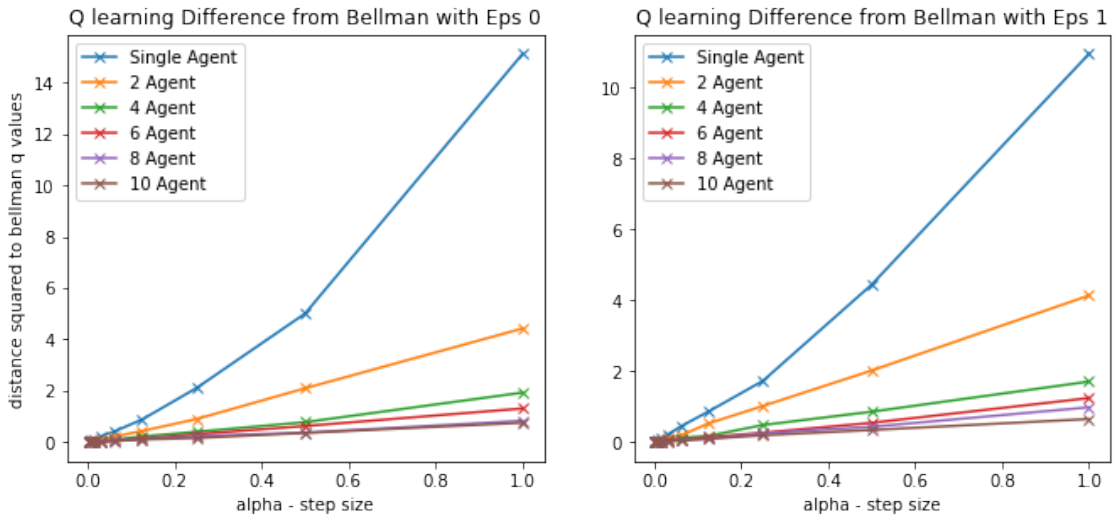
graphing

May 1, 2022

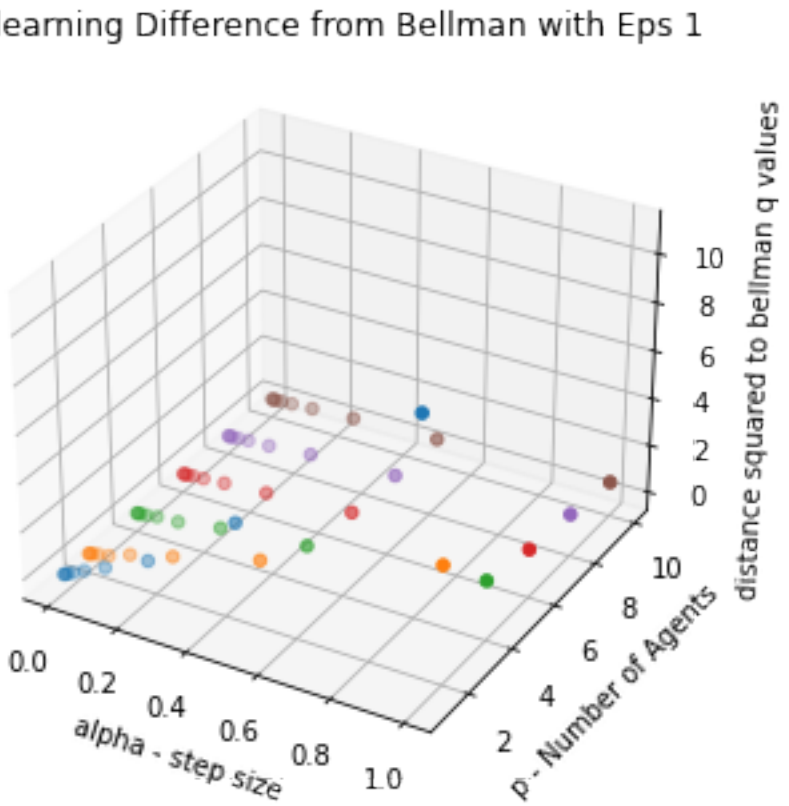
```
epsilons [0, 1]
convN 10
dicount 0.6
syncBackups 1
stochasticPolicy True
envSeed 42
alpha 1
fedPs [2, 4, 6, 8, 10]
```

This graph is a comparison of learned models to the bellman model. A higher number of agents and smaller step size results in the lowest error to bellman. This trend appears linear on a log log plot for a variety of different numbers of agents.

