Discussion Papers in
Management

The Best Banking Strategy When Playing the Weakest Link

L.C. Thomas

Number M02-6
May 2002
ISSN 1356-3548

The Best Banking Strategy When Playing the Weakest Link

L.C.Thomas

School of Management

University of Southampton

Southampton U.K.

Abstract

The paper uses dynamic programming to investigate when contestants should bank their current winnings in the TV quiz show , the Weakest Link. It obtains the optimal strategy for the team as a whole and then looks at two possible reasons why the contestants tend to use other strategies in reality.

Keywords

Dynamic programming, recreation

The “Weakest Link” has proved to be a successful television game show in a number of countries. In each round the team of contestants are asked questions in turn and the amount the team can win goes up as the chain of correctly answered questions increases. If a question is answered incorrectly, then all the money built up in the chain is lost. Each contestant can decide to “bank” the money before hearing their next question and then the chain starts from 0 again. There is an upper limit on the amount that can be won in each round and the length of time of the round, (as well as the number of contestants) goes down from round to round.   The decision of which team member is voted off at each round – “the weakest link”- gives the game its name but there is little science, some psychology and lots of prejudice in that decision. The decision of when to bank though can be treated rationally though the contestants rarely seem to make such rational decisions.

In the next section the problem of when to bank is formulated as a dynamic programming problem and the optimal strategy identified. The formulation works for all the versions of the game but we give the results for the UK tea-time version of the game where the chain’s value goes £0, £20, £50, £100, £200, £300, £450, £600, £800 and £1000. We then examine the myopic strategy of deciding whether or not to bank by considering what maximises the value after one more question. This could be considered the strategy that is individually rational. Finally we examine what happens if a contestant is arrogant (or modest) and believes they are more (or less) likely to answer a question correctly than the other contestants.

OPTIMAL STRATEGY

Suppose the maximum you can earn in each round is M and the value of banking after a chain of i correct answers is c_i where $0=c_o < c_1 < c_2 < c_3 < c_4 < c_5 < c_6 < c_7 < c_8 < c_9 = M$. Assume that there are N questions available in the current round. The rounds are of a given time, and the length of time to ask and answer a question are closely distributed around 7.5 secs. Hence it is reasonable to assume that there are a known number of questions in each round. To solve the problem one needs to estimate p, the probability of the team answering each question . We assume this is the same for each question asked and that the team members have a good estimate for it when making their decisions about banking.

Thus at any point in a round, the state of the game is given by (n, i, b) where n is the number of questions still to be asked in that round, i is the number of questions correctly answered so far in the current chain and b is the amount banked in the round. The objective is to maximise the expected amount banked by the end of the round. Let that be V(n,i,b) if the current state is (n,i,b). Then V(n,i, b) satisfies the following optimality equation[1],

$$V(n,i,b) = \max\begin{cases} pV(n-1,1,\min(M,b+c_i)) + (1-p)V(n-1,0,\min(M,b+c_i)) : \mathit{Banking} \\ pV(n-1,i+1,b) + (1-p)V(n-1,0,b) : \mathit{Not}\ \mathrm{B}\mathit{anking} \end{cases} \quad (1)$$

where $V(0,i,b) = \min(M, b+c_i)$ for $0 \le c_i \le M$

The first term is the result of banking now so that $b+c_i$ is in the bank (provided it does not exceed M) and then the next question is correctly answered with probability p or incorrectly answered with probability 1-p. The second term is what happens if the there is no banking at this stage. With probability p the run increases to i+1 correctly answered questions while with probability 1-p the next answer is wrong and the chain reverts to 0 correctly answered questions.

Once can solve this equation using a spreadsheet and the results for the UK example are as follows

Example: UK where $c_o = 0$, $c_1=20$, $c_2=50$, $c_3=100$, $c_4=200$, $c_5=300$, $c_6=450$, $c_7=600$, $c_8=800$, $c_9 = M=1000$. N= 25 for (180 sec rounds) down to 15 for (110sec round)

The expected payoffs in the cases N=15,20 and 25 are given below as p varies.

