the range, the impulses in both directions will balance each other. In this way it becomes understandable that learning will not push the process towards the optimum.

3. Design of the New Experiment

In the following, we shall report the results of our own experiments. Our set-up was analogous to that of BALL et al., with the following differences. Our experiments were extended over 100 periods rather than only 20, because one might think that the lack of convergence to the optimum in the study of BALL et al. were due to an insufficient number of periods. The value v was uniformly distributed over the integers [u,...,99], with $u = 1$, $u = 11$, and $u = 21$ for 18 subjects each. In the case $u = 1$, the optimal profit is reached at two bids: $x = 1$ and also $x = 2$. We experimented with different minimum values since it seemed to be desirable not to restrict the study to conditions under which the optimum is extreme. For $u = 11$, the optimum is reached at $x = 21$ and $x = 22$, and for $u = 21$ at $x = 41$ and $x = 42$.

The subjects started with initial assets of 250 and received a fixed income of 20 in each period in addition to the income from the auction. This fixed income had the purpose to prevent bankruptcies caused by losses due to the winner's curse. The money payoff was DM 1.50 for 100 experimental money units. Subjects earned payoffs in the range from DM 8.48 to DM 36.65 for a time of about one hour.

The experiments were run in the *Laboratorium für experimentelle Wirtschaftsforschung* at the University of Bonn. The subjects were students at the University of Bonn, mostly majoring in economics or law. The experiment was computerised, with software developed using *RatImage* (ABBINK and SADRIEH 1995). Each session began with an introductory talk. A translation of the written instructions is reproduced in the appendix. The instructions were read aloud and explained in detail. After the introduction, the subjects were seated in cubicles, visually separated from one another by curtains.

4. Confirmation of Learning Direction Theory

If a subject has made a bid which turns out to be equal to the value, ex-post rationality does not indicate a change in either direction. Learning direction theory makes predictions for the case that bid and value differ. There are two such experience conditions.

1. Overpayment: $x > v$ tendency towards lower x.
2. Lost opportunity: $x < v$ tendency towards higher x.

In both cases, the best bid in the last period would have been $x = v$. Therefore, ex-post rationality indicates a decrease in the case of overpayment and an increase in the case of lost opportunity, if the bid is changed at all.

Let r be the fraction of changes indicated by learning direction theory within all bid changes, and let $\bar{r}$ be the mean of the fractions r over all subjects. Learning direction theory predicts that this mean is greater than expected under some random theory.

$\bar{r} = 0.5$ is not an appropriate random hypothesis. A uniform random bid distribution yields an expected value $E(r) = 0.67$ for $u = 1$. This is due to the fact that in the case of a high bid in the last period, it is on the one hand highly probable that next period's random bid will be lower and on the other hand also highly probable that last period's value was below last period's bid. The case of a low bid in the last period is analogous. In view of this, there is a higher probability of 0.5 of moving in the direction indicated by ex-post rationality by a random bid.

The problem arises how to construct an adequate comparison theory against which learning direction theory can be tested in a meaningful way. Our answer to this question is based on a null hypothesis which assumes that there is no relationship between bid and values, or, more precisely, between the bid change and the experienced situation in the preceding period. Our approach to derive testable consequences from this supposed absence of a relationship is a construction involving a randomisation of the 100 values with which a subject was confronted during the experiment. In this randomisation the bid sequence is kept fixed.

Of course, one could randomise in a different way, with values kept fixed and bids permuted. However, this seems to be less adequate since often bids are not changed and therefore one obtains completely different distributions of bid changes if the bids are permuted. Since learning direction theory in concerned with bid changes rather than bids, an unchanged bid sequence together with randomly permuted value sequences seems to be the better point of departure for a comparison theory.

Let p be the expected value of r under a random permutation of the values in the 100 periods with the bid sequence kept fixed. For every subject, we determined the *surplus*

$$s = r - p$$

of the observed r over the comparison value p. Let $\bar{s}$ be the expected value of s for the population of subjects. Our null hypothesis is that $\bar{s} = 0$. It is clear that this should be the case if there is no relationship between bids and values. The alternative hypothesis according to learning direction theory is $\bar{s} > 0$. The Wilcoxon test applied to the individual surpluses s yields the following result.