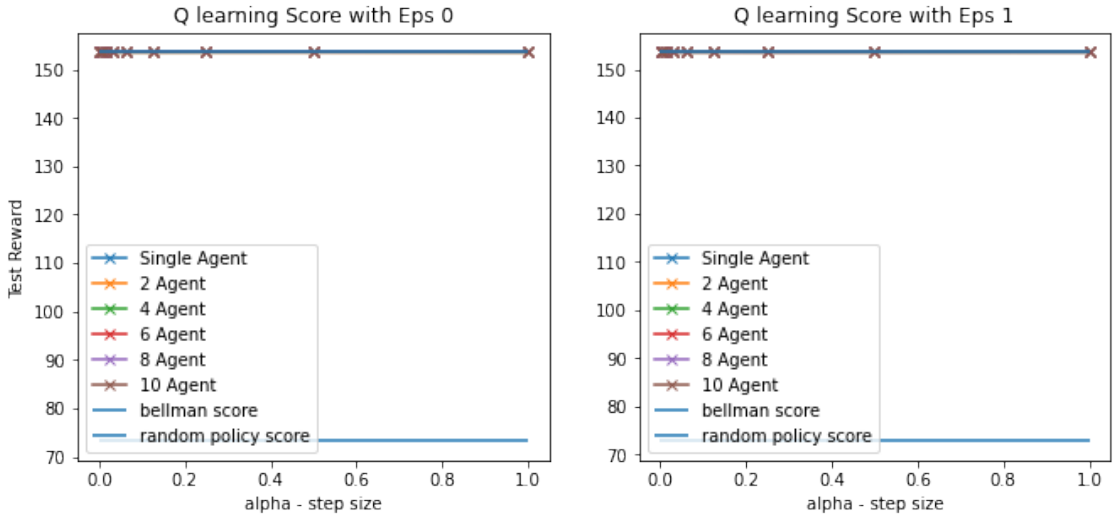


3d version of above plot by using number of agents as an additional axis instead of different series. This appears to be a plane on a log log log plot.

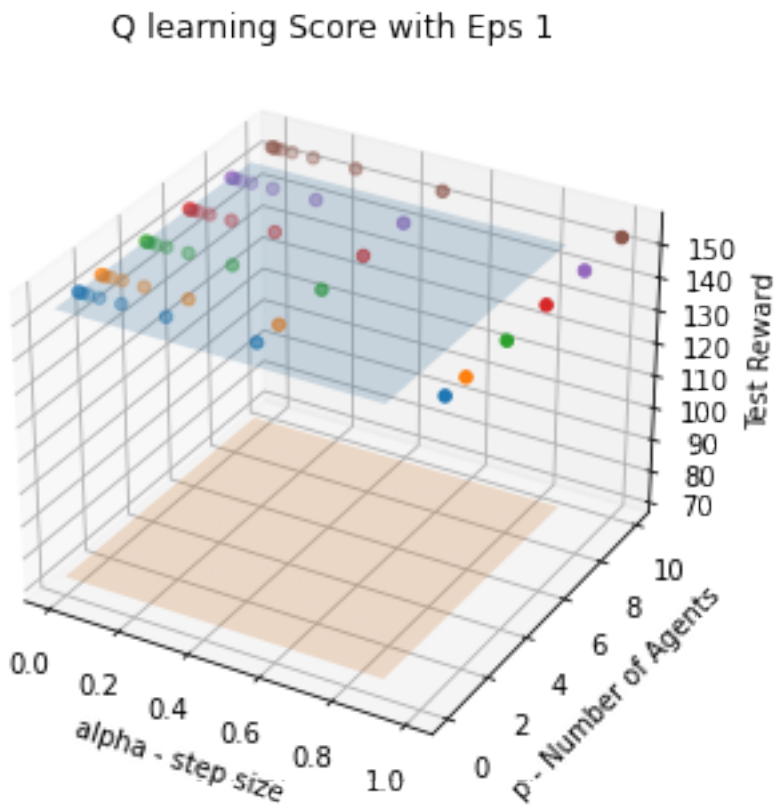
```
<Figure size 720x360 with 0 Axes>
```



