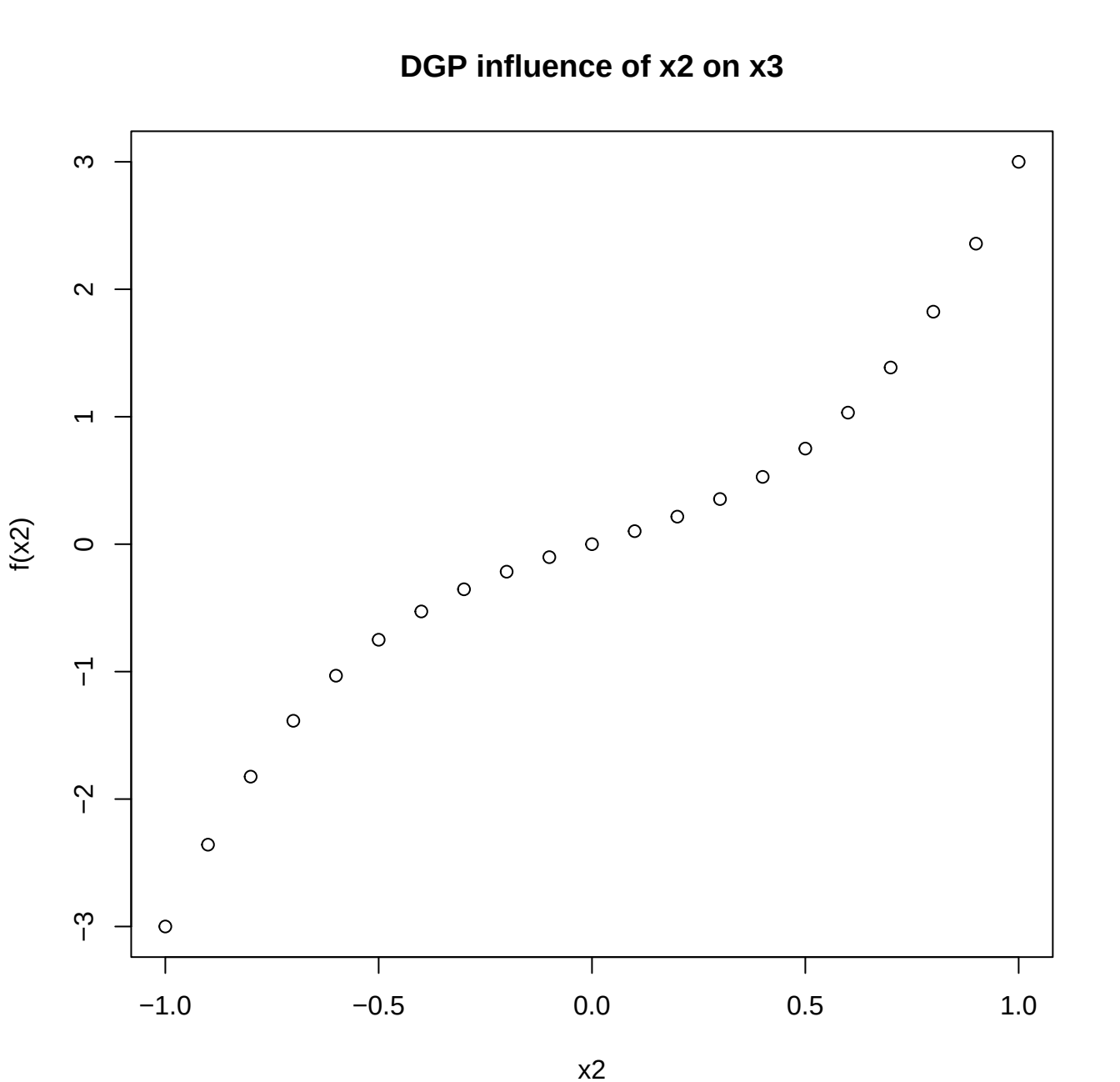
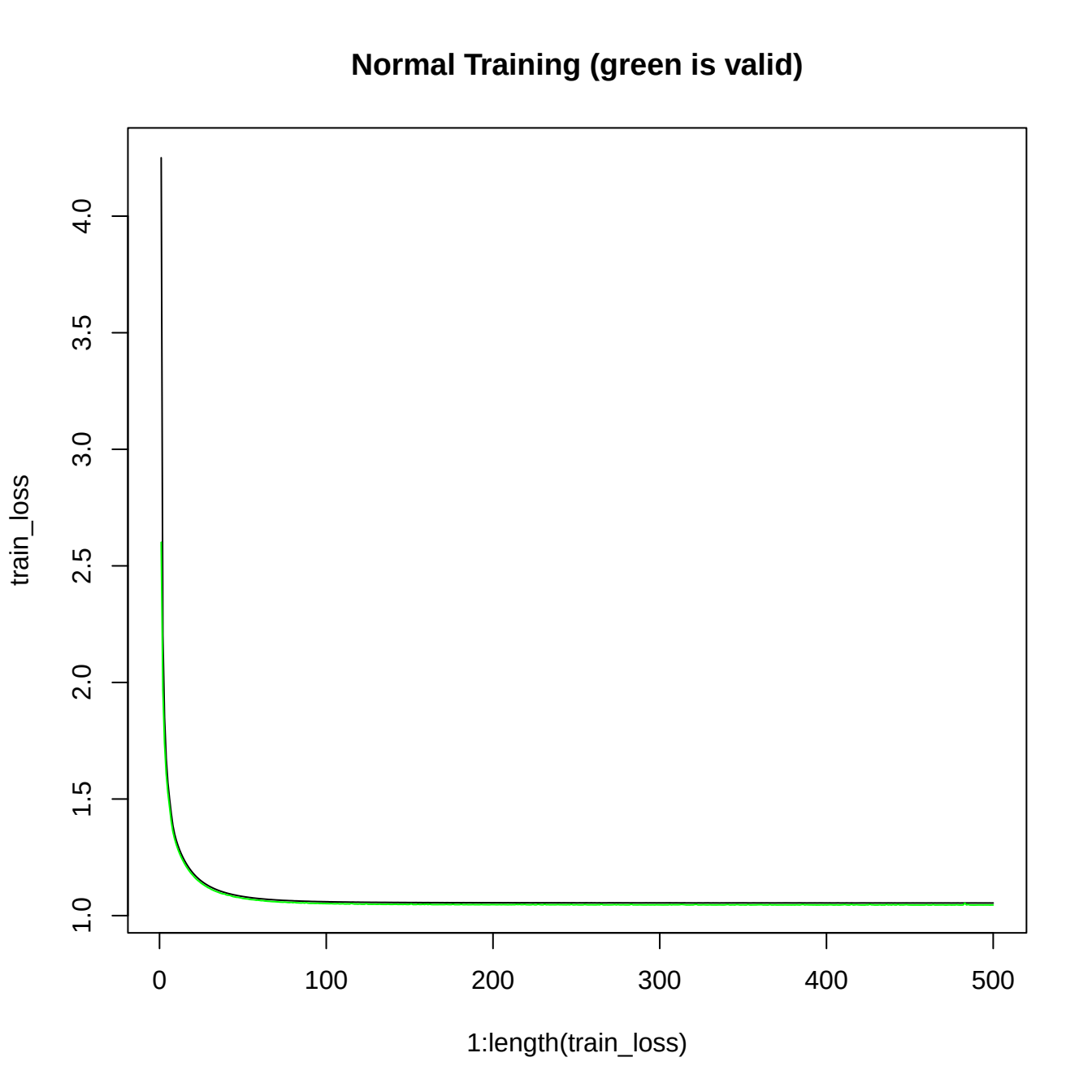