Result 1. *For $u = 1$, $u = 11$, and $u = 21$ separately, the null hypothesis $\bar{s} = 0$ is rejected with 0.5% one-tailed significance in favour of $\bar{s} > 0$.*

The average bids for $u = 1$, $u = 11$, and $u = 21$ are shown in figure 1. It can be seen that there is no obvious tendency of convergence towards the optimum. However, rank correlations of bid and period number for individual subjects are more often negative than positive. This effect is significant according to the Wilcoxon test on the 5% level (one-sided) in the case $u = 1$, but not for $u = 11$ (one-tail $p = 0.077$) and $u = 21$ (one-tail $p = 0.152$). If there is a tendency towards the optimum, then it must be a very weak one.

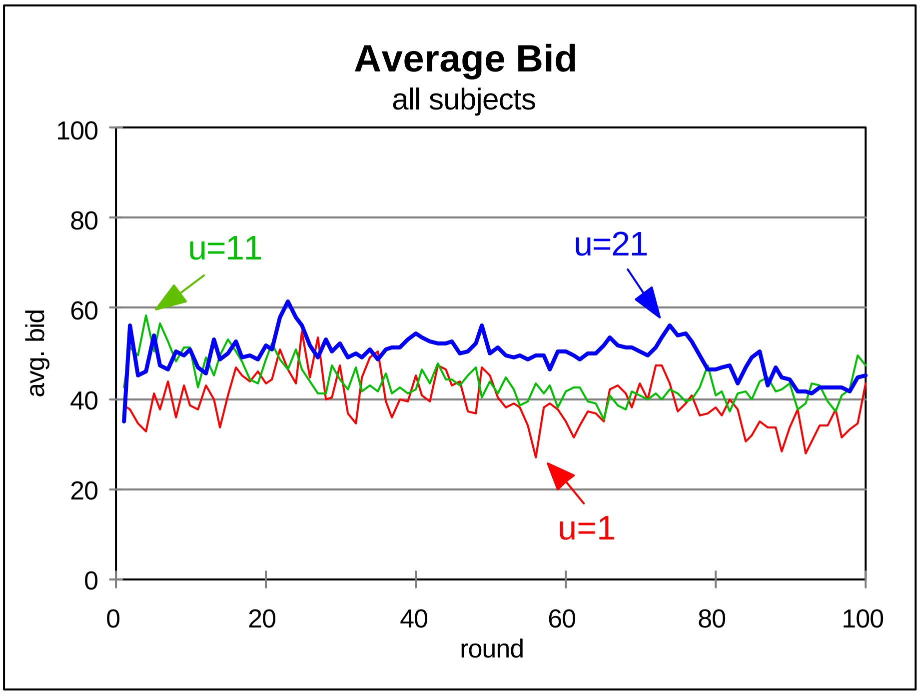

Figure 1

5. Other Influences on Behaviour

After the experiment the subjects had to fill a questionnaire in which the following questions were asked:

(1) Imagine you would commission somebody to make bids in your place. What instructions would you give to this person?

(2) What are your reasons for these instructions?

(3) Comments

The answers to these questionnaires revealed influences on behaviour very different from learning direction theory. Of course, the instructions do not necessarily agree what the person actually did. Sometimes the subjects even expressed that their own behaviour was wrong and something different should be done. Not every subject clearly described the rationale of her or his instructions. Some answers were quite explicit and very suggestive for our analysis, but others were not very useful.

In his book on the psychology of stochastic thinking, SCHOLZ (1987) makes a distinction between an intuitive and an analytic style of approaching decision tasks. The intuitive decision maker does not make any calculations but relies on his feelings, whereas the analytic decision maker tries to find a way to calculate the decision. Of course, the same person can sometimes be in an intuitive and sometimes in an analytical mode of decision making, and may switch between the two styles within an experiment.

Presumably, an intuitive decision maker cannot tell how he arrives at his decision processes, since they are hidden below the level of consciousness. Analytical decision makers can describe the way in which they calculate their decision, but they may not be able to express themselves clearly.

The answers to the questionnaire suggest types of behaviour, but in view of the lack of clarity of many of them and because of the difference between instructions and own behaviour we categorise the subjects not on the basis of their questionnaires, but on the basis of their modal bids. We distinguish the categories listed in table 2. The table shows the modal bids for five of the categories and the number of subjects in each category. The first three categories are analytical. In some cases, the subjects did not exactly follow the principles underlying the modal bids, but rounded to the next number divisible by five, or made a computational mistake clearly visible in their comments. We now explain the categories and the underlying principles.

