

Figure 11: Breakdown of the metrics reported in Tier 4, with respect to different contextual factors. The Leak metric shows the ratio of cases where there is a leakage, and the  $\sim$ Item shows missing action item. Lower is better for all values. Top row is the results without privacy prompts, and the bottom row is the results with the privacy prompts (the ones presented in the main body of the paper).