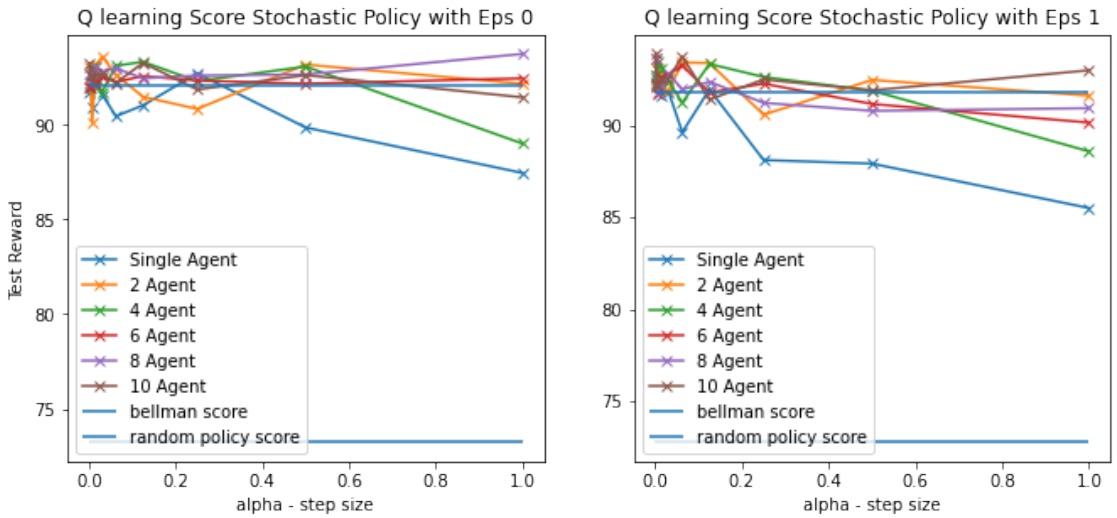
Test Reward when evaluated with a deterministic policy. Everything appears to be the same as bellman. This may present an issue with using reward to learn and get closer to bellman values for MDPs



