```
Last login: Mon Mar 30 10:20:44 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-3-215-134-165.compute-1.amazonaws.com
Welcome to Ubuntu 18.04.3 LTS (GNU/Linux 4.15.0-1060-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

  System information as of Mon Mar 30 14:50:10 UTC 2020

  System load:  12.3               Processes:           211
  Usage of /:   56.9% of 15.45GB   Users logged in:     0
  Memory usage: 1%                 IP address for ens5: 172.31.9.80
  Swap usage:   0%

 * Kubernetes 1.18 GA is now available! See https://microk8s.io for docs or
   install it with:

     sudo snap install microk8s --channel=1.18 --classic

 * Multipass 1.1 adds proxy support for developers behind enterprise
   firewalls. Rapid prototyping for cloud operations just got easier.

     https://multipass.run/

 * Canonical Livepatch is available for installation.
   - Reduce system reboots and improve kernel security. Activate at:
     https://ubuntu.com/livepatch

50 packages can be updated.
0 updates are security updates.


*** System restart required ***
Last login: Mon Mar 30 14:20:48 2020 from 107.13.161.147
ubuntu@ip-172-31-9-80:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1
ubuntu@ip-172-31-9-80:~$ python EC2.py
10:51, 03/30; num of cores:16

Basic setting:[sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, lam, simple, M_in_R] = [5, 5, 5, 0.2, 1, 1, 1e-05, True, True]


---------------------------------------
[pattern_seed, T, sd_R] = [0, 336, 5]

max(u_O) =  156.6
O_threshold = -3
means of Order:

141.6 107.8 121.0

155.7 144.5 81.8

120.3 96.5 97.5

target policy:

1 1 0

1 1 1

0 0 0

number of reward locations:  5
O_threshold = -2
target policy:

0 0 1

0 0 0

0 1 0

number of reward locations:  2
O_threshold = -1
target policy:

0 1 0

0 0 0

0 0 0

number of reward locations:  1
O_threshold = 90
target policy:

1 1 1 1 1
```

```
0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
O_threshold = 100
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
Process Process-5:
Process Process-6:
Process Process-7:
Process Process-8:
Process Process-9:
Process Process-10:
Process Process-11:
Process Process-12:
Process Process-13:
Traceback (most recent call last):
Process Process-14:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
Traceback (most recent call last):
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
Traceback (most recent call last):
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
Process Process-15:
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
Traceback (most recent call last):
IndexError: list index out of range
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Process Process-16:
```

```
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
IndexError: list index out of range
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
IndexError: list index out of range
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
Traceback (most recent call last):
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
```

```
    self.run()
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
IndexError: list index out of range
^CProcess Process-2:
Process Process-3:
Process Process-4:
Process Process-1:
Traceback (most recent call last):
  File "EC2.py", line 70, in <module>
    file = file, print_flag_target = False
  File "/home/ubuntu/simu_funs.py", line 62, in simu
    value_reps = rep_seeds(once, OPE_rep_times)
  File "/home/ubuntu/_uti_basic.py", line 119, in rep_seeds
    return list(map(fun, range(rep_times)))
  File "/home/ubuntu/simu_funs.py", line 58, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 190, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 130, in V_DR
    r = arr(parmap(getOneRegionValue, range(N), n_cores))
  File "/home/ubuntu/_uti_basic.py", line 74, in parmap
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/_uti_basic.py", line 74, in <listcomp>
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-9-80:~$ ^C
ubuntu@ip-172-31-9-80:~$ python EC2.py
10:51, 03/30; num of cores:16

Basic setting:[sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, lam, simple, M_in_R] = [5, 5, 5, 0.2, 1, 1, 1e-05, True, True]


---------------------------------------
[pattern_seed, T, sd_R] = [0, 336, 5]

max(u_O) =  156.6
O_threshold = -3
means of Order:

141.6 107.8 121.0

155.7 144.5 81.8

120.3 96.5 97.5

target policy:

1 1 0

1 1 1

0 0 0

number of reward locations:  5
O_threshold = -2
target policy:

0 0 1

0 0 0

0 1 0

number of reward locations:  2
O_threshold = -1
target policy:

0 1 0

0 0 0

0 0 0

number of reward locations:  1
O_threshold = 90
target policy:

1 1 1 1 1
```

```
0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
O_threshold = 100
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
