Average Escalation Score Over Time (Invasion Scenario) -GPT-4 **GPT-3.5** Claude-2.0 -Llama-2-Chat -GPT-4-Base 20 Escalation Score ← 15 10 5 10 12

Time t [Days]