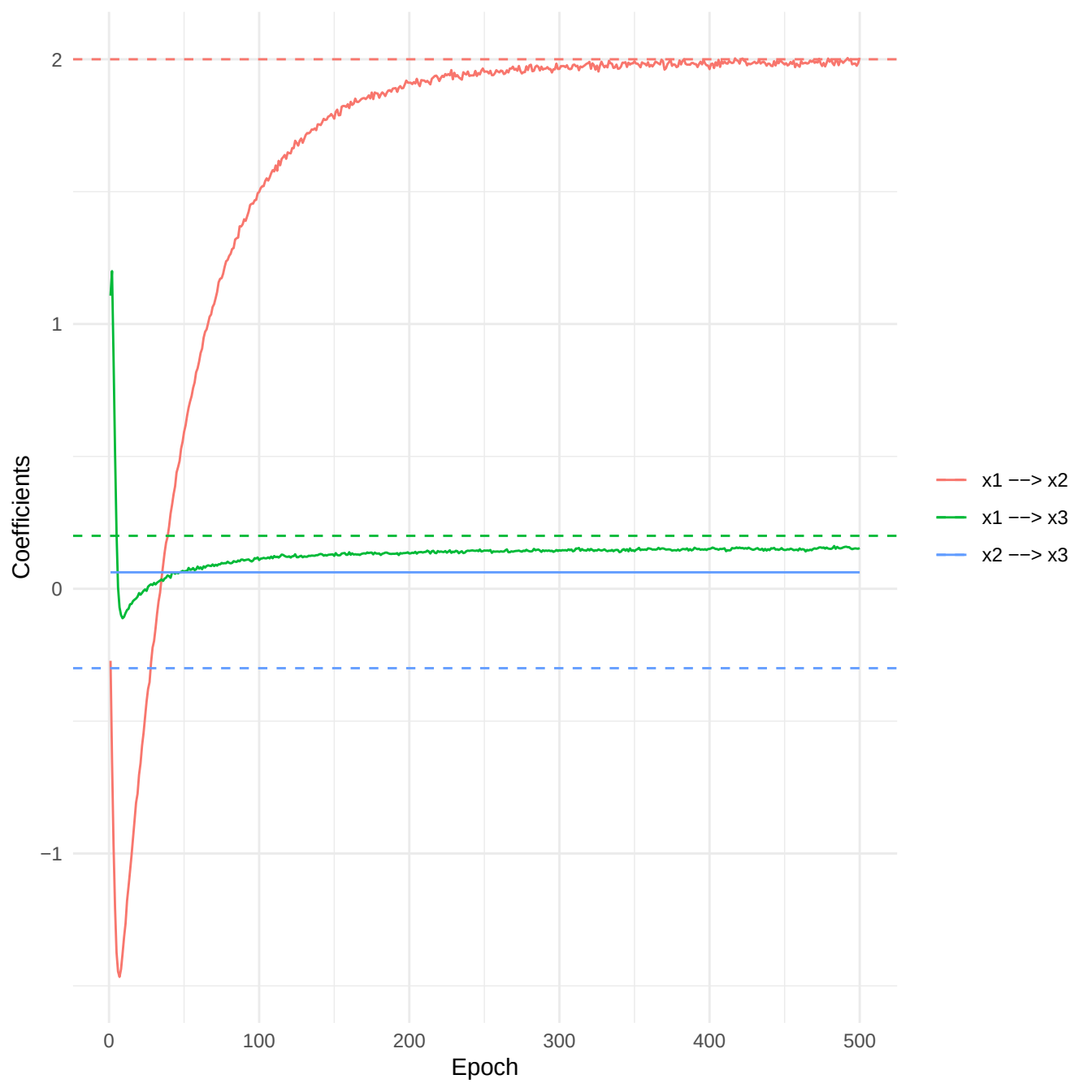
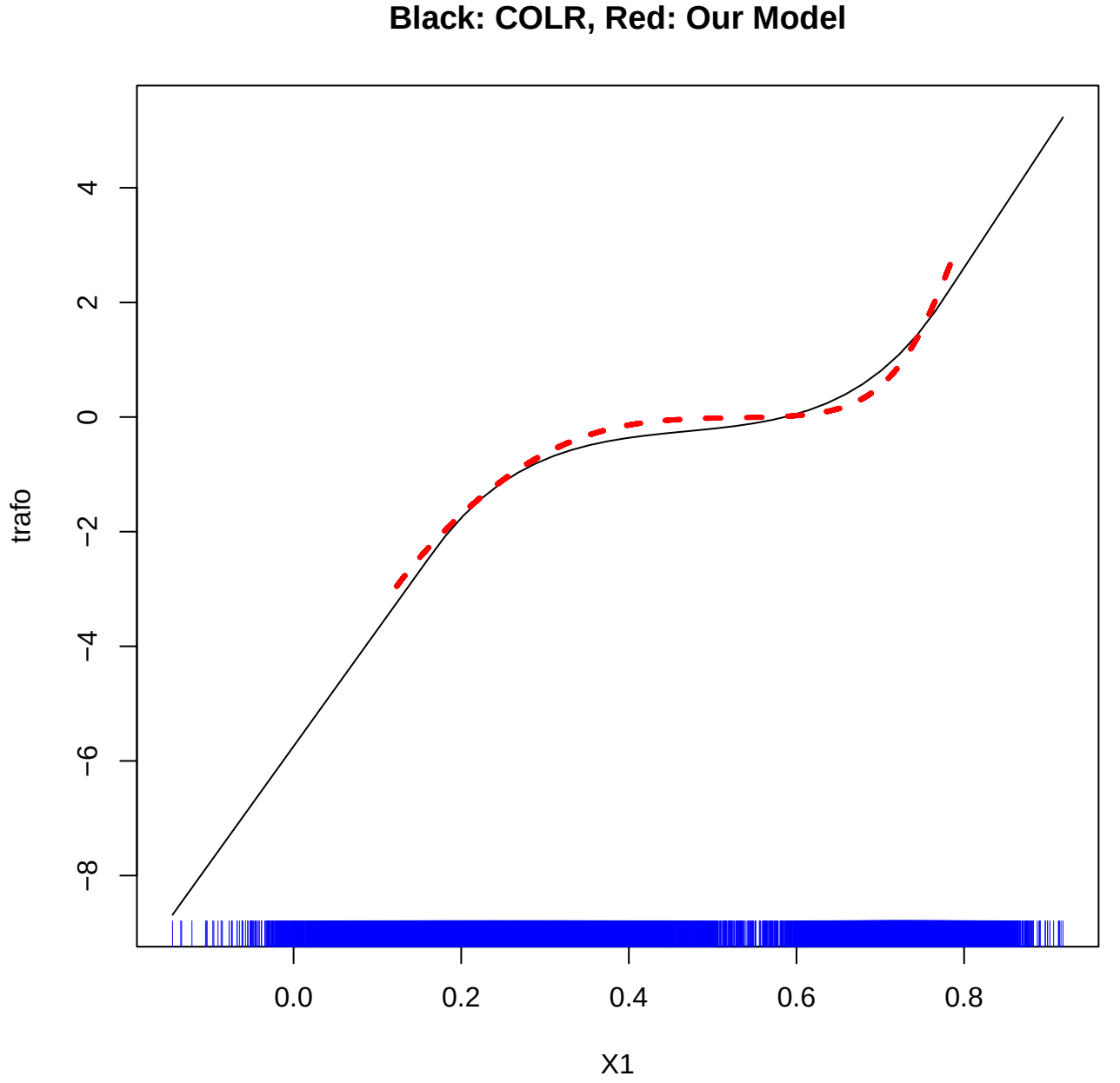