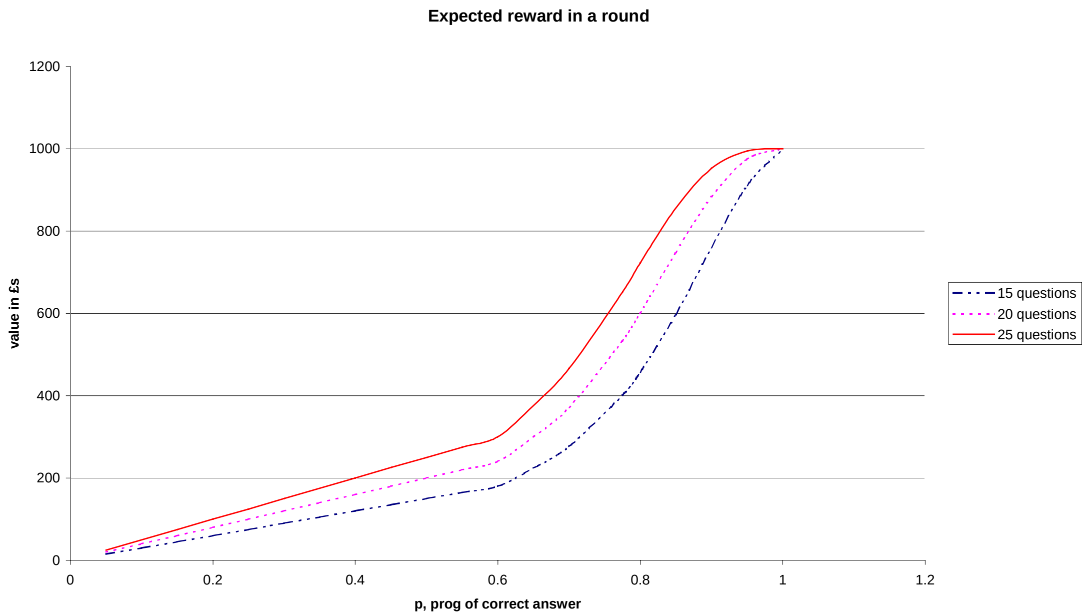
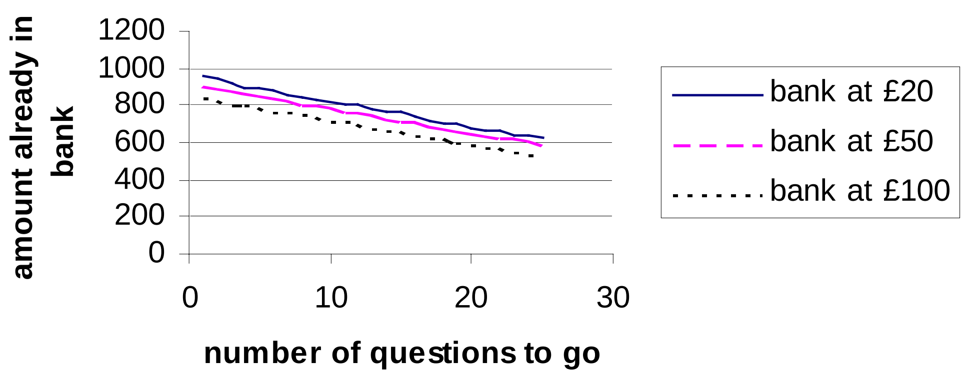


