Experiment 2

February 19, 2018

Differences from Experiment 1

The only difference between this experiment and experiment 1 is that we give principal 2 a "head-start" in the sense of letting them be the only firm in the market for the first 1000 periods. Thus, the principals are effectively given 1000 free observations for learning and the agents are given 1000 free observations in order to form a reputation score for principal 2. The question this simulation then attempts to answer is if *information advantage* can be viewed as a barrier to entry. Namely, it attempts to begin to answer if the first mover advantage manifests itself in this model by giving the incumbent firm a better prior than any entrants (who we assume come in with only "fake" priors). Under what assumptions about the agent rationality (response functions or memory sizes) and what algorithms do we see the entering firm able to gain any market share?

Simulation Details

Considered $K = 3$, $T = 5005$, $N = 60$. Report statistics at $t = 1000, 3000, 5000$
The Bandit priors that were considered:

- Uniform: Draw the mean rewards for the arms from [0.25, 0.75]
- "HeavyTail": We took the mean rewards to be randomly drawn from Beta($\alpha = 0.6, \beta = 0.6$). With this distribution it was likely to have arms that were at the extremes (close to 1 and close to 0) but also some of the arms with intermediate value means.
- Needle-in-haystack
  1. Medium - 9 arms with mean 0.50, 1 arm with mean 0.55 (+ 0.05)
  2. High - 9 arms with mean 0.50, 1 arm with mean 0.70 (+ 0.20)

Algorithms considered:

1. ThompsonSampling with priors of $Beta(1, 1)$ for every arm.
2. DynamicGreedy with priors of $Beta(1, 1)$ for every arm
3. Bayesian Dynamic ϵ-greedy with priors of $Beta(1, 1)$ for every arm and $\epsilon = 0.05$

Agent Algorithms considered:

