

ζ -R-max agent policy with alpha = 1.2