Process Process-5:
Process Process-6:
Process Process-7:
Process Process-8:
Process Process-9:
Process Process-10:
Traceback (most recent call last):
Process Process-11:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
Process Process-12:
Traceback (most recent call last):
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
Process Process-13:
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
Process Process-14:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
Traceback (most recent call last):
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Process Process-15:
Process Process-16:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
```

```
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
Traceback (most recent call last):
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
```

```
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Process Process-2:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-1:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-4:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-3:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
^CTraceback (most recent call last):
  File "EC2.py", line 70, in <module>
    file = file, print_flag_target = False
  File "/home/ubuntu/simu_funs.py", line 62, in simu
    value_reps = rep_seeds(once, OPE_rep_times)
  File "/home/ubuntu/_uti_basic.py", line 119, in rep_seeds
    return list(map(fun, range(rep_times)))
  File "/home/ubuntu/simu_funs.py", line 58, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 190, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 130, in V_DR
    r = arr(parmap(getOneRegionValue, range(N), n_cores))
  File "/home/ubuntu/_uti_basic.py", line 74, in parmap
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/_uti_basic.py", line 74, in <listcomp>
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-9-80:~$ python EC2.py
10:53, 03/30; num of cores:16

Basic setting:[sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, lam, simple, M_in_R] = [5, 5, 5, 0.2, 1, 1, 1e-05, True, True]
```

```
--------------------------------------
[pattern_seed, T, sd_R] = [0, 336, 5]

max(u_O) =  156.6
O_threshold = -3
means of Order:

141.6 107.8 121.0

155.7 144.5 81.8

120.3 96.5 97.5

target policy:

1 1 0

1 1 1

0 0 0

number of reward locations:  5
O_threshold = -2
target policy:

0 0 1

0 0 0

0 1 0

number of reward locations:  2
O_threshold = -1
target policy:

0 1 0

0 0 0

0 0 0

number of reward locations:  1
O_threshold = 90
target policy:

1 1 1 1 1

0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
O_threshold = 100
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
Process Process-5:
Process Process-6:
Process Process-7:
Process Process-8:
Process Process-9:
Traceback (most recent call last):
Process Process-10:
Process Process-11:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
```

```
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Process Process-12:
Traceback (most recent call last):
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
Process Process-13:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
Traceback (most recent call last):
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
IndexError: list index out of range
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
Process Process-14:
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
IndexError: list index out of range
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Traceback (most recent call last):
Process Process-15:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
Process Process-16:
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
```

```
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
IndexError: list index out of range
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
Traceback (most recent call last):
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
IndexError: list index out of range
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-1:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-3:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-4:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
```

```
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-2:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
^C^XTraceback (most recent call last):
  File "EC2.py", line 70, in <module>
    file = file, print_flag_target = False
  File "/home/ubuntu/simu_funs.py", line 62, in simu
    value_reps = rep_seeds(once, OPE_rep_times)
  File "/home/ubuntu/_uti_basic.py", line 119, in rep_seeds
    return list(map(fun, range(rep_times)))
  File "/home/ubuntu/simu_funs.py", line 58, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 190, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 130, in V_DR