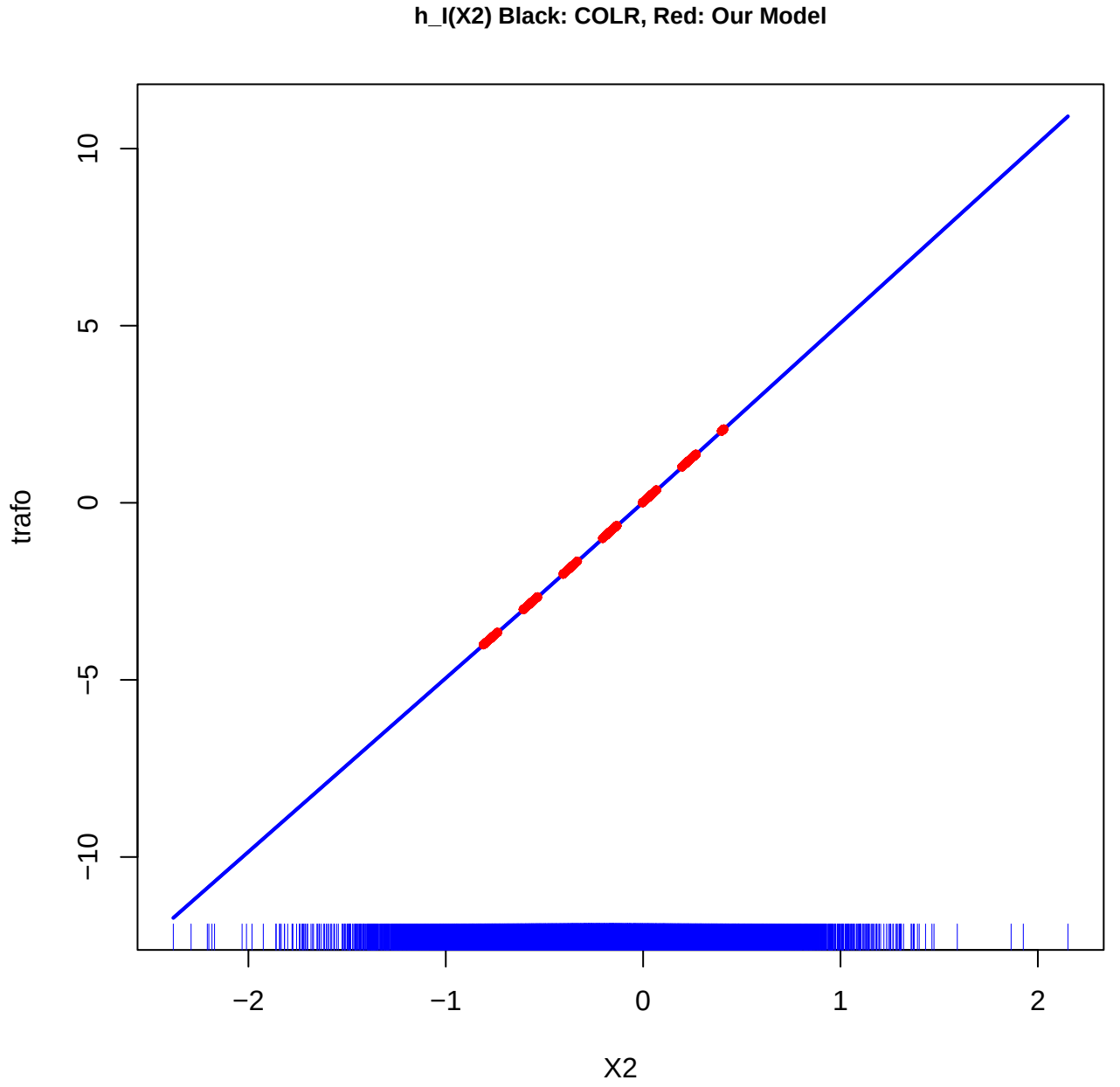
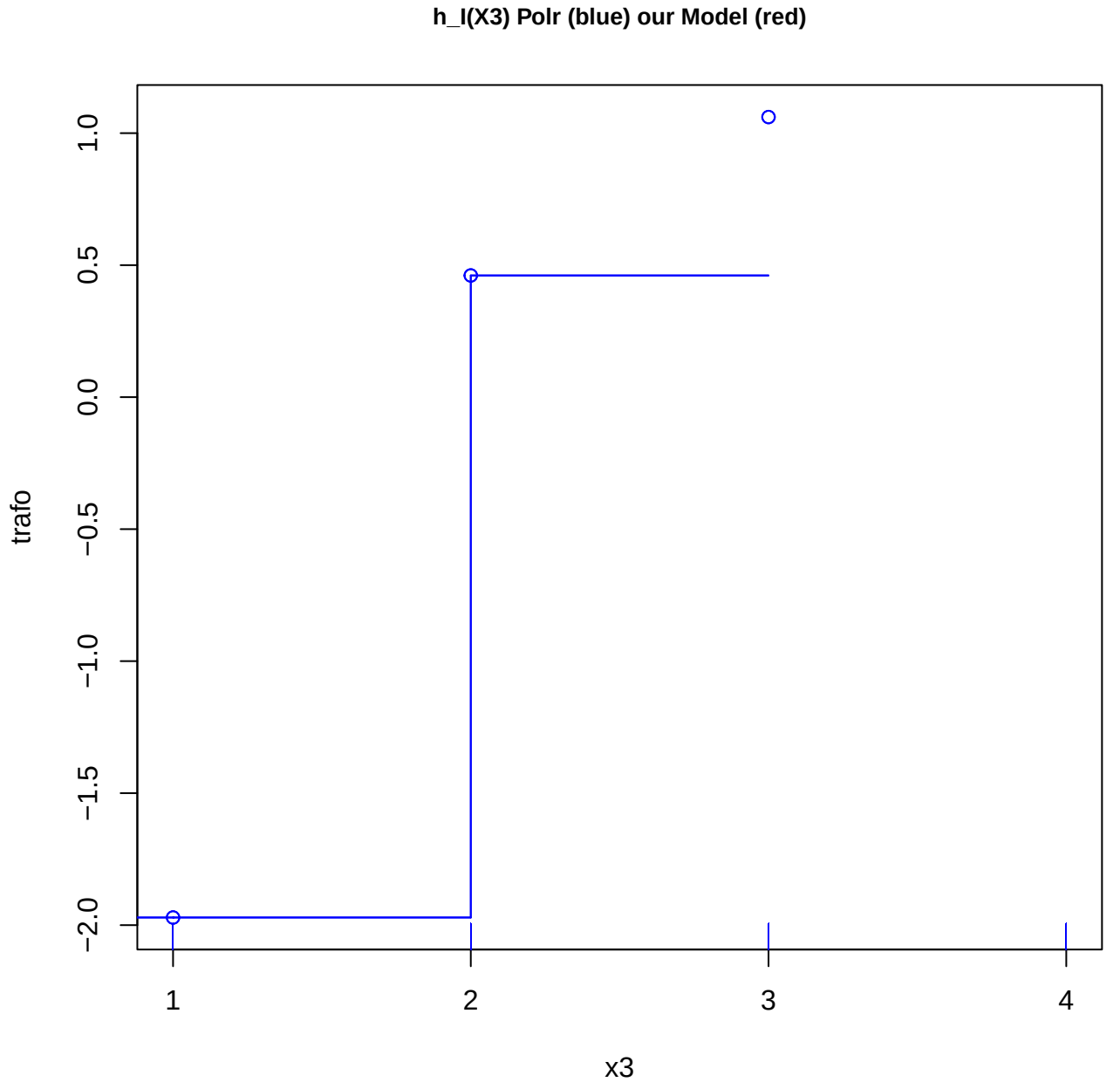

