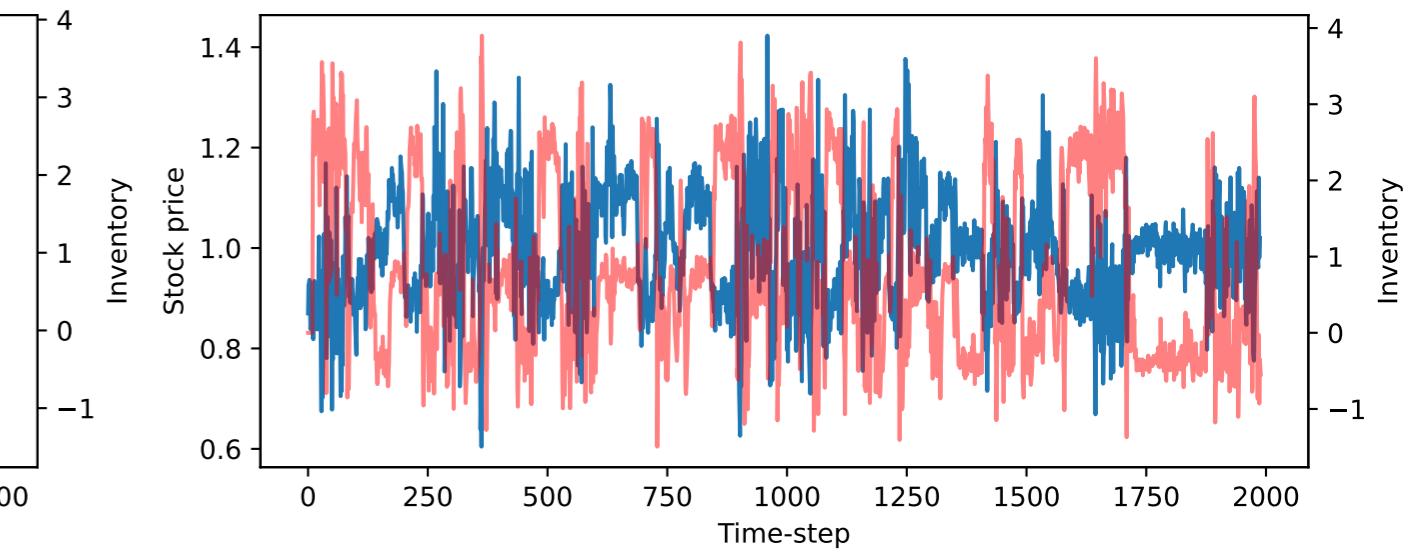
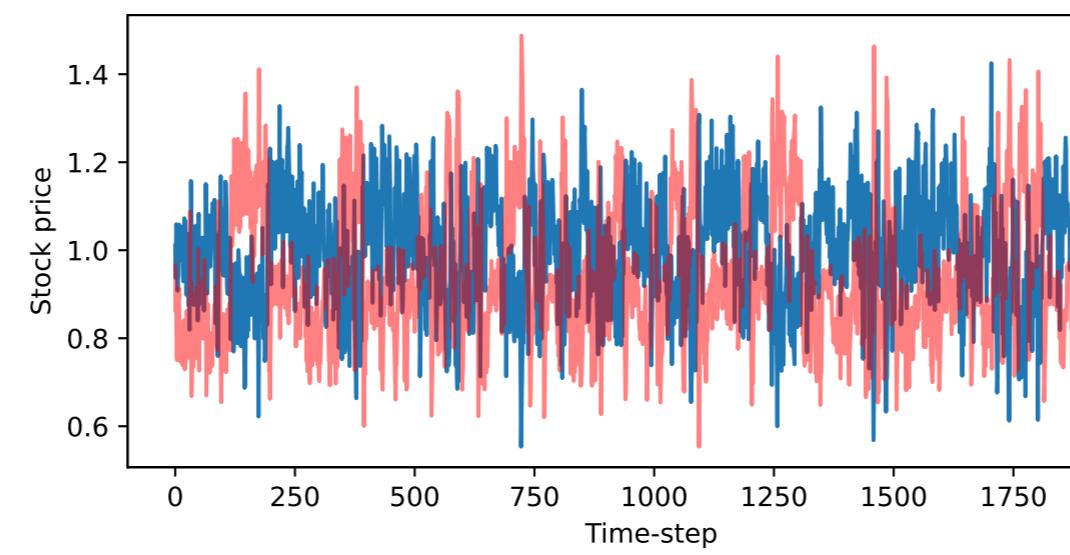
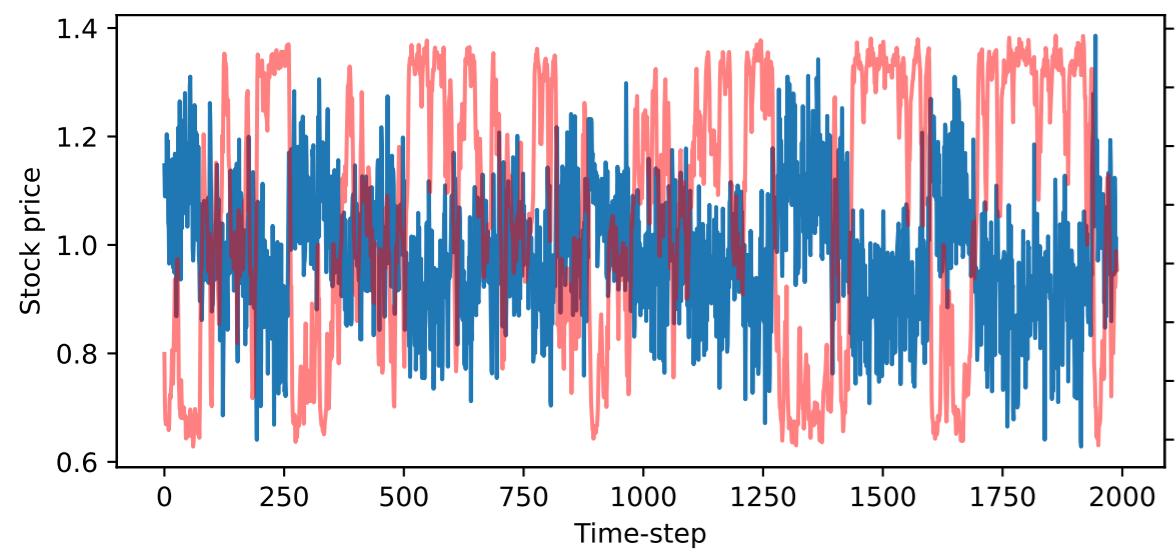
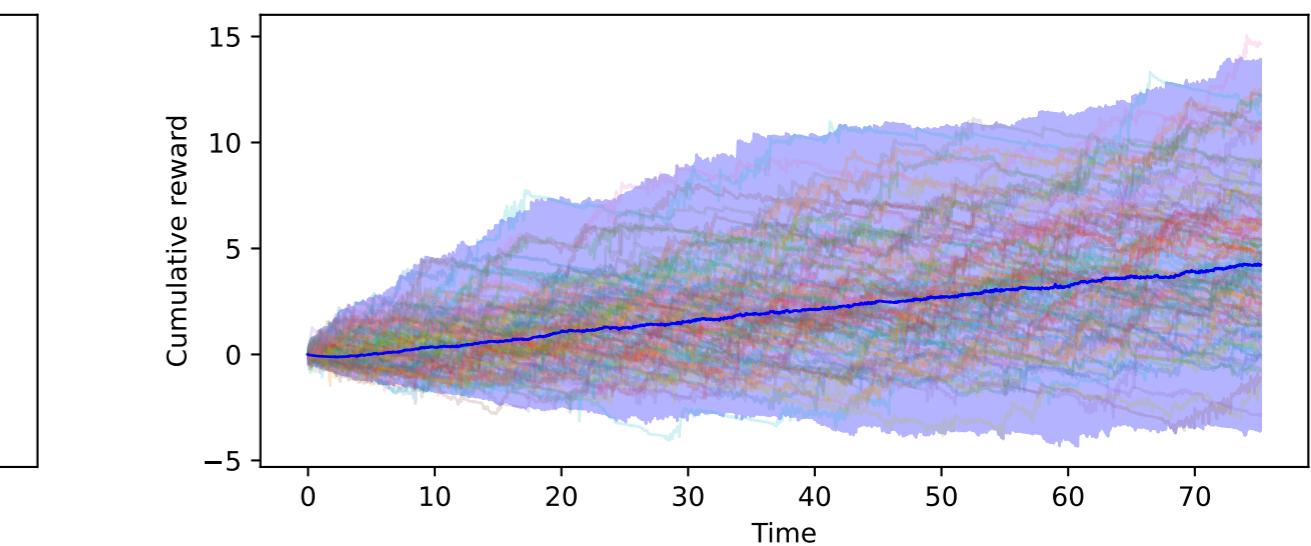
