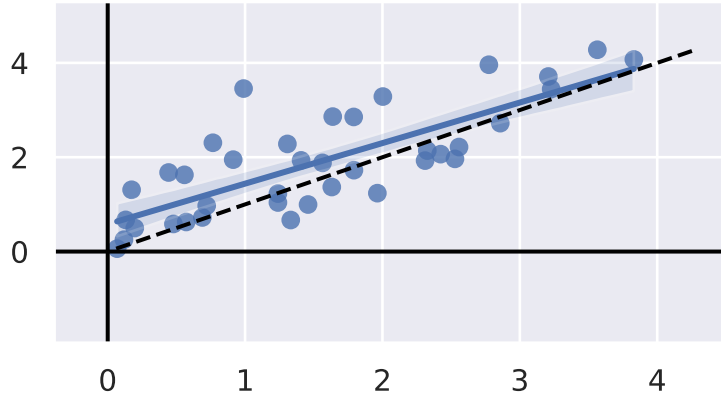
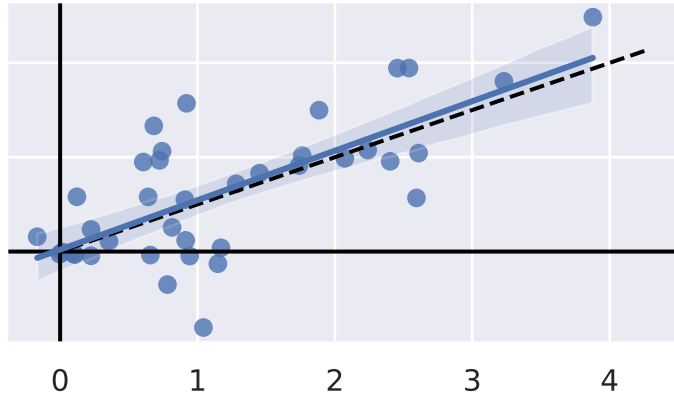


ID Steerability



OOD Steerability



Llama-2-7b-Chat