DGP influence of x2 on x3
f(x2)
3
2
1
0
−1
−2
−3
−1.0
−0.5
0.0
0.5
1.0
x2
summerof24/runs/triangle_structured_mixed/run_small_net/triangle_mixed_DPG2x3+x_Mode

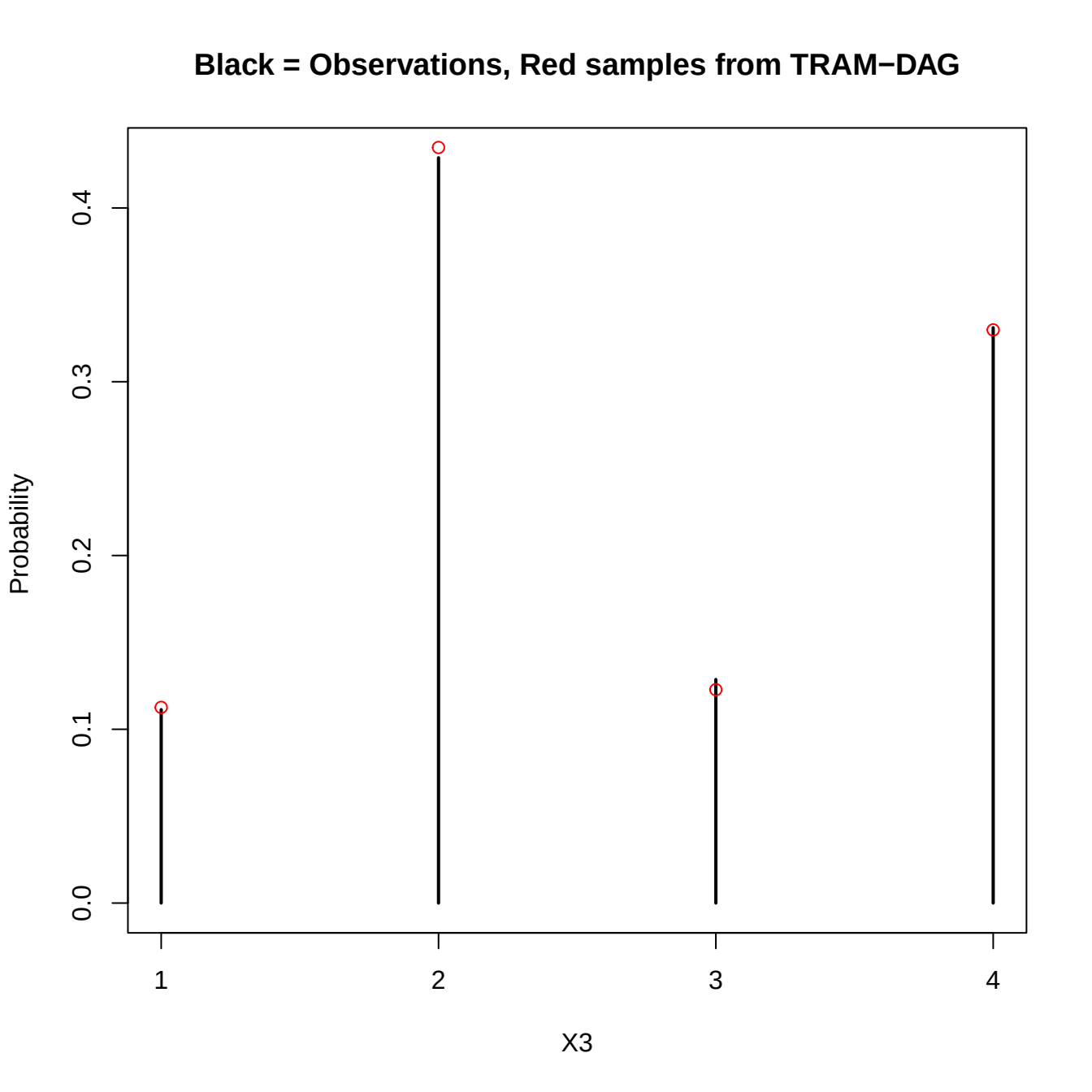

Normal Training (green is valid)
train_loss
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0
100
200
300
400
500
1:length(train_loss)

