rewards per episode when stochastic=0 and noise=0