```
std:[[1.72, 1.7, 1.3]][[0.67, 0.0, 0.3]]
MSE:[[7.47, 7.46, 7.73]][[13.32, 58.7, 5.88]]
MSE(-DR):[[0.0, -0.01, 0.26]][[5.85, 51.23, -1.59]]
***
==============
******************* THIS SETTING IS GOOD *******************
[[ 0.65  0.68  1.17   1.35 58.15  5.29]
 [ 2.27  2.32  3.1    4.15 57.71  4.85]
 [ 2.18  2.21  2.96   5.28 56.7   3.84]
 [ 7.21  7.18  7.56  13.2  58.65  5.78]]


[[ 0.77  0.82  1.32   1.4  58.18  5.33]
 [ 2.15  2.23  3.24   4.25 57.73  4.89]
 [ 2.15  2.2   3.07   5.34 56.72  3.89]
 [ 7.31  7.3   7.6   13.26 58.67  5.83]]


[[ 1.34  1.39  1.58   1.53 58.2   5.39]
 [ 2.42  2.52  3.42   4.39 57.76  4.94]
 [ 2.35  2.4   3.21   5.43 56.74  3.93]
 [ 7.47  7.46  7.73  13.32 58.7   5.88]]


time spent until now: 117.2 mins

19:37, 04/12
ubuntu@ip-172-31-9-175:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
19:47, 04/12; num of cores:16
median_u_O_u_D_other_C

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, t_func] = [16, None, None, 20, 0.5, 1.5, 10, None]

[thre_range, sd_R_range, day_range, penalty_range]:  [[90, 95, 120], [0, 15, 30], [7], [[0.0001, 5e-05], [0.0001, 5e-05]]]

----------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 0]

max(u_O) =  145.8
O_threshold = 90
number of reward locations:  14
O_threshold = 95
number of reward locations:  12
O_threshold = 120
number of reward locations:  3
target 1 in 3 DONE!
target 2 in 3 DONE!
target 3 in 3 DONE!

----------------------------------------
Value of Behaviour policy:52.865
O_threshold = 90
MC for this TARGET:[62.904, 0.077]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.61, 1.33, -0.48]][[1.91, -62.9, -10.04]]
std:[[0.3, 0.3, 0.36]][[0.3, 0.0, 0.23]]
MSE:[[1.64, 1.36, 0.6]][[1.93, 62.9, 10.04]]
MSE(-DR):[[0.0, -0.28, -1.04]][[0.29, 61.26, 8.4]]
**
==============
O_threshold = 95
MC for this TARGET:[59.142, 0.082]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[2.59, 2.4, 1.12]][[2.48, -59.14, -6.28]]
std:[[0.48, 0.47, 0.38]][[0.27, 0.0, 0.23]]
MSE:[[2.63, 2.45, 1.18]][[2.49, 59.14, 6.28]]
MSE(-DR):[[0.0, -0.18, -1.45]][[-0.14, 56.51, 3.65]]
==============
O_threshold = 120
MC for this TARGET:[58.647, 0.054]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.15, -7.14, -7.54]][[-13.2, -58.65, -5.78]]
std:[[0.74, 0.74, 0.48]][[0.32, 0.0, 0.23]]
MSE:[[7.19, 7.18, 7.56]][[13.2, 58.65, 5.78]]
MSE(-DR):[[0.0, -0.01, 0.37]][[6.01, 51.46, -1.41]]
***
==============
[[ 1.64  1.36  0.6    1.93 62.9  10.04]
 [ 2.63  2.45  1.18   2.49 59.14  6.28]
 [ 7.19  7.18  7.56  13.2  58.65  5.78]]


time spent until now: 29.2 mins

20:16, 04/12

----------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 15]
```

```
max(u_O) =  145.8
O_threshold = 90
number of reward locations:  14
O_threshold = 95
number of reward locations:  12
O_threshold = 120
number of reward locations:   3
target 1 in 3 DONE!
target 2 in 3 DONE!
target 3 in 3 DONE!


----------------------------------------
Value of Behaviour policy:52.843
O_threshold = 90
MC for this TARGET:[62.928, 0.188]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.85, 1.6, -0.3]][[2.0, -62.93, -10.08]]
std:[[0.5, 0.5, 0.45]][[0.42, 0.0, 0.21]]
MSE:[[1.92, 1.68, 0.54]][[2.04, 62.93, 10.08]]
MSE(-DR):[[0.0, -0.24, -1.38]][[0.12, 61.01, 8.16]]
**
==============
O_threshold = 95
MC for this TARGET:[59.165, 0.186]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[2.77, 2.59, 1.24]][[2.54, -59.16, -6.32]]
std:[[0.51, 0.5, 0.44]][[0.35, 0.0, 0.21]]
MSE:[[2.82, 2.64, 1.32]][[2.56, 59.16, 6.32]]
MSE(-DR):[[0.0, -0.18, -1.5]][[-0.26, 56.34, 3.5]]
==============
O_threshold = 120
MC for this TARGET:[58.671, 0.183]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.22, -7.23, -7.56]][[-13.25, -58.67, -5.83]]
std:[[1.01, 0.98, 0.69]][[0.38, 0.0, 0.21]]
MSE:[[7.29, 7.3, 7.59]][[13.26, 58.67, 5.83]]
MSE(-DR):[[0.0, 0.01, 0.3]][[5.97, 51.38, -1.46]]
***
==============
[[ 1.64  1.36  0.6   1.93 62.9  10.04]
 [ 2.63  2.45  1.18  2.49 59.14  6.28]
 [ 7.19  7.18  7.56 13.2  58.65  5.78]]


[[ 1.92  1.68  0.54  2.04 62.93 10.08]
 [ 2.82  2.64  1.32  2.56 59.16  6.32]
 [ 7.29  7.3   7.59 13.26 58.67  5.83]]


time spent until now: 58.5 mins

20:45, 04/12

----------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 30]

max(u_O) =  145.8
O_threshold = 90
number of reward locations:  14
O_threshold = 95
number of reward locations:  12
O_threshold = 120
number of reward locations:   3
target 1 in 3 DONE!
target 2 in 3 DONE!
target 3 in 3 DONE!


----------------------------------------
Value of Behaviour policy:52.822
O_threshold = 90
MC for this TARGET:[62.951, 0.354]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[2.15, 1.88, -0.1]][[2.08, -62.95, -10.13]]
std:[[0.86, 0.9, 0.69]][[0.71, 0.0, 0.3]]
MSE:[[2.32, 2.08, 0.7]][[2.2, 62.95, 10.13]]
MSE(-DR):[[0.0, -0.24, -1.62]][[-0.12, 60.63, 7.81]]
==============
O_threshold = 95
MC for this TARGET:[59.189, 0.349]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[2.99, 2.8, 1.38]][[2.6, -59.19, -6.37]]
std:[[0.99, 0.91, 0.73]][[0.61, 0.0, 0.3]]
MSE:[[3.15, 2.94, 1.56]][[2.67, 59.19, 6.38]]
MSE(-DR):[[0.0, -0.21, -1.59]][[-0.48, 56.04, 3.23]]
==============
O_threshold = 120
MC for this TARGET:[58.695, 0.353]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.25, -7.26, -7.62]][[-13.3, -58.7, -5.87]]
```