

# Competing Bandits Overview

May 29, 2018

## Simulation Details

Considered  $K = 3, 10$ ,  $Memory = 100$  (unless otherwise noted)

**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. High -  $K - 1$  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. Dynamic  $\epsilon$ -greedy with priors of  $Beta(1, 1)$  for every arm and  $\epsilon = 0.05$

**Agent Algorithms considered:**

1. HardMax
2. HardMaxWithRandom ( $\epsilon = 0.05$ )
3. SoftMax ( $\alpha = 30$ )

**Memory Sizes**

1. 100

## Competing Bandits Simulation Procedure

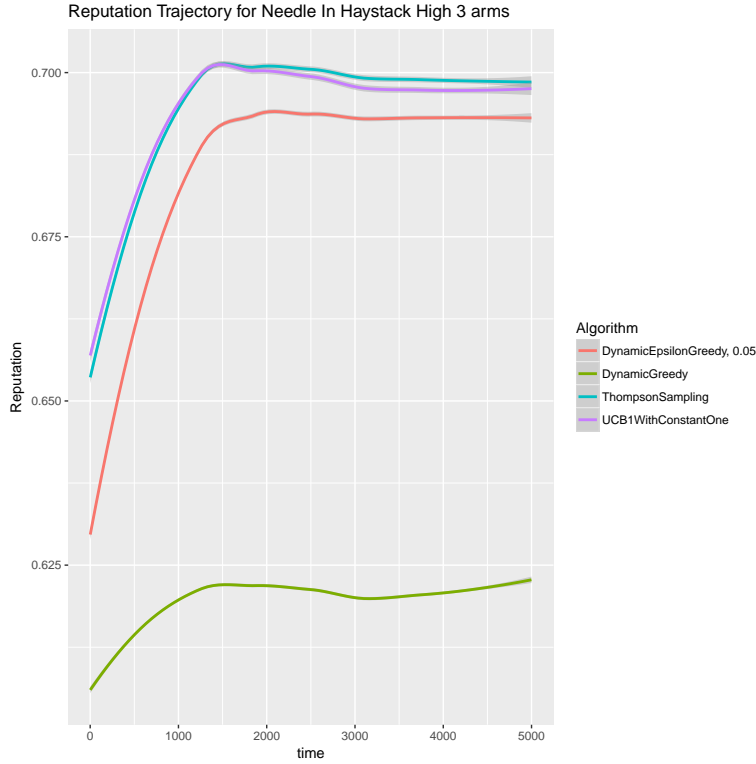
---

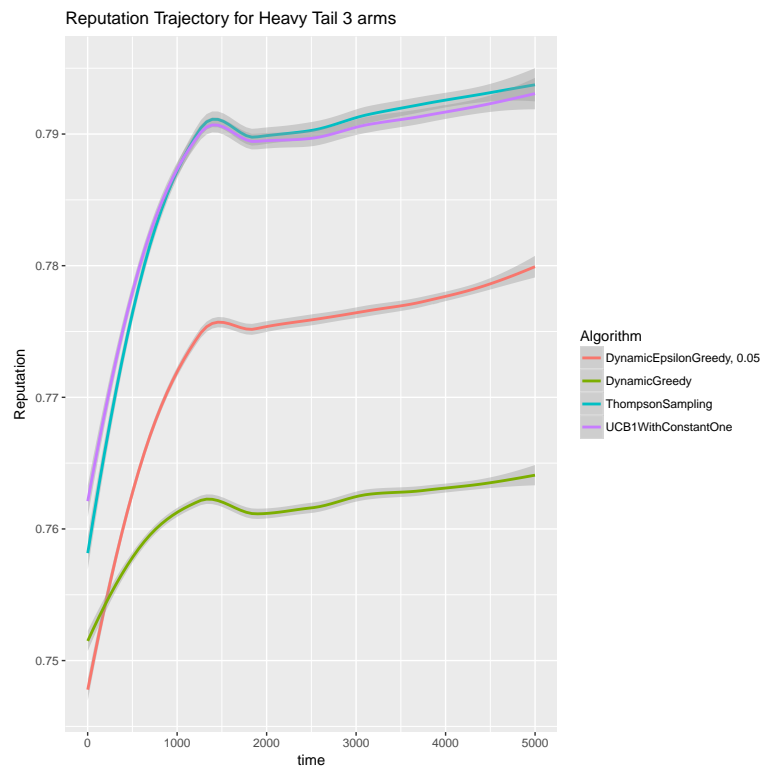
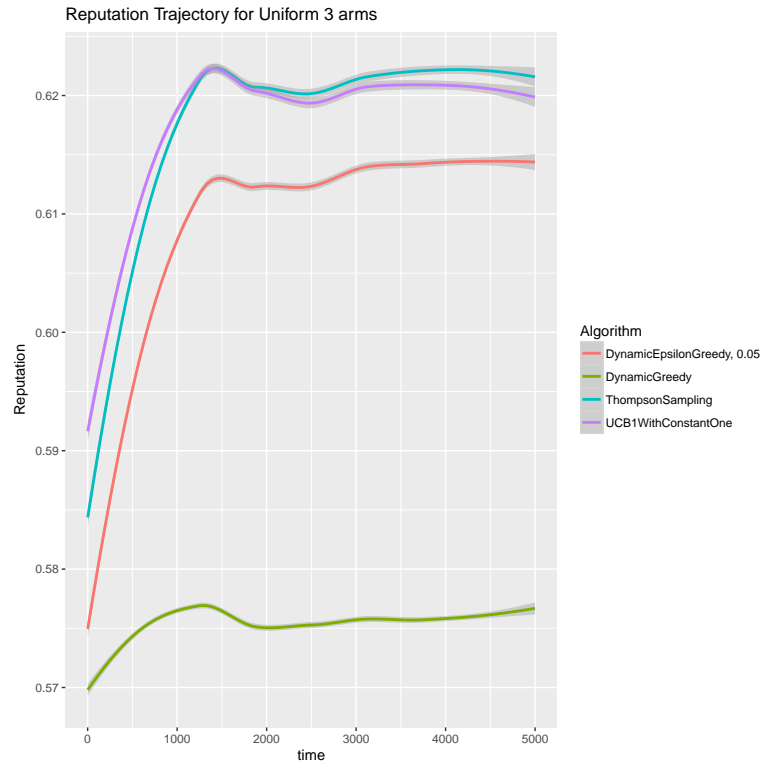
```
1: for Each prior  $p$  do
2:   Generate true distribution from  $p$  (except for needle-in-haystack, just use  $p$  itself)
3:   Generate  $T \times K$  realizations for the arms
4:   for Each agent algorithm  $agentalg$  do
5:     for Each principal algorithm pair  $principalalg1, principalalg2$  do
6:       for  $N$  simulations do
7:         Give principal 2  $X$  free observations (the agents also get these observations)
8:         Give the agents  $k$  observations from each principal
9:         Run simulation for  $T$  periods
10:      end for
11:    end for
12:  end for
13: end for
```

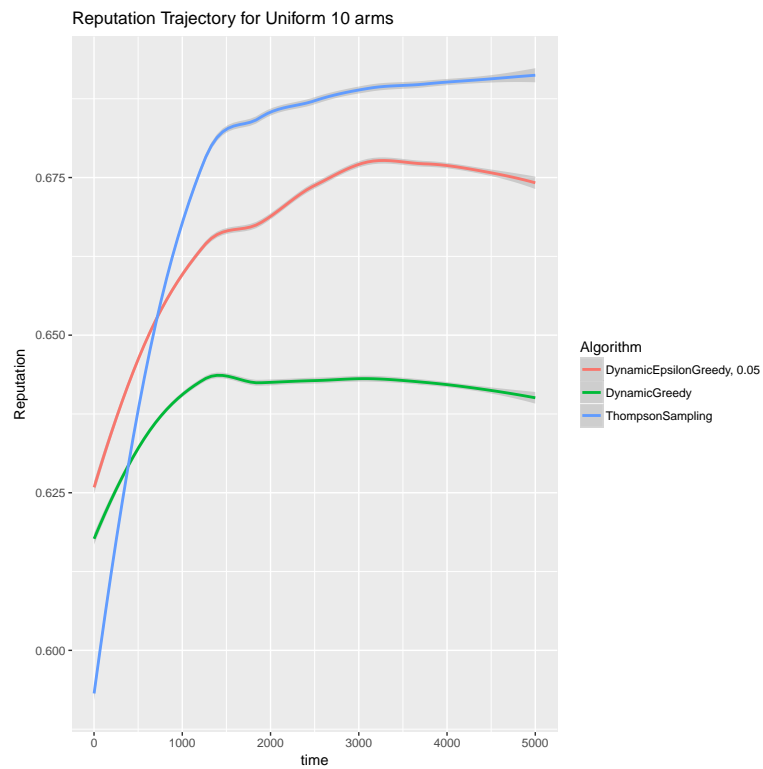
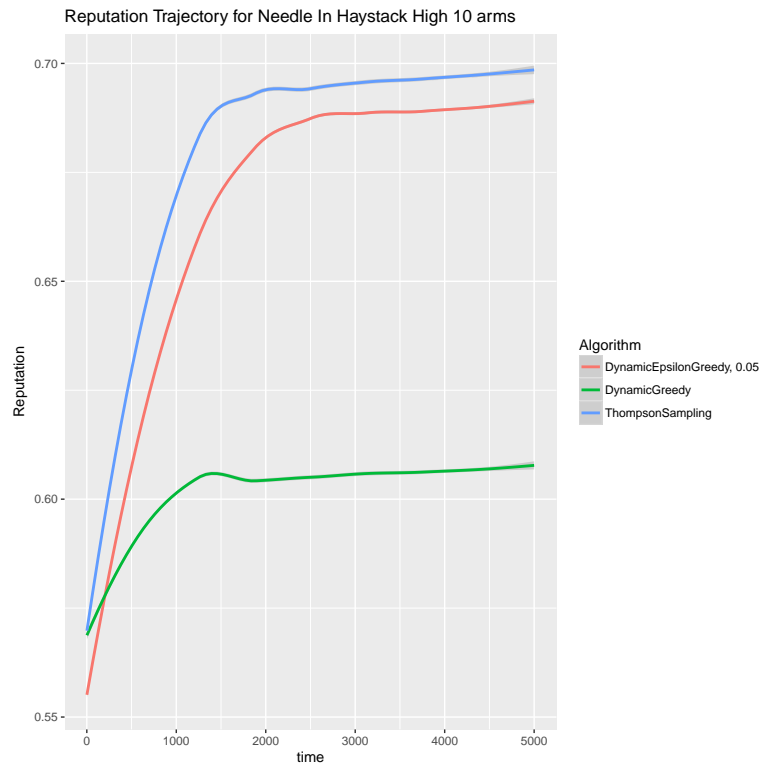
---

