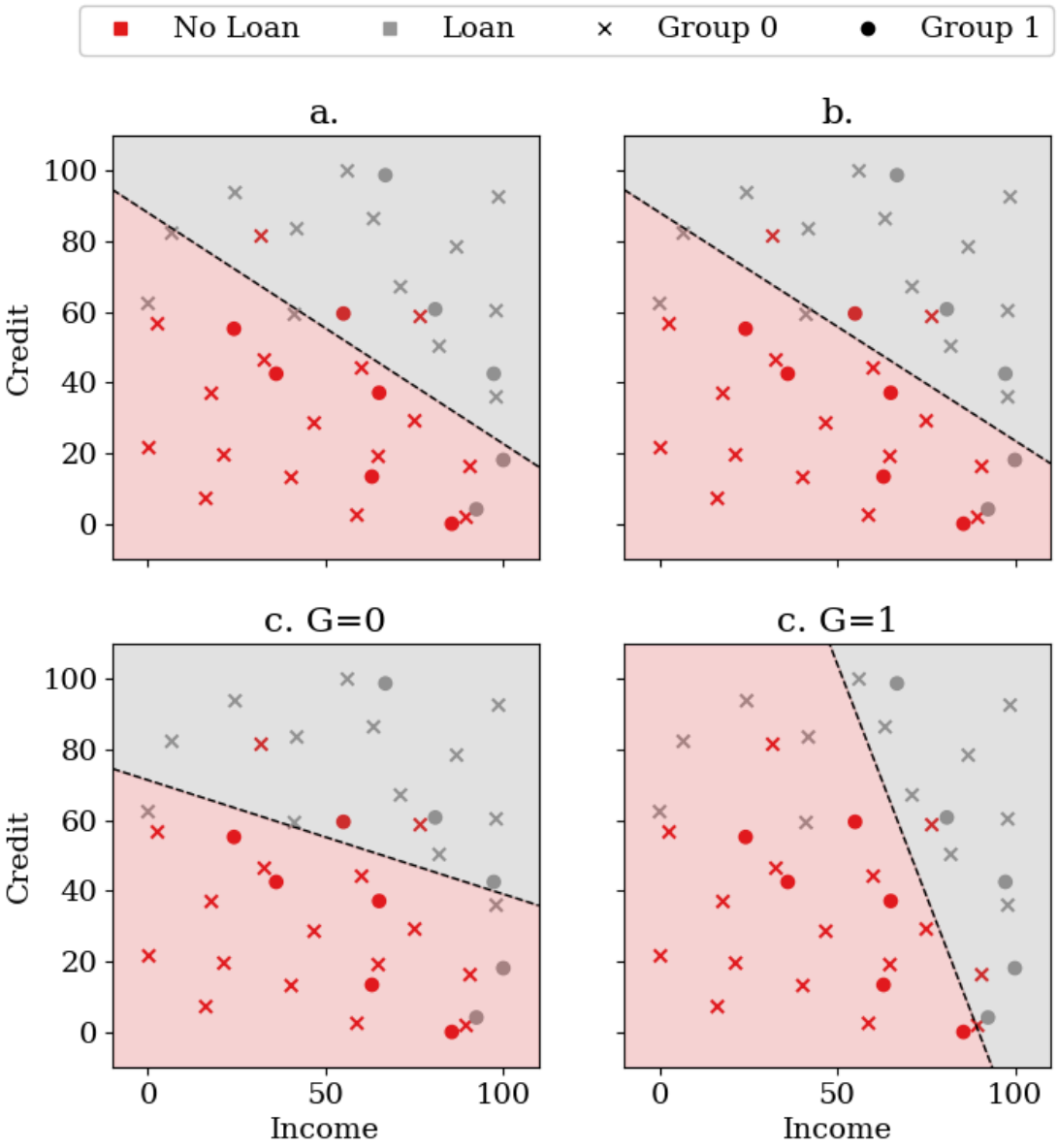

a.

i. In terms of demographic parity the classifier does a fair job on assigning a positive label for each group. Based on the demographic parity metric, group 0 only has 0.62% higher chance of being provided a loan. However in terms of equalized odds, the negative label inaccuracy for group 1 is 10.92% higher than group 0. Which is a significant difference!

ii. I would want to be group 0 because that would mean that I will have a 10.92% higher chance of getting a loan that I should not get.

b.

i. The only difference is that the group type is also added as a feature.

ii. Decision boundary and fairness metrics are equivalent to part a. This does not surprise me because the demographic category of the loan seeker should not have any effect on the output of the model.

c.

i. The one on the left hand side is trained to only make predictions on group 0, the one on the rhs is trained to make predictions only for group 1.

ii. I would like to be classified by part c. “G=0” classifier, because that has 1 more ‘x’ label on the positive part.

iii. What I find interesting from part b to part c is in part c. “G=1” classifier. It seems that the model puts significantly more importance on the income of group 1 compared to their credit score.