

Mean of effective $\bar{\delta}$ upon value convergence with TD RPE

Two-Timescale Negative integrator, optimal policy, stationary environment

Smoothed across 30 runs

