Total rewards for each learning rate, inverse temperature pair
for a simulation ran 100 times.