ζ-EB optimal policy - last 500 steps -0.9 -1 -0.9 -0.8 -0.6 -0.9 -0.8 -1.1 -0.6 0.5 -0.8 -1 -1.2 -0.8 -0.8 -0.7 -1.1 -0.9 -0.6 -0.8 -0.7 -0.7 -0.7 -0.9 -0.6 1.5 -0.6 -0.7 -1 -0.9 -0.8 -1.3 -0.8 -1.2 beta 2.0 -0.6 -0.7 -1 -0.6 -1.1 -1.1 2.5 -0.7 -1.1 -1 -0.6 -0.8 3.0 -0.6 -0.6 -0.8 3.5 -0.8 -0.6 -0.8 -0.9 -1.2 -12 2.5 0.1 0.5 1.0 1.5 2.0 3.0 3.5 4.0

alpha