Question 2

0.2	WALL	→	2
↑	←	↑→	↑
↑	↑←	WALL	↑
↑	↑←	-0.4	↑

Gamma = 0.1
[0.2, 0.0, -0.042, 2.0]
[-0.188, -0.219, -0.206, -0.042]
[-0.219, -0.222, 0.0, -0.208]
[-0.222, -0.222, -0.4, -0.222]

0.2	WALL	→	2
→	→	↑→	↑
→	↑	WALL	↑
↑	↑	-0.4	↑

Gamma = 0.99
[0.2, 0.0, 1.696, 2.0]
[0.79, 1.151, 1.455, 1.696]
[0.588, 0.854, 0.0, 1.425]
[0.345, 0.47, -0.4, 0.987]

Question 2

0.2	WALL	↓	2
↓→	←→	↑	←
↑↓→	←↑↓	WALL	↑
↑→	↑	-0.4	↑

Step cost = 2
[0.2, 0.0, 101.73, 2.0]
[101.73, 101.73, 101.73, 101.726]
[101.73, 101.73, 0.0, 101.726]
[101.73, 101.73, -0.4, 101.722]

0.2	WALL	→	2
→	→	↑→	↑
↑	↑	WALL	↑
↑	↑	-0.4	↑

Step cost = -0.4
[0.2, 0.0, 1.419, 2.0]
[-0.138, 0.38, 0.959, 1.419]
[-0.586, -0.175, 0.0, 0.902]
[-1.037, -0.682, -0.4, 0.305]

Question 2

0.2	WALL	→	2
↑	→	↑→	↑
↑	↑	WALL	↑
↑	→	-0.4	↑

Step cost = -0.5
[0.2, 0.0, 1.281, 2.0]
[-0.38, -0.004, 0.712, 1.281]
[-0.964, -0.667, 0.0, 0.641]
[-1.514, -0.981, -0.4, -0.036]

0.2	WALL	→	2
↑	←	↑→	↑
↑	←↑	WALL	↑
→	→	-0.4	←

Step cost = -2
[0.2, 0.0, -0.79, 2.0]
[-2.601, -5.104, -2.998, -0.79]
[-5.104, -5.56, 0.0, -3.27]
[-5.56, -3.178, -0.4, -2.93]

Question 2

Observations:
i)If step cost is positive ,it will never try to reach the goal state because it can increase the reward by moving around instead of reaching the goal state.

ii)If step cost is negative and greater than negative goal ,it tries to go to the goal state which has the positive value.

iii)If step cost is negative and even less than negative goal value it tries to go to goal state Which is nearer instead of going to goal state with positive value and farther.