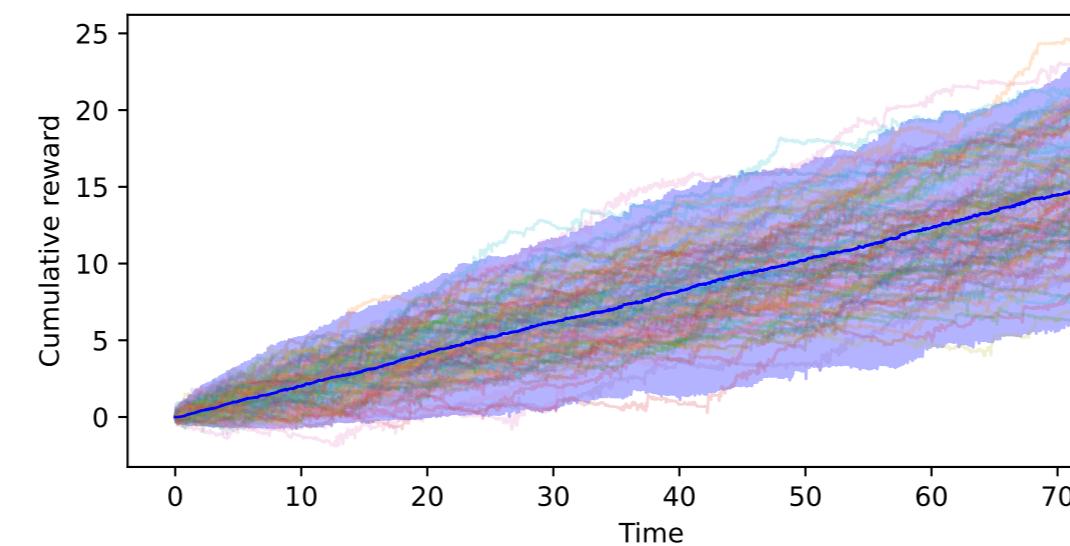
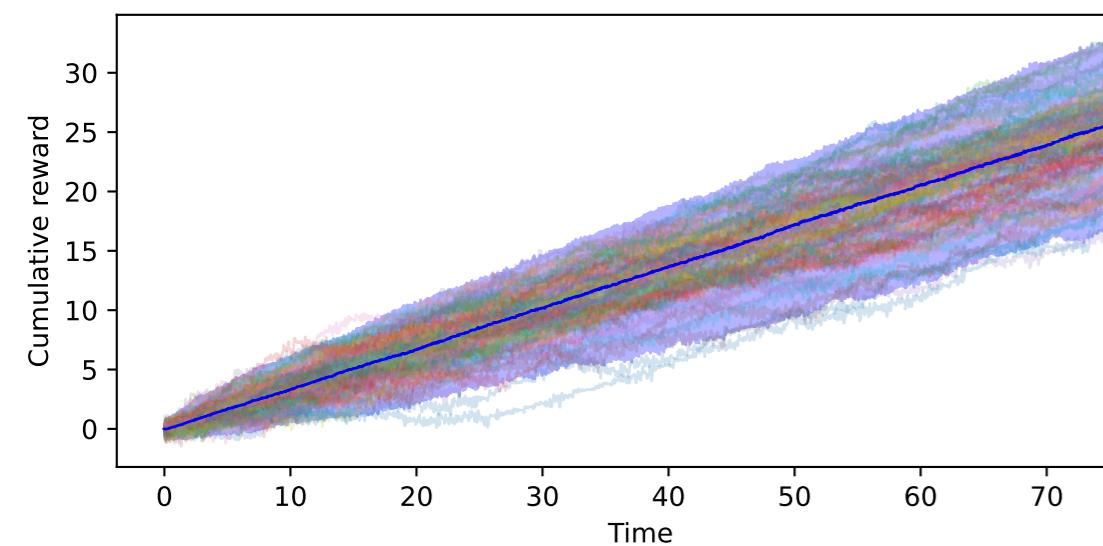
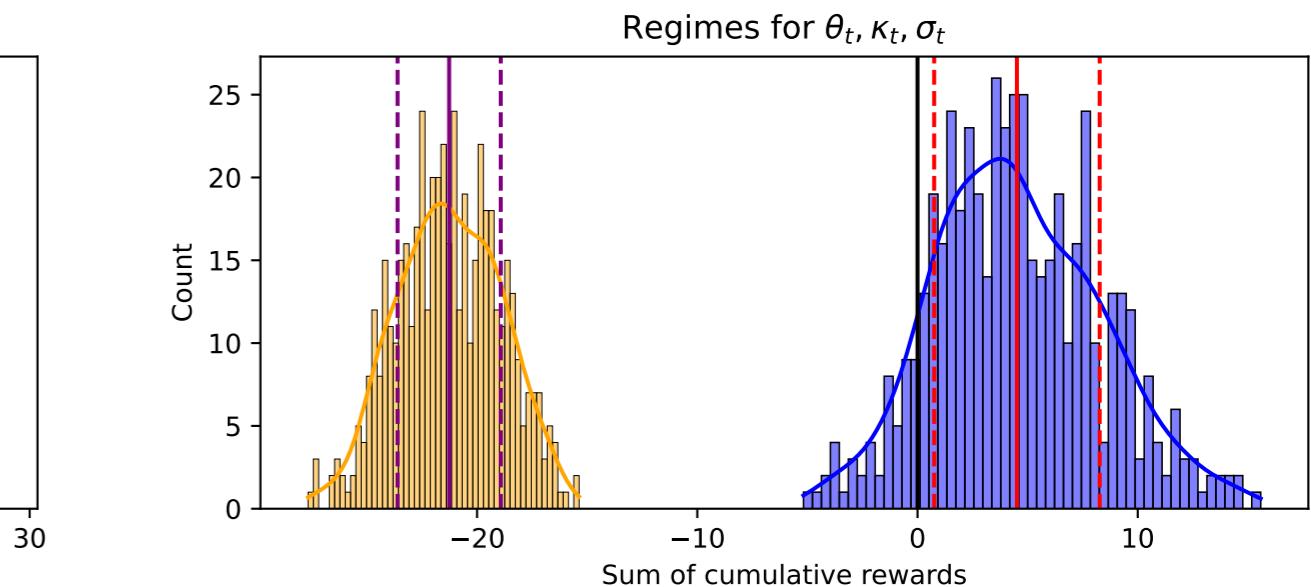
