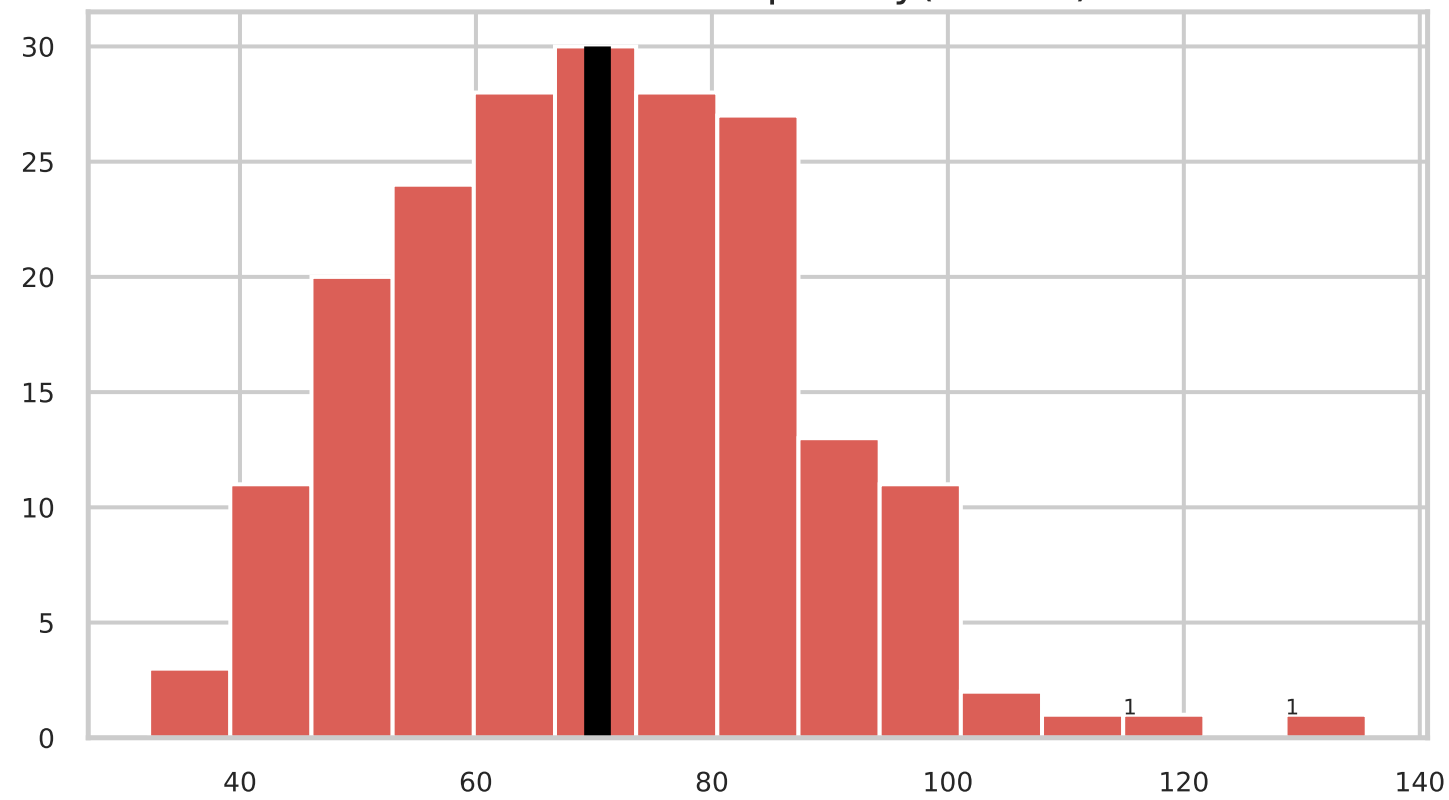


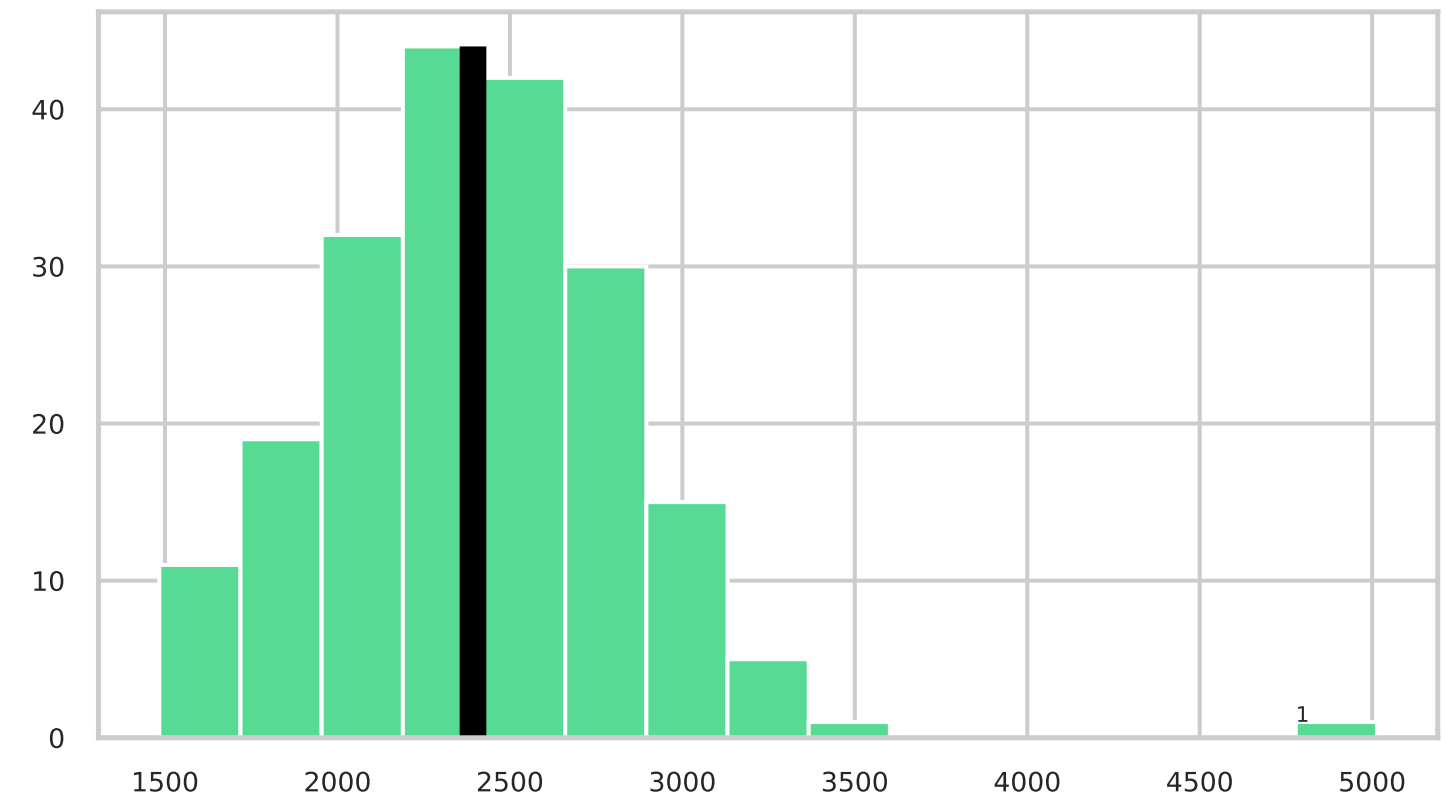
# Histogram of regrets for different multi-players bandit algorithms

9 arms:  $[B(0.1), B(0.2), B(0.3), B(0.4)^*, B(0.5)^*, B(0.6)^*, B(0.7)^*, B(0.8)^*, B(0.9)^*]$