(1) ***Optimisers:*** A subject is categorised as an optimiser if its modal bid maximises expected payoff.

(2) ***Loss avoiders:*** These subjects want to avoid losses and, under this constraint, make a bid as high as possible. In the case $u = 1$, there is no difference between optimisation and loss avoidance. We did not count subjects under the condition $u = 1$ as loss avoiders.

(3) ***Asset conservers:*** These subjects want to avoid a decrease of their assets and, under this constraint, make a bid as high as possible. Here, it is important that a subject receives an additional fixed income of 20 each period and therefore ca make a loss of up to 20 without decreasing her or his asset.

(4) ***Gamblers:*** These subjects are interested in high gains. The highest possible gain can be obtained for a bid of 99, in the case that the value is 99. Clearly, subjects with a modal bid of 99 must be extremely risk-loving and therefore may be called "gamblers".

(5) ***Refusers:*** These subjects had modal bids of 0 under the conditions $u = 11$ and $u = 21$ in spite of the fact that they could have risk-less gains like the loss avoiders. A zero bid can be interpreted as a refusal to enter the auction.

(6) ***Adapters:*** This is the residual category of all subjects not in the five other categories.

Table 2: Categories, modal bids of the categories, and number of subjects in each category

category	u = 1		u = 11		u = 21	
	modal bid	number of subjects	modal bid	number of subjects	modal bid	number of subjects
Optimisers	0, 2	4	20, 22	3	40, 42	3
Loss avoiders	–	–	16	1	29, 31	2
Asset conservers	20	2	34, 35, 36	5	50	1
Gamblers	99	2	99	0	99	0
Refusers	–	–	0	0	0	2

The assignment of subjects to categories is based on modal bids. Thus, an optimiser may be somebody who learned to optimise only relatively late in the game. Therefore, we do not assert that a subject in one of the categories is exclusively guided by the underlying principle. We cannot even assert that the underlying principle was the most important determinant of somebody classified in one of the categories. However, the modal bid suggest that the principle was probably a substantial influence. One can expect that learning direction theory applies more to the adapters than to the members of the other groups, and this is in fact the case. We test this by an application of the Mann-Whitney U-test to the surpluses $s = r - p$.

Result 2. *The null hypothesis that the surpluses of adapters and subjects in other categories come from the same population can be rejected on the one-tailed significance level of 1% in favour of the hypothesis that adapters conform more to learning direction theory than other subjects.*

This result leaves room for the possibility that learning direction theory to some extent also influences the behaviour of subjects in the first five categories. In fact, the mean surplus is positive for all these categories with the exception of that of the gamblers. Both gamblers have negative surpluses. An explanation in terms of the gambler's fallacy suggests itself. A gambler expects that winning will become the more probable the longer a series of losses becomes. Suppose that last period's bid was high, but the value was low. Then, a subject influenced by the gambler's fallacy expects that now the probability of a high value is greater than before. This may be a motivation to increase the bid contrary to what is expected according to learning direction theory. Similarly, the gambler's fallacy may induce the lowering of a bid after a high value. Obviously, the influences of direction learning and the gambler's fallacy go in opposite direction.

In some cases the experimenters had the impression that subjects wanted to take an analytic approach, but were unable to find one, and therefore had to remain among the adapters. This suggests that people with better analytical knowledge may be more likely to take an analytical approach.

The subjects were asked whether they had taken a game theory course or not. Three subjects did not reply this question. They were not optimisers. For the other subjects, table 3 shows that among those who had taken a game theory course there were relatively more optimisers than among the subjects who had not.

Result 3. *Subjects with game-theory knowledge are more likely to be optimisers. The Fisher exact test yields a one-tailed significance level of 0.5% for table 3.*

Table 3. Influence of game theory knowledge[1]

		Optimisers: yes	Optimisers: no
Game theory course	yes	7	8
	no	3	33

[1] Three subjects did not answer the question about the game theory course. They were all non-optimisers.