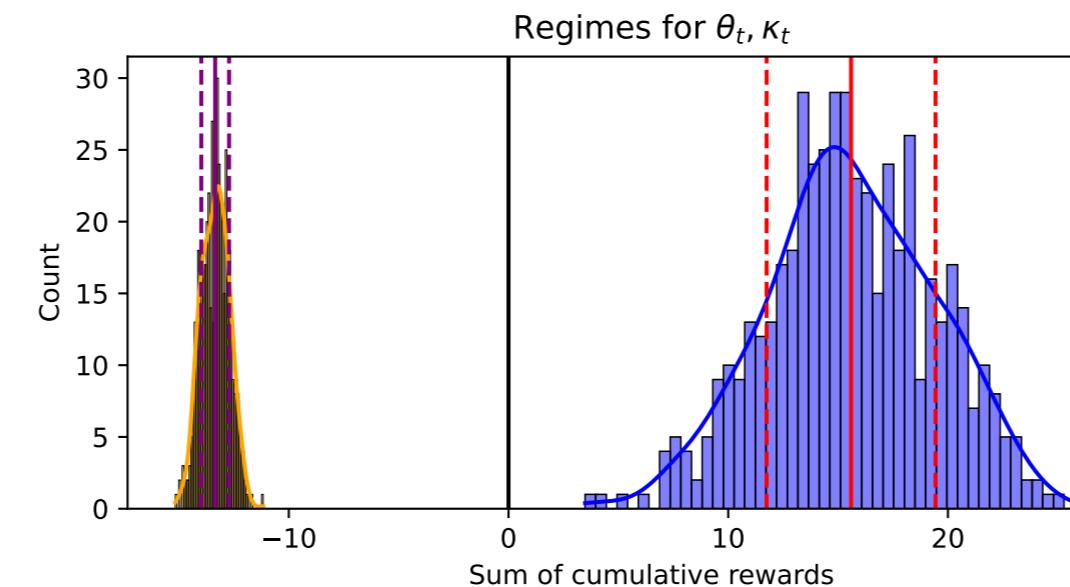
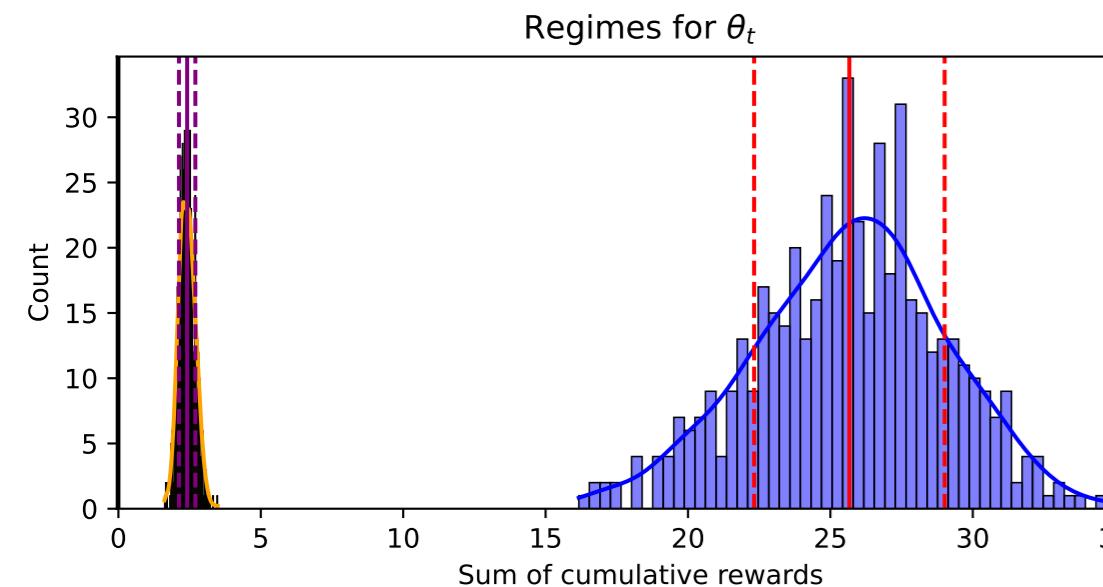


Histograms and plot of cumulative rewards for 1000 test episodes



For regimes of just  $\theta_t$  we used a lookback window of 10 time-steps, and 5 layers of 20 nodes. For regimes of  $\theta_t, \kappa_t$  we used 6 layers of 20 nodes.

For regimes of  $\theta_t, \kappa_t, \sigma_t$  we used 6 layers of 20 nodes and a lookback window of 20 time-steps