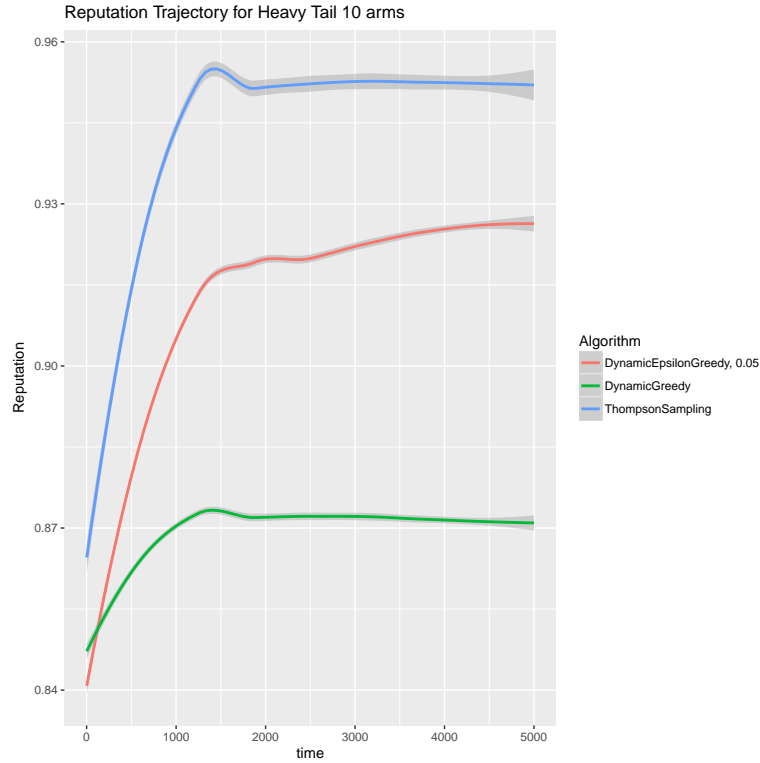
## Preliminary Calibration

First, let's look at the preliminary simulation results on the instances we considered. In these simulations we simply run the selected bandit algorithm on the selected prior (no competition). What is plotted here is the reputation score for a memory size of 100 for each of the priors and each of the algorithms (recall that reputation score is the average of the last 100 observations):



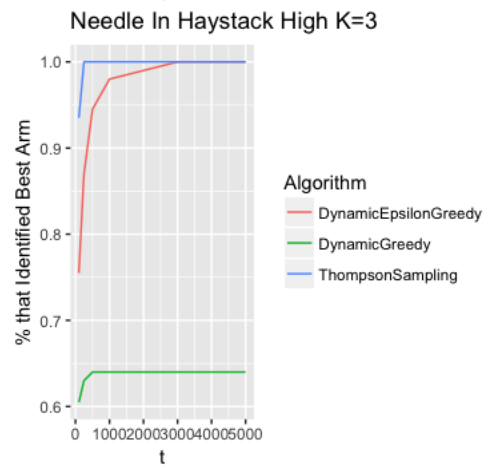


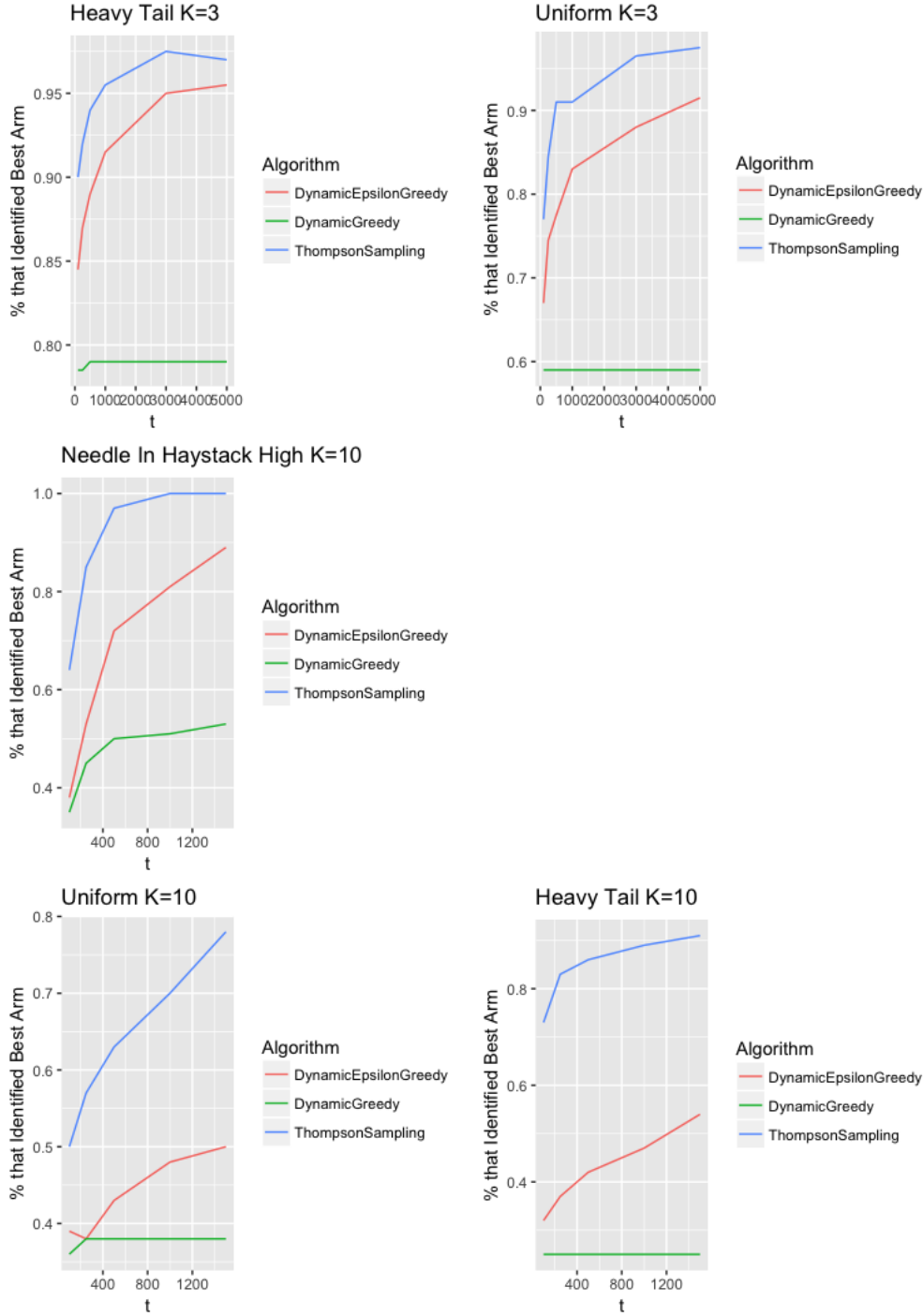




One thing which may be problematic with these plots is that for Heavy Tail and Uniform the means for the instances are randomly drawn for each simulation but these plots just take averages over all of the simulations.

Some additional preliminary calibration plots. These plots show the % of simulations over  $N = 200$  that identified the best arm over time (based on looking at the posterior distributions at a given time):





## Simultaneous Entry

The first set of results we will consider are  $X = 0$  and  $k = 5$  so that there is simultaneous entry. Recall to read the tables:

1. In bold is the average market share for principal 1 and with a 95% confidence interval
2. The var term is the variance of the market shares
3. The “share” line means the % of simulations that ended up with one principal getting more than 90% of the market.

**Question of Interest:** What algorithms should win under different agent response models?

**Overall Conjecture:** For the agent response functions with randomness (HMR and SoftMax), the algorithm performance in the later rounds of the single agent case should dictate which algorithm does better in the long-run in the competing bandits game. For the agent response functions without randomness (HM), the algorithm that does better in the very short-run of the single agent case should do better in the long-run in the competing bandits game.

**Conjecture:** HardMaxWithRandom and SoftMax should eventually lead to the better algorithm winning as long as, on the instance we are considering there is a gap in the long-term reward of the better algorithms.

**Why?** Looking at the preliminary calibration plots, since there is some randomness in the choice rule of the agents, each principal should always be getting some free observations (our parameters such as the  $\alpha$  in SoftMax and  $\epsilon$  in HardMaxWithRandom tune how many expected observations a principal should get for a fixed  $T$ ). Let’s fix attention on the case of ThompsonSampling vs Bayesian/Dynamic Greedy. Looking at each of the preliminary plots we see that ThompsonSampling beats Dynamic / Bayesian Greedy for sufficiently large  $T$ . Thus, if we calibrate the time horizon long enough so that the principal playing Thompson Sampling should, in expectation, get sufficiently many samples to be on the part of the curve where learning is almost done then it should be able to accumulate a higher reputation score (by having better arm selection than DynamicGreedy) and eventually get more of the market.

However, this is not always the case. For Heavy Tail with  $K = 3$  we have that  $DG > DEG$ . This does not hold however for  $K = 10$ . This is possibly because learning is “easy” for DG in  $K = 3$  (it identifies the best arm very quickly and in many cases - see preliminary calibration plots). As well, since DEG engages in non-adaptive exploration and Heavy Tail possibly has some very bad arms (i.e. means close to 0), this non-adaptive exploration could have large effects on reputation. Learning appears to be difficult when  $K = 10$  and the reputation gap appears to be slightly higher between DEG and DG but it is unclear what leads to the results we see.