Figure 2 shows the average bid of the adapters for $u = 1$, $u = 11$, and $u = 21$. Figure 2 looks similar to figure 1, with the difference that fluctuations are wider in figure 2. This is due to the smaller number of subjects.

For the adapters we also obtain predominantly negative rank correlations of bid and period number in the condition $u = 1$ (for 9 subjects out of 10). According to the Wilcoxon test, the result is significant at the 2.5% level (one-sided). This is not the case for $u = 11$ and $u = 21$. Learning direction theory excludes a tendency towards the optimum, but not necessarily a weak negative or a positive trend over 100 periods.

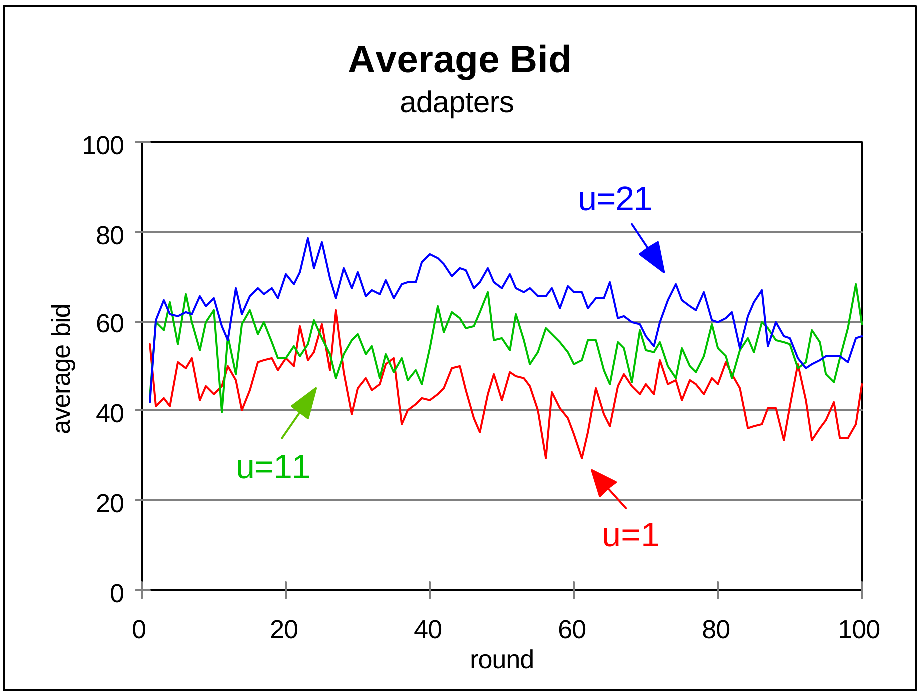

Figure 2

6. Impulse Balance Theory

In this section, we want to propose *impulse balance theory*, an attempt to derive rough quantitative predictions about the long-run consequences of direction learning without a detailed learning theory. Ex-post rationality results in an upward or downward change of the parameter. Impulse balance theory is concerned with situations in which the feedback at the end of a period permits a clear conclusion about what would have been the best possible choice of the parameter in the last period, given what is known by hindsight. We call this choice the ex-post rational choice.

It is plausible to assume that the expected change is the greater the greater the payoff is which could have been gained in the last period by the ex-post rational parameter choice. This is the basic idea behind impulse balance theory. However, this idea needs to be modified by treating gains and losses differently, because it is well-known that individuals tend to be much more sensitive to losses than to gains (see e.g. KAHNEMAN and TVERSKY 1979, BENARTZI and THALER 1995).

One could try to build a full-fledged learning theory on this approach, but it seems to be better to avoid the modelling details connected to this. Therefore, impulse balance theory only determines a single parameter value, the *impulse balance point*, at which expected upward and downward impulses are equally strong. It is plausible to conjecture that the bulk of the stationary distribution of the learning process will be placed in the vicinity of the impulse balance point. The acceptance of impulse balance theory amounts to the assumption that this is correct.