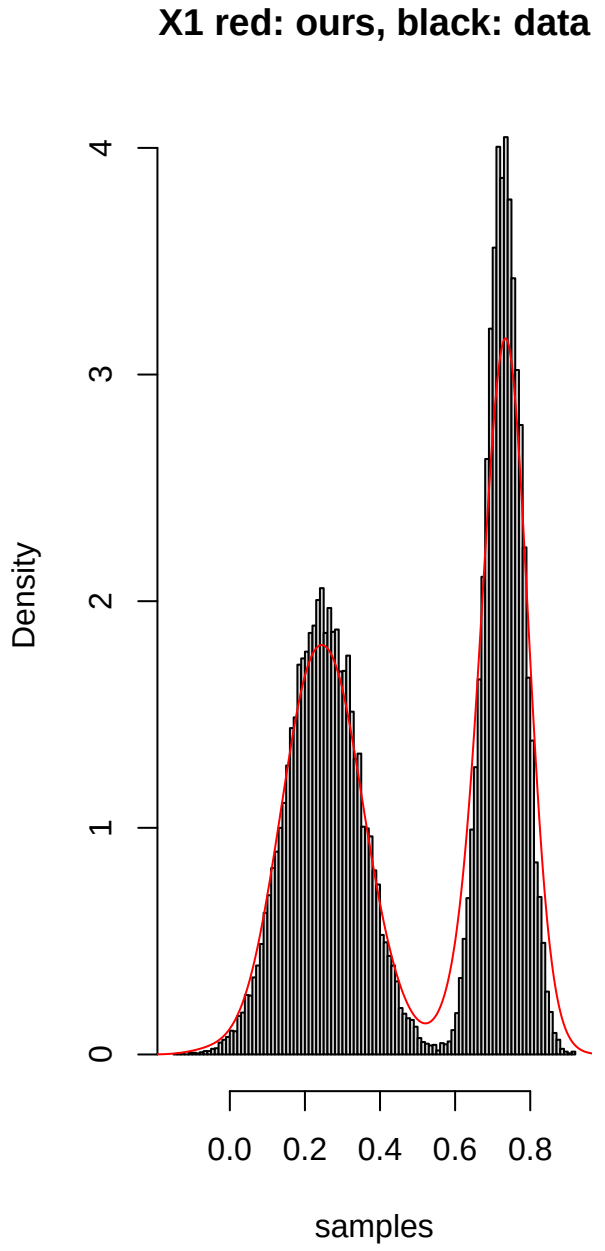

2
1
0
−1
Coefficients
0
100
200
300
400
500
Epoch
x1 −−> x2
x1 −−> x3
x2 −−> x3

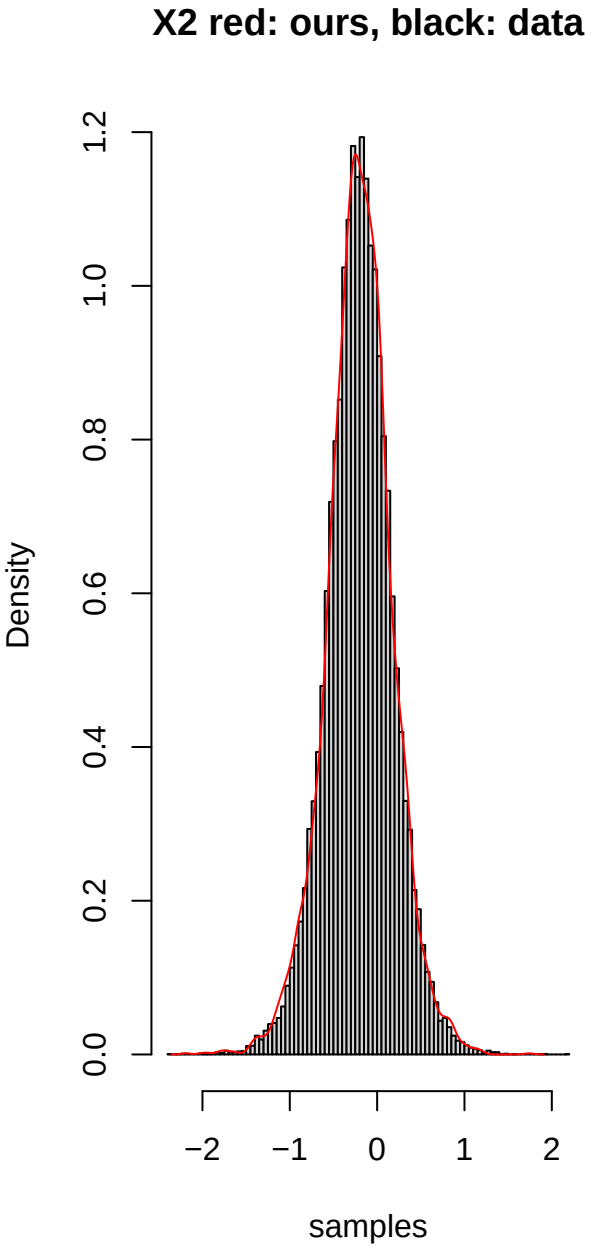

Black: COLR, Red: Our Model
trafo
4
2
0
−2
−4
−6
−8
0.0
0.2
0.4
0.6
0.8
X1

