```
22:36, 04/02; num of cores:96
[377/1864]

one attempt for final result with SD_R trend
Basic setting:[T, rep_times, sd_O, sd_D, sd_R, sd_u_O, w_O, w_A,
[M_in_R, mean_reversion, poisO, simple, u_O_u_D]] = [None, 96, 10, 10,
None, 0.3, 0.5, 1, [True, False, True, False, 10]]


-------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 0]

max(u_O) =  197.9
O_threshold = 80
number of reward locations:  15
O_threshold = 90
number of reward locations:  12
O_threshold = 100
number of reward locations:  8
O_threshold = 110
number of reward locations:  6
1 -th region DONE with time cost 0.0 mins
6 -th region DONE with time cost 0.0 mins
11 -th region DONE with time cost 0.0 mins
16 -th region DONE with time cost 0.0 mins
21 -th region DONE with time cost 0.0 mins
1 -th region DONE with time cost 0.0 mins
6 -th region DONE with time cost 0.0 mins
11 -th region DONE with time cost 0.0 mins
16 -th region DONE with time cost 0.0 mins
21 -th region DONE with time cost 0.0 mins
1 -th region DONE with time cost 0.0 mins
6 -th region DONE with time cost 0.0 mins
11 -th region DONE with time cost 0.0 mins
16 -th region DONE with time cost 0.0 mins
21 -th region DONE with time cost 0.0 mins
1 -th region DONE with time cost 0.0 mins
6 -th region DONE with time cost 0.0 mins
11 -th region DONE with time cost 0.0 mins
16 -th region DONE with time cost 0.0 mins
21 -th region DONE with time cost 0.0 mins

-------------------------------------
Value of Behaviour policy:60.795
O_threshold = 80
MC for this TARGET:[70.898, 0.049]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.83, 0.61, -0.68]][[1.49, -402888.91, -10.1]]
std:[[0.3, 0.31, 0.17]][[0.19, 4433309.11, 0.12]]
MSE:[[0.88, 0.68, 0.7]][[1.5, 4451578.28, 10.1]]
```

```
MSE(-DR):[[0.0, -0.2, -0.18]][[0.62, 4451577.4, 9.22]]
**
==============
O_threshold = 90
MC for this TARGET:[69.381, 0.055]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.15, -0.03, -0.74]][[-0.34, -6142.2, -8.59]]
std:[[0.26, 0.26, 0.17]][[0.18, 368095.17, 0.12]]
MSE:[[0.3, 0.26, 0.76]][[0.38, 368146.41, 8.59]]
MSE(-DR):[[0.0, -0.04, 0.46]][[0.08, 368146.11, 8.29]]
***
MC-based ATE = -1.52
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.68, -0.64, -0.06]][[-1.82, 396746.7, 1.52]]
std:[[0.29, 0.3, 0.17]][[0.07, 4421357.57, 0.0]]
MSE:[[0.74, 0.71, 0.18]][[1.82, 4439122.74, 1.52]]
MSE(-DR):[[0.0, -0.03, -0.56]][[1.08, 4439122.0, 0.78]]
==============
O_threshold = 100
MC for this TARGET:[68.929, 0.054]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.88, -3.01, -3.17]][[-4.6, 44582.66, -8.13]]
std:[[0.29, 0.3, 0.16]][[0.19, 341239.45, 0.12]]
MSE:[[2.89, 3.02, 3.17]][[4.6, 344139.47, 8.13]]
