Political Stability (GPT-4) **Prompt Ablation** -Unablated 60 -+-No Messaging ──No History −♦−No Past Actions ---Shutdown When Nuked 50 -X-No Goals −♦−Action Autonomy Political Stability ─X—Low-Stakes Simulation 40 30 20 10 6 10 12 Time t [Days]