acc values wrt train set bias level: (OULADstem, unbiased test set) 1.0 0.9 0.8 0.7 label bias label+reweighing 0.6 label+massaging label+FTU selection bias 0.5 select+reweighing select+massaging select+FTU 0.4 0.6 0.0 0.2 0.8 Bias intensity (β_l for label, p_{ij} for selection)