A downward impulse results if the bid x is greater than the value v. In this case, we speak of an *overpayment*. If x is greater than $1.5v$, then a loss results from the bid. Upward impulses are foregone payoffs of $0.5v$ in the case that the value is greater than the bid x. Accordingly, we introduce the following notation.

$$a_-(x,v) = \max(x - v, 0) \qquad \text{(overpayment)}$$

$$a_L(x,v) = \max(x - 1.5v, 0) \qquad \text{(loss)}$$

$$a_+(x,v) = \begin{cases} 0.5v & \textit{for} \quad v > x \\ 0 & \textit{for} \quad v \le x \end{cases} \qquad \text{(foregone payoff)}$$

Impulse balance theory concerns the expected values of these variables. The symbol E denotes the expectation operator. We use the following notation.

$$\mathrm{A}_-(x) = \mathrm{E}(a_-(x, v))$$

$$\mathrm{A}_\mathrm{L}(x) = \mathrm{E}(a_\mathrm{L}(x, v))$$

$$\mathrm{A}_+(x) = \mathrm{E}(a_+(x, v))$$

The impulse balance point $\tilde{x}$ is defined by

$$\mathrm{A}_-(\tilde{x}) + \mathrm{A}_\mathrm{L}(\tilde{x}) = \mathrm{A}_+(\tilde{x}) \qquad \text{(impulse balance equation)}$$

where $\tilde{x}$ is a real number. The bids can take only integer values, but this does not mean that we have to restrict the impulse balance point in the same way, since it has the meaning of a point in whose vicinity the stationary bid distribution is concentrated. For each of our conditions $u = 1$, $u = 11$, and $u = 21$ we obtain a uniquely determined impulse balance point. The values are shown in table 4.

It can be seen easily that the impulse balance point is in fact uniquely determined. The left side of the impulse balance equation is monotonically increasing in $\tilde{x}$ and the right hand side is monotonically decreasing in $\tilde{x}$. Moreover, the left hand side is smaller than the right hand side at $\tilde{x} = 0$ and greater than the right hand side at $\tilde{x} = 99$.

Table 4 also shows the average mean bids of adapters for the three treatments over all periods. It also shows the estimated standard error of the average mean bid counting each mean bid of an adapter as one observation. Note that this is the estimated standard error of the sample mean and not of the individual observation.

It can be seen that in all three conditions the deviation of the impulse balance point from the average mean bid of adapters is smaller in absolute value than one standard error. Obviously, impulse balance theory fits the data quite well.

Table 4. Comparison of impulse balance theory with the data

u	number of adapters	impulse balance point $\tilde{x}$	average mean bid of adapters	estimated standard error of mean
1	10	44.5	44.5	3.0
11	9	57.0	54.7	2.8
21	10	65.2	64.1	2.1

7. Monotonic Ladder Processes and Impulse Balance Theory

In this section, we want to present a very simple mathematical result which throws light on impulse balance theory. This result concerns special Markov chains which we call *monotonic ladder processes*. Consider a Markov chain with a state space 1..n with $n > 1$. We use the following notation.

p_i probability for the transition from i to $i + 1$ for $i = 1, ..., n–1$

q_i probability for the transition from $i + 1$ to i for $i = 2, ..., n$

The Markov chain is a *ladder process* if all p_i and q_i are positive and all transition probabilities from i to j with $*i – j* > 1$ are zero.

Figure 3 shows a graphical representation of a ladder process.

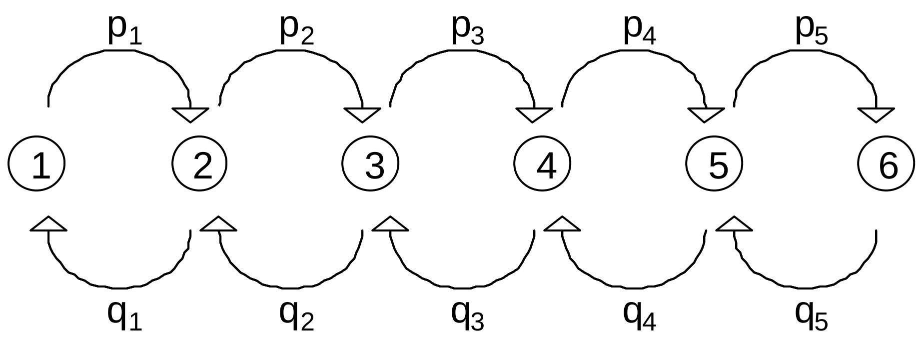

Figure 3

A ladder process is *monotonic* if we have

$$p_{i+1} < p_i \text{ and } q_{i+1} > q_i \text{ for } i = 1, ..., n - 1$$

Consider an experiment in which the subjects can move from one state to an neighbouring state only. We may for example think of a simplified version of the bidding experiment described in this paper, in which only the 21 bids 0, 5, ..., 100 are permitted, and bids can be changed only by 5. Suppose that the probabilities p_i and q_i are proportional to the corresponding expected impulses and have the properties required by the definition of a monotonic ladder process. For the simplified version of our experiment this would be the case. Then, the following definition adapts the impulse balance theory to the situation under consideration:

Definition: k is the left impulse balance point if $\frac{p_k}{q_k} \geq 1 \geq \frac{p_{k+1}}{q_{k+1}}$ holds.

Under the assumptions made above, the impulse balance equation would be exactly satisfied for $p_k / q_k = 1$. However, in general such a k does not exist. The left impulse balance point is the greatest k at which the expected upward impulse is at least as strong as the expected downward impulse. One could also define a right impulse balance point in the same way, but it turns out that the left impulse balance point is more closely connected to the stationary distribution of the process. This is shown by the following theorem.

Theorem: *Every monotonic ladder process has a unique left impulse balance point k. This k is a mode of a stationary distribution of the process, the only one unless* $p_{k+1} / q_{k+1} = 1$. *In this border case both k and k + 1 are modes.*

Proof: Let x_i be the stationary probability of state *i*. We must have

$$x_1 p_1 = x_2 q_1$$

since the outflow from state 1 must be equal to the inflow to state 1. We show by induction that

$$x_i p_i = x_{i+1} q_i$$

holds for $i = 1, ..., n$. We have seen that the statement is true for $i = 1$. Assume that it is true for $i = 2, ..., m$ with $m < n$. The equality of inflow and outflow at state $m + 1$ requires

$$x_m p_m + x_{m+2} q_m = x_{m+1} p_{m+1} + x_{m+1} q_m$$

or equivalently

$$x_{m+2} q_m - x_{m+1} q_m = x_{m+1} p_{m+1} - x_m p_m$$

By assumption the right hand side equals zero. This shows that the induction assumption also holds for $i = m + 1$, and therefore for $i = 1, ..., n$.

It follows that we have

$$x_{i+1} = \frac{p_i}{q_i} x_i$$

This together with the monotonicity properties of the p_i and q_i yields the conclusion that x_i increases until x_k is reached. For $p_{k+1} / q_{k+1} < 1$ it decreases from there on. In the exceptional case $p_{k+1} / q_{k+1} = 1$ we have $x_k = x_{k+1}$ and x_i decreases from there on. This completes the proof.

The theorem suggests that maybe also for more complex cases the impulse balance point is a prediction for the mode of the stationary distribution. However, the mode is a much less stable statistic than the average. Therefore, we compare the impulse balance point with average mean bids rather than average modes. If the stationary distribution is single-peaked and approximately symmetric, then there will be no great difference between the average and the mode. The average modes in the experiment are 44.5 for $u = 1$, and 57.5 for $u = 11$, and 67.6 for $u = 21$ (adapters only).

8. Summary and Conclusions

Our experiments clearly suggest an explanation of the winner's curse phenomenon in the simple SAMUELSON-BAZERMAN task by learning direction theory. The same explanation may also be valid in other contexts. As we have seen, the behaviour of the subjects conforms to the predictions of learning direction theory. However, it can be seen that for a sizeable proportion of the subject pool other influences on behaviour can be identified. Our categorisation of subjects by modal bids has shown that the adapters who presumably are mainly influenced by direction learning in fact conform more closely to learning direction theory.

Impulse balance theory is an attempt to make rough quantitative predictions on the basis of learning direction theory without fully specifying the learning behaviour. The theory determines an *impulse balance point* which is near to the average bids of the adapters for all three experimental conditions.

Finally, we have looked at a class of simple Markov chains, the *monotonic ladder processes*. It has been shown that for processes of this kind arising from situations to which impulse balance theory applies, the impulse balance point is connected to the mode of the stationary distribution.

Impulse balance theory, if it turns out to be of predictive value for other experimental situations, could be an interesting tool of behavioural analysis for many theoretically important problems.

References

ABBINK, Klaus, and Abdolkarim SADRIEH (1995): RatImage – Research Assistance Toolbox for Computer-Aided Human Behavior Experiments. SFB Discussion Paper B-325, University of Bonn.

