ζ-EB optimal policy - last 500 steps 0.5 -0.6 -1 -0.8 -0.6 -0.7 -0.9 -1.1 -0.8 -1.2 -0.7 -0.8 -1 3.5 -0.8 -0.7 -0.6 -1.1 -0.9 -0.7 -0.9 -0.7 -0.9 4.5 -0.9 -0.7 -0.9 -0.9 -0.8 -0.8 -1.1 -0.8 -1.2 0.001 0.01 0.1 1 10 prior LP