

$\zeta$ -R-max agent policy with alpha = 2.7

