Q5 Optimal policy (R1) 0 1 2 3 4 5 6