1. HardMax
2. HardMaxWithRandom
3. SoftMax

Memory Sizes

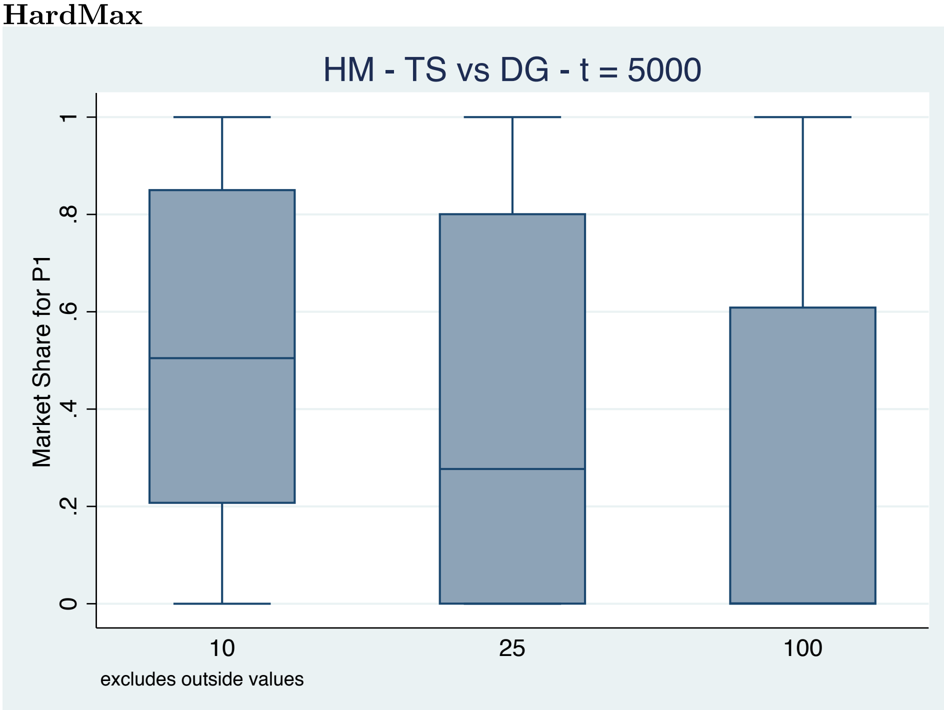
1. 10
2. 25
3. 100

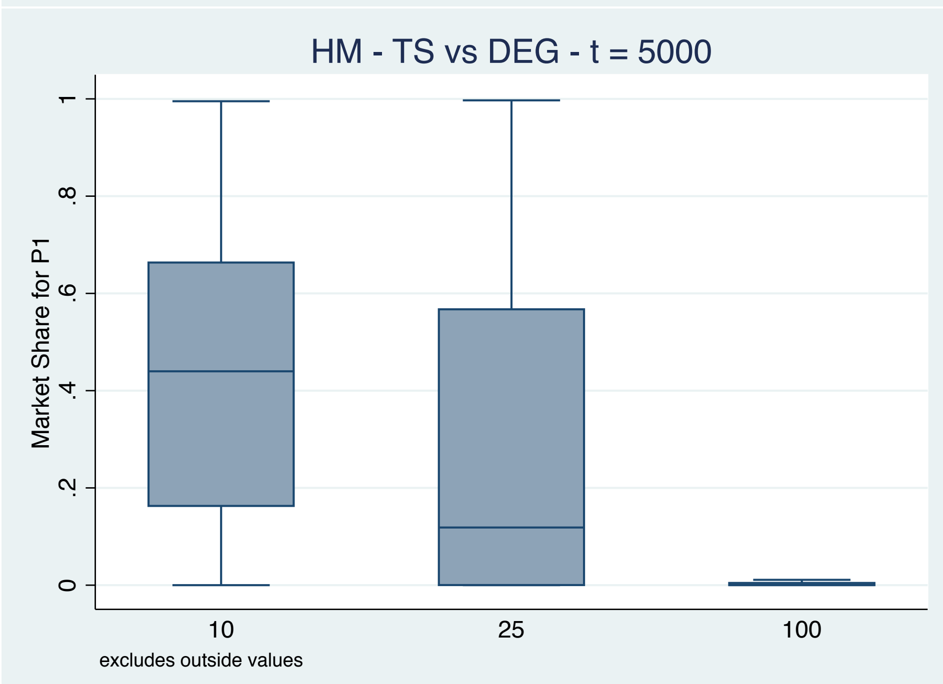
The simulation procedure is the only thing that has changed compared to Experiment 1: **Simulation Procedure**

```
for Each prior p do
    for Each agent algorithm agentalg do
        for Each principal algorithm pair principalalg1, principalalg2 do
            for N simulations do
                Generate true distribution from p (except for needle-in-haystack, just use p itself)
                Give the agents 5 observations from each principal
                Give principal 2 1000 free observations (and give the agents this information too)
                Run simulation for T periods
            end for
        end for
    end for
end for
```