MSE(-DR):[[0.0, 0.13, 0.28]][[1.71, 344136.58, 5.24]]
***
MC-based ATE = -1.97
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.71, -3.62, -2.49]][[-6.09, 447471.57, 1.97]]
std:[[0.35, 0.36, 0.19]][[0.13, 4467667.26, 0.0]]
MSE:[[3.73, 3.64, 2.5]][[6.09, 4490020.22, 1.97]]
MSE(-DR):[[0.0, -0.09, -1.23]][[2.36, 4490016.49, -1.76]]
==============
O_threshold = 110
MC for this TARGET:[70.47, 0.052]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.63, -6.71, -6.62]][[-8.62, -75989.81, -9.68]]
std:[[0.38, 0.39, 0.22]][[0.17, 784795.7, 0.12]]
MSE:[[6.64, 6.72, 6.62]][[8.62, 788466.07, 9.68]]
MSE(-DR):[[0.0, 0.08, -0.02]][[1.98, 788459.43, 3.04]]
**
MC-based ATE = -0.43
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.46, -7.32, -5.94]][[-10.11, 326899.09, 0.43]]
std:[[0.43, 0.44, 0.25]][[0.14, 4499885.04, 0.0]]
MSE:[[7.47, 7.33, 5.95]][[10.11, 4511743.39, 0.43]]
MSE(-DR):[[0.0, -0.14, -1.52]][[2.64, 4511735.92, -7.04]]
==============
[[8.8000e-01 6.8000e-01 7.0000e-01 1.5000e+00 4.4516e+06 1.0100e+01]
 [3.0000e-01 2.6000e-01 7.6000e-01 3.8000e-01 3.6815e+05 8.5900e+00]
```

```
 [2.8900e+00 3.0200e+00 3.1700e+00 4.6000e+00 3.4414e+05 8.1300e+00]
 [6.6400e+00 6.7200e+00 6.6200e+00 8.6200e+00 7.8847e+05 9.6800e+00]]


time spent until now: 173.4 mins




--------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 5]
--------------------------------------
[240/1864]
Value of Behaviour policy:60.792
O_threshold = 80
MC for this TARGET:[70.896, 0.064]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.82, 0.6, -0.69]][[1.48, -821167.61, -10.1]]
std:[[0.39, 0.4, 0.21]][[0.21, 8264001.62, 0.12]]
MSE:[[0.91, 0.72, 0.72]][[1.49, 8304699.82, 10.1]]
MSE(-DR):[[0.0, -0.19, -0.19]][[0.58, 8304698.91, 9.19]]
**
==============
O_threshold = 90
MC for this TARGET:[69.379, 0.07]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.14, -0.03, -0.73]][[-0.34, -13404.44, -8.59]]
std:[[0.31, 0.31, 0.21]][[0.2, 317190.25, 0.12]]
MSE:[[0.34, 0.31, 0.76]][[0.39, 317473.36, 8.59]]
MSE(-DR):[[0.0, -0.03, 0.42]][[0.05, 317473.02, 8.25]]
***
MC-based ATE = -1.52
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.68, -0.64, -0.04]][[-1.82, 807763.17, 1.52]]
std:[[0.35, 0.36, 0.17]][[0.08, 8264967.16, 0.0]]
MSE:[[0.76, 0.73, 0.17]][[1.82, 8304346.06, 1.52]]
MSE(-DR):[[0.0, -0.03, -0.59]][[1.06, 8304345.3, 0.76]]
==============
O_threshold = 100
MC for this TARGET:[68.927, 0.065]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.86, -2.98, -3.18]][[-4.6, 46389.1, -8.14]]
std:[[0.36, 0.36, 0.18]][[0.21, 291322.38, 0.12]]
MSE:[[2.88, 3.0, 3.19]][[4.6, 294992.67, 8.14]]
MSE(-DR):[[0.0, 0.12, 0.31]][[1.72, 294989.79, 5.26]]
***
MC-based ATE = -1.97