    r = arr(parmap(getOneRegionValue, range(N), n_cores))
  File "/home/ubuntu/_uti_basic.py", line 74, in parmap
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/_uti_basic.py", line 74, in <listcomp>
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-9-80:~$ python EC2.py
10:54, 03/30; num of cores:16

Basic setting:[sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, lam, simple, M_in_R] = [5, 5, 5, 0.2, 1, 1, 1e-05, True, True]


---------------------------------------
[pattern_seed, T, sd_R] = [0, 336, 5]

max(u_O) =  156.6
O_threshold = -3
1
2
3
4
5
6
7
8
9
means of Order:

141.6 107.8 121.0

155.7 144.5 81.8

120.3 96.5 97.5

target policy:

1 1 0

1 1 1

0 0 0

number of reward locations:  5
O_threshold = -2
1
2
3
4
5
6
7
8
9
```

```
target policy:

0 0 1

0 0 0

0 1 0

number of reward locations:  2
O_threshold = -1
1
2
3
4
5
6
7
8
9
target policy:

0 1 0

0 0 0

0 0 0

number of reward locations:  1
O_threshold = 90
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
target policy:

1 1 1 1 1

0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
O_threshold = 100
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
```

```
22
23
24
25
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
Process Process-5:
Process Process-6:
Process Process-8:
Process Process-7:
Process Process-9:
Process Process-10:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
Process Process-11:
Process Process-12:
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
Traceback (most recent call last):
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
Traceback (most recent call last):
Process Process-13:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
Process Process-14:
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
IndexError: list index out of range
IndexError: list index out of range
Process Process-15:
Traceback (most recent call last):
Process Process-16:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
```

```
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
Traceback (most recent call last):
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 160, in getRegionData
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
Traceback (most recent call last):
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
Traceback (most recent call last):
IndexError: list index out of range
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
  File "/home/ubuntu/main.py", line 160, in <listcomp>
    T_ait_1_pi = Ta([pi[a[0]](a[1][t + 1][0], random_choose = True) for a in data_neigh])
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
```

```
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
IndexError: list index out of range
Process Process-3:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-4:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-1:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
Process Process-2:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 62, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 42, in getOneRegionValue
    tuples_i = getRegionData(data[i], i, data_neigh, pi, Ts, Ta, mean_field = True, time_dependent = time_dependent)
  File "/home/ubuntu/main.py", line 159, in getRegionData
    pi_Sit_1 = pi[i](S_it1, random_choose = True)
IndexError: list index out of range
^CTraceback (most recent call last):
  File "EC2.py", line 70, in <module>
    file = file, print_flag_target = False
  File "/home/ubuntu/simu_funs.py", line 62, in simu
    value_reps = rep_seeds(once, OPE_rep_times)
  File "/home/ubuntu/_uti_basic.py", line 119, in rep_seeds
    return list(map(fun, range(rep_times)))
  File "/home/ubuntu/simu_funs.py", line 58, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 190, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 130, in V_DR
    r = arr(parmap(getOneRegionValue, range(N), n_cores))
  File "/home/ubuntu/_uti_basic.py", line 74, in parmap
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/_uti_basic.py", line 74, in <listcomp>
    sent = [q_in.put((i, x)) for i, x in enumerate(X)]
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-9-80:~$ python EC2.py
10:55, 03/30; num of cores:16

Basic setting:[sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, lam, simple, M_in_R] = [5, 5, 5, 0.2, 1, 1, 1e-05, True, True]


----------------------------------------
[pattern_seed, T, sd_R] = [0, 336, 5]

max(u_O) =  156.6
O_threshold = -3
means of Order:
```

```
141.6 107.8 121.0 155.7 144.5

81.8 120.3 96.5 97.5 108.0

102.4 133.1 115.8 101.9 108.7

106.3 134.1 95.5 105.9 83.9

59.7 113.4 118.3 85.8 156.6

target policy:

1 1 0 1 1

1 0 0 0 0

0 0 1 0 1

1 0 1 0 0

0 1 0 0 1

number of reward locations:  11
O_threshold = -2
target policy:

0 0 1 0 0

0 0 1 0 0

1 1 0 1 0

1 1 0 1 0

1 0 0 0 0

number of reward locations:  9