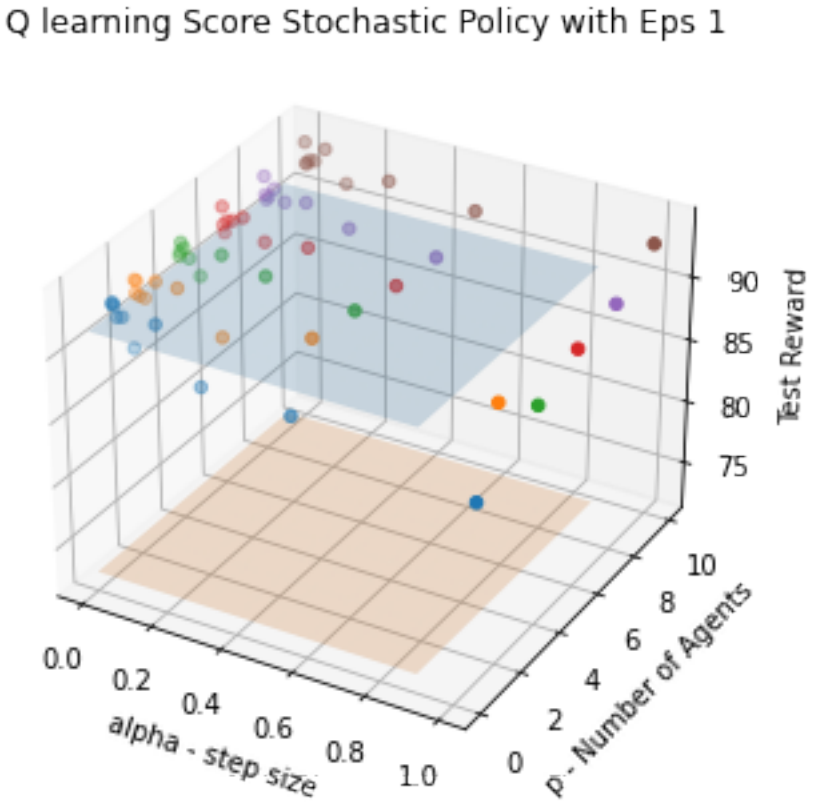
<Figure size 720x360 with 0 Axes>



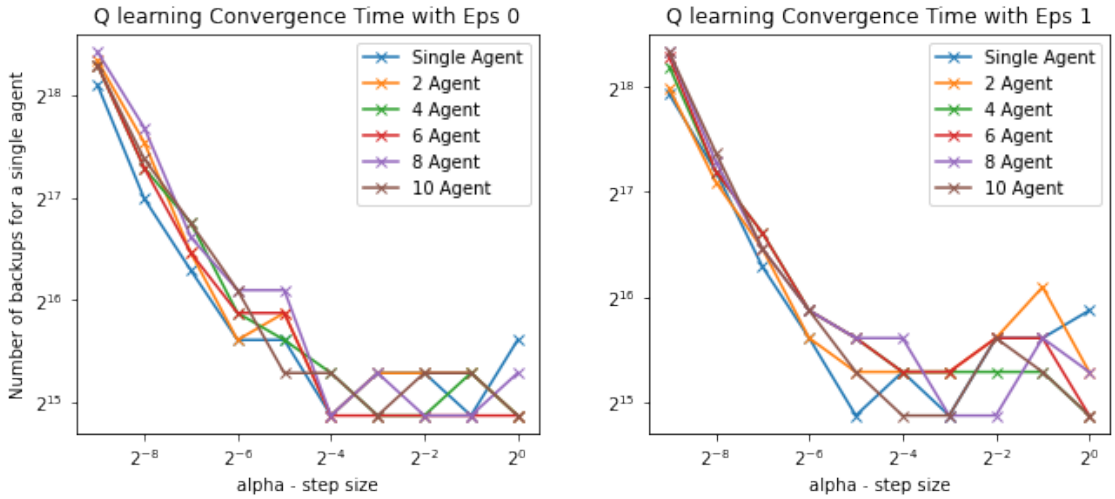
Test Reward when evaluated with a stochastic policy. This is the omly graph where epsilon 0 is noticably different than epsilon 1. Here the Q learning can get a higher score than the bellman Q values. Shows a rough increase in reward for more agents and smaller step sizes.



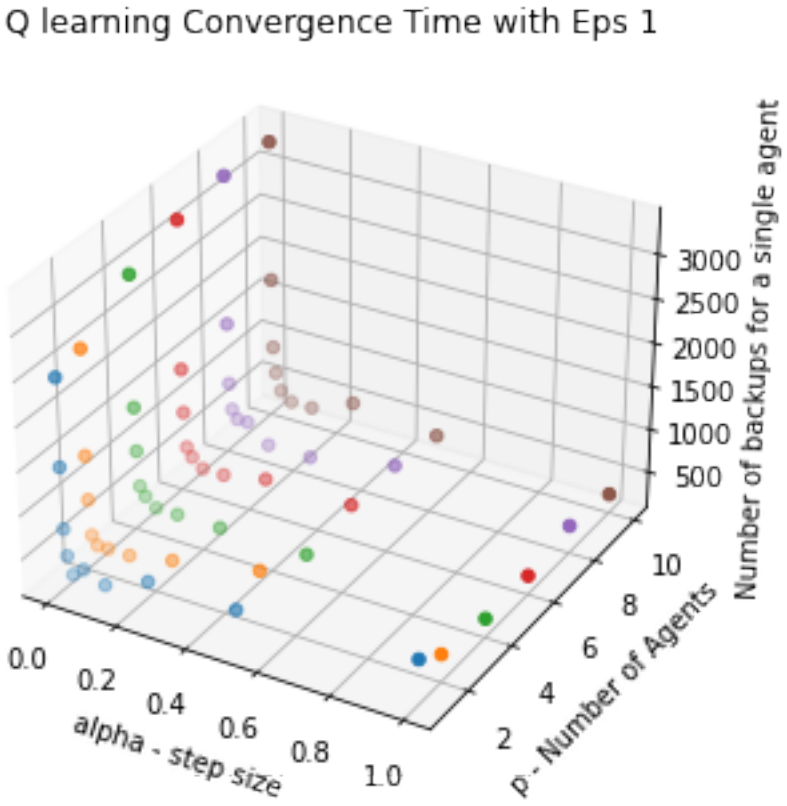
```
<Figure size 720x360 with 0 Axes>
```