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.68, -3.59, -2.49]][[-6.09, 867556.71, 1.97]]
std:[[0.39, 0.41, 0.19]][[0.13, 8329347.43, 0.0]]
```

```
MSE:[[3.7, 3.61, 2.5]][[6.09, 8374406.44, 1.97]]
MSE(-DR):[[0.0, -0.09, -1.2]][[2.39, 8374402.74, -1.73]]
==============
O_threshold = 110
MC for this TARGET:[70.468, 0.059]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.6, -6.68, -6.61]][[-8.62, -76304.45, -9.68]]
std:[[0.43, 0.45, 0.25]][[0.19, 773969.3, 0.12]]
MSE:[[6.61, 6.7, 6.61]][[8.62, 777721.57, 9.68]]
MSE(-DR):[[0.0, 0.09, 0.0]][[2.01, 777714.96, 3.07]]
***
MC-based ATE = -0.43
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.42, -7.28, -5.93]][[-10.11, 744863.15, 0.43]]
std:[[0.5, 0.52, 0.29]][[0.15, 8306070.46, 0.0]]
MSE:[[7.44, 7.3, 5.94]][[10.11, 8339402.11, 0.43]]
MSE(-DR):[[0.0, -0.14, -1.5]][[2.67, 8339394.67, -7.01]]
==============
[[8.8000e-01 6.8000e-01 7.0000e-01 1.5000e+00 4.4516e+06 1.0100e+01]
 [3.0000e-01 2.6000e-01 7.6000e-01 3.8000e-01 3.6815e+05 8.5900e+00]
 [2.8900e+00 3.0200e+00 3.1700e+00 4.6000e+00 3.4414e+05 8.1300e+00]
 [6.6400e+00 6.7200e+00 6.6200e+00 8.6200e+00 7.8847e+05 9.6800e+00]]


[[9.1000e-01 7.2000e-01 7.2000e-01 1.4900e+00 8.3047e+06 1.0100e+01]
 [3.4000e-01 3.1000e-01 7.6000e-01 3.9000e-01 3.1747e+05 8.5900e+00]
 [2.8800e+00 3.0000e+00 3.1900e+00 4.6000e+00 2.9499e+05 8.1400e+00]
 [6.6100e+00 6.7000e+00 6.6100e+00 8.6200e+00 7.7772e+05 9.6800e+00]]


time spent until now: 347.1 mins

---------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 10]

---------------------------------------
Value of Behaviour policy:60.789
O_threshold = 80
MC for this TARGET:[70.894, 0.091]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.81, 0.59, -0.69]][[1.48, -701236.44, -10.11]]
std:[[0.52, 0.53, 0.29]][[0.25, 7210245.9, 0.14]]
MSE:[[0.96, 0.79, 0.75]][[1.5, 7244265.21, 10.11]]
MSE(-DR):[[0.0, -0.17, -0.21]][[0.54, 7244264.25, 9.15]]
**
==============
O_threshold = 90
MC for this TARGET:[69.377, 0.097]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.13, -0.03, -0.74]][[-0.35, -25526.03, -8.59]]
```

```
std:[[0.43, 0.43, 0.28]][[0.25, 321127.75, 0.14]]
MSE:[[0.45, 0.43, 0.79]][[0.43, 322140.67, 8.59]]
MSE(-DR):[[0.0, -0.02, 0.34]][[-0.02, 322140.22, 8.14]]
MC-based ATE = -1.52
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.68, -0.63, -0.06]][[-1.83, 675710.41, 1.52]]
std:[[0.45, 0.47, 0.22]][[0.09, 7208241.18, 0.0]]
MSE:[[0.82, 0.79, 0.23]][[1.83, 7239842.92, 1.52]]
MSE(-DR):[[0.0, -0.03, -0.59]][[1.01, 7239842.1, 0.7]]
==============
O_threshold = 100
MC for this TARGET:[68.925, 0.09]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.84, -2.96, -3.17]][[-4.61, 59315.0, -8.14]]