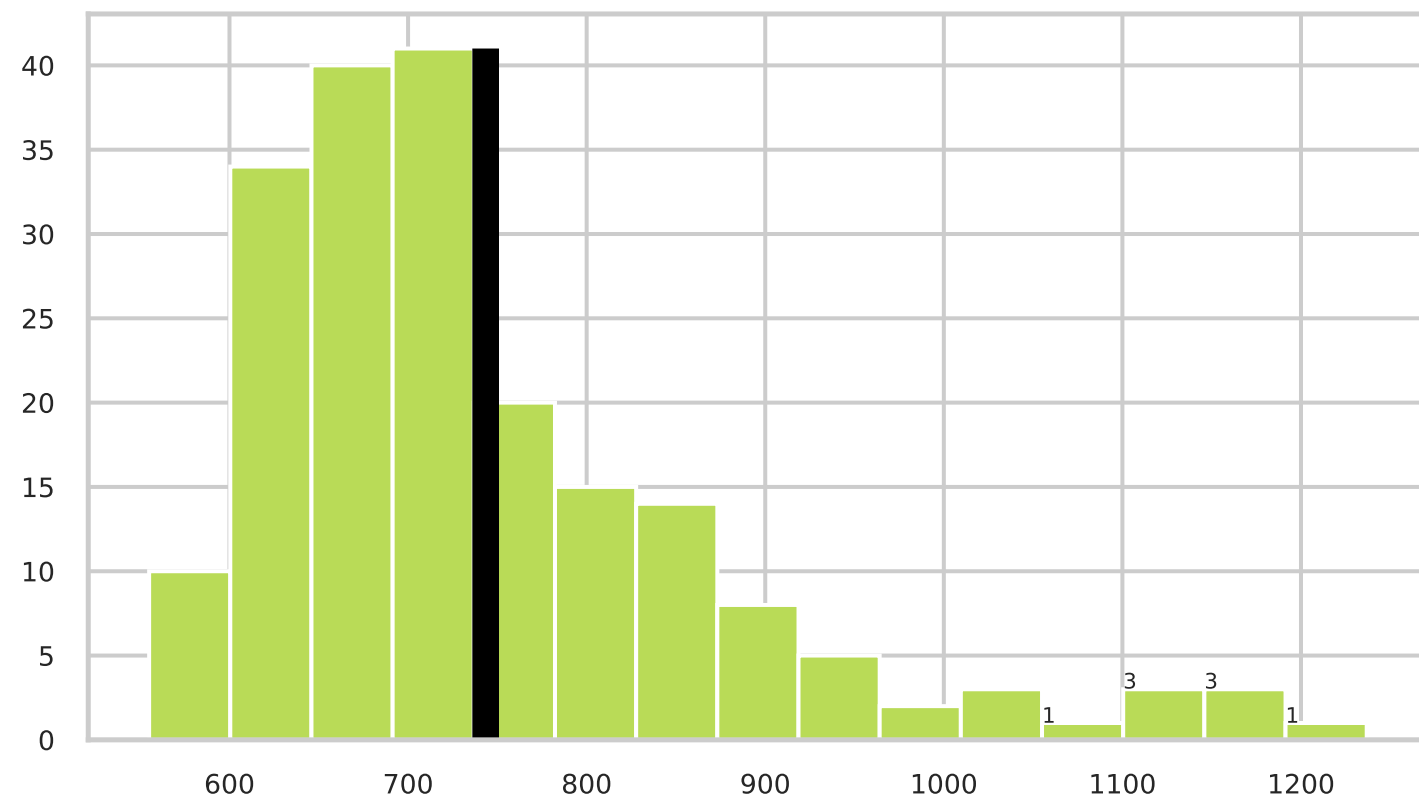
CentralizedMultiplePlay(kl-UCB)