Figure 1: expected reward each round as function of p

The optimal strategy obviously varies depending on the probability, p, of getting the questions correct. If p=0.6 (and any lower value of p), the optimal strategy is to bank at £20 in a chain (i.e. after each correctly answered question) no matter how much is already banked or how many more questions to go. If p= 0.7, then one should always bank when the chain reaches £200 if there is £520 or less in the bank already. Figure 2 shows what happens thereafter. The dotted line is the boundary between banking at £200 and banking at £100 in the chain; the dashed line is the boundary between banking at £100 and £50 and the complete line is the boundary between banking at £50 and £20 in the chain. One starts banking at £100 if there is £530 in the bank and there are 25 questions to go. The amount needed in the bank before one banks at £100 increases as the number of questions left decreases. Thus one first banks at £100 with one question left if there is £840 or more in the bank already. As the amount already in the bank increases one will at some point bank when the chain reaches only £50. This occurs when there is £590 already in the bank if there are 25 questions to go but goes up to £910 when there is only 1 question to go. If the amount in the bank increases even more then one will eventually bank with only £20 in a chain,. The lower limits for this is £620 in the bank already if there are 25 questions to go and £960 if there is only 1 question to go.

Figure 2: Optimal strategy



This structure of banking earlier in the chain the more one has in the bank also occurs when p=0.8 but in this case the point at which one stops the chain varies from £800 to £20. One banks at £800 if one has £0 already in the bank and only 3 questions left while if one only has 1 question left one would bank at £800 if the current amount is £230 or less. With 25 questions to go, the following strategy is optimal. One would bank at £300 if one had £60 or less in the bank already; one would bank at £200 if one already had between £70 and £500 in the bank ; one would bank at £100 if between £510 and £580 in the bank; one would bank at £50 if one had between £590 and £630 and thereafter one would bank at £20.

One might expect the banking strategy to have some obvious monotone properties . Formally if the game is at stage (n,i,b,) (n questions to go i correct answers in the chain so far and b in the bank) one might expect that if one banks at (n,i,b) one might also want to bank at (n, i+1, b), so one gets more conservative the longer the chain. One might also think one banking at (n,i,b) means one wants to bank at (n+1,i,b) in that having an extra question to go should not make you want to take more risks. One might also expect that if one banks at (n,i,b) one would also want to bank at (n,i,b+a) so the more one has in the bank already the more conservatively one should play . In fact the first two conjectures are incorrect as the following two counter examples show

Example 1: Not conservative in i

With p=.6, then at state (1, 2,500) one banks and so expects to get 562, (i.e. 550 + 0.6(20) +0.4(0)) , rather than 560, (i.e.500 +0.6 (100)). At (1,3,500) one does not bank since one expects 620 (i.e.500 + 0.6(200)), rather than the 612 (i.e. 600 +0.6(20)) if one banks).

Example 2: Not conservative in n.

Again with p=0.6 as described above one should bank at (1, 2,500). At (2, 2,500) though one should not bank and the expected reward is 576.8 (500 + .36(200)+ .24(20) +.4(0)) as opposed to 574 (550 + .6(20) +.6(20)) if one banks.

SELFISHNESS AND THE MYOPIC POLICY

An alternative way of deciding when to bank is to work out whether or not to bank at this question assuming that the team would bank at the next question. This is not an optimal policy but usually does very well. It is termed a myopic policy because it is optimal in the shortsighted way of trying to maximise the reward over the next question. It is also the policy that a selfish contestant might use if they argue as follows. What is important is not to be picked out by the other contestants as the weakest link. Thus one does not want to lose the team large amounts. Hence the objective should be to maximise the expected gain to the team from the next question. So in both cases the analysis goes as follows.

If i questions have been correctly answered so far and p is the belief that one can answer the question correctly, then not banking now wil give an expected reward at the end of the question of $pc_{i+1} +(1-p)0$, while banking now will give $c_i +pc_1 +(1-p)0$. Thus one should bank if one believes $c_i +pc_1 > pc_{i+1}$ or

$$p < c_i /(c_{i+1}-c_1) \qquad (2)$$

Consider what this suggests in the following three examples

Example: UK early evening.

