ζ-EB optimal policy - last 500 steps -0.8 -0.9 -0.8 -0.9 -0.9 -1.2 -0.7 -0.7 -0.6 -0.8 -0.9 -0.8 -0.9 -0.9 -0.7 ლ − -0.8 -0.8 -0.7 -1.2 -0.7 2 / -0.7 -0.7 beta -0.5 0 -0.7 15 -0.6 -0.6 2.5 3.5 0.1 0.5 1.0 1.5 2.0 3.0 4.0 alpha