In the case of I must lose

MC Method:

Pbar = 8

P0 = 10;

D0 = 5;

DelP = 1; pu = 1/2; pd=1-pu;

DelD = 1; p =1/2;Pover = 9;

N = 5000

‘q0 = 1, Q0 = --9.9461

‘q1 = 11 6 1 1 2 0.250000000000000 -16.9484000000000

‘q1 = 11 4 1 1 2 0.250000000000000 -10.9356000000000

‘q1 = 9 6 1 1 2 0.250000000000000 -6.94960000000000

‘q1 = 9 4 1 1 2 0.250000000000000 -4.95080000000000

Q-learning:

N = 1000

‘q2 = [[ 0., 0., 10., 5., 0.]]])

Q2 = [ 0. , 0. , 10. , 5. ,

0. , -9.66025485],

‘q3=[[[ 0., 0., 11., 6., 0.],

[ 0., 0., 11., 4., 0.],

[ 0., 0., 9., 6., 0.],

[ 0., 0., 9., 4., 0.],

Q3 = [ 0. , 0. , 9. , 4. ,

0. , -4.002002 ],

[ 0. , 0. , 9. , 4. ,

1. , -4.00700701],

[[ 0. , 0. , 11. , 6. ,

1. , -17.98998999],

[[ 0. , 0. , 9. , 6. ,

0. , -5.99199199],

[ 0. , 0. , 9. , 6. ,

1. , -5.996997 ],

[[ 0. , 0. , 11. , 4. ,

0. , -12. ],

[ 0. , 0. , 11. , 4. ,

1. , -12.00500501],

[ 0. , 0. , 11. , 4. ,

2. , -12.01001001],