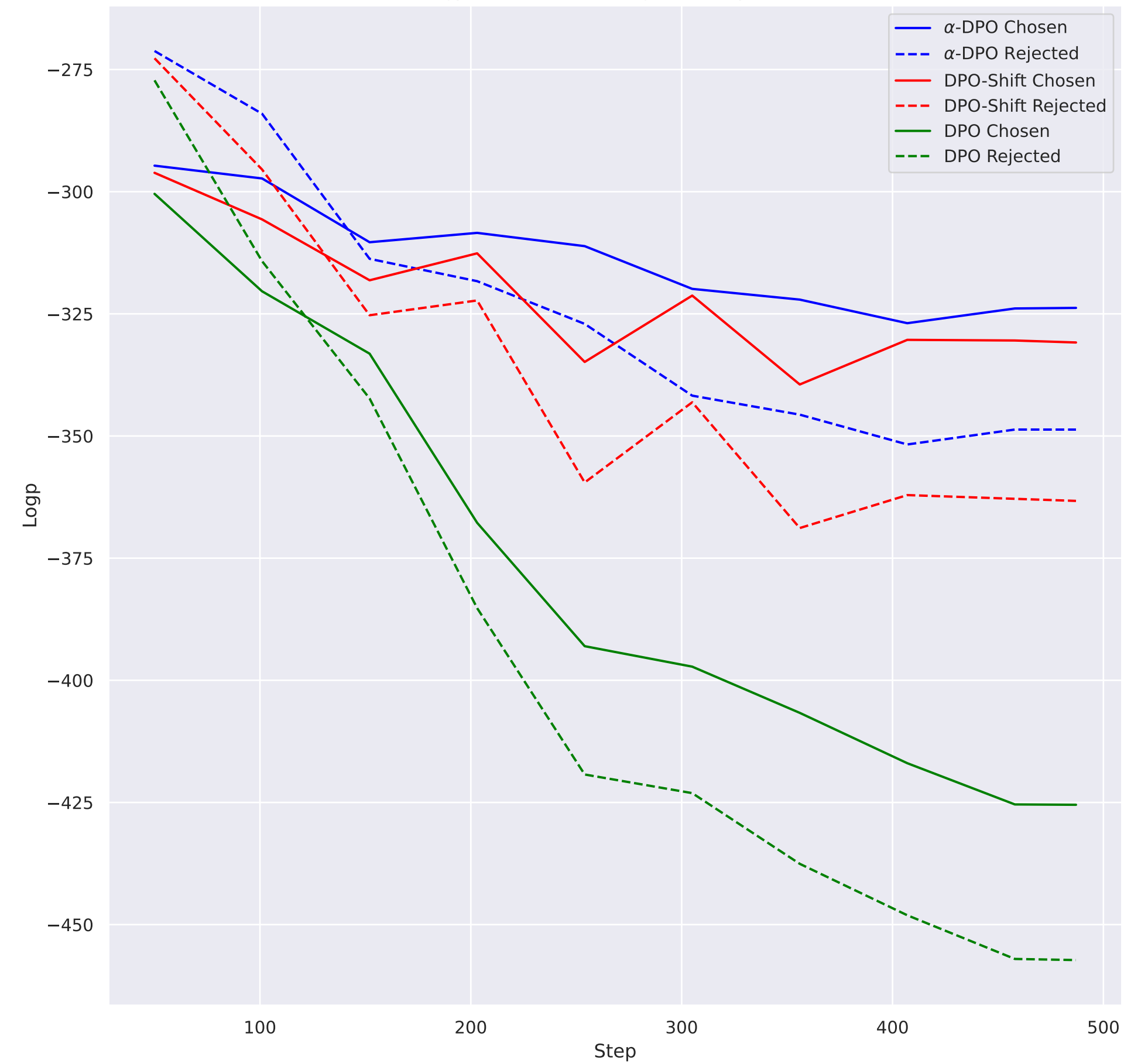


logp for Chosen and Rejected responses



Reward Margin