AKERLOF, G.A. (1970): The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*. 84:3, 488-500.

BALL, S.B., H. BAZERMAN, and J.S. CARROLL (1991): An Evaluation of Learning in the Bilateral Winner's Curse. *Organizational Behavior and Human Decision Processes*, 48, 1-22.

BENARTZI, S., and R.H. THALER, (1995): Myopic Loss Aversion and the Equity Premium Puzzle. *Quarterly Journal of Economics*, 110:1, 73-92.

BERNINGHAUS, S., and K.-M. EHRHART (1998): Time Horizon and Equilibrium Selection in Tacit Coordination Games: Experimental Results. *Journal of Economic Behavior & Organization*, 37:2, 231-248.

BUSH, R.R., and R. MOSTELLER (1955): Stochastic Models of Learning. Wiley, New York.

CAPEN, E.C., R.V. CLAPP, and W.M. CAMPBELL (1971): Competitive Bidding in High-Risk Situations. *Journal of Petroleum Technology*, 23, 641-653.

CASON, T., and D. FRIEDMAN (1997): Price Formation in Single Call Markets. *Econometrica* 65:2, 311-345.

GROSSKOPF, B., I. EREV, and E. YECHIAM (2000): Foregone with the Wind. Mimeo, Universitat Pompeu Fabra.

KAGEL, J.H. (1995): Auctions. In J.H. KAGEL and A.E. ROTH: The Handbook of Experimental Economics. Princeton University Press.

KAGEL, J.H., and D. LEVIN (1999): Common Value Auctions with Insider Information. *Econometrica*, 67:5, 1219-1238.

KAHNEMAN, D., and A. TVERSKY (1979): Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47:2. 263-291.

KUON, B. (1993): Two-Person Bargaining with Incomplete Information. *Lecture Notes in Economics and Mathematical Systems*, 412. Springer-Verlag Berlin.

MITZKEWITZ, M., and R. NAGEL (1993): Envy, Greed, and Anticipation in Ultimatum Games with Incomplete Information. *International Journal of Game Theory*, 22, 171-198.

NAGEL, R. (1996): Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, 85:5, 1313-1326.

ROTH, A., and I. EREV (1995): Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior*, 8, 164-212.

RYLL, W. (1996): Litigation and Settlement in a Game with Incomplete Information. *Lecture Notes in Economics and Mathematical Systems*, 440. Springer-Verlag Berlin.

SADRIEH, A. (1998): The alternating double auction market: A game theoretic and experimental investigation. *Lecture Notes in Economics and Mathematical Systems*, 466. Springer-Verlag, Berlin.

SAMUELSON, W.F., and M.H. BAZERMAN (1985): The Winner's Curse in Bilateral Negotiations. In V.L. SMITH (ed.): Research in Experimental Economics. JAI Press, Greenwich.

SCHOLZ, R.W. (1987): Cognitive strategies in stochastic thinking. Reidel, Dordrecht.

SELTEN, R., and J. BUCHTA (1999); Experimenta Sealed-Bid First Price Auctions with Directly Observed Bid Functions. In D. BUDESCU, I. EREV, and R. ZWICK (eds.): Games and Human Behavior: Essays in the Honor of Amnon Rapoport. Lawrenz Associates Mahwah NJ.

SELTEN, R., and R. STOECKER (1986): End Behavior in Sequences of Finite Prisoners' Dilemma Supergames: A Learning Theory Approach. *Journal of Economic Behavior and Organization*, 7, 47-70.

Appendix: The Written Instructions

Duration

The experiment lasts for **100 rounds**.

Structure

The experiment is an **individual decision experiment**. The decisions of the participants to **not** mutually influence each other.

Decision

In each round you submit a **bid** out of the range from **0 and 99**. Then the Machine randomly draws a **value** from 1 to 99, where all values are equally likely. If your bid is **lower** than the value, you do not receive anything. If your bid is **equal to** or **greater than** the value, you receive **1.5 times the value minus your bid**

Lump sum payments

You receive an initial endowment of 250 talers. In the beginning of each further round a lump sum payment of 20 talers is credited to your talers account.

Exchange rate

100 talers correspond to **DM 1.50**.

Good success!