O_threshold = -1
target policy:

0 1 0 0 0

0 0 0 0 1

0 1 0 1 0

1 0 1 0 0

1 1 0 1 1

number of reward locations:  10
O_threshold = 90
target policy:

1 1 1 1 1

0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
O_threshold = 100
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
1 2 3 4 5 1 2 3 4 5
--------------------------------------
Value of Behaviour policy:90.868
O_threshold = -3
MC for this TARGET:[93.502, 0.114]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.53, -1.63, -1.67]][[-1.37, -1.39, -1.42]][[-93.5, -93.5, -93.5]][[-1.77, -2.63]]
std:[[0.01, 0.03, 0.04]][[0.08, 0.03, 0.0]][[0.0, 0.0, 0.0]][[0.06, 0.05]]
MSE:[[1.53, 1.63, 1.67]][[1.37, 1.39, 1.42]][[93.5, 93.5, 93.5]][[1.77, 2.63]]
```

```
MSE(-DR):[[0.0, 0.1, 0.14]][[-0.16, -0.14, -0.11]][[91.97, 91.97, 91.97]][[0.24, 1.1]]
==============


O_threshold = -2
MC for this TARGET:[90.639, 0.102]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.17, 0.11, 0.14]][[-0.73, -0.77, -0.75]][[-90.64, -90.64, -90.64]][[0.08, 0.23]]
std:[[0.45, 0.46, 0.17]][[0.0, 0.02, 0.1]][[0.0, 0.0, 0.0]][[0.18, 0.05]]
MSE:[[0.48, 0.47, 0.22]][[0.73, 0.77, 0.76]][[90.64, 90.64, 90.64]][[0.2, 0.24]]
MSE(-DR):[[0.0, -0.01, -0.26]][[0.25, 0.29, 0.28]][[90.16, 90.16, 90.16]][[-0.28, -0.24]]
better than DR_NO_MARL
MC-based ATE = -2.86
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[1.7, 1.74, 1.81]][[0.63, 0.62, 0.67]][[2.86, 2.86, 2.86]][1.85]
std:[[0.46, 0.48, 0.21]][[0.08, 0.05, 0.11]][[0.0, 0.0, 0.0]][0.24]
MSE:[[1.76, 1.8, 1.82]][[0.64, 0.62, 0.68]][[2.86, 2.86, 2.86]][1.87]
MSE(-DR):[[0.0, 0.04, 0.06]][[-1.12, -1.14, -1.08]][[1.1, 1.1, 1.1]][0.11]
==============


O_threshold = -1
MC for this TARGET:[91.349, 0.104]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.58, -1.62, -2.01]][[-1.95, -1.92, -1.94]][[-91.35, -91.35, -91.35]][[-2.05, -0.48]]
std:[[0.12, 0.13, 0.01]][[0.16, 0.11, 0.15]][[0.0, 0.0, 0.0]][[0.02, 0.05]]
MSE:[[1.58, 1.63, 2.01]][[1.96, 1.92, 1.95]][[91.35, 91.35, 91.35]][[2.05, 0.48]]
MSE(-DR):[[0.0, 0.05, 0.43]][[0.38, 0.34, 0.37]][[89.77, 89.77, 89.77]][[0.47, -1.1]]
******
MC-based ATE = -2.15
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-0.05, 0.01, -0.35]][[-0.58, -0.53, -0.52]][[2.15, 2.15, 2.15]][-0.28]
std:[[0.11, 0.11, 0.03]][[0.08, 0.09, 0.15]][[0.0, 0.0, 0.0]][0.04]
MSE:[[0.12, 0.11, 0.35]][[0.59, 0.54, 0.54]][[2.15, 2.15, 2.15]][0.28]
MSE(-DR):[[0.0, -0.01, 0.23]][[0.47, 0.42, 0.42]][[2.03, 2.03, 2.03]][0.16]
******
==============


O_threshold = 90
MC for this TARGET:[95.681, 0.108]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.76, 0.69, 0.64]][[1.03, 0.95, 0.87]][[-95.68, -95.68, -95.68]][[0.58, -4.81]]
std:[[0.02, 0.02, 0.11]][[0.06, 0.05, 0.1]][[0.0, 0.0, 0.0]][[0.06, 0.05]]
MSE:[[0.76, 0.69, 0.65]][[1.03, 0.95, 0.88]][[95.68, 95.68, 95.68]][[0.58, 4.81]]
MSE(-DR):[[0.0, -0.07, -0.11]][[0.27, 0.19, 0.12]][[94.92, 94.92, 94.92]][[-0.18, 4.05]]
better than DR_NO_MARL
MC-based ATE = 2.18
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[2.29, 2.32, 2.31]][[2.4, 2.34, 2.29]][[-2.18, -2.18, -2.18]][2.34]
std:[[0.03, 0.0, 0.15]][[0.02, 0.02, 0.1]][[0.0, 0.0, 0.0]][0.12]
MSE:[[2.29, 2.32, 2.31]][[2.4, 2.34, 2.29]][[2.18, 2.18, 2.18]][2.34]
MSE(-DR):[[0.0, 0.03, 0.02]][[0.11, 0.05, 0.0]][[-0.11, -0.11, -0.11]][0.05]
******
==============


O_threshold = 100
MC for this TARGET:[96.88, 0.109]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-0.13, -0.16, -0.3]][[0.31, 0.19, 0.15]][[-96.88, -96.88, -96.88]][[-0.33, -6.01]]
std:[[0.35, 0.22, 0.25]][[0.13, 0.11, 0.18]][[0.0, 0.0, 0.0]][[0.12, 0.05]]
MSE:[[0.37, 0.27, 0.39]][[0.34, 0.22, 0.23]][[96.88, 96.88, 96.88]][[0.35, 6.01]]
MSE(-DR):[[0.0, -0.1, 0.02]][[-0.03, -0.15, -0.14]][[96.51, 96.51, 96.51]][[-0.02, 5.64]]
MC-based ATE = 3.38
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[1.4, 1.47, 1.37]][[1.68, 1.58, 1.58]][[-3.38, -3.38, -3.38]][1.44]
std:[[0.36, 0.25, 0.29]][[0.05, 0.09, 0.18]][[0.0, 0.0, 0.0]][0.18]
MSE:[[1.45, 1.49, 1.4]][[1.68, 1.58, 1.59]][[3.38, 3.38, 3.38]][1.45]
MSE(-DR):[[0.0, 0.04, -0.05]][[0.23, 0.13, 0.14]][[1.93, 1.93, 1.93]][0.0]
better than DR_NO_MARL
==============


time spent until now: 3.7 mins


----------------------------------------
[pattern_seed, T, sd_R] = [0, 672, 5]

max(u_O) =  156.6
O_threshold = -3
means of Order:

141.6 107.8 121.0 155.7 144.5

81.8 120.3 96.5 97.5 108.0

102.4 133.1 115.8 101.9 108.7
```

```
106.3 134.1 95.5 105.9 83.9

59.7 113.4 118.3 85.8 156.6

target policy:

1 1 0 1 1

1 0 0 0 0

0 0 1 0 1

1 0 1 0 0

0 1 0 0 1

number of reward locations:  11
O_threshold = -2
target policy:

0 0 1 0 0

0 0 1 0 0

1 1 0 1 0

1 1 0 1 0

1 0 0 0 0

number of reward locations:  9
O_threshold = -1
target policy:

0 1 0 0 0

0 0 0 0 1

0 1 0 1 0

1 0 1 0 0

1 1 0 1 1