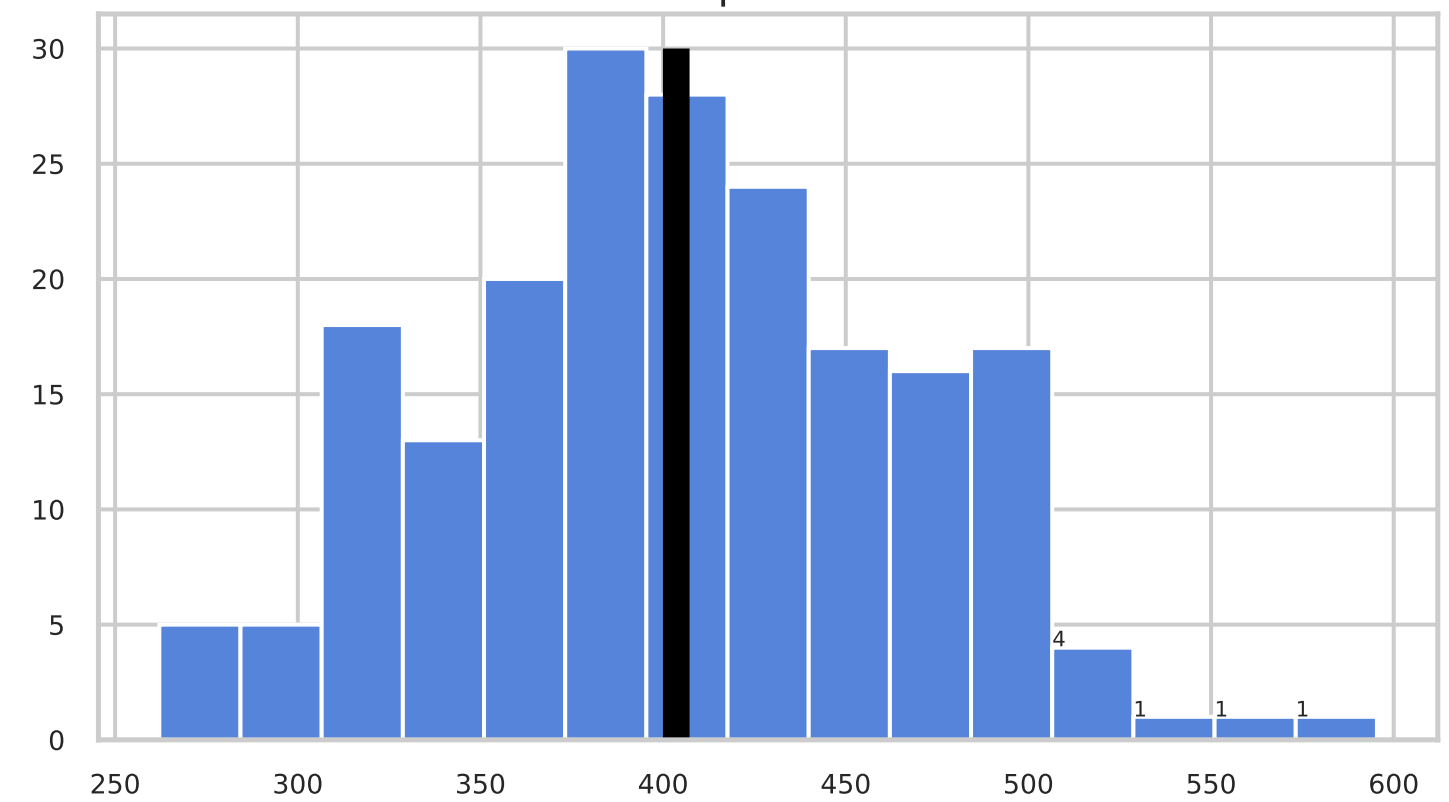
RhoRand-kl-UCB



Selfish-kl-UCB



MCTopM-kl-UCB



Regret value  $R_T$  at the end of simulation, for  $T=10000$

Number of observations, 200 repetitions