**Conjecture:** Under HardMax, the early rounds should dictate almost everything since we suppose the algorithms start with little initial information and thus the difference between the algorithms should not make much of a difference. Since the rewards are randomly drawn, lucky early round draws may dictate the course of the game. Thus a conjecture is that the mean market share will be 50/50.

This also does not hold. See the puzzles section after the results.

Results for t= 2000 Heavy Tail Memory= 100 K= 3

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.42</b> +/- 0.06 Var: 0.23 Share: 93 %	<b>0.45</b> +/- 0.06 Var: 0.24 Share: 96 %	<b>0.57</b> +/- 0.06 Var: 0.23 Share: 92 %			
HMR	<b>0.47</b> +/- 0.05 Var: 0.16 Share: 64 %	<b>0.44</b> +/- 0.05 Var: 0.17 Share: 70 %	<b>0.54</b> +/- 0.05 Var: 0.14 Share: 52 %			
SM	<b>0.57</b> +/- 0.02 Var: 0.035 Share: 3.2 %	<b>0.51</b> +/- 0.03 Var: 0.042 Share: 2.8 %	<b>0.53</b> +/- 0.03 Var: 0.053 Share: 4.4 %			

Results for t= 2000 Heavy Tail Memory= 100 K= 10

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.33</b> +/- 0.05 Var: 0.22 Share: 100 %	<b>0.34</b> +/- 0.05 Var: 0.22 Share: 99 %	<b>0.53</b> +/- 0.05 Var: 0.22 Share: 86 %			
HMR	<b>0.35</b> +/- 0.05 Var: 0.19 Share: 92 %	<b>0.43</b> +/- 0.05 Var: 0.18 Share: 79 %	<b>0.6</b> +/- 0.04 Var: 0.14 Share: 56 %			
SM	<b>0.61</b> +/- 0.03 Var: 0.057 Share: 5 %	<b>0.56</b> +/- 0.03 Var: 0.075 Share: 5.3 %	<b>0.54</b> +/- 0.02 Var: 0.048 Share: 1.3 %			

Results for t= 15000 Heavy Tail Memory= 100 K= 3

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.43</b> +/- 0.06 Var: 0.23 Share: 95 %	<b>0.47</b> +/- 0.06 Var: 0.24 Share: 95 %	<b>0.56</b> +/- 0.06 Var: 0.22 Share: 89 %			
HMR	<b>0.67</b> +/- 0.02 Var: 0.034 Share: 14 %	<b>0.55</b> +/- 0.03 Var: 0.053 Share: 16 %	<b>0.6</b> +/- 0.04 Var: 0.082 Share: 29 %			
SM	<b>0.6</b> +/- 0.01 Var: 0.0066 Share: 0 %	<b>0.56</b> +/- 0.02 Var: 0.022 Share: 0.4 %	<b>0.52</b> +/- 0.03 Var: 0.041 Share: 2.4 %			



Results for t= 15000 Heavy Tail Memory= 100 K= 10

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.33</b> +/- 0.05 Var: 0.22 Share: 100 %	<b>0.35</b> +/- 0.05 Var: 0.22 Share: 98 %	<b>0.52</b> +/- 0.05 Var: 0.23 Share: 91 %			
HMR	<b>0.73</b> +/- 0.02 Var: 0.038 Share: 28 %	<b>0.67</b> +/- 0.03 Var: 0.088 Share: 36 %	<b>0.53</b> +/- 0.04 Var: 0.12 Share: 40 %			
SM	<b>0.73</b> +/- 0.01 Var: 0.011 Share: 0.33 %	<b>0.68</b> +/- 0.02 Var: 0.038 Share: 4 %	<b>0.46</b> +/- 0.03 Var: 0.065 Share: 2 %			

Results for t= 2000 Uniform Memory= 100 K= 3

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.5</b> +/- 0.06 Var: 0.23 Share: 93 %	<b>0.45</b> +/- 0.06 Var: 0.24 Share: 95 %	<b>0.5</b> +/- 0.06 Var: 0.23 Share: 90 %			
HMR	<b>0.48</b> +/- 0.05 Var: 0.14 Share: 51 %	<b>0.51</b> +/- 0.05 Var: 0.14 Share: 52 %	<b>0.5</b> +/- 0.04 Var: 0.12 Share: 40 %			
SM	<b>0.53</b> +/- 0.03 Var: 0.048 Share: 2 %	<b>0.52</b> +/- 0.03 Var: 0.054 Share: 2.4 %	<b>0.49</b> +/- 0.03 Var: 0.056 Share: 2.8 %			

Results for t= 2000 Uniform Memory= 100 K= 10

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.49</b> +/- 0.06 Var: 0.24 Share: 97 %	<b>0.48</b> +/- 0.06 Var: 0.24 Share: 96 %	<b>0.51</b> +/- 0.05 Var: 0.22 Share: 88 %			
HMR	<b>0.44</b> +/- 0.05 Var: 0.17 Share: 71 %	<b>0.48</b> +/- 0.05 Var: 0.17 Share: 71 %	<b>0.49</b> +/- 0.04 Var: 0.13 Share: 47 %			
SM	<b>0.41</b> +/- 0.03 Var: 0.061 Share: 1.7 %	<b>0.46</b> +/- 0.03 Var: 0.07 Share: 4 %	<b>0.47</b> +/- 0.03 Var: 0.051 Share: 2.3 %			

Results for t= 15000 Uniform Memory= 100 K= 3

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.52</b> +/- 0.06 Var: 0.22 Share: 88 %	<b>0.47</b> +/- 0.06 Var: 0.23 Share: 90 %	<b>0.47</b> +/- 0.06 Var: 0.23 Share: 89 %			
HMR	<b>0.56</b> +/- 0.02 Var: 0.036 Share: 7.2 %	<b>0.59</b> +/- 0.03 Var: 0.055 Share: 18 %	<b>0.48</b> +/- 0.03 Var: 0.067 Share: 19 %			
SM	<b>0.55</b> +/- 0.01 Var: 0.0088 Share: 0 %	<b>0.57</b> +/- 0.02 Var: 0.026 Share: 0.8 %	<b>0.45</b> +/- 0.02 Var: 0.035 Share: 0.8 %			

Results for t= 15000 Uniform Memory= 100 K= 10

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.49</b> +/- 0.06 Var: 0.25 Share: 98 %	<b>0.51</b> +/- 0.06 Var: 0.24 Share: 95 %	<b>0.5</b> +/- 0.05 Var: 0.22 Share: 87 %			
HMR	<b>0.53</b> +/- 0.03 Var: 0.069 Share: 15 %	<b>0.61</b> +/- 0.03 Var: 0.079 Share: 23 %	<b>0.42</b> +/- 0.03 Var: 0.074 Share: 25 %			
SM	<b>0.54</b> +/- 0.01 Var: 0.016 Share: 0 %	<b>0.59</b> +/- 0.02 Var: 0.036 Share: 0.33 %	<b>0.42</b> +/- 0.02 Var: 0.038 Share: 1 %			

Results for t= 2000 Needle In Haystack High Memory= 100 K= 3

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.52</b> +/- 0.06 Var: 0.24 Share: 96 %	<b>0.49</b> +/- 0.06 Var: 0.24 Share: 94 %	<b>0.43</b> +/- 0.06 Var: 0.23 Share: 92 %			
HMR	<b>0.48</b> +/- 0.05 Var: 0.16 Share: 62 %	<b>0.59</b> +/- 0.05 Var: 0.14 Share: 56 %	<b>0.48</b> +/- 0.05 Var: 0.15 Share: 58 %			
SM	<b>0.55</b> +/- 0.03 Var: 0.057 Share: 3.2 %	<b>0.53</b> +/- 0.03 Var: 0.063 Share: 2 %	<b>0.45</b> +/- 0.03 Var: 0.075 Share: 6.4 %			

Results for t= 2000 Needle In Haystack High Memory= 100 K= 10

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.53</b> +/- 0.06 Var: 0.24 Share: 96 %	<b>0.66</b> +/- 0.05 Var: 0.21 Share: 92 %	<b>0.45</b> +/- 0.05 Var: 0.22 Share: 85 %			
HMR	<b>0.6</b> +/- 0.05 Var: 0.18 Share: 80 %	<b>0.67</b> +/- 0.04 Var: 0.14 Share: 67 %	<b>0.43</b> +/- 0.04 Var: 0.15 Share: 57 %			
SM	<b>0.55</b> +/- 0.03 Var: 0.091 Share: 5.3 %	<b>0.57</b> +/- 0.03 Var: 0.093 Share: 5.3 %	<b>0.45</b> +/- 0.03 Var: 0.082 Share: 5.7 %			

Results for t= 15000 Needle In Haystack High Memory= 100 K= 3

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.54</b> +/- 0.06 Var: 0.24 Share: 94 %	<b>0.5</b> +/- 0.06 Var: 0.23 Share: 92 %	<b>0.42</b> +/- 0.06 Var: 0.23 Share: 92 %			
HMR	<b>0.57</b> +/- 0.02 Var: 0.04 Share: 11 %	<b>0.64</b> +/- 0.03 Var: 0.067 Share: 33 %	<b>0.44</b> +/- 0.04 Var: 0.09 Share: 34 %			
SM	<b>0.54</b> +/- 0.01 Var: 0.006 Share: 0 %	<b>0.59</b> +/- 0.02 Var: 0.034 Share: 0.8 %	<b>0.4</b> +/- 0.03 Var: 0.046 Share: 2 %			

Results for t= 15000 Needle In Haystack High Memory= 100 K= 10

	TS vs DEG	TS vs DG	DG vs DEG	TS vs TS	DEG vs DEG	DG vs DG
HM	<b>0.54</b> +/- 0.06 Var: 0.25 Share: 99 %	<b>0.7</b> +/- 0.05 Var: 0.2 Share: 95 %	<b>0.4</b> +/- 0.05 Var: 0.22 Share: 92 %			
HMR	<b>0.61</b> +/- 0.04 Var: 0.11 Share: 45 %	<b>0.68</b> +/- 0.04 Var: 0.1 Share: 52 %	<b>0.38</b> +/- 0.04 Var: 0.13 Share: 59 %			
SM	<b>0.58</b> +/- 0.02 Var: 0.026 Share: 0 %	<b>0.65</b> +/- 0.03 Var: 0.052 Share: 0.67 %	<b>0.38</b> +/- 0.03 Var: 0.072 Share: 1 %			

**Puzzles:** There are two puzzles with respect to HardMax. Why does Needle in a Haystack ( $K = 10$ ) lead to  $TS > DG$  and why does Heavy Tail (both  $K = 3$  and  $K = 10$ ) lead to  $DG, DEG > TS$ ? Clearly the original conjecture is flawed in some way. What are things that stand out? Under Heavy Tail we see that the difference in algorithm performance without competition is not that large (looking at reputation plots) and that  $DG$  identifies the best arm in a large fraction of runs ( $\approx 0.8$ ). Under Needle in a Haystack we see that the difference between algorithm performance is large and that  $TS$  identifies the best arm with only a few samples, but what in particular leads it to win under competition with the HardMax response function?

