ζ-EB optimal policy - last 500 steps