Results

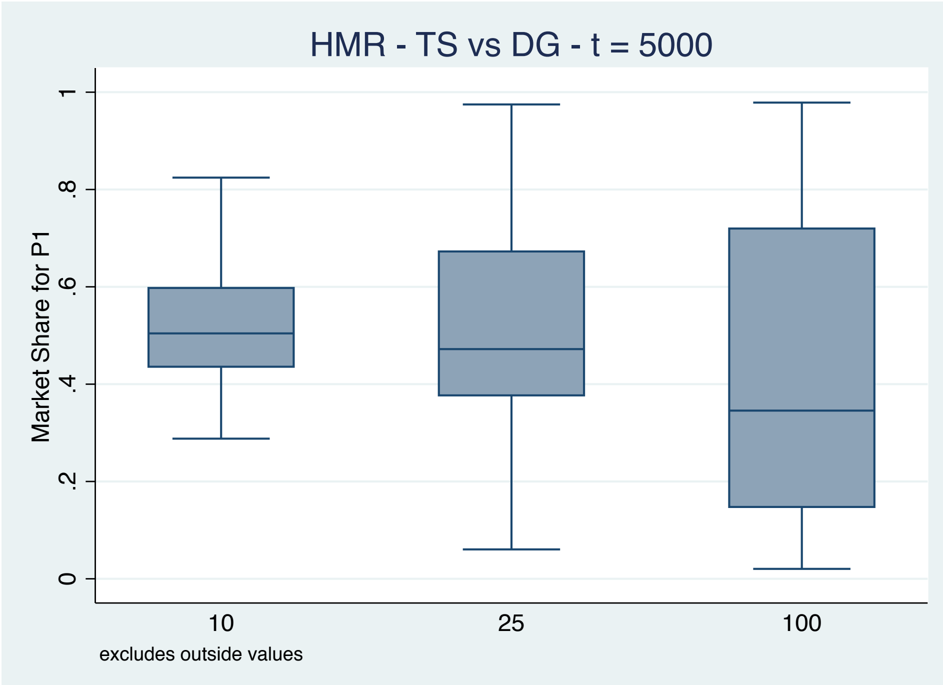
HardMax

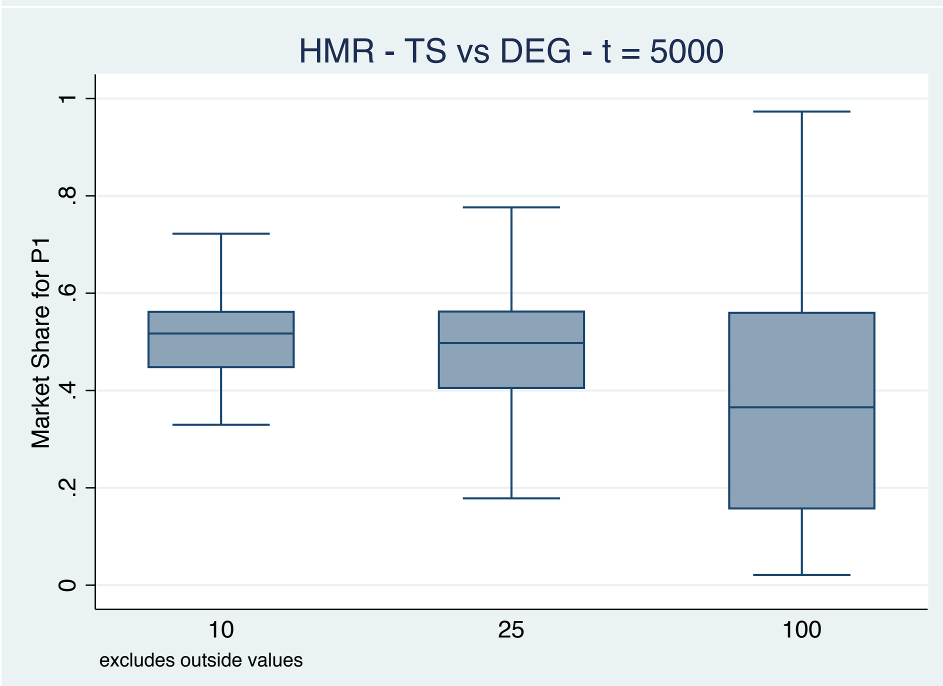




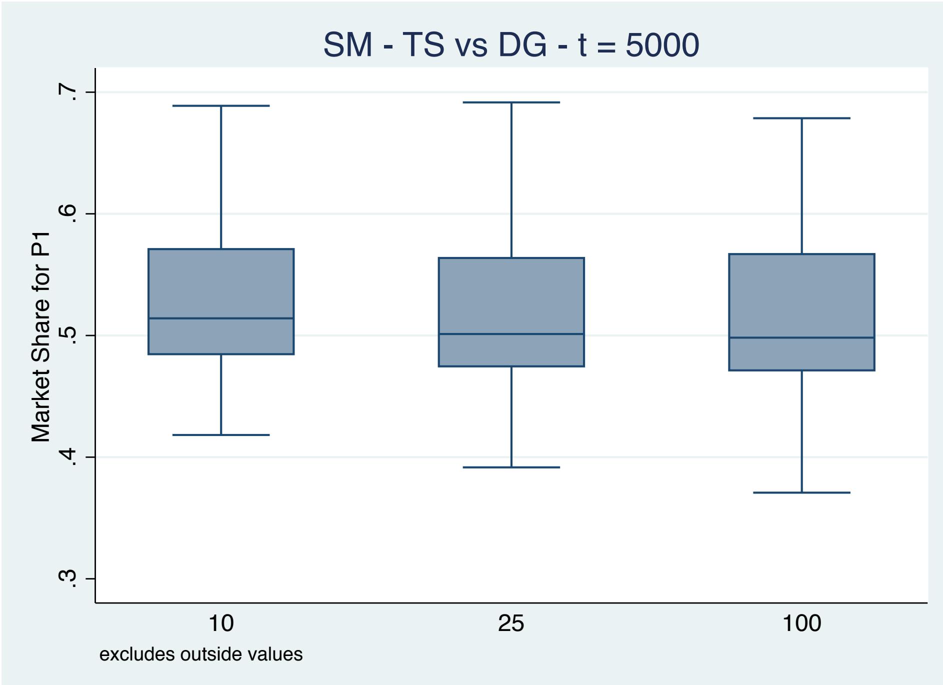
The main interesting result here is that for ThompsonSampling vs DEG, with low memory sizes we have that ThompsonSampling almost catches up to DEG, but as we increase memory size TS gets almost no market share! Should re-run this experiment with higher N to validate this.