std:[[0.49, 0.49, 0.25]][[0.26, 415692.53, 0.14]]
MSE:[[2.88, 3.0, 3.18]][[4.62, 419903.02, 8.14]]
MSE(-DR):[[0.0, 0.12, 0.3]][[1.74, 419900.14, 5.26]]
***
MC-based ATE = -1.97
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.66, -3.55, -2.49]][[-6.08, 760551.44, 1.97]]
std:[[0.53, 0.54, 0.26]][[0.15, 7301984.57, 0.0]]
MSE:[[3.7, 3.59, 2.5]][[6.08, 7341486.03, 1.97]]
MSE(-DR):[[0.0, -0.11, -1.2]][[2.38, 7341482.33, -1.73]]
==============
O_threshold = 110
MC for this TARGET:[70.467, 0.083]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.58, -6.66, -6.6]][[-8.63, -53894.02, -9.68]]
std:[[0.57, 0.59, 0.33]][[0.24, 682018.31, 0.14]]
MSE:[[6.6, 6.69, 6.61]][[8.63, 684144.39, 9.68]]
MSE(-DR):[[0.0, 0.09, 0.01]][[2.03, 684137.79, 3.08]]
***
MC-based ATE = -0.43
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.39, -7.25, -5.91]][[-10.11, 647342.42, 0.43]]
std:[[0.67, 0.7, 0.38]][[0.17, 7242293.64, 0.0]]
MSE:[[7.42, 7.28, 5.92]][[10.11, 7271166.99, 0.43]]
MSE(-DR):[[0.0, -0.14, -1.5]][[2.69, 7271159.57, -6.99]]
==============
[[8.8000e-01 6.8000e-01 7.0000e-01 1.5000e+00 4.4516e+06 1.0100e+01]
 [3.0000e-01 2.6000e-01 7.6000e-01 3.8000e-01 3.6815e+05 8.5900e+00]
 [2.8900e+00 3.0200e+00 3.1700e+00 4.6000e+00 3.4414e+05 8.1300e+00]
 [6.6400e+00 6.7200e+00 6.6200e+00 8.6200e+00 7.8847e+05 9.6800e+00]]


[[9.1000e-01 7.2000e-01 7.2000e-01 1.4900e+00 8.3047e+06 1.0100e+01]
 [3.4000e-01 3.1000e-01 7.6000e-01 3.9000e-01 3.1747e+05 8.5900e+00]
 [2.8800e+00 3.0000e+00 3.1900e+00 4.6000e+00 2.9499e+05 8.1400e+00]
 [6.6100e+00 6.7000e+00 6.6100e+00 8.6200e+00 7.7772e+05 9.6800e+00]]
```

```
[[9.6000e-01 7.9000e-01 7.5000e-01 1.5000e+00 7.2443e+06 1.0110e+01]
 [4.5000e-01 4.3000e-01 7.9000e-01 4.3000e-01 3.2214e+05 8.5900e+00]
 [2.8800e+00 3.0000e+00 3.1800e+00 4.6200e+00 4.1990e+05 8.1400e+00]
 [6.6000e+00 6.6900e+00 6.6100e+00 8.6300e+00 6.8414e+05 9.6800e+00]]


--------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 20]

max(u_O) =  197.9

--------------------------------------
Value of Behaviour policy:60.783
O_threshold = 80
MC for this TARGET:[70.89, 0.157]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.78, 0.57, -0.7]][[1.47, -510451.1, -10.11]]
std:[[0.85, 0.86, 0.5]][[0.36, 5361129.22, 0.19]]
MSE:[[1.15, 1.03, 0.86]][[1.51, 5385375.27, 10.11]]
MSE(-DR):[[0.0, -0.12, -0.29]][[0.36, 5385374.12, 8.96]]
**
==============
O_threshold = 90
MC for this TARGET:[69.373, 0.161]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.12, -0.05, -0.74]][[-0.36, -39657.59, -8.59]]
std:[[0.73, 0.72, 0.48]][[0.38, 443462.55, 0.19]]
MSE:[[0.74, 0.72, 0.88]][[0.52, 445232.25, 8.59]]