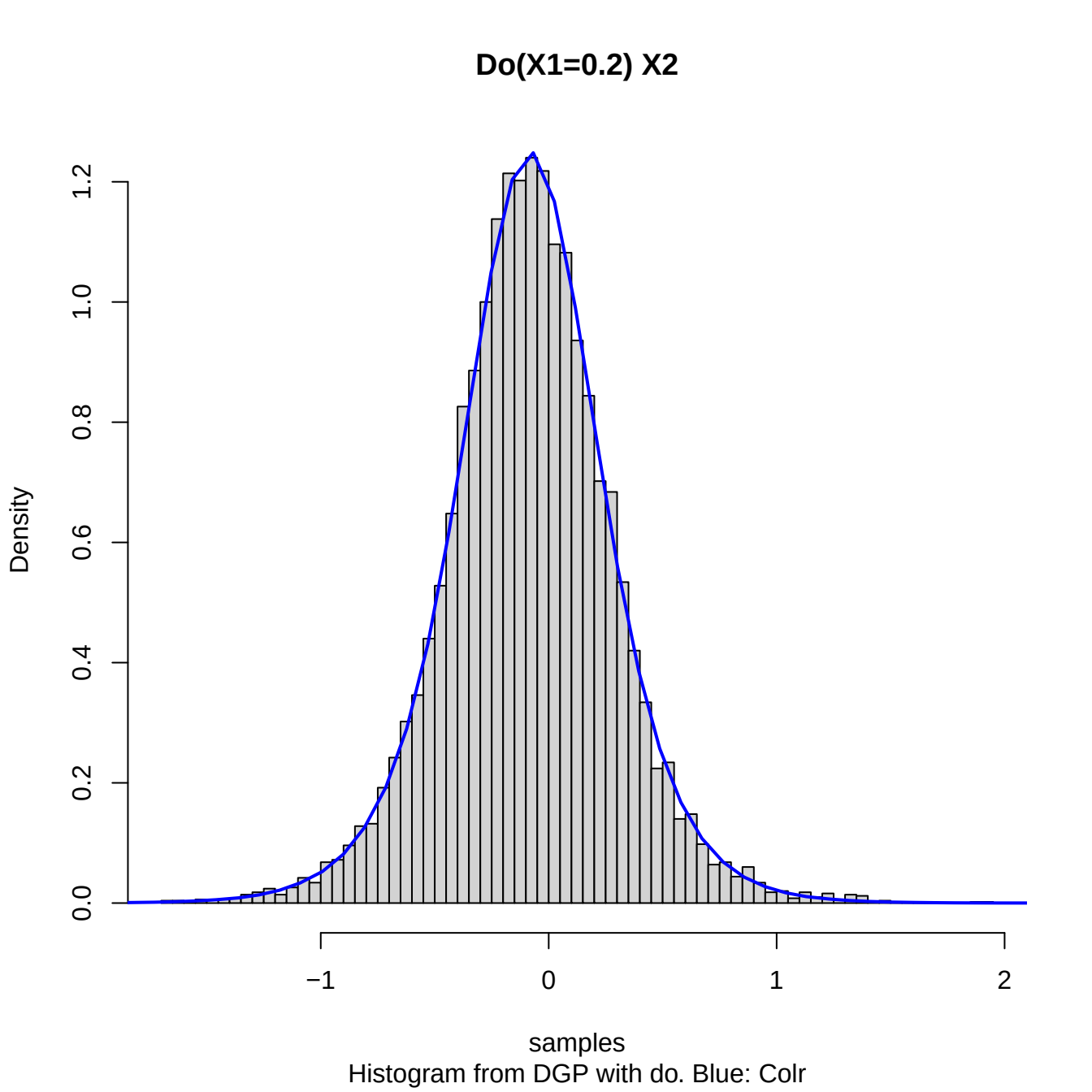

h_I(X2) Black: COLR, Red: Our Model
trafo
10
5
0
−5
−10
−2
−1
0
1
2
X2

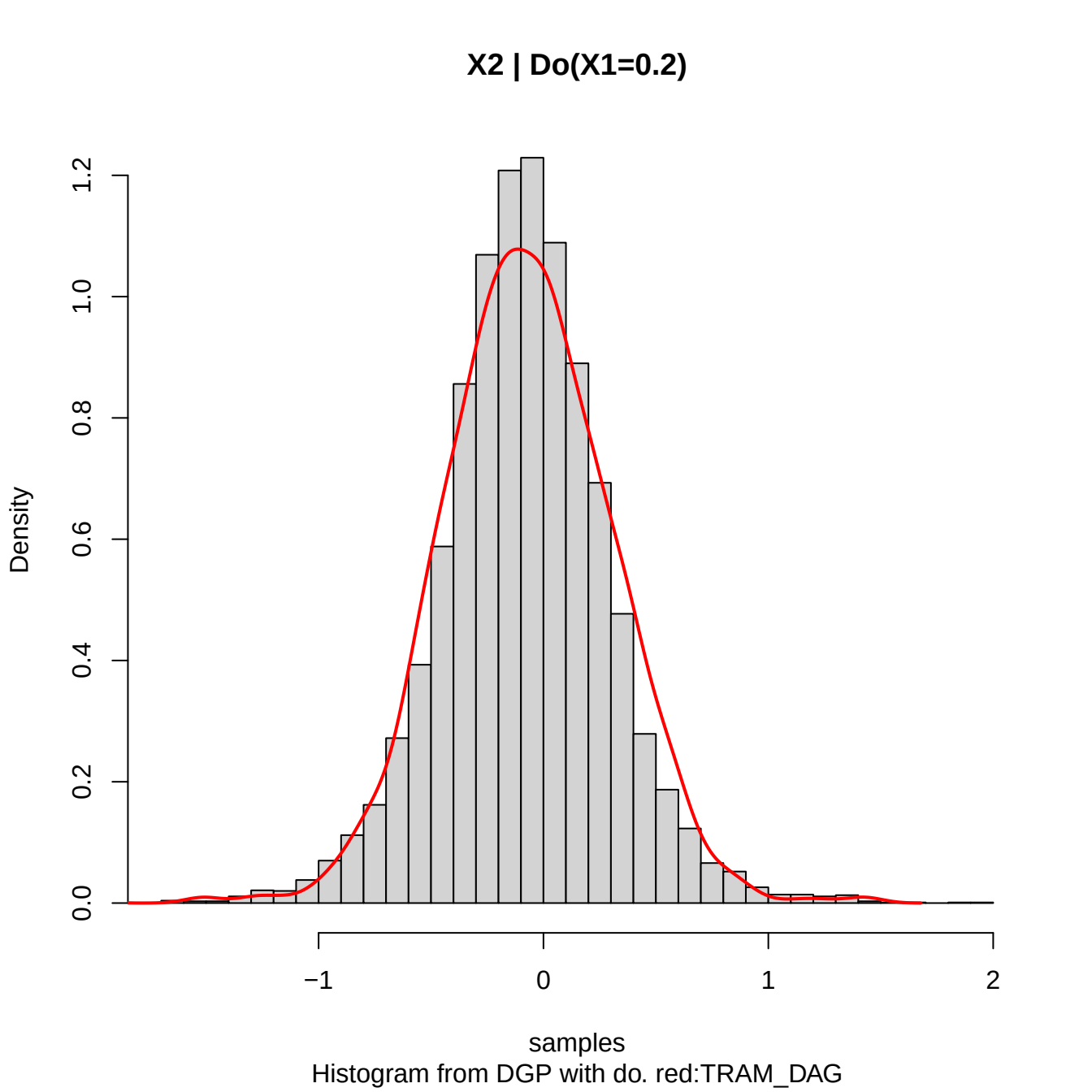

h_I(X3) Polr (blue) our Model (red)
trafo
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
1
2
3
4
x3