In this c_1=20, c_2=50, c_3=100, c_4=200, c_5=300, c_6=450, c_7=600, c_8=800, c_9 =M=1000. Thus one should bank at £20 if one believes $p<0.66$. The other limits are bank at £50 if $p<0.625$, bank at £100 if $p< 0.55$; bank at £200 if $p<0.714$; bank at £300 if $p<0.697$; bank at £450 if $p<0.775$; bank at £600 if $p<0.769$ and bank at £800 if $p<0.816$. So in fact one should only bank at £20 (if one believes $p<0.66$), at £200 (if one believes $0.66<p<0.714$), at £450 (if one believes $0.714<p<0.775$) or £800 (if one believes $0,775<p<0,816$). This is very different behaviour from that displayed by most of the contestants who tend to bank at £50 and £100.

Example: UK late evening.

In this case c_1=100, c_2=250, c_3=500, c_4=2000, c_5=2500, c_6=2250, c_7=3000, c_8=4000, c_9 =M=5000. Since all these levels are five times the early evening equivalent, the results are exactly the same and contestant should only bank at £100, £1000, £2250 or £4000 depending on their beliefs about getting a question correct.

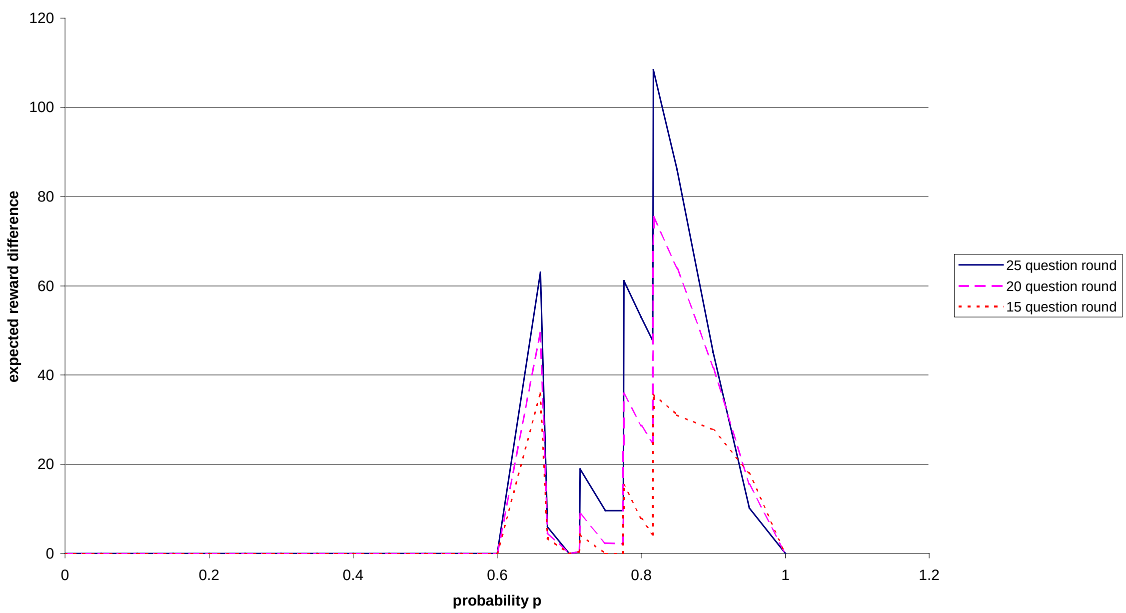
Example US

In this case c_1=\$1000, c_2=\$2500, c_3=\$5000, c_4=\$10000, c_5=\$25000, c_6=\$50000, c_7=\$75000, c_8 =M=\$125000.

In this case one should only bank at \$1000 (if one believes $p<0.66$) or \$50000 (if one believes $0.666<p<0.6756$). Otherwise one goes all the way to \$125000.

