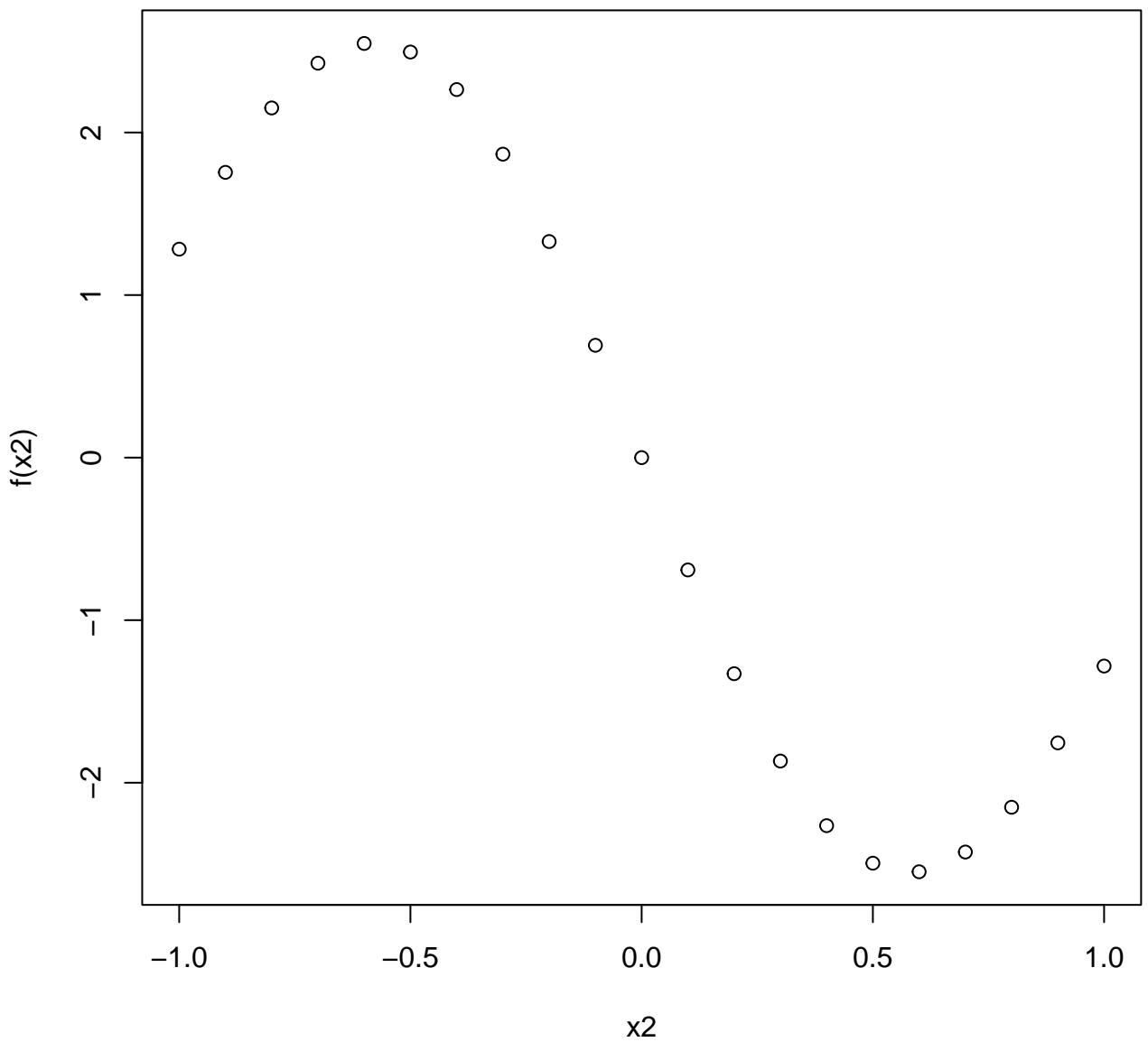