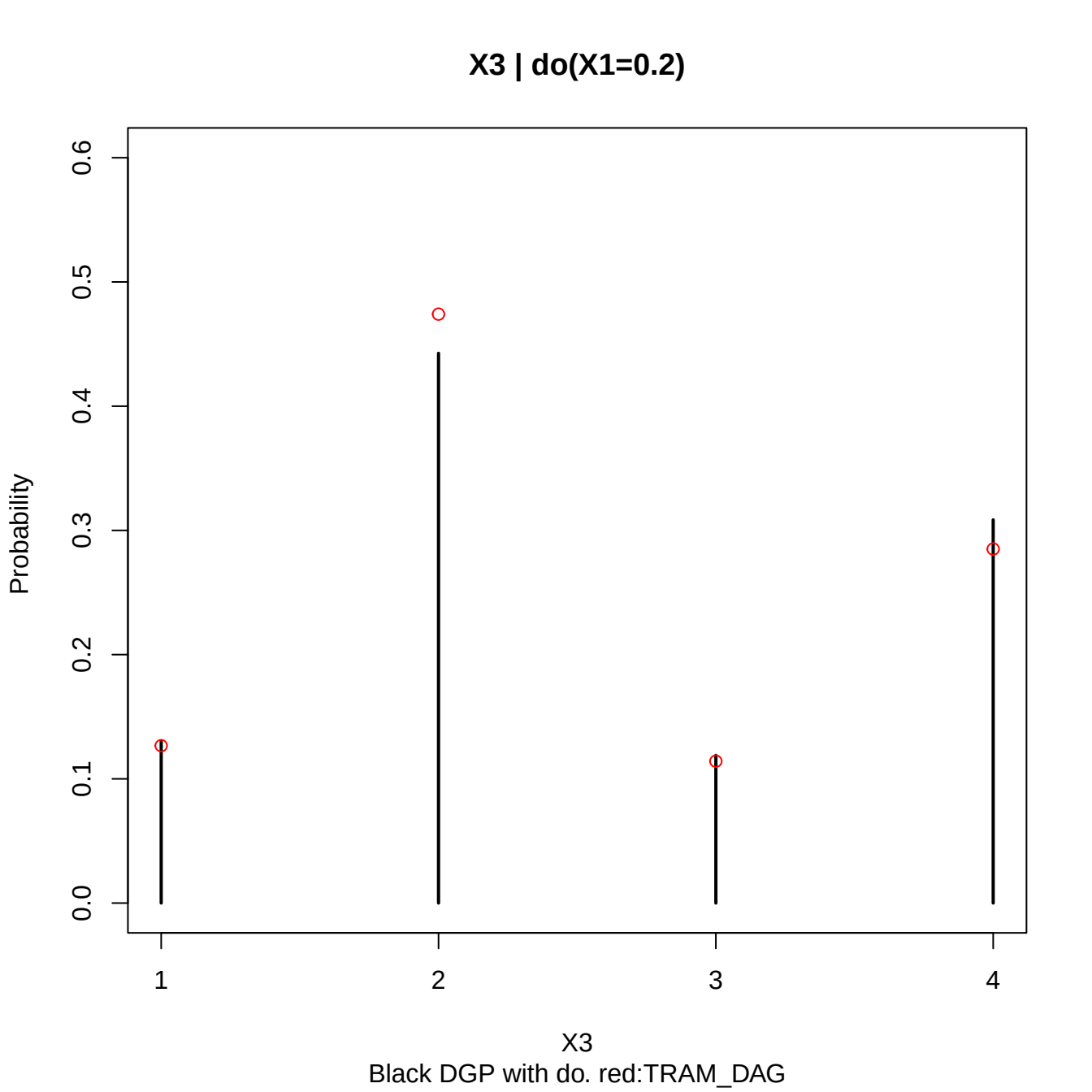

Black = Observations, Red samples from TRAM−DAG
Probability
0.0
0.1
0.2
0.3
0.4
1
2
3
4
X3