In general, some properties of the instances and the algorithms seem to drive who wins when. We do not have consistency as in the single instance case where  $TS > DEG > DG$ , but it's unclear under what conditions we have who wins. Should think more about the properties of the instances.

## Incumbent Experiment

Now, we change  $X = 200$ . In words, we give principal 2  $X$  free observations (the “incumbent”) and this updates her information set and reputation amongst the agents. In this setting what should the incumbent do? What should the entrant do? Does this depend on the agent model? I would conjecture that when we have sufficient randomness in the agent model the same conclusions as before should hold - namely that the entrant, even if getting beat upon entry, should eventually get enough “free” observations that she can catch up and take more of the market. However, the time horizon, in contrast to the simultaneous entry case, seems like it may be way longer! So far, most of the simulations have simply looked at  $t = 2000$ , but it may make sense to look at longer time horizons due to the conjecture from above.

What about with HardMax? The only intuition I have is for the incumbent. For the entrant, it is not clear to me what the entrant should do. For the incumbent, in our simulations we fix a  $X$  when the entrant will come in and the incumbent should want to maximize both the information that she accrues before then and the reputation that she amongst the agents. However, this seems exactly like the standard exploration exploitation dilemma where reputation is simply the cumulative reward and “smart” algorithms like Thompson Sampling ought to be better than the “dumb” algorithms at maximizing information gain while at the same time optimizing cumulative reward. Thus, I expect that Thompson Sampling should be in a better starting position at the start of the competing bandits game (we can view the incumbent experiment as being the simultaneous entry game where the incumbent starts with a head start in terms of information and reputation) and thus control more of the market. This seems to be the case in the reported simulations, both with large  $X = 200$  and  $X = 100$ . Below are the result for  $X = 200$  ( $X = 100$  is in the appendix figures)

To read the tables: the columns represent the algorithm played by the incumbent and the row represents the algorithm played by the entrant. For instance, looking at the second cell in the first row we have that the Incumbent played Dynamic (Bayesian) Epsilon Greedy and the entrant played Thompson Sampling and the reported mean is the average market share for the entrant.

In this version of the simulations we additionally have simulations where the information, reputation, or both are artificially erased from the incumbent. The timing of events is as follows:

1. The incumbent gets  $X$  free observations

2. Either information gains are erased (the posterior beliefs of the incumbent are reset to the fake priors) or the reputation gains are erased (the agents' reputation score for the incumbent is reset) or both
3. Both the incumbent and the entrant get  $k = 5$  free observations

## Results for HardMax t = 2000 Heavy Tail Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.069</b> +/- 0.024 0.046 (0.04, 0.055) Extreme Shares: 92 %	<b>0.11</b> +/- 0.031 0.072 (0.062, 0.085) Extreme Shares: 90 %	<b>0.1</b> +/- 0.032 0.081 (0.069, 0.095) Extreme Shares: 94 %
Dynamic Epsilon Greedy	<b>0.096</b> +/- 0.025 0.047 (0.041, 0.056) Extreme Shares: 81 %	<b>0.14</b> +/- 0.032 0.081 (0.069, 0.095) Extreme Shares: 80 %	<b>0.15</b> +/- 0.033 0.084 (0.072, 0.099) Extreme Shares: 78 %
DynamicGreedy	<b>0.12</b> +/- 0.029 0.065 (0.055, 0.076) Extreme Shares: 79 %	<b>0.19</b> +/- 0.038 0.11 (0.094, 0.13) Extreme Shares: 80 %	<b>0.19</b> +/- 0.035 0.093 (0.08, 0.11) Extreme Shares: 72 %

## Results for HardMax t = 2000 Heavy Tail Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.11</b> +/- 0.033 0.086 (0.074, 0.1) Extreme Shares: 96 %	<b>0.2</b> +/- 0.043 0.14 (0.12, 0.17) Extreme Shares: 93 %	<b>0.19</b> +/- 0.042 0.14 (0.12, 0.17) Extreme Shares: 95 %
Dynamic Epsilon Greedy	<b>0.18</b> +/- 0.038 0.11 (0.098, 0.14) Extreme Shares: 85 %	<b>0.22</b> +/- 0.042 0.14 (0.12, 0.16) Extreme Shares: 84 %	<b>0.27</b> +/- 0.046 0.16 (0.14, 0.19) Extreme Shares: 81 %
DynamicGreedy	<b>0.2</b> +/- 0.041 0.13 (0.11, 0.15) Extreme Shares: 84 %	<b>0.3</b> +/- 0.048 0.18 (0.15, 0.21) Extreme Shares: 85 %	<b>0.26</b> +/- 0.043 0.14 (0.12, 0.17) Extreme Shares: 77 %

## Results for HardMax t = 2000 Heavy Tail Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.051</b> +/- 0.021 0.034 (0.029, 0.04) Extreme Shares: 94 %	<b>0.13</b> +/- 0.036 0.099 (0.085, 0.12) Extreme Shares: 93 %	<b>0.14</b> +/- 0.037 0.11 (0.093, 0.13) Extreme Shares: 95 %
Dynamic Epsilon Greedy	<b>0.12</b> +/- 0.029 0.065 (0.056, 0.077) Extreme Shares: 80 %	<b>0.14</b> +/- 0.031 0.074 (0.064, 0.088) Extreme Shares: 79 %	<b>0.15</b> +/- 0.034 0.09 (0.077, 0.11) Extreme Shares: 82 %
DynamicGreedy	<b>0.19</b> +/- 0.04 0.12 (0.11, 0.15) Extreme Shares: 86 %	<b>0.23</b> +/- 0.041 0.13 (0.11, 0.15) Extreme Shares: 76 %	<b>0.18</b> +/- 0.034 0.087 (0.075, 0.1) Extreme Shares: 71 %

## Results for HardMaxWithRandom t = 2000 Heavy Tail Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.14</b> +/- 0.023 0.042 (0.036, 0.05) Extreme Shares: 70 %	<b>0.26</b> +/- 0.033 0.087 (0.074, 0.1) Extreme Shares: 54 %	<b>0.26</b> +/- 0.038 0.11 (0.098, 0.13) Extreme Shares: 67 %
Dynamic Epsilon Greedy	<b>0.19</b> +/- 0.026 0.053 (0.045, 0.062) Extreme Shares: 57 %	<b>0.3</b> +/- 0.032 0.078 (0.066, 0.092) Extreme Shares: 41 %	<b>0.3</b> +/- 0.035 0.093 (0.08, 0.11) Extreme Shares: 48 %
DynamicGreedy	<b>0.27</b> +/- 0.032 0.077 (0.066, 0.091) Extreme Shares: 49 %	<b>0.4</b> +/- 0.037 0.11 (0.091, 0.13) Extreme Shares: 42 %	<b>0.35</b> +/- 0.031 0.076 (0.065, 0.089) Extreme Shares: 29 %

## Results for HardMaxWithRandom t = 2000 Heavy Tail Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.17</b> +/- 0.029 0.065 (0.056, 0.077) Extreme Shares: 72 %	<b>0.3</b> +/- 0.036 0.099 (0.085, 0.12) Extreme Shares: 53 %	<b>0.31</b> +/- 0.041 0.13 (0.11, 0.15) Extreme Shares: 63 %
Dynamic Epsilon Greedy	<b>0.23</b> +/- 0.031 0.073 (0.062, 0.086) Extreme Shares: 57 %	<b>0.31</b> +/- 0.034 0.092 (0.079, 0.11) Extreme Shares: 45 %	<b>0.32</b> +/- 0.037 0.11 (0.091, 0.13) Extreme Shares: 50 %
DynamicGreedy	<b>0.3</b> +/- 0.034 0.087 (0.075, 0.1) Extreme Shares: 46 %	<b>0.44</b> +/- 0.039 0.12 (0.099, 0.14) Extreme Shares: 41 %	<b>0.38</b> +/- 0.035 0.096 (0.083, 0.11) Extreme Shares: 35 %

## Results for HardMaxWithRandom t = 2000 Heavy Tail Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.21</b> +/- 0.032 0.081 (0.069, 0.096) Extreme Shares: 66 %	<b>0.28</b> +/- 0.037 0.1 (0.089, 0.12) Extreme Shares: 60 %	<b>0.29</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 64 %
Dynamic Epsilon Greedy	<b>0.25</b> +/- 0.033 0.087 (0.074, 0.1) Extreme Shares: 57 %	<b>0.34</b> +/- 0.034 0.091 (0.078, 0.11) Extreme Shares: 40 %	<b>0.31</b> +/- 0.035 0.097 (0.083, 0.12) Extreme Shares: 49 %
DynamicGreedy	<b>0.3</b> +/- 0.036 0.099 (0.085, 0.12) Extreme Shares: 52 %	<b>0.4</b> +/- 0.037 0.11 (0.093, 0.13) Extreme Shares: 40 %	<b>0.35</b> +/- 0.033 0.086 (0.074, 0.1) Extreme Shares: 35 %

