FDR for a1false-a2true-b1false-c1false, reward function 1