Y1: performance of single uncalibrated vs calibrated model on subgroup-by-insurance: government uncalibrated-one-model-for-all calibrated-one-model-for-all 100 -89.7 90.7 86.1 80 69.2 Score(%) 60 -55.4 <sub>54.4</sub> 46.2 39.7 40 20 f1 precision recall accuracy