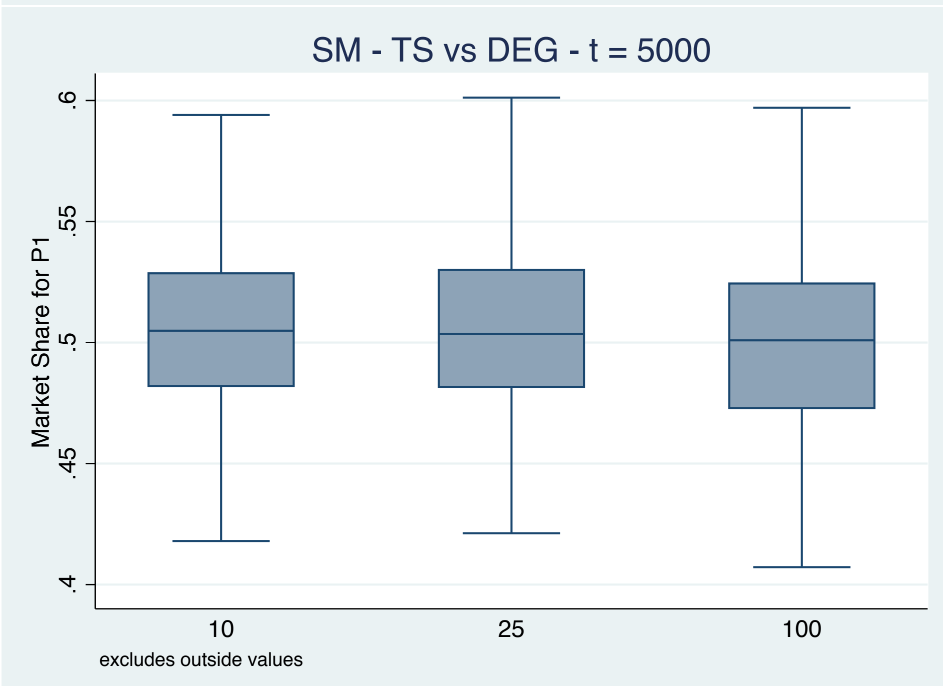
HardMaxWithRandom



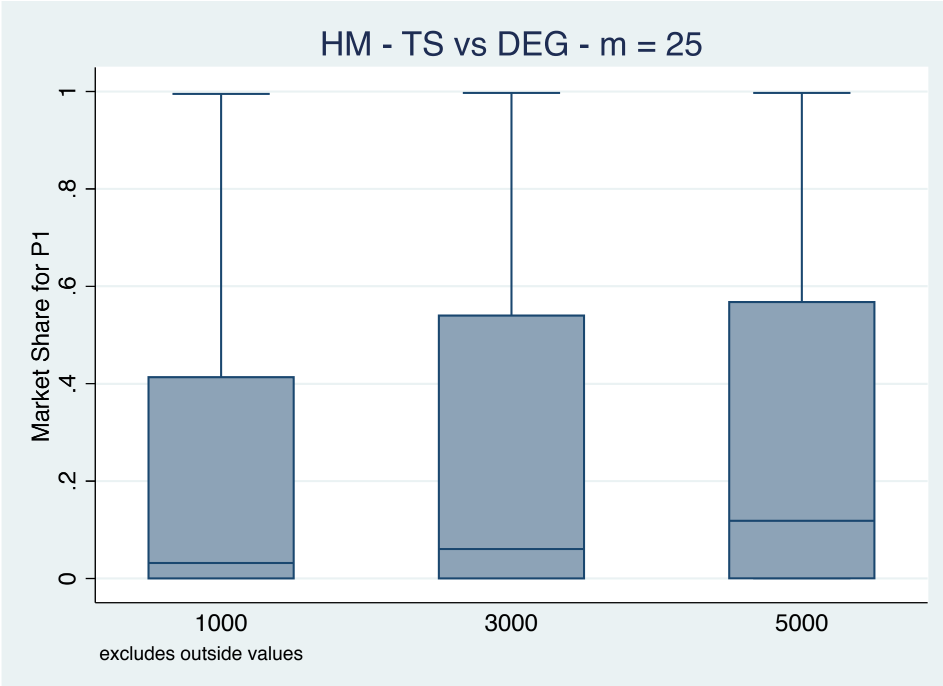


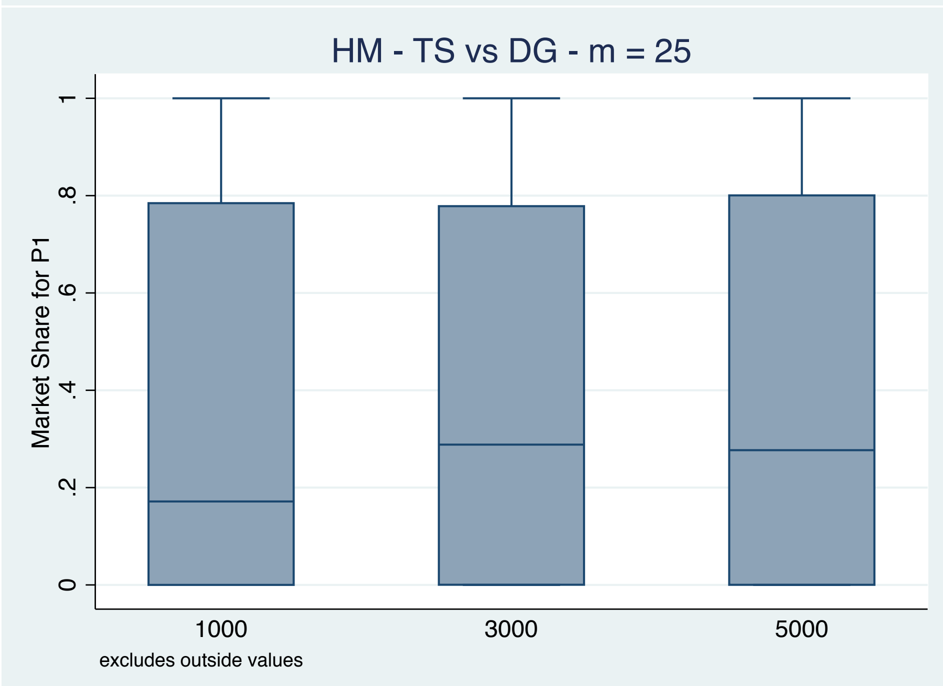
SoftMax


SM - TS vs DG - t = 5000
Market Share for P1
.7
.6
.5
.4
.3
10
25
100
excludes outside values


SM - TS vs DEG - t = 5000
Market Share for P1
.6
.55
.5
.45
.4
10
25
100
excludes outside values

Zooming in on the HardMax results and looking at them over time we get (for memory = 25):





Memory = 100

Take note of the scale on the y-axis on this graph.