Soft Power (GPT-3.5) **Prompt Ablation** Unablated -+-No Messaging 60 ──No History −♦−No Past Actions ——Shutdown When Nuked 50 -X-No Goals −♦−Action Autonomy ─X—Low-Stakes Simulation Soft Power 40 30 20 10 10 12 14 Time t [Days]