## Behavior Policies: Mean Track Loss Probability 0: myopic, innov. 4: myopic, prob. 0.35 1: non-myopic, innov. 5: non-myopic, prob. 2: myopic, innov. & RI 6: myopic, prob. & RI 3: non-myopic, innov. & RI 7: non-myopic, prob. & RI 0.30 Mean track loss probability 0.25 0.20 0.15 0.10 0.05 0.00 Training episodes