Action probabilities after 0 episodes