MSE(-DR):[[0.0, -0.02, 0.14]][[-0.22, 445231.51, 7.85]]
MC-based ATE = -1.52
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.67, -0.62, -0.04]][[-1.83, 470793.52, 1.52]]
std:[[0.73, 0.74, 0.38]][[0.14, 5352228.12, 0.0]]
MSE:[[0.99, 0.97, 0.38]][[1.84, 5372894.23, 1.52]]
MSE(-DR):[[0.0, -0.02, -0.61]][[0.85, 5372893.24, 0.53]]
==============
O_threshold = 100
MC for this TARGET:[68.921, 0.154]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.81, -2.93, -3.18]][[-4.61, 69371.52, -8.14]]
std:[[0.78, 0.78, 0.41]][[0.4, 425865.6, 0.19]]
MSE:[[2.92, 3.03, 3.21]][[4.63, 431478.76, 8.14]]
MSE(-DR):[[0.0, 0.11, 0.29]][[1.71, 431475.84, 5.22]]
***
MC-based ATE = -1.97
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.59, -3.5, -2.48]][[-6.08, 579822.63, 1.97]]
std:[[0.88, 0.9, 0.45]][[0.22, 5425690.6, 0.0]]
```

```
MSE:[[3.7, 3.61, 2.52]][[6.08, 5456584.35, 1.97]]
MSE(-DR):[[0.0, -0.09, -1.18]][[2.38, 5456580.65, -1.73]]
==============
O_threshold = 110
MC for this TARGET:[70.463, 0.146]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.55, -6.62, -6.58]][[-8.64, -37262.62, -9.68]]
std:[[0.94, 0.95, 0.55]][[0.37, 720562.35, 0.19]]
MSE:[[6.62, 6.69, 6.6]][[8.65, 721525.19, 9.68]]
MSE(-DR):[[0.0, 0.07, -0.02]][[2.03, 721518.57, 3.06]]
**
MC-based ATE = -0.43
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.33, -7.2, -5.88]][[-10.12, 473188.48, 0.43]]
std:[[1.15, 1.17, 0.64]][[0.26, 5404456.53, 0.0]]
MSE:[[7.42, 7.29, 5.91]][[10.12, 5425132.05, 0.43]]
MSE(-DR):[[0.0, -0.13, -1.51]][[2.7, 5425124.63, -6.99]]
==============

==============
[[8.8000e-01 6.8000e-01 7.0000e-01 1.5000e+00 4.4516e+06 1.0100e+01]
 [3.0000e-01 2.6000e-01 7.6000e-01 3.8000e-01 3.6815e+05 8.5900e+00]
 [2.8900e+00 3.0200e+00 3.1700e+00 4.6000e+00 3.4414e+05 8.1300e+00]
 [6.6400e+00 6.7200e+00 6.6200e+00 8.6200e+00 7.8847e+05 9.6800e+00]]


[[9.1000e-01 7.2000e-01 7.2000e-01 1.4900e+00 8.3047e+06 1.0100e+01]
 [3.4000e-01 3.1000e-01 7.6000e-01 3.9000e-01 3.1747e+05 8.5900e+00]
 [2.8800e+00 3.0000e+00 3.1900e+00 4.6000e+00 2.9499e+05 8.1400e+00]
 [6.6100e+00 6.7000e+00 6.6100e+00 8.6200e+00 7.7772e+05 9.6800e+00]]


[[9.6000e-01 7.9000e-01 7.5000e-01 1.5000e+00 7.2443e+06 1.0110e+01]
 [4.5000e-01 4.3000e-01 7.9000e-01 4.3000e-01 3.2214e+05 8.5900e+00]
 [2.8800e+00 3.0000e+00 3.1800e+00 4.6200e+00 4.1990e+05 8.1400e+00]
 [6.6000e+00 6.6900e+00 6.6100e+00 8.6300e+00 6.8414e+05 9.6800e+00]]


[[1.1500e+00 1.0300e+00 8.6000e-01 1.5100e+00 5.3854e+06 1.0110e+01]
 [7.4000e-01 7.2000e-01 8.8000e-01 5.2000e-01 4.4523e+05 8.5900e+00]
 [2.9200e+00 3.0300e+00 3.2100e+00 4.6300e+00 4.3148e+05 8.1400e+00]
 [6.6200e+00 6.6900e+00 6.6000e+00 8.6500e+00 7.2153e+05 9.6800e+00]]
```