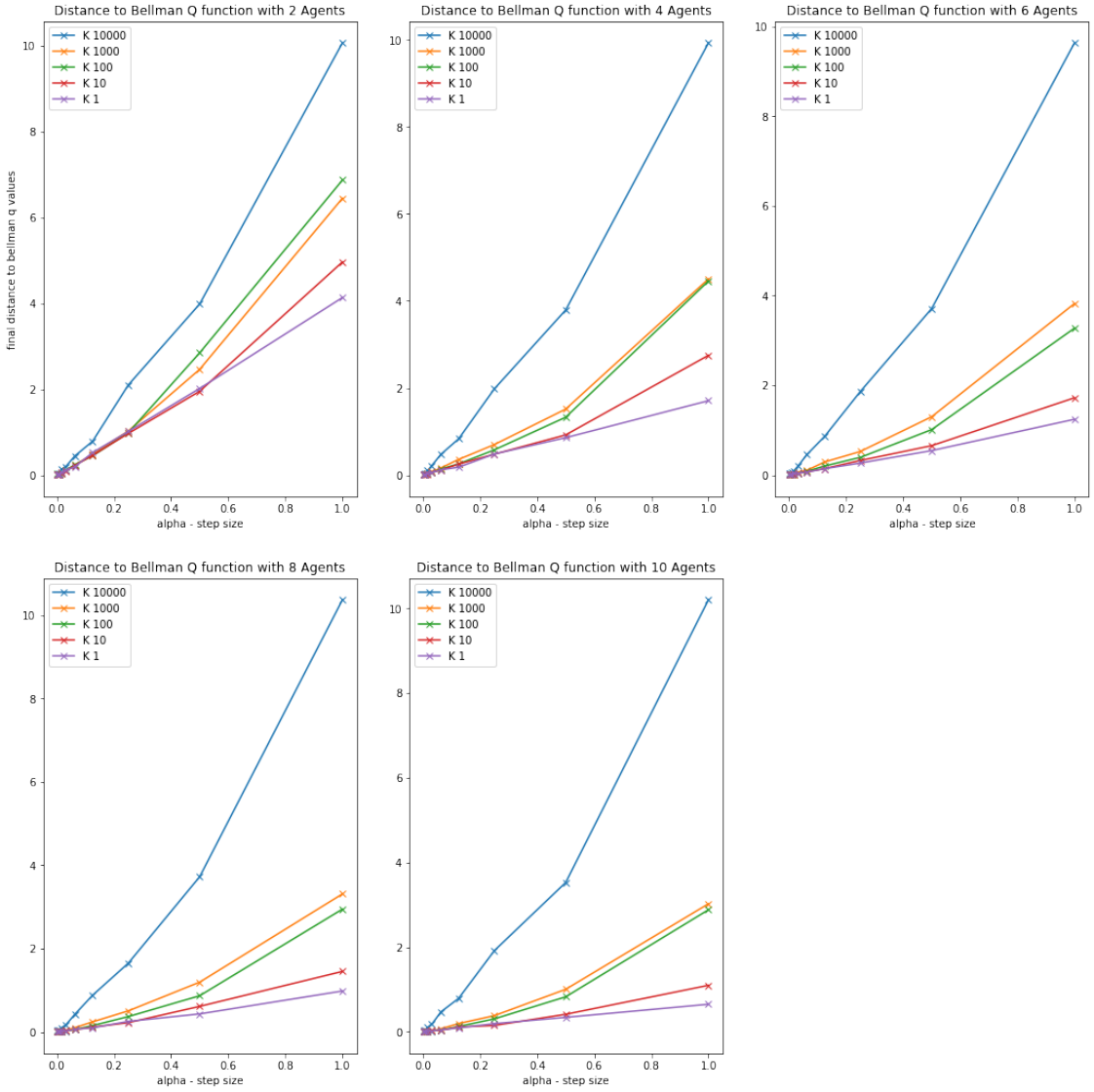
Convergence time, measured in model backups for a single agent (not all agents). This appears to show no speedup