number of reward locations:  10
O_threshold = 90
target policy:

1 1 1 1 1

0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
O_threshold = 100
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
1 2 3 4 5 1 2 3 4 5
----------------------------------------
Value of Behaviour policy:90.884
O_threshold = -3
MC for this TARGET:[93.5, 0.075]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.26, -1.33, -1.59]][[-1.41, -1.48, -1.53]][[-93.5, -93.5, -93.5]][[-1.66, -2.62]]
std:[[0.07, 0.04, 0.06]][[0.02, 0.02, 0.02]][[0.0, 0.0, 0.0]][[0.08, 0.02]]
MSE:[[1.26, 1.33, 1.59]][[1.41, 1.48, 1.53]][[93.5, 93.5, 93.5]][[1.66, 2.62]]
MSE(-DR):[[0.0, 0.07, 0.33]][[0.15, 0.22, 0.27]][[92.24, 92.24, 92.24]][[0.4, 1.36]]
******
==============


O_threshold = -2
```

```
MC for this TARGET:[90.639, 0.08]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.37, 0.38, 0.34]][[-0.62, -0.63, -0.62]][[-90.64, -90.64, -90.64]][[0.35, 0.25]]
std:[[0.13, 0.13, 0.05]][[0.07, 0.03, 0.12]][[0.0, 0.0, 0.0]][[0.05, 0.02]]
MSE:[[0.39, 0.4, 0.34]][[0.62, 0.63, 0.63]][[90.64, 90.64, 90.64]][[0.35, 0.25]]
MSE(-DR):[[0.0, 0.01, -0.05]][[0.23, 0.24, 0.24]][[90.25, 90.25, 90.25]][[-0.04, -0.14]]
better than DR_NO_MARL
MC-based ATE = -2.86
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[1.63, 1.71, 1.93]][[0.78, 0.85, 0.9]][[2.86, 2.86, 2.86]][2.0]
std:[[0.2, 0.17, 0.1]][[0.06, 0.05, 0.1]][[0.0, 0.0, 0.0]][0.13]
MSE:[[1.64, 1.72, 1.93]][[0.78, 0.85, 0.91]][[2.86, 2.86, 2.86]][2.0]
MSE(-DR):[[0.0, 0.08, 0.29]][[-0.86, -0.79, -0.73]][[1.22, 1.22, 1.22]][0.36]
==============


O_threshold = -1
MC for this TARGET:[91.35, 0.077]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.88, -1.88, -2.08]][[-1.82, -1.82, -1.82]][[-91.35, -91.35, -91.35]][[-2.08, -0.47]]
std:[[0.34, 0.3, 0.21]][[0.04, 0.01, 0.08]][[0.0, 0.0, 0.0]][[0.17, 0.02]]
MSE:[[1.91, 1.9, 2.09]][[1.82, 1.82, 1.82]][[91.35, 91.35, 91.35]][[2.09, 0.47]]
MSE(-DR):[[0.0, -0.01, 0.18]][[-0.09, -0.09, -0.09]][[89.44, 89.44, 89.44]][[0.18, -1.44]]
MC-based ATE = -2.15
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-0.61, -0.55, -0.49]][[-0.42, -0.34, -0.29]][[2.15, 2.15, 2.15]][-0.43]
std:[[0.27, 0.26, 0.26]][[0.02, 0.02, 0.06]][[0.0, 0.0, 0.0]][0.25]
MSE:[[0.67, 0.61, 0.55]][[0.42, 0.34, 0.3]][[2.15, 2.15, 2.15]][0.5]
MSE(-DR):[[0.0, -0.06, -0.12]][[-0.25, -0.33, -0.37]][[1.48, 1.48, 1.48]][-0.17]
==============


O_threshold = 90
MC for this TARGET:[95.687, 0.08]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[1.09, 1.02, 0.78]][[0.69, 0.61, 0.67]][[-95.69, -95.69, -95.69]][[0.71, -4.8]]
std:[[0.26, 0.26, 0.2]][[0.12, 0.07, 0.09]][[0.0, 0.0, 0.0]][[0.2, 0.02]]
MSE:[[1.12, 1.05, 0.81]][[0.7, 0.61, 0.68]][[95.69, 95.69, 95.69]][[0.74, 4.8]]
MSE(-DR):[[0.0, -0.07, -0.31]][[-0.42, -0.51, -0.44]][[94.57, 94.57, 94.57]][[-0.38, 3.68]]
MC-based ATE = 2.19
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[2.35, 2.35, 2.37]][[2.1, 2.09, 2.2]][[-2.19, -2.19, -2.19]][2.36]
std:[[0.33, 0.3, 0.15]][[0.1, 0.09, 0.07]][[0.0, 0.0, 0.0]][0.12]
MSE:[[2.37, 2.37, 2.37]][[2.1, 2.09, 2.2]][[2.19, 2.19, 2.19]][2.36]
MSE(-DR):[[0.0, 0.0, 0.0]][[-0.27, -0.28, -0.17]][[-0.18, -0.18, -0.18]][-0.01]
==============


O_threshold = 100
MC for this TARGET:[96.882, 0.081]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.15, 0.05, -0.28]][[0.17, 0.05, 0.08]][[-96.88, -96.88, -96.88]][[-0.38, -6.0]]
std:[[0.14, 0.12, 0.12]][[0.1, 0.05, 0.1]][[0.0, 0.0, 0.0]][[0.1, 0.02]]
MSE:[[0.21, 0.13, 0.3]][[0.2, 0.07, 0.13]][[96.88, 96.88, 96.88]][[0.39, 6.0]]
MSE(-DR):[[0.0, -0.08, 0.09]][[-0.01, -0.14, -0.08]][[96.67, 96.67, 96.67]][[0.18, 5.79]]
MC-based ATE = 3.38
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[1.41, 1.38, 1.31]][[1.57, 1.53, 1.61]][[-3.38, -3.38, -3.38]][1.28]
std:[[0.21, 0.16, 0.06]][[0.08, 0.07, 0.08]][[0.0, 0.0, 0.0]][0.02]
MSE:[[1.43, 1.39, 1.31]][[1.57, 1.53, 1.61]][[3.38, 3.38, 3.38]][1.28]
MSE(-DR):[[0.0, -0.04, -0.12]][[0.14, 0.1, 0.18]][[1.95, 1.95, 1.95]][-0.15]
better than DR_NO_MARL
==============


time spent until now: 7.8 mins

ubuntu@ip-172-31-9-80:~$ python EC2.py
11:03, 03/30; num of cores:16

Basic setting:[sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, lam, simple, M_in_R] = [5, 5, 5, 0.2, 1, 1, 1e-05, True, True]


----------------------------------------
[pattern_seed, T, sd_R] = [0, 336, 5]

max(u_O) =  156.6
O_threshold = 100
means of Order:

141.6 107.8 121.0 155.7 144.5

81.8 120.3 96.5 97.5 108.0

102.4 133.1 115.8 101.9 108.7

106.3 134.1 95.5 105.9 83.9

59.7 113.4 118.3 85.8 156.6
```

```
target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations:  18
O_threshold = -4
target policy:

1 1 1 1 1

0 1 0 0 0

1 0 1 1 0

1 0 0 0 1

0 1 1 1 0

number of reward locations:  14
O_threshold = -3
target policy:

1 1 0 1 1

1 0 0 0 0

