SandingEnvDifficult

Legend:
- ppo
- ppo_extension

X-axis: Timestep (×10^5)
Y-axis: Average Reward