## Results for SoftMax t = 2000 Heavy Tail Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.39</b> +/- 0.015 0.016 (0.014, 0.019) Extreme Shares: 0.33 %	<b>0.51</b> +/- 0.018 0.026 (0.022, 0.03) Extreme Shares: 0.67 %	<b>0.48</b> +/- 0.024 0.046 (0.039, 0.054) Extreme Shares: 5 %
Dynamic Epsilon Greedy	<b>0.34</b> +/- 0.017 0.022 (0.019, 0.026) Extreme Shares: 5 %	<b>0.43</b> +/- 0.018 0.024 (0.02, 0.028) Extreme Shares: 3 %	<b>0.43</b> +/- 0.023 0.042 (0.036, 0.05) Extreme Shares: 1.3 %
DynamicGreedy	<b>0.4</b> +/- 0.02 0.031 (0.027, 0.037) Extreme Shares: 5.7 %	<b>0.49</b> +/- 0.023 0.039 (0.034, 0.046) Extreme Shares: 4.3 %	<b>0.47</b> +/- 0.016 0.02 (0.018, 0.024) Extreme Shares: 1 %

## Results for SoftMax t = 2000 Heavy Tail Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.39</b> +/- 0.015 0.017 (0.015, 0.021) Extreme Shares: 0.33 %	<b>0.52</b> +/- 0.019 0.027 (0.023, 0.032) Extreme Shares: 0.67 %	<b>0.47</b> +/- 0.025 0.048 (0.041, 0.057) Extreme Shares: 1.7 %
Dynamic Epsilon Greedy	<b>0.34</b> +/- 0.018 0.024 (0.021, 0.029) Extreme Shares: 2.3 %	<b>0.44</b> +/- 0.019 0.027 (0.023, 0.032) Extreme Shares: 0.67 %	<b>0.41</b> +/- 0.023 0.04 (0.034, 0.048) Extreme Shares: 2.3 %
DynamicGreedy	<b>0.41</b> +/- 0.021 0.034 (0.029, 0.04) Extreme Shares: 4.3 %	<b>0.5</b> +/- 0.022 0.039 (0.033, 0.046) Extreme Shares: 3.3 %	<b>0.46</b> +/- 0.018 0.024 (0.021, 0.028) Extreme Shares: 1.3 %

## Results for SoftMax t = 2000 Heavy Tail Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.42</b> +/- 0.015 0.018 (0.015, 0.021) Extreme Shares: 0 %	<b>0.53</b> +/- 0.02 0.031 (0.027, 0.037) Extreme Shares: 0.67 %	<b>0.48</b> +/- 0.024 0.043 (0.037, 0.051) Extreme Shares: 3.3 %
Dynamic Epsilon Greedy	<b>0.36</b> +/- 0.02 0.03 (0.026, 0.035) Extreme Shares: 4.7 %	<b>0.45</b> +/- 0.019 0.029 (0.024, 0.034) Extreme Shares: 2 %	<b>0.43</b> +/- 0.023 0.042 (0.036, 0.049) Extreme Shares: 0.67 %
DynamicGreedy	<b>0.42</b> +/- 0.022 0.037 (0.032, 0.044) Extreme Shares: 5 %	<b>0.5</b> +/- 0.023 0.041 (0.035, 0.049) Extreme Shares: 3.3 %	<b>0.48</b> +/- 0.017 0.023 (0.02, 0.027) Extreme Shares: 0.67 %



## Results for HardMax t = 2000 Uniform Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.11</b> +/- 0.029 0.067 (0.058, 0.08) Extreme Shares: 86 %	<b>0.17</b> +/- 0.035 0.097 (0.084, 0.12) Extreme Shares: 79 %	<b>0.21</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 77 %
Dynamic Epsilon Greedy	<b>0.15</b> +/- 0.032 0.082 (0.07, 0.096) Extreme Shares: 80 %	<b>0.18</b> +/- 0.035 0.092 (0.079, 0.11) Extreme Shares: 74 %	<b>0.19</b> +/- 0.036 0.1 (0.087, 0.12) Extreme Shares: 76 %
DynamicGreedy	<b>0.15</b> +/- 0.033 0.082 (0.071, 0.097) Extreme Shares: 79 %	<b>0.18</b> +/- 0.035 0.097 (0.083, 0.11) Extreme Shares: 78 %	<b>0.17</b> +/- 0.032 0.08 (0.069, 0.094) Extreme Shares: 71 %

## Results for HardMax t = 2000 Uniform Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.34</b> +/- 0.05 0.2 (0.17, 0.23) Extreme Shares: 87 %	<b>0.32</b> +/- 0.049 0.19 (0.16, 0.22) Extreme Shares: 86 %	<b>0.32</b> +/- 0.049 0.18 (0.16, 0.22) Extreme Shares: 85 %
Dynamic Epsilon Greedy	<b>0.32</b> +/- 0.048 0.18 (0.16, 0.21) Extreme Shares: 85 %	<b>0.33</b> +/- 0.048 0.18 (0.15, 0.21) Extreme Shares: 81 %	<b>0.3</b> +/- 0.046 0.16 (0.14, 0.19) Extreme Shares: 76 %
DynamicGreedy	<b>0.34</b> +/- 0.05 0.19 (0.17, 0.23) Extreme Shares: 84 %	<b>0.32</b> +/- 0.047 0.17 (0.15, 0.21) Extreme Shares: 80 %	<b>0.31</b> +/- 0.047 0.17 (0.15, 0.2) Extreme Shares: 78 %

## Results for HardMax t = 2000 Uniform Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.19</b> +/- 0.039 0.12 (0.099, 0.14) Extreme Shares: 83 %	<b>0.18</b> +/- 0.038 0.11 (0.098, 0.14) Extreme Shares: 85 %	<b>0.2</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 82 %
Dynamic Epsilon Greedy	<b>0.2</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 80 %	<b>0.23</b> +/- 0.04 0.12 (0.11, 0.15) Extreme Shares: 76 %	<b>0.23</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 74 %
DynamicGreedy	<b>0.2</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 81 %	<b>0.2</b> +/- 0.037 0.11 (0.093, 0.13) Extreme Shares: 77 %	<b>0.23</b> +/- 0.038 0.11 (0.096, 0.13) Extreme Shares: 70 %

## Results for HardMaxWithRandom t = 2000 Uniform Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.25</b> +/- 0.03 0.07 (0.06, 0.083) Extreme Shares: 48 %	<b>0.31</b> +/- 0.035 0.094 (0.08, 0.11) Extreme Shares: 45 %	<b>0.35</b> +/- 0.037 0.11 (0.091, 0.13) Extreme Shares: 48 %
Dynamic Epsilon Greedy	<b>0.3</b> +/- 0.032 0.08 (0.069, 0.095) Extreme Shares: 41 %	<b>0.35</b> +/- 0.035 0.094 (0.08, 0.11) Extreme Shares: 39 %	<b>0.39</b> +/- 0.033 0.087 (0.075, 0.1) Extreme Shares: 30 %
DynamicGreedy	<b>0.3</b> +/- 0.033 0.087 (0.074, 0.1) Extreme Shares: 42 %	<b>0.31</b> +/- 0.031 0.075 (0.065, 0.089) Extreme Shares: 38 %	<b>0.38</b> +/- 0.031 0.073 (0.063, 0.087) Extreme Shares: 28 %

## Results for HardMaxWithRandom t = 2000 Uniform Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.32</b> +/- 0.036 0.1 (0.087, 0.12) Extreme Shares: 49 %	<b>0.36</b> +/- 0.036 0.1 (0.088, 0.12) Extreme Shares: 42 %	<b>0.42</b> +/- 0.04 0.12 (0.11, 0.15) Extreme Shares: 48 %
Dynamic Epsilon Greedy	<b>0.35</b> +/- 0.036 0.098 (0.084, 0.12) Extreme Shares: 42 %	<b>0.36</b> +/- 0.036 0.1 (0.087, 0.12) Extreme Shares: 43 %	<b>0.39</b> +/- 0.036 0.1 (0.086, 0.12) Extreme Shares: 37 %
DynamicGreedy	<b>0.35</b> +/- 0.035 0.094 (0.08, 0.11) Extreme Shares: 39 %	<b>0.39</b> +/- 0.037 0.11 (0.09, 0.12) Extreme Shares: 39 %	<b>0.4</b> +/- 0.035 0.093 (0.079, 0.11) Extreme Shares: 33 %

## Results for HardMaxWithRandom t = 2000 Uniform Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.34</b> +/- 0.036 0.1 (0.086, 0.12) Extreme Shares: 46 %	<b>0.33</b> +/- 0.037 0.11 (0.092, 0.13) Extreme Shares: 49 %	<b>0.38</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 46 %
Dynamic Epsilon Greedy	<b>0.33</b> +/- 0.036 0.1 (0.086, 0.12) Extreme Shares: 48 %	<b>0.36</b> +/- 0.036 0.1 (0.087, 0.12) Extreme Shares: 40 %	<b>0.39</b> +/- 0.036 0.099 (0.085, 0.12) Extreme Shares: 37 %
DynamicGreedy	<b>0.35</b> +/- 0.036 0.098 (0.084, 0.12) Extreme Shares: 42 %	<b>0.36</b> +/- 0.036 0.098 (0.084, 0.12) Extreme Shares: 40 %	<b>0.41</b> +/- 0.034 0.088 (0.076, 0.1) Extreme Shares: 28 %

## Results for SoftMax t = 2000 Uniform Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.42</b> +/- 0.019 0.029 (0.025, 0.034) Extreme Shares: 1.3 %	<b>0.46</b> +/- 0.022 0.037 (0.032, 0.044) Extreme Shares: 1 %	<b>0.47</b> +/- 0.026 0.053 (0.045, 0.062) Extreme Shares: 0.33 %
Dynamic Epsilon Greedy	<b>0.4</b> +/- 0.022 0.038 (0.033, 0.045) Extreme Shares: 4.7 %	<b>0.45</b> +/- 0.021 0.035 (0.03, 0.042) Extreme Shares: 1.7 %	<b>0.47</b> +/- 0.024 0.044 (0.037, 0.052) Extreme Shares: 1.3 %
DynamicGreedy	<b>0.41</b> +/- 0.023 0.039 (0.034, 0.047) Extreme Shares: 4 %	<b>0.44</b> +/- 0.023 0.04 (0.034, 0.047) Extreme Shares: 3.3 %	<b>0.48</b> +/- 0.023 0.039 (0.034, 0.046) Extreme Shares: 0.67 %