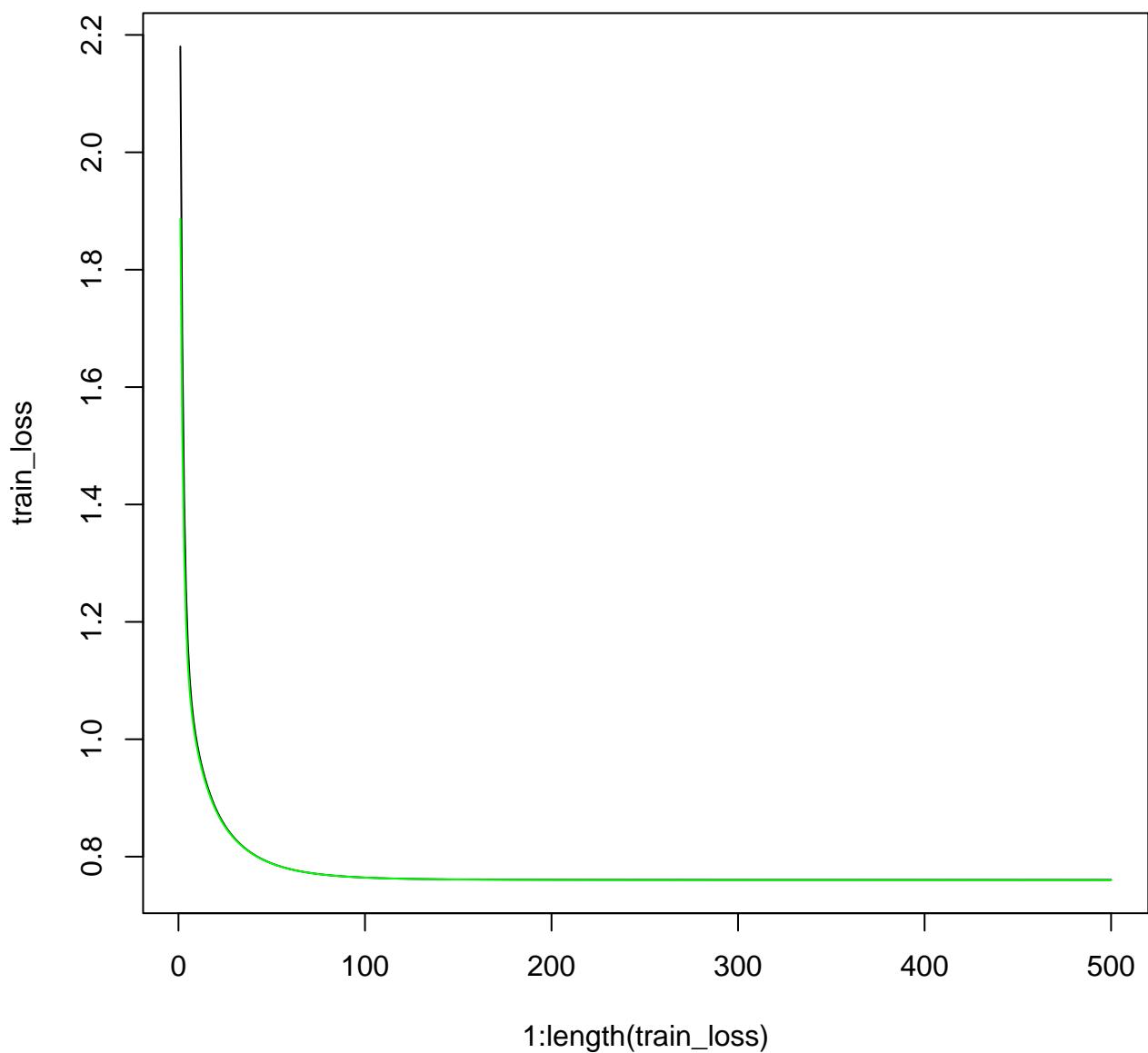


