Tableshift: Food Stamps constant 0.75 top all features top causal features out-of-domain accuracy 0.70 top arguably causal features 0.65 0.60 0.55 0.50 0.50 0.55 0.60 0.65 0.70 0.75 in-domain accuracy