## Results for SoftMax t = 2000 Uniform Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.44</b> +/- 0.019 0.029 (0.025, 0.034) Extreme Shares: 0 %	<b>0.46</b> +/- 0.022 0.036 (0.031, 0.043) Extreme Shares: 1 %	<b>0.49</b> +/- 0.025 0.047 (0.04, 0.055) Extreme Shares: 2.3 %
Dynamic Epsilon Greedy	<b>0.42</b> +/- 0.019 0.029 (0.025, 0.034) Extreme Shares: 0.33 %	<b>0.46</b> +/- 0.023 0.041 (0.035, 0.049) Extreme Shares: 2 %	<b>0.49</b> +/- 0.024 0.043 (0.037, 0.051) Extreme Shares: 1 %
DynamicGreedy	<b>0.42</b> +/- 0.024 0.043 (0.037, 0.051) Extreme Shares: 4.3 %	<b>0.45</b> +/- 0.024 0.045 (0.038, 0.053) Extreme Shares: 4.7 %	<b>0.46</b> +/- 0.022 0.036 (0.031, 0.042) Extreme Shares: 2 %

## Results for SoftMax t = 2000 Uniform Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.46</b> +/- 0.02 0.03 (0.026, 0.035) Extreme Shares: 0.33 %	<b>0.47</b> +/- 0.022 0.038 (0.033, 0.045) Extreme Shares: 0.67 %	<b>0.51</b> +/- 0.026 0.053 (0.046, 0.063) Extreme Shares: 1.7 %
Dynamic Epsilon Greedy	<b>0.45</b> +/- 0.024 0.043 (0.037, 0.051) Extreme Shares: 2.7 %	<b>0.44</b> +/- 0.022 0.038 (0.033, 0.045) Extreme Shares: 3.3 %	<b>0.48</b> +/- 0.023 0.041 (0.035, 0.048) Extreme Shares: 0.67 %
DynamicGreedy	<b>0.46</b> +/- 0.026 0.053 (0.046, 0.063) Extreme Shares: 3.7 %	<b>0.46</b> +/- 0.023 0.041 (0.035, 0.048) Extreme Shares: 2.7 %	<b>0.48</b> +/- 0.023 0.041 (0.035, 0.048) Extreme Shares: 1.7 %

## Results for HardMax t = 2000 Needle In Haystack High Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.068</b> +/- 0.023 0.04 (0.035, 0.048) Extreme Shares: 89 %	<b>0.14</b> +/- 0.035 0.096 (0.082, 0.11) Extreme Shares: 89 %	<b>0.23</b> +/- 0.044 0.15 (0.13, 0.18) Extreme Shares: 87 %
Dynamic Epsilon Greedy	<b>0.08</b> +/- 0.024 0.044 (0.038, 0.052) Extreme Shares: 87 %	<b>0.14</b> +/- 0.033 0.086 (0.074, 0.1) Extreme Shares: 86 %	<b>0.2</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 80 %
DynamicGreedy	<b>0.093</b> +/- 0.028 0.061 (0.052, 0.072) Extreme Shares: 90 %	<b>0.12</b> +/- 0.032 0.078 (0.067, 0.092) Extreme Shares: 86 %	<b>0.2</b> +/- 0.038 0.11 (0.094, 0.13) Extreme Shares: 76 %

## Results for HardMax t = 2000 Needle In Haystack High Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.31</b> +/- 0.05 0.2 (0.17, 0.23) Extreme Shares: 93 %	<b>0.28</b> +/- 0.049 0.18 (0.16, 0.22) Extreme Shares: 92 %	<b>0.34</b> +/- 0.051 0.2 (0.17, 0.24) Extreme Shares: 90 %
Dynamic Epsilon Greedy	<b>0.27</b> +/- 0.047 0.17 (0.15, 0.2) Extreme Shares: 88 %	<b>0.3</b> +/- 0.049 0.19 (0.16, 0.22) Extreme Shares: 90 %	<b>0.37</b> +/- 0.051 0.2 (0.17, 0.24) Extreme Shares: 86 %
DynamicGreedy	<b>0.26</b> +/- 0.046 0.17 (0.14, 0.2) Extreme Shares: 89 %	<b>0.33</b> +/- 0.05 0.19 (0.17, 0.23) Extreme Shares: 88 %	<b>0.34</b> +/- 0.048 0.18 (0.15, 0.21) Extreme Shares: 80 %

## Results for HardMax t = 2000 Needle In Haystack High Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.14</b> +/- 0.035 0.094 (0.081, 0.11) Extreme Shares: 90 %	<b>0.21</b> +/- 0.044 0.15 (0.13, 0.18) Extreme Shares: 94 %	<b>0.31</b> +/- 0.047 0.17 (0.15, 0.21) Extreme Shares: 82 %
Dynamic Epsilon Greedy	<b>0.15</b> +/- 0.036 0.099 (0.085, 0.12) Extreme Shares: 88 %	<b>0.2</b> +/- 0.04 0.13 (0.11, 0.15) Extreme Shares: 85 %	<b>0.22</b> +/- 0.04 0.13 (0.11, 0.15) Extreme Shares: 78 %
DynamicGreedy	<b>0.15</b> +/- 0.037 0.11 (0.091, 0.13) Extreme Shares: 93 %	<b>0.15</b> +/- 0.034 0.092 (0.079, 0.11) Extreme Shares: 86 %	<b>0.19</b> +/- 0.036 0.1 (0.086, 0.12) Extreme Shares: 73 %

## Results for HardMaxWithRandom t = 2000 Needle In Haystack High Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.18</b> +/- 0.026 0.053 (0.046, 0.063) Extreme Shares: 61 %	<b>0.26</b> +/- 0.035 0.095 (0.081, 0.11) Extreme Shares: 60 %	<b>0.44</b> +/- 0.044 0.15 (0.13, 0.18) Extreme Shares: 61 %
Dynamic Epsilon Greedy	<b>0.23</b> +/- 0.031 0.072 (0.062, 0.085) Extreme Shares: 57 %	<b>0.3</b> +/- 0.034 0.092 (0.079, 0.11) Extreme Shares: 51 %	<b>0.45</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 42 %
DynamicGreedy	<b>0.26</b> +/- 0.033 0.083 (0.071, 0.098) Extreme Shares: 54 %	<b>0.3</b> +/- 0.035 0.094 (0.08, 0.11) Extreme Shares: 51 %	<b>0.39</b> +/- 0.035 0.097 (0.083, 0.11) Extreme Shares: 36 %

## Results for HardMaxWithRandom t = 2000 Needle In Haystack High Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.3</b> +/- 0.036 0.1 (0.086, 0.12) Extreme Shares: 55 %	<b>0.34</b> +/- 0.041 0.13 (0.11, 0.15) Extreme Shares: 59 %	<b>0.49</b> +/- 0.044 0.15 (0.13, 0.18) Extreme Shares: 61 %
Dynamic Epsilon Greedy	<b>0.32</b> +/- 0.037 0.11 (0.091, 0.13) Extreme Shares: 51 %	<b>0.36</b> +/- 0.038 0.11 (0.095, 0.13) Extreme Shares: 49 %	<b>0.47</b> +/- 0.039 0.12 (0.1, 0.14) Extreme Shares: 43 %
DynamicGreedy	<b>0.34</b> +/- 0.038 0.11 (0.095, 0.13) Extreme Shares: 55 %	<b>0.35</b> +/- 0.038 0.11 (0.094, 0.13) Extreme Shares: 50 %	<b>0.41</b> +/- 0.036 0.098 (0.084, 0.12) Extreme Shares: 37 %

## Results for HardMaxWithRandom t = 2000 Needle In Haystack High Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.28</b> +/- 0.039 0.12 (0.099, 0.14) Extreme Shares: 64 %	<b>0.38</b> +/- 0.044 0.15 (0.13, 0.17) Extreme Shares: 65 %	<b>0.43</b> +/- 0.044 0.15 (0.13, 0.17) Extreme Shares: 57 %
Dynamic Epsilon Greedy	<b>0.3</b> +/- 0.038 0.11 (0.095, 0.13) Extreme Shares: 58 %	<b>0.34</b> +/- 0.038 0.11 (0.098, 0.14) Extreme Shares: 53 %	<b>0.45</b> +/- 0.04 0.12 (0.1, 0.14) Extreme Shares: 43 %
DynamicGreedy	<b>0.3</b> +/- 0.036 0.1 (0.088, 0.12) Extreme Shares: 55 %	<b>0.34</b> +/- 0.038 0.11 (0.094, 0.13) Extreme Shares: 52 %	<b>0.39</b> +/- 0.035 0.094 (0.081, 0.11) Extreme Shares: 36 %

## Results for SoftMax t = 2000 Needle In Haystack High Info + Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.37</b> +/- 0.018 0.026 (0.022, 0.03) Extreme Shares: 1 %	<b>0.43</b> +/- 0.024 0.044 (0.038, 0.052) Extreme Shares: 3 %	<b>0.53</b> +/- 0.032 0.078 (0.067, 0.092) Extreme Shares: 4.3 %
Dynamic Epsilon Greedy	<b>0.35</b> +/- 0.022 0.036 (0.031, 0.043) Extreme Shares: 8.7 %	<b>0.42</b> +/- 0.025 0.047 (0.041, 0.056) Extreme Shares: 5.7 %	<b>0.54</b> +/- 0.028 0.059 (0.051, 0.07) Extreme Shares: 2.3 %
DynamicGreedy	<b>0.35</b> +/- 0.024 0.045 (0.038, 0.053) Extreme Shares: 9 %	<b>0.4</b> +/- 0.026 0.052 (0.045, 0.062) Extreme Shares: 8 %	<b>0.47</b> +/- 0.024 0.045 (0.038, 0.053) Extreme Shares: 4.7 %