0 0 1 0 1

1 0 1 0 0

0 1 0 0 1

number of reward locations:  11
O_threshold = -2
target policy:

0 0 1 0 0

0 0 1 0 0

1 1 0 1 0

1 1 0 1 0

1 0 0 0 0

number of reward locations:  9
O_threshold = -1
target policy:

0 1 0 0 0

0 0 0 0 1

0 1 0 1 0

1 0 1 0 0

1 1 0 1 1

number of reward locations:  10
O_threshold = 90
target policy:

1 1 1 1 1

0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
1 2 3 4 5 6 1 2 3 4 5 6
--------------------------------------
Value of Behaviour policy:90.868
O_threshold = 100
MC for this TARGET:[96.88, 0.109]
```

```
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-0.13, -0.16, -0.3]][[0.31, 0.19, 0.15]][[-96.88, -96.88, -96.88]][[-0.34, -6.01]]
std:[[0.31, 0.22, 0.25]][[0.13, 0.11, 0.19]][[0.0, 0.0, 0.0]][[0.17, 0.05]]
MSE:[[0.34, 0.27, 0.39]][[0.34, 0.22, 0.24]][[96.88, 96.88, 96.88]][[0.38, 6.01]]
MSE(-DR):[[0.0, -0.07, 0.05]][[0.0, -0.12, -0.1]][[96.54, 96.54, 96.54]][[0.04, 5.67]]
******
==============


O_threshold = -4
MC for this TARGET:[92.821, 0.113]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-0.08, -0.05, -0.42]][[0.43, 0.38, 0.34]][[-92.82, -92.82, -92.82]][[-0.39, -1.95]]
std:[[0.01, 0.02, 0.02]][[0.04, 0.03, 0.03]][[0.0, 0.0, 0.0]][[0.06, 0.05]]
MSE:[[0.08, 0.05, 0.42]][[0.43, 0.38, 0.34]][[92.82, 92.82, 92.82]][[0.39, 1.95]]
MSE(-DR):[[0.0, -0.03, 0.34]][[0.35, 0.3, 0.26]][[92.74, 92.74, 92.74]][[0.31, 1.87]]
******
MC-based ATE = -4.06
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[0.05, 0.11, -0.11]][[0.11, 0.19, 0.19]][[4.06, 4.06, 4.06]][-0.06]
std:[[0.29, 0.25, 0.27]][[0.09, 0.08, 0.16]][[0.0, 0.0, 0.0]][0.22]
MSE:[[0.29, 0.27, 0.29]][[0.14, 0.21, 0.25]][[4.06, 4.06, 4.06]][0.23]
MSE(-DR):[[0.0, -0.02, 0.0]][[-0.15, -0.08, -0.04]][[3.77, 3.77, 3.77]][-0.06]
==============


O_threshold = -3
MC for this TARGET:[93.502, 0.114]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.54, -1.63, -1.69]][[-1.36, -1.39, -1.42]][[-93.5, -93.5, -93.5]][[-1.78, -2.63]]
std:[[0.01, 0.03, 0.04]][[0.08, 0.03, 0.01]][[0.0, 0.0, 0.0]][[0.07, 0.05]]
MSE:[[1.54, 1.63, 1.69]][[1.36, 1.39, 1.42]][[93.5, 93.5, 93.5]][[1.78, 2.63]]
MSE(-DR):[[0.0, 0.09, 0.15]][[-0.18, -0.15, -0.12]][[91.96, 91.96, 91.96]][[0.24, 1.09]]
MC-based ATE = -3.38
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-1.42, -1.47, -1.39]][[-1.68, -1.58, -1.56]][[3.38, 3.38, 3.38]][-1.45]
std:[[0.3, 0.25, 0.29]][[0.05, 0.09, 0.2]][[0.0, 0.0, 0.0]][0.24]
MSE:[[1.45, 1.49, 1.42]][[1.68, 1.58, 1.57]][[3.38, 3.38, 3.38]][1.47]
MSE(-DR):[[0.0, 0.04, -0.03]][[0.23, 0.13, 0.12]][[1.93, 1.93, 1.93]][0.02]
better than DR_NO_MARL
==============


O_threshold = -2
MC for this TARGET:[90.639, 0.102]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.17, 0.11, 0.11]][[-0.74, -0.77, -0.76]][[-90.64, -90.64, -90.64]][[0.04, 0.23]]
std:[[0.48, 0.46, 0.19]][[0.0, 0.02, 0.1]][[0.0, 0.0, 0.0]][[0.16, 0.05]]
MSE:[[0.51, 0.47, 0.22]][[0.74, 0.77, 0.77]][[90.64, 90.64, 90.64]][[0.16, 0.24]]
MSE(-DR):[[0.0, -0.04, -0.29]][[0.23, 0.26, 0.26]][[90.13, 90.13, 90.13]][[-0.35, -0.27]]
better than DR_NO_MARL
MC-based ATE = -6.24
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[0.3, 0.27, 0.41]][[-1.06, -0.96, -0.9]][[6.24, 6.24, 6.24]][0.38]
std:[[0.17, 0.23, 0.06]][[0.13, 0.14, 0.08]][[0.0, 0.0, 0.0]][0.0]
MSE:[[0.34, 0.35, 0.41]][[1.07, 0.97, 0.9]][[6.24, 6.24, 6.24]][0.38]
MSE(-DR):[[0.0, 0.01, 0.07]][[0.73, 0.63, 0.56]][[5.9, 5.9, 5.9]][0.04]
******
==============


O_threshold = -1
MC for this TARGET:[91.349, 0.104]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.61, -1.62, -2.04]][[-1.95, -1.92, -1.94]][[-91.35, -91.35, -91.35]][[-2.05, -0.48]]
std:[[0.13, 0.13, 0.02]][[0.16, 0.11, 0.15]][[0.0, 0.0, 0.0]][[0.01, 0.05]]
MSE:[[1.62, 1.63, 2.04]][[1.96, 1.92, 1.95]][[91.35, 91.35, 91.35]][[2.05, 0.48]]
MSE(-DR):[[0.0, 0.01, 0.42]][[0.34, 0.3, 0.33]][[89.73, 89.73, 89.73]][[0.43, -1.14]]
******
MC-based ATE = -5.53
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-1.48, -1.46, -1.73]][[-2.26, -2.11, -2.08]][[5.53, 5.53, 5.53]][-1.72]
std:[[0.44, 0.36, 0.23]][[0.03, 0.0, 0.04]][[0.0, 0.0, 0.0]][0.15]
MSE:[[1.54, 1.5, 1.75]][[2.26, 2.11, 2.08]][[5.53, 5.53, 5.53]][1.73]
MSE(-DR):[[0.0, -0.04, 0.21]][[0.72, 0.57, 0.54]][[3.99, 3.99, 3.99]][0.19]
******
==============


O_threshold = 90
MC for this TARGET:[95.681, 0.108]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.75, 0.69, 0.6]][[1.04, 0.95, 0.88]][[-95.68, -95.68, -95.68]][[0.54, -4.81]]
std:[[0.0, 0.02, 0.09]][[0.06, 0.05, 0.11]][[0.0, 0.0, 0.0]][[0.06, 0.05]]
MSE:[[0.75, 0.69, 0.61]][[1.04, 0.95, 0.89]][[95.68, 95.68, 95.68]][[0.54, 4.81]]
MSE(-DR):[[0.0, -0.06, -0.14]][[0.29, 0.2, 0.14]][[94.93, 94.93, 94.93]][[-0.21, 4.06]]
better than DR_NO_MARL
MC-based ATE = -1.2
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
```

