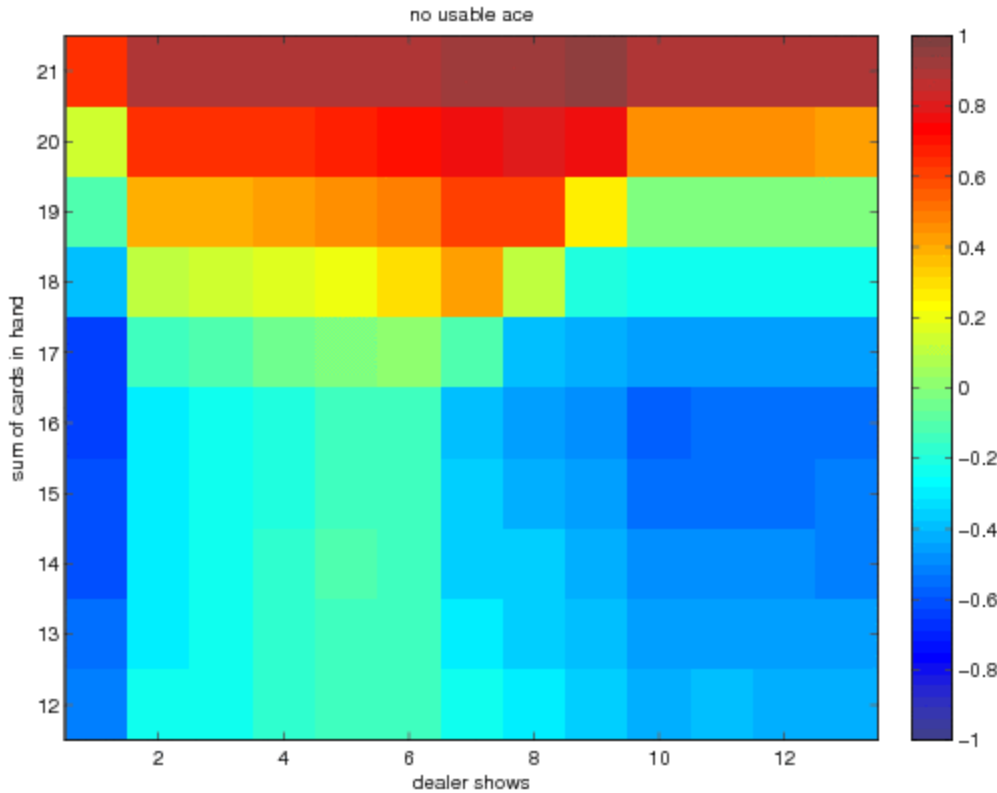
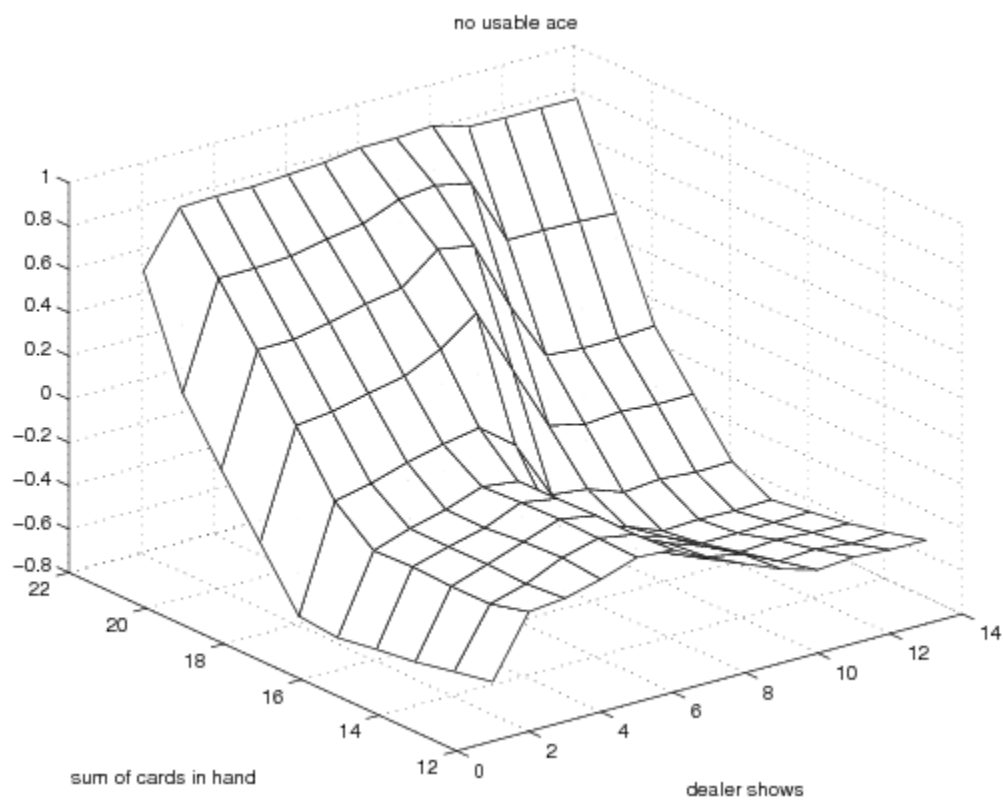
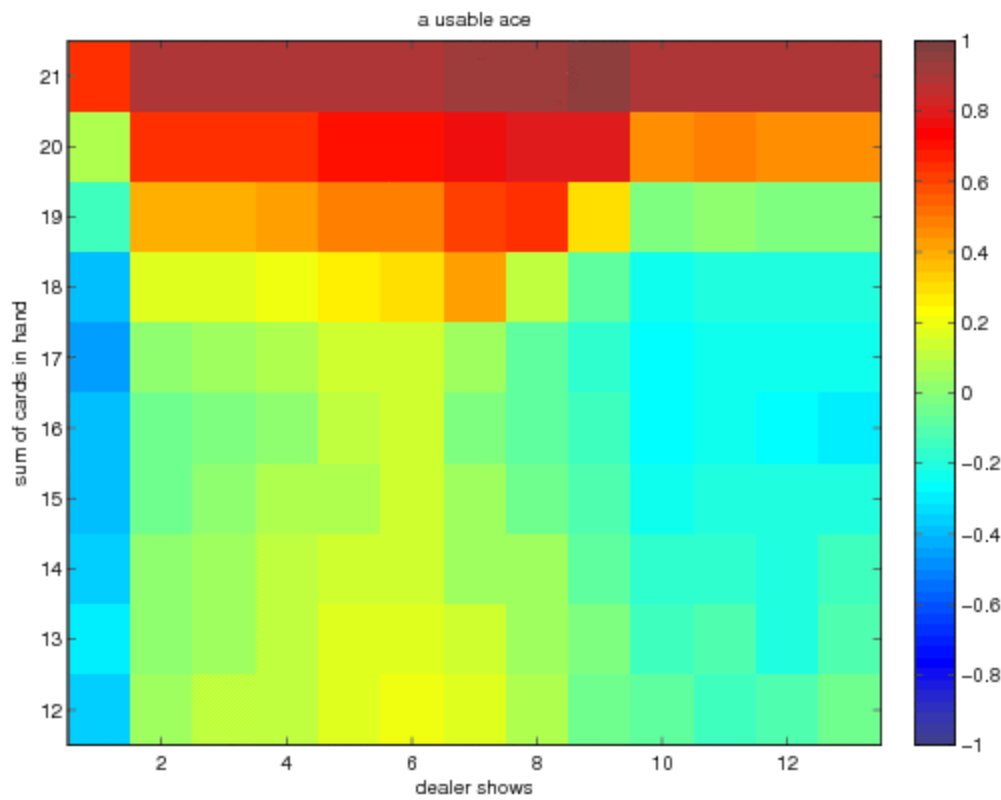
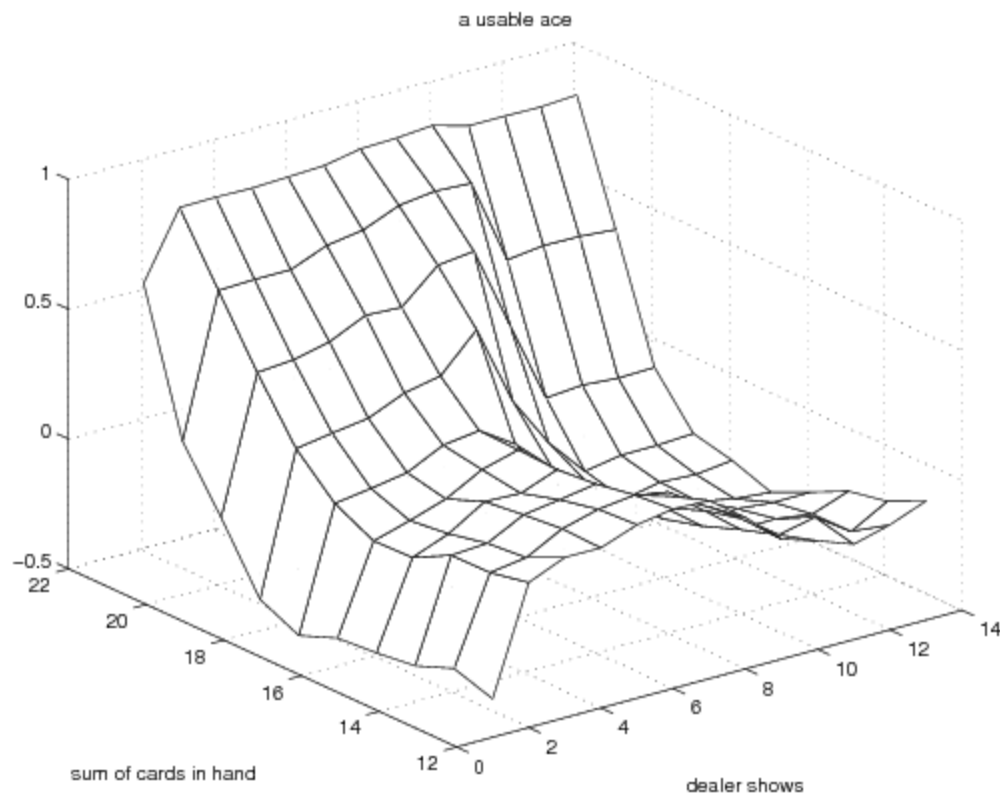


The final optimal state value function and deterministic policy obtained when using exploring starts monte carlo approximation. In this implementation I took the "dealer card" showing component of the state to be a number between 1 and 13, while the book choose to collapse this to between 1 and 10. The results are equivalent since the cards 11-13 all have the same face value (of ten). The optimal policy and value function correspondingly have the same values over those ranges (as they should). The optimal policy agrees with that shown in the book. These results were obtained with 5 million monte carlo trials. We first present the optimal state value function and then the optimal policy. The optimal state value function looks like







The optimal policies look like