## Results for SoftMax t = 2000 Needle In Haystack High Info

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.38</b> +/- 0.019 0.028 (0.024, 0.033) Extreme Shares: 1.3 %	<b>0.45</b> +/- 0.024 0.046 (0.039, 0.054) Extreme Shares: 3 %	<b>0.53</b> +/- 0.031 0.076 (0.065, 0.09) Extreme Shares: 7.3 %
Dynamic Epsilon Greedy	<b>0.38</b> +/- 0.023 0.042 (0.036, 0.049) Extreme Shares: 4.7 %	<b>0.44</b> +/- 0.026 0.051 (0.044, 0.06) Extreme Shares: 3.7 %	<b>0.53</b> +/- 0.027 0.058 (0.05, 0.069) Extreme Shares: 5.7 %
DynamicGreedy	<b>0.38</b> +/- 0.027 0.055 (0.047, 0.065) Extreme Shares: 8.3 %	<b>0.41</b> +/- 0.027 0.058 (0.049, 0.068) Extreme Shares: 8.7 %	<b>0.5</b> +/- 0.025 0.048 (0.042, 0.057) Extreme Shares: 4 %

## Results for SoftMax t = 2000 Needle In Haystack High Rep

	Thompson Sampling	Dynamic Epsilon Greedy	DynamicGreedy
Thompson Sampling	<b>0.42</b> +/- 0.023 0.04 (0.034, 0.047) Extreme Shares: 0.67 %	<b>0.49</b> +/- 0.026 0.053 (0.045, 0.063) Extreme Shares: 2 %	<b>0.54</b> +/- 0.031 0.073 (0.063, 0.086) Extreme Shares: 2.7 %
Dynamic Epsilon Greedy	<b>0.42</b> +/- 0.026 0.05 (0.043, 0.06) Extreme Shares: 4 %	<b>0.46</b> +/- 0.026 0.052 (0.044, 0.061) Extreme Shares: 3 %	<b>0.5</b> +/- 0.028 0.06 (0.051, 0.071) Extreme Shares: 3 %
DynamicGreedy	<b>0.39</b> +/- 0.029 0.066 (0.057, 0.078) Extreme Shares: 8.7 %	<b>0.41</b> +/- 0.026 0.053 (0.045, 0.062) Extreme Shares: 5.7 %	<b>0.48</b> +/- 0.026 0.052 (0.044, 0.061) Extreme Shares: 3.3 %

What next? Have re-run this for lower values of  $X$  and higher  $T$  (3000) and seen qualitatively similar results (see Appendix). Some potential follow-up questions of interest are:

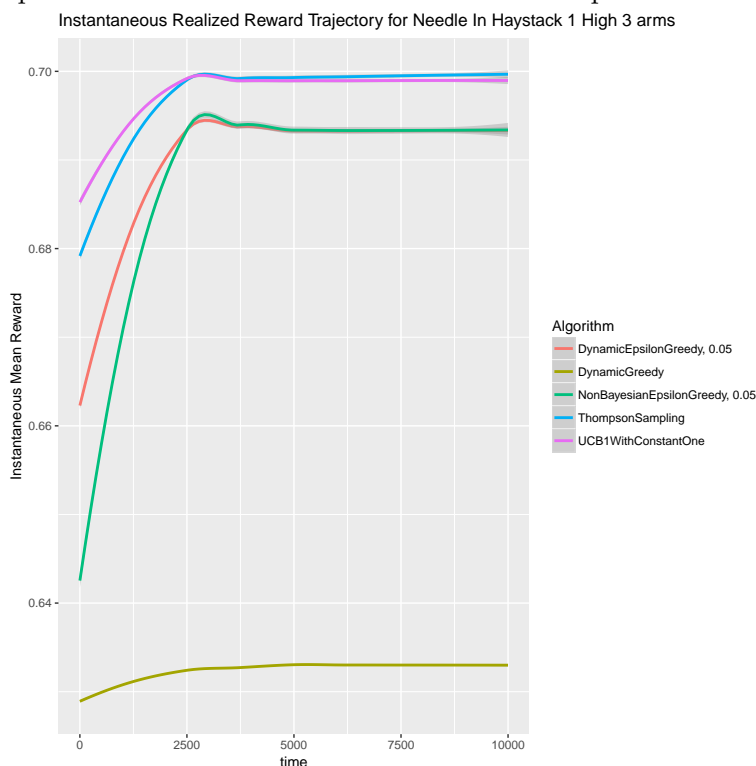
1. Even for HardMax it seems that the entrant gets a non-zero market share (the average seems to be between .05 and .25 usually). What do the market share trajectories look like? Are there a lot of simulations where the entrant simply gets zero and some where the entrant seems to get lucky upon entry and has an increasing market share over time?
2. Run for a larger time horizon for the models with agent randomness to see if better algorithms win in the long-run.
3. What results are surprising?

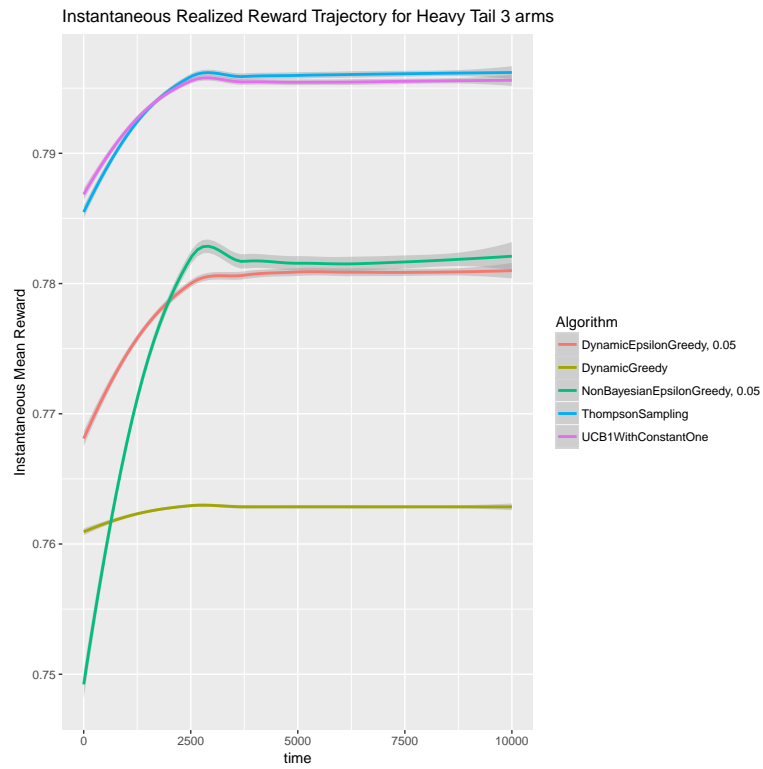
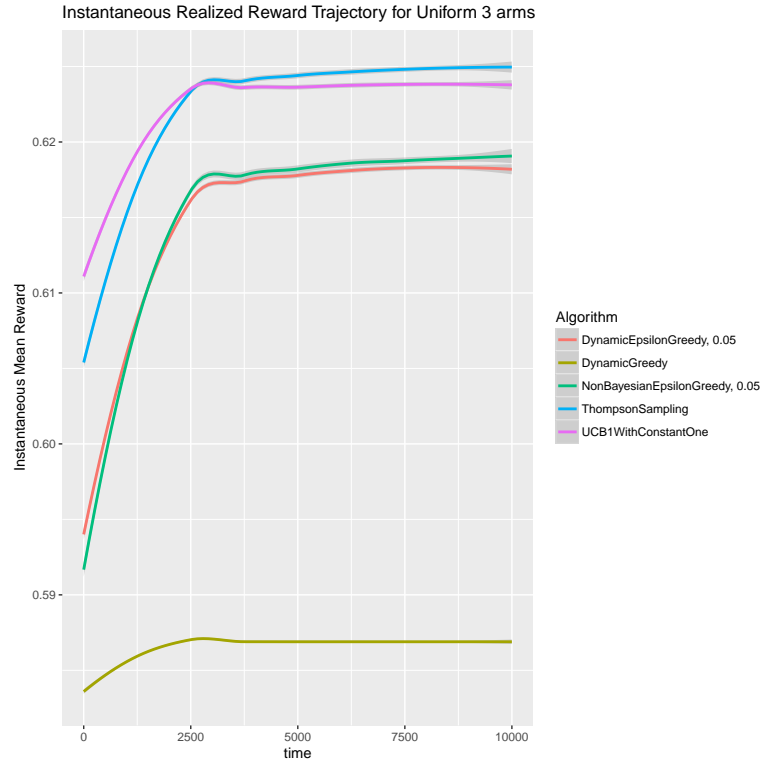
## Appendix

This contains some figures which are useful, but seem second-order to me.

### Preliminary Simulations

The plots for the realized rewards for each of the priors:







## Simultaneous Entry

The following is the result of running simulations where instead of the realized reward being utilized to form the reputation score, the TRUE mean of the selected arm was included in the reputation score.