bias:[[0.88, 0.85, 0.9]][[0.72, 0.76, 0.74]][[1.2, 1.2, 1.2]][0.87]
std:[[0.31, 0.25, 0.16]][[0.06, 0.06, 0.08]][[0.0, 0.0, 0.0]][0.1]
MSE:[[0.93, 0.89, 0.91]][[0.72, 0.76, 0.74]][[1.2, 1.2, 1.2]][0.88]
MSE(-DR):[[0.0, -0.04, -0.02]][[-0.21, -0.17, -0.19]][[0.27, 0.27, 0.27]][-0.05]
==========

time spent until now: 4.3 mins

[pattern_seed, T, sd_R] = [0, 672, 5]

max(u_O) = 156.6
O_threshold = 100
means of Order:

141.6 107.8 121.0 155.7 144.5

81.8 120.3 96.5 97.5 108.0

102.4 133.1 115.8 101.9 108.7

106.3 134.1 95.5 105.9 83.9

59.7 113.4 118.3 85.8 156.6

target policy:

1 1 1 1 1

0 1 0 0 1

1 1 1 1 1

1 1 0 1 0

0 1 1 0 1

number of reward locations: 18
O_threshold = -4
target policy:

1 1 1 1 1

0 1 0 0 0

1 0 1 1 0

1 0 0 0 1

0 1 1 1 0

number of reward locations: 14
O_threshold = -3
target policy:

1 1 0 1 1

1 0 0 0 0

0 0 1 0 1

1 0 1 0 0

0 1 0 0 1

number of reward locations: 11
O_threshold = -2
target policy:

0 0 1 0 0

0 0 1 0 0

1 1 0 1 0

1 1 0 1 0

1 0 0 0 0

number of reward locations: 9
O_threshold = -1
target policy:

0 1 0 0 0

0 0 0 0 1

```
0 1 0 1 0

1 0 1 0 0

1 1 0 1 1

number of reward locations:  10
O_threshold = 90
target policy:

1 1 1 1 1

0 1 1 1 1

1 1 1 1 1

1 1 1 1 0

0 1 1 0 1

number of reward locations:  21
1 2 3 4 5 6 1 2 3 4 5 6
----------------------------------------
Value of Behaviour policy:90.884
O_threshold = 100
MC for this TARGET:[96.882, 0.081]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.16, 0.05, -0.25]][[0.18, 0.05, 0.1]][[-96.88, -96.88, -96.88]][[-0.36, -6.0]]
std:[[0.12, 0.12, 0.09]][[0.1, 0.05, 0.11]][[0.0, 0.0, 0.0]][[0.08, 0.02]]
MSE:[[0.2, 0.13, 0.27]][[0.21, 0.07, 0.15]][[96.88, 96.88, 96.88]][[0.37, 6.0]]
MSE(-DR):[[0.0, -0.07, 0.07]][[0.01, -0.13, -0.05]][[96.68, 96.68, 96.68]][[0.17, 5.8]]
******
==============


O_threshold = -4
MC for this TARGET:[92.822, 0.079]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-0.09, -0.14, -0.38]][[0.09, 0.04, 0.05]][[-92.82, -92.82, -92.82]][[-0.43, -1.94]]
std:[[0.18, 0.12, 0.34]][[0.18, 0.15, 0.18]][[0.0, 0.0, 0.0]][[0.28, 0.02]]
MSE:[[0.2, 0.18, 0.51]][[0.2, 0.16, 0.19]][[92.82, 92.82, 92.82]][[0.51, 1.94]]
MSE(-DR):[[0.0, -0.02, 0.31]][[0.0, -0.04, -0.01]][[92.62, 92.62, 92.62]][[0.31, 1.74]]
******
MC-based ATE = -4.06
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-0.25, -0.19, -0.13]][[-0.09, -0.0, -0.05]][[4.06, 4.06, 4.06]][-0.07]
std:[[0.3, 0.24, 0.43]][[0.08, 0.1, 0.07]][[0.0, 0.0, 0.0]][0.37]
MSE:[[0.39, 0.31, 0.45]][[0.12, 0.1, 0.09]][[4.06, 4.06, 4.06]][0.38]
MSE(-DR):[[0.0, -0.08, 0.06]][[-0.27, -0.29, -0.3]][[3.67, 3.67, 3.67]][-0.01]
==============


O_threshold = -3
MC for this TARGET:[93.5, 0.075]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.25, -1.33, -1.58]][[-1.41, -1.48, -1.53]][[-93.5, -93.5, -93.5]][[-1.66, -2.62]]
std:[[0.08, 0.04, 0.05]][[0.02, 0.02, 0.03]][[0.0, 0.0, 0.0]][[0.09, 0.02]]
MSE:[[1.25, 1.33, 1.58]][[1.41, 1.48, 1.53]][[93.5, 93.5, 93.5]][[1.66, 2.62]]
MSE(-DR):[[0.0, 0.08, 0.33]][[0.16, 0.23, 0.28]][[92.25, 92.25, 92.25]][[0.41, 1.37]]
******
MC-based ATE = -3.38
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-1.41, -1.38, -1.33]][[-1.59, -1.53, -1.63]][[3.38, 3.38, 3.38]][-1.3]
std:[[0.21, 0.16, 0.04]][[0.08, 0.07, 0.08]][[0.0, 0.0, 0.0]][0.0]
MSE:[[1.43, 1.39, 1.33]][[1.59, 1.53, 1.63]][[3.38, 3.38, 3.38]][1.3]
MSE(-DR):[[0.0, -0.04, -0.1]][[0.16, 0.1, 0.2]][[1.95, 1.95, 1.95]][-0.13]
better than DR_NO_MARL
==============


O_threshold = -2
MC for this TARGET:[90.639, 0.08]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[0.39, 0.38, 0.38]][[-0.62, -0.63, -0.63]][[-90.64, -90.64, -90.64]][[0.36, 0.25]]
std:[[0.12, 0.13, 0.04]][[0.06, 0.03, 0.12]][[0.0, 0.0, 0.0]][[0.04, 0.02]]
MSE:[[0.41, 0.4, 0.38]][[0.62, 0.63, 0.64]][[90.64, 90.64, 90.64]][[0.36, 0.25]]
MSE(-DR):[[0.0, -0.01, -0.03]][[0.21, 0.22, 0.23]][[90.23, 90.23, 90.23]][[-0.05, -0.16]]
better than DR_NO_MARL
MC-based ATE = -6.24
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[0.23, 0.33, 0.62]][[-0.8, -0.68, -0.73]][[6.24, 6.24, 6.24]][0.72]
std:[[0.0, 0.01, 0.13]][[0.04, 0.02, 0.01]][[0.0, 0.0, 0.0]][0.13]
MSE:[[0.23, 0.33, 0.63]][[0.8, 0.68, 0.73]][[6.24, 6.24, 6.24]][0.73]
MSE(-DR):[[0.0, 0.1, 0.4]][[0.57, 0.45, 0.5]][[6.01, 6.01, 6.01]][0.5]
******
==============


O_threshold = -1
```

```
MC for this TARGET:[91.35, 0.077]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[-1.88, -1.88, -2.1]][[-1.81, -1.82, -1.81]][[-91.35, -91.35, -91.35]][[-2.09, -0.47]]
std:[[0.33, 0.3, 0.17]][[0.04, 0.01, 0.08]][[0.0, 0.0, 0.0]][[0.15, 0.02]]
MSE:[[1.91, 1.9, 2.11]][[1.81, 1.82, 1.81]][[91.35, 91.35, 91.35]][[2.1, 0.47]]
MSE(-DR):[[0.0, -0.01, 0.2]][[-0.1, -0.09, -0.1]][[89.44, 89.44, 89.44]][[0.19, -1.44]]
MC-based ATE = -5.53
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[-2.04, -1.93, -1.85]][[-1.99, -1.87, -1.91]][[5.53, 5.53, 5.53]][-1.73]
std:[[0.45, 0.42, 0.26]][[0.06, 0.06, 0.03]][[0.0, 0.0, 0.0]][0.23]
MSE:[[2.09, 1.98, 1.87]][[1.99, 1.87, 1.91]][[5.53, 5.53, 5.53]][1.75]
MSE(-DR):[[0.0, -0.11, -0.22]][[-0.1, -0.22, -0.18]][[3.44, 3.44, 3.44]][-0.34]
==============


O_threshold = 90
MC for this TARGET:[95.687, 0.08]
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR/QV/IS]_NO_MF; [DR2, V_behav]
bias:[[1.09, 1.02, 0.77]][[0.69, 0.61, 0.66]][[-95.69, -95.69, -95.69]][[0.7, -4.8]]
std:[[0.24, 0.26, 0.18]][[0.11, 0.07, 0.08]][[0.0, 0.0, 0.0]][[0.19, 0.02]]
MSE:[[1.12, 1.05, 0.79]][[0.7, 0.61, 0.66]][[95.69, 95.69, 95.69]][[0.73, 4.8]]
MSE(-DR):[[0.0, -0.07, -0.33]][[-0.42, -0.51, -0.46]][[94.57, 94.57, 94.57]][[-0.39, 3.68]]
MC-based ATE = -1.2
   [DR/QV/IS]; [DR/QV/IS]_NO_MARL; [DR2]
bias:[[0.93, 0.97, 1.02]][[0.52, 0.56, 0.57]][[1.2, 1.2, 1.2]][1.06]
std:[[0.12, 0.14, 0.09]][[0.01, 0.02, 0.03]][[0.0, 0.0, 0.0]][0.11]
MSE:[[0.94, 0.98, 1.02]][[0.52, 0.56, 0.57]][[1.2, 1.2, 1.2]][1.07]
MSE(-DR):[[0.0, 0.04, 0.08]][[-0.42, -0.38, -0.37]][[0.26, 0.26, 0.26]][0.13]
==============


time spent until now: 9.0 mins

ubuntu@ip-172-31-9-80:~$
```