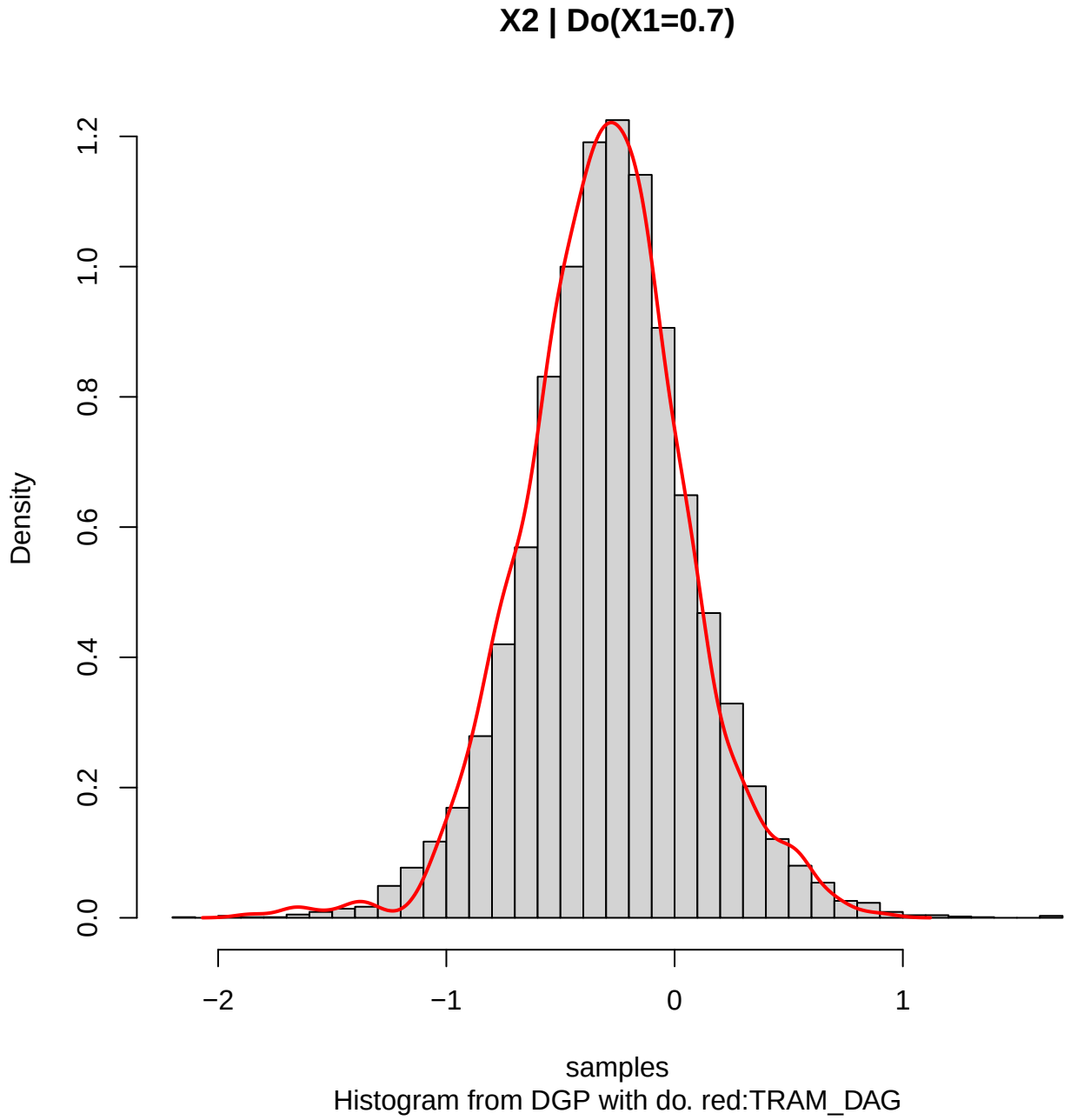

X1 red: ours, black: data
Density
0
1
2
3
4
0.0
0.2
0.4
0.6
0.8
samples

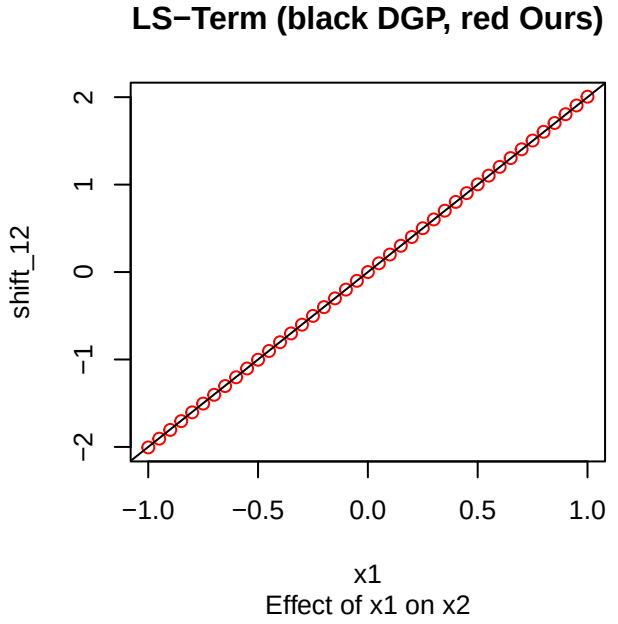

X2 red: ours, black: data
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−2
−1
0
1
2
samples

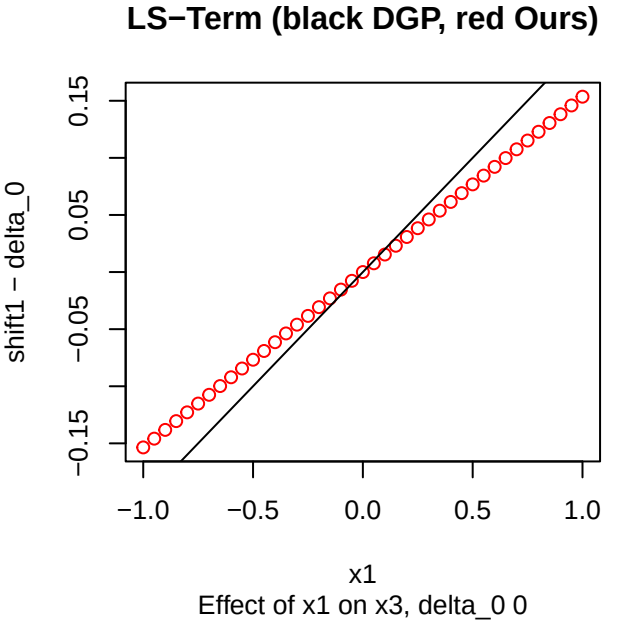

Do(X1=0.2) X2
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−1
0
1
2
samples
Histogram from DGP with do. Blue: Colr

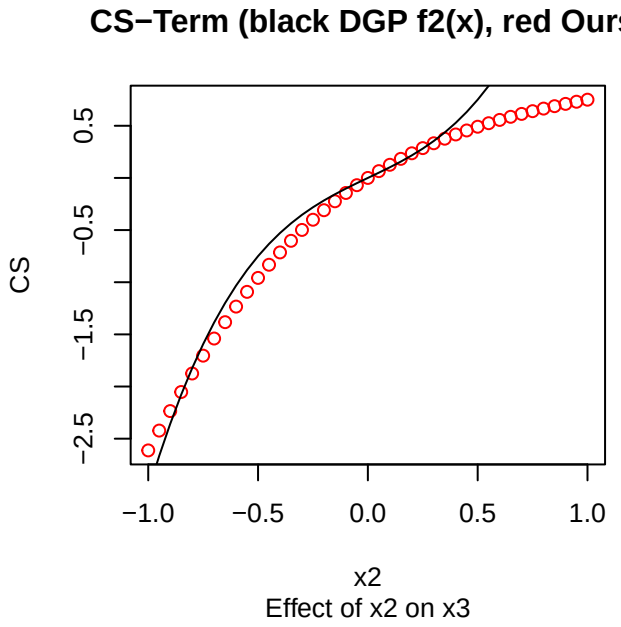

X2 | Do(X1=0.2)
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−1
0
1
2
samples
Histogram from DGP with do. red:TRAM_DAG


X3 | do(X1=0.2)
Probability
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1
2
3
4
X3
Black DGP with do. red:TRAM_DAG


X2 | Do(X1=0.7)
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−2
−1
0
1
samples
Histogram from DGP with do. red:TRAM_DAG


LS−Term (black DGP, red Ours)
shift_12
2
1
0
−1
−2
−1.0
−0.5
0.0
0.5
1.0
x1
Effect of x1 on x2


LS−Term (black DGP, red Ours)
shift1 − delta_0
0.15
0.05
−0.05
−0.15
−1.0
−0.5
0.0
0.5
1.0
x1
Effect of x1 on x3, delta_0 0


CS−Term (black DGP f2(x), red Ours)
CS
0.5
−0.5
−1.5
−2.5
−1.0
−0.5
0.0
0.5
1.0
x2
Effect of x2 on x3