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

One-Timescale Pos/Neg integrator, optimal policy, stationary environment

Smoothed across 30 runs