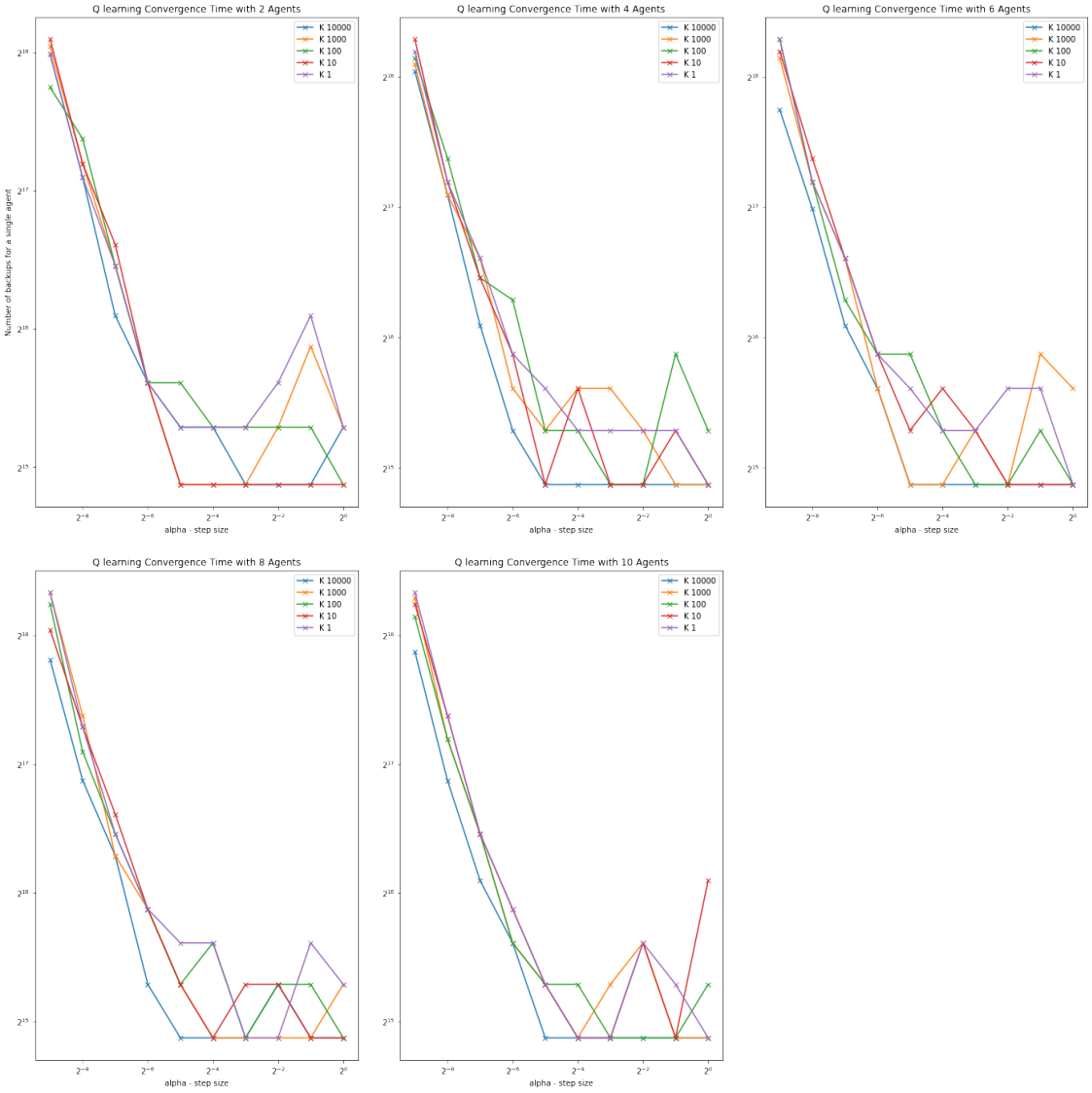


```
<Figure size 720x360 with 0 Axes>
```



0


Distance to Bellman Q function with 2 Agents
Distance to Bellman Q function with 4 Agents
Distance to Bellman Q function with 6 Agents
Distance to Bellman Q function with 8 Agents
Distance to Bellman Q function with 10 Agents
K 10000
K 1000
K 100
K 10
K 1
final distance to bellman q values
alpha - step size


Q learning Convergence Time with 2 Agents
Q learning Convergence Time with 4 Agents
Q learning Convergence Time with 6 Agents
Q learning Convergence Time with 8 Agents
Q learning Convergence Time with 10 Agents
K 10000
K 1000
K 100
K 10
K 1
Number of backups for a single agent
alpha - step size