The question arises of how good is this strategy? One can calculate the value under this strategy by using equation 1 but where we use the banking policy specified rather than the banking decision that gives the maximum value (i.e. if in the UK example if we believe p=0.7 then we bank at £200 and do no bank before). The difference between the expected reward under the optimal policy and this myopic policy is shown in Figure 2 for the case of 15, 20 and 25 question rounds and all the various choices of probability p. For low p, the myopic policy of banking at £20 is also the optimal policy and there is no difference in expected total rewards. The major differences occur when the myopic policy changes from one banking level, so for example at p=0.816 it suggests banking at £800 while at p=0.817 it suggests banking at £1000.

Fig 3: Difference between optimal and myopic policies



ARROGANCE

It is clear to any addict of the show that the contestants rarely follow the above banking strategies. There may be many reasons for this, including that they have little idea of what percentage of questions are being correctly answered, but one other reason springs to mind. A contestant might believe they have a different probability of answering questions correctly than the other team members. This arrogance (or modesty) might lead them to a different banking strategy.

Given that the state of the game is (n, i, b) and one believes the team answers each question correctly with probability p, what is the probability of answering the next question correctly that makes one decide not to bank. We call this the arrogance probability a(n,i,b,p), and it is the probability value which means it is equally valid to bank or not bank. If one believes one's chance of getting the question correct is less than this value one should bank and if one believes it is more than this value then one should not bank. Assuming V(n,i,b,p) is the solution to equation(1) with probability p used, then a(n,i,b,p) satisfies

$$\begin{aligned} &\mathrm{a}(n,i,b,p)V(n-1,i+1,b,p)+(1-\mathrm{a}(n,i,b,p))V(n-1,0,b,p)= \\ &\mathrm{a}(\mathrm{n},\mathrm{i},\mathrm{b},\mathrm{p})\mathrm{V}(\mathrm{n}-1,1,\mathrm{b}+\mathrm{c}_{\mathrm{i}},p)+(1-\mathrm{a}(n,i,b,p))V(n-1,0,b+c_i,p)) \quad (3) \end{aligned}$$

Then

$$\mathrm{a}(n,i,b,p)=\frac{V(n-1,0,b+c_i,p)-V(n-1,0,b,p)}{V(n-1,i+1,b,p)+V(n-1,0,b+c_i,p)-V(n-1,0,b,p)-V(n-1,1,b+c_i,p)}$$

For p=0.7 and n=15, Table 1 gives the values of a(15,i,b,0.7) as i and b vary, (na means it is not appropriate to have a value in this case as the total one can get in the round already exceeds M).

Current value of run (c_i)	Amount already banked (b)						
	0	150	300	450	600	750	900
20	0.515	0.515	0.519	0.547	0.568	0.689	0.992
50	0.532	0.532	0.536	0.567	0.604	0.852	1.000
100	0.578	0.578	0.580	0.602	0.626	0.988	na
200	0.733	0.737	0.753	0.783	0.954	1.000	na
300	0.712	0.725	0.762	0.888	1.000	na	na
450	0.798	0.829	0.927	1.000	na	na	na
600	0.831	0.946	1.000	na	na	na	na
800	0.991	1.000	na	na	na	na	na

Table 1: value of a(15, i, b, 0.7)

The table shows that if the current run is only at £20 then someone using this strategy is willing to risk trying the question even if they feel they only have a slightly over 50% chance of getting it correct. On the other hand if the current run is worth £450 they will only risk going for the question if they are 80% or more sure of getting it correct. Essentially they can be modest and go for the question without banking even if they believe their chance of getting it right is below the other contestant's average of 0.7 if there is £100 or less in the current run. If there is more than this available in the current run, they should be arrogant and only risk not banking if they believe they are much more likely to get it right than the others. This form of the result holds for almost all the choices of n and p that are likely to arise.

CONCLUSIONS

Dynamic programming analysis shows that the optimal banking strategy is quite different from that displayed by the contestants. Whether this difference is due to contestants' selfishness, arrogance or inability to assess the chances of answering questions correctly is open to debate.

REFERENCES

1. Puterman M.L. (1994), Markov Decision Processes, Wiley, New York