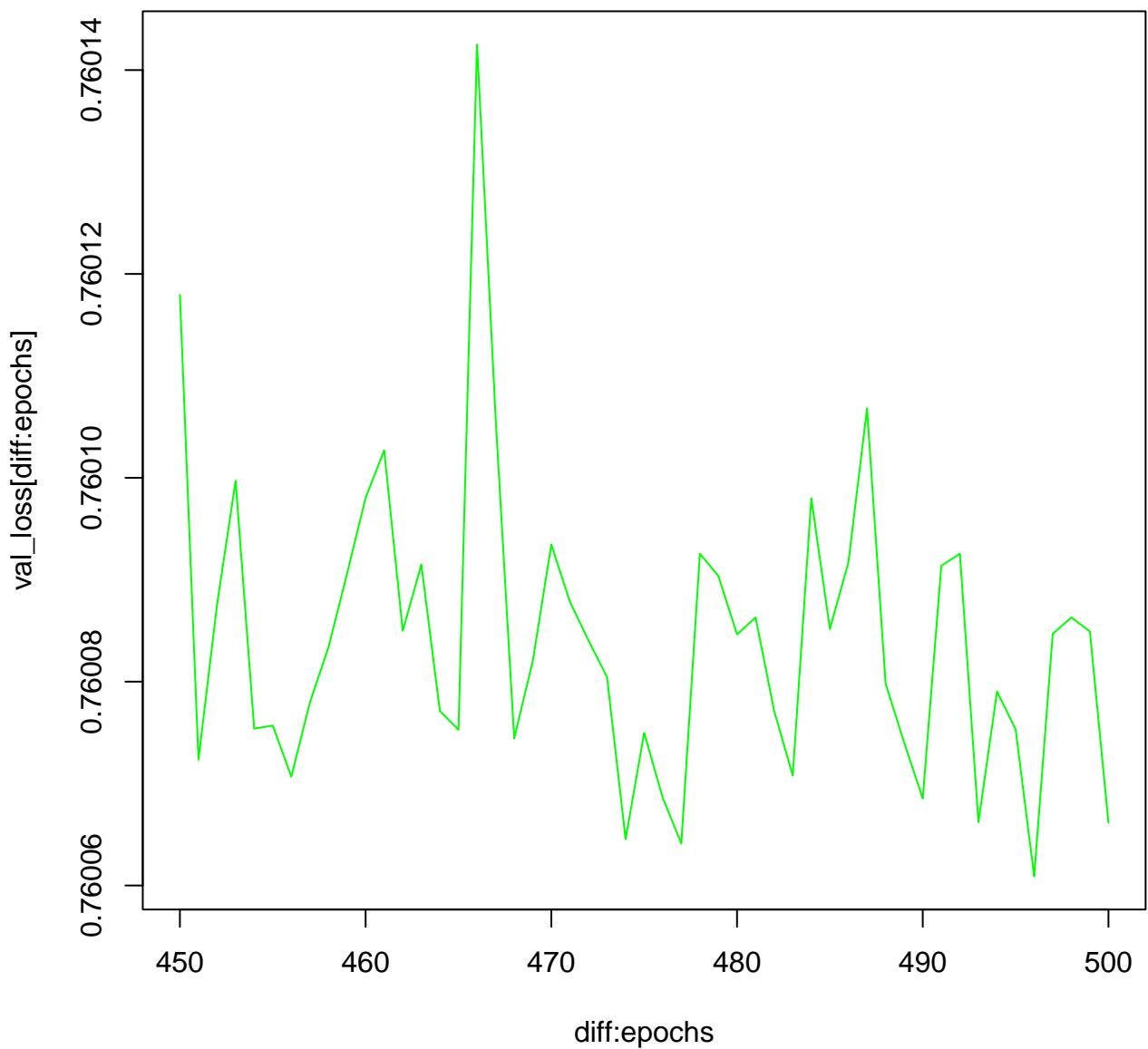
DGP influence of x_2 on x_3

