

See Figure 7 for CDF plots of forecasting accuracy by condition for each question. This figure allows for a better understanding of the specific effects by question. For example, it shows that the majority of the accuracy advantage that the noisy LLM augmentation condition enjoys over the other two conditions is due to having less predictions that were at the winsorized bound.

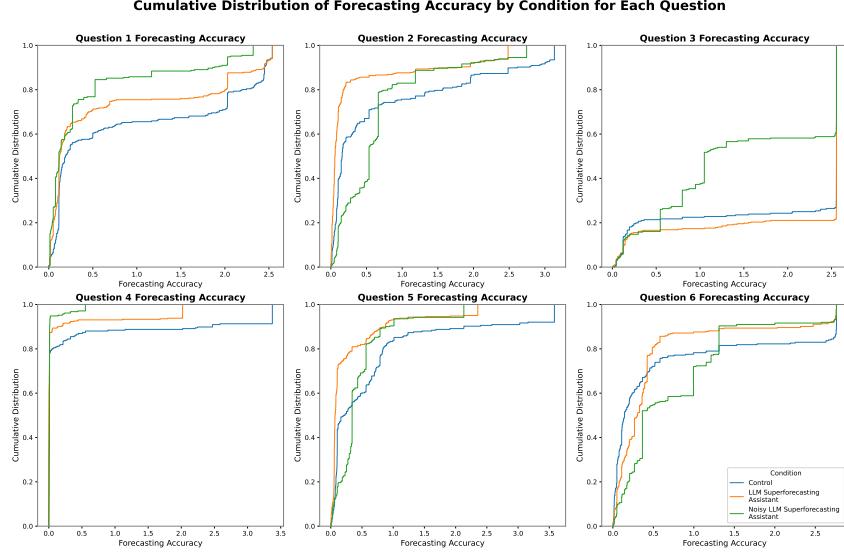


Figure 7: CDF plots of forecasting accuracy by condition for each question.