Results for t= 8000 Heavy Tail Memory= 100

	<b>TS vs BEG</b>	<b>TS vs BG</b>	<b>BG vs BEG</b>	<b>TS vs TS</b>	<b>BEG vs BEG</b>	<b>BG vs BG</b>
EU	<b>0.35</b> +/- 0.07 Var: 0.23 Share: 100 %	<b>0.33</b> +/- 0.07 Var: 0.22 Share: 100 %	<b>0.57</b> +/- 0.07 Var: 0.25 Share: 100 %			

Results for t= 8000 Uniform Memory= 100

	<b>TS vs BEG</b>	<b>TS vs BG</b>	<b>BG vs BEG</b>	<b>TS vs TS</b>	<b>BEG vs BEG</b>	<b>BG vs BG</b>
EU	<b>0.43</b> +/- 0.07 Var: 0.24 Share: 100 %	<b>0.37</b> +/- 0.07 Var: 0.23 Share: 100 %	<b>0.55</b> +/- 0.07 Var: 0.25 Share: 100 %			

Results for t= 8000 Needle In Haystack High Memory= 100

	<b>TS vs BEG</b>	<b>TS vs BG</b>	<b>BG vs BEG</b>	<b>TS vs TS</b>	<b>BEG vs BEG</b>	<b>BG vs BG</b>
EU	<b>0.52</b> +/- 0.07 Var: 0.25 Share: 100 %	<b>0.62</b> +/- 0.07 Var: 0.24 Share: 100 %	<b>0.47</b> +/- 0.07 Var: 0.25 Share: 100 %			

## Asymmetric Entry

The following is the result of running the incumbent experiment with  $X = 100$

## Results for HardMax t = 1000 Needle In Haystack High

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.14</b> +/- 0.027 0.097 (0.086, 0.11) Extreme Shares: 88 %	<b>0.2</b> +/- 0.032 0.13 (0.12, 0.15) Extreme Shares: 88 %	<b>0.24</b> +/- 0.035 0.15 (0.14, 0.18) Extreme Shares: 86 %
DynamicEpsilonGreedy	<b>0.13</b> +/- 0.027 0.097 (0.086, 0.11) Extreme Shares: 92 %	<b>0.21</b> +/- 0.033 0.14 (0.13, 0.16) Extreme Shares: 89 %	<b>0.25</b> +/- 0.035 0.16 (0.14, 0.18) Extreme Shares: 86 %
DynamicGreedy	<b>0.12</b> +/- 0.025 0.078 (0.069, 0.088) Extreme Shares: 90 %	<b>0.2</b> +/- 0.032 0.14 (0.12, 0.15) Extreme Shares: 89 %	<b>0.25</b> +/- 0.035 0.15 (0.14, 0.18) Extreme Shares: 85 %

## Results for HardMaxWithRandom t = 1000 Needle In Haystack High

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.21</b> +/- 0.027 0.094 (0.083, 0.11) Extreme Shares: 73 %	<b>0.29</b> +/- 0.032 0.13 (0.12, 0.15) Extreme Shares: 73 %	<b>0.34</b> +/- 0.034 0.15 (0.14, 0.17) Extreme Shares: 71 %
DynamicEpsilonGreedy	<b>0.24</b> +/- 0.028 0.1 (0.093, 0.12) Extreme Shares: 69 %	<b>0.3</b> +/- 0.031 0.13 (0.11, 0.14) Extreme Shares: 68 %	<b>0.42</b> +/- 0.034 0.15 (0.13, 0.17) Extreme Shares: 64 %
DynamicGreedy	<b>0.27</b> +/- 0.029 0.11 (0.098, 0.13) Extreme Shares: 68 %	<b>0.29</b> +/- 0.03 0.12 (0.11, 0.14) Extreme Shares: 65 %	<b>0.37</b> +/- 0.033 0.14 (0.12, 0.16) Extreme Shares: 62 %

## Results for SoftMax t = 1000 Needle In Haystack High

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.44</b> +/- 0.0085 0.0093 (0.0082, 0.011) Extreme Shares: 0 %	<b>0.48</b> +/- 0.012 0.019 (0.017, 0.022) Extreme Shares: 0 %	<b>0.53</b> +/- 0.016 0.033 (0.029, 0.038) Extreme Shares: 0 %
DynamicEpsilonGreedy	<b>0.43</b> +/- 0.012 0.018 (0.016, 0.021) Extreme Shares: 0 %	<b>0.47</b> +/- 0.016 0.031 (0.028, 0.036) Extreme Shares: 0 %	<b>0.51</b> +/- 0.017 0.038 (0.034, 0.043) Extreme Shares: 0 %
DynamicGreedy	<b>0.42</b> +/- 0.014 0.025 (0.022, 0.029) Extreme Shares: 0 %	<b>0.45</b> +/- 0.015 0.03 (0.026, 0.034) Extreme Shares: 0 %	<b>0.49</b> +/- 0.018 0.041 (0.036, 0.047) Extreme Shares: 0 %

## Results for HardMax t = 1000 Heavy Tail

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.056</b> +/- 0.018 0.04 (0.036, 0.046) Extreme Shares: 94 %	<b>0.11</b> +/- 0.025 0.083 (0.074, 0.094) Extreme Shares: 92 %	<b>0.11</b> +/- 0.025 0.08 (0.071, 0.091) Extreme Shares: 91 %
DynamicEpsilonGreedy	<b>0.094</b> +/- 0.021 0.058 (0.052, 0.066) Extreme Shares: 87 %	<b>0.16</b> +/- 0.027 0.096 (0.085, 0.11) Extreme Shares: 83 %	<b>0.15</b> +/- 0.026 0.088 (0.078, 0.1) Extreme Shares: 83 %
DynamicGreedy	<b>0.15</b> +/- 0.027 0.094 (0.083, 0.11) Extreme Shares: 85 %	<b>0.22</b> +/- 0.032 0.14 (0.12, 0.15) Extreme Shares: 82 %	<b>0.18</b> +/- 0.027 0.097 (0.086, 0.11) Extreme Shares: 76 %

## Results for HardMaxWithRandom t = 1000 Heavy Tail

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.13</b> +/- 0.02 0.053 (0.047, 0.061) Extreme Shares: 80 %	<b>0.21</b> +/- 0.027 0.092 (0.082, 0.1) Extreme Shares: 74 %	<b>0.2</b> +/- 0.027 0.095 (0.084, 0.11) Extreme Shares: 77 %
DynamicEpsilonGreedy	<b>0.19</b> +/- 0.023 0.07 (0.062, 0.079) Extreme Shares: 69 %	<b>0.28</b> +/- 0.028 0.1 (0.088, 0.11) Extreme Shares: 57 %	<b>0.24</b> +/- 0.027 0.092 (0.082, 0.1) Extreme Shares: 62 %
DynamicGreedy	<b>0.23</b> +/- 0.027 0.092 (0.082, 0.11) Extreme Shares: 65 %	<b>0.34</b> +/- 0.03 0.12 (0.1, 0.13) Extreme Shares: 52 %	<b>0.29</b> +/- 0.029 0.11 (0.094, 0.12) Extreme Shares: 58 %

## Results for SoftMax t = 1000 Heavy Tail

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.45</b> +/- 0.0063 0.0051 (0.0045, 0.0058) Extreme Shares: 0 %	<b>0.49</b> +/- 0.008 0.0083 (0.0074, 0.0095) Extreme Shares: 0 %	<b>0.49</b> +/- 0.012 0.018 (0.016, 0.02) Extreme Shares: 0 %
DynamicEpsilonGreedy	<b>0.43</b> +/- 0.0084 0.0091 (0.008, 0.01) Extreme Shares: 0 %	<b>0.46</b> +/- 0.0091 0.011 (0.0094, 0.012) Extreme Shares: 0 %	<b>0.47</b> +/- 0.012 0.019 (0.016, 0.021) Extreme Shares: 0.2 %
DynamicGreedy	<b>0.45</b> +/- 0.0099 0.013 (0.011, 0.014) Extreme Shares: 0.2 %	<b>0.48</b> +/- 0.011 0.017 (0.015, 0.019) Extreme Shares: 0 %	<b>0.48</b> +/- 0.01 0.014 (0.012, 0.016) Extreme Shares: 0 %

## Results for HardMax t = 1000 Uniform

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.17</b> +/- 0.028 0.099 (0.088, 0.11) Extreme Shares: 81 %	<b>0.17</b> +/- 0.029 0.11 (0.095, 0.12) Extreme Shares: 86 %	<b>0.21</b> +/- 0.031 0.13 (0.11, 0.15) Extreme Shares: 83 %
DynamicEpsilonGreedy	<b>0.2</b> +/- 0.03 0.12 (0.1, 0.13) Extreme Shares: 79 %	<b>0.19</b> +/- 0.029 0.11 (0.099, 0.13) Extreme Shares: 81 %	<b>0.24</b> +/- 0.033 0.14 (0.12, 0.16) Extreme Shares: 80 %
DynamicGreedy	<b>0.18</b> +/- 0.029 0.11 (0.097, 0.12) Extreme Shares: 80 %	<b>0.21</b> +/- 0.03 0.11 (0.1, 0.13) Extreme Shares: 76 %	<b>0.24</b> +/- 0.032 0.13 (0.12, 0.15) Extreme Shares: 77 %

## Results for HardMaxWithRandom t = 1000 Uniform

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.25</b> +/- 0.026 0.087 (0.077, 0.098) Extreme Shares: 57 %	<b>0.32</b> +/- 0.031 0.12 (0.11, 0.14) Extreme Shares: 57 %	<b>0.33</b> +/- 0.03 0.12 (0.1, 0.13) Extreme Shares: 55 %
DynamicEpsilonGreedy	<b>0.29</b> +/- 0.028 0.1 (0.091, 0.12) Extreme Shares: 55 %	<b>0.33</b> +/- 0.029 0.11 (0.097, 0.12) Extreme Shares: 51 %	<b>0.35</b> +/- 0.03 0.12 (0.1, 0.13) Extreme Shares: 51 %
DynamicGreedy	<b>0.3</b> +/- 0.029 0.11 (0.095, 0.12) Extreme Shares: 55 %	<b>0.32</b> +/- 0.029 0.11 (0.096, 0.12) Extreme Shares: 52 %	<b>0.36</b> +/- 0.03 0.12 (0.1, 0.13) Extreme Shares: 49 %

## Results for SoftMax t = 1000 Uniform

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	<b>0.46</b> +/- 0.0077 0.0076 (0.0067, 0.0086) Extreme Shares: 0 %	<b>0.47</b> +/- 0.0099 0.013 (0.011, 0.014) Extreme Shares: 0 %	<b>0.5</b> +/- 0.012 0.02 (0.017, 0.022) Extreme Shares: 0 %
DynamicEpsilonGreedy	<b>0.46</b> +/- 0.01 0.013 (0.011, 0.015) Extreme Shares: 0 %	<b>0.47</b> +/- 0.011 0.015 (0.013, 0.017) Extreme Shares: 0 %	<b>0.49</b> +/- 0.013 0.022 (0.02, 0.025) Extreme Shares: 0 %
DynamicGreedy	<b>0.45</b> +/- 0.011 0.016 (0.014, 0.018) Extreme Shares: 0 %	<b>0.47</b> +/- 0.012 0.019 (0.017, 0.021) Extreme Shares: 0 %	<b>0.49</b> +/- 0.013 0.023 (0.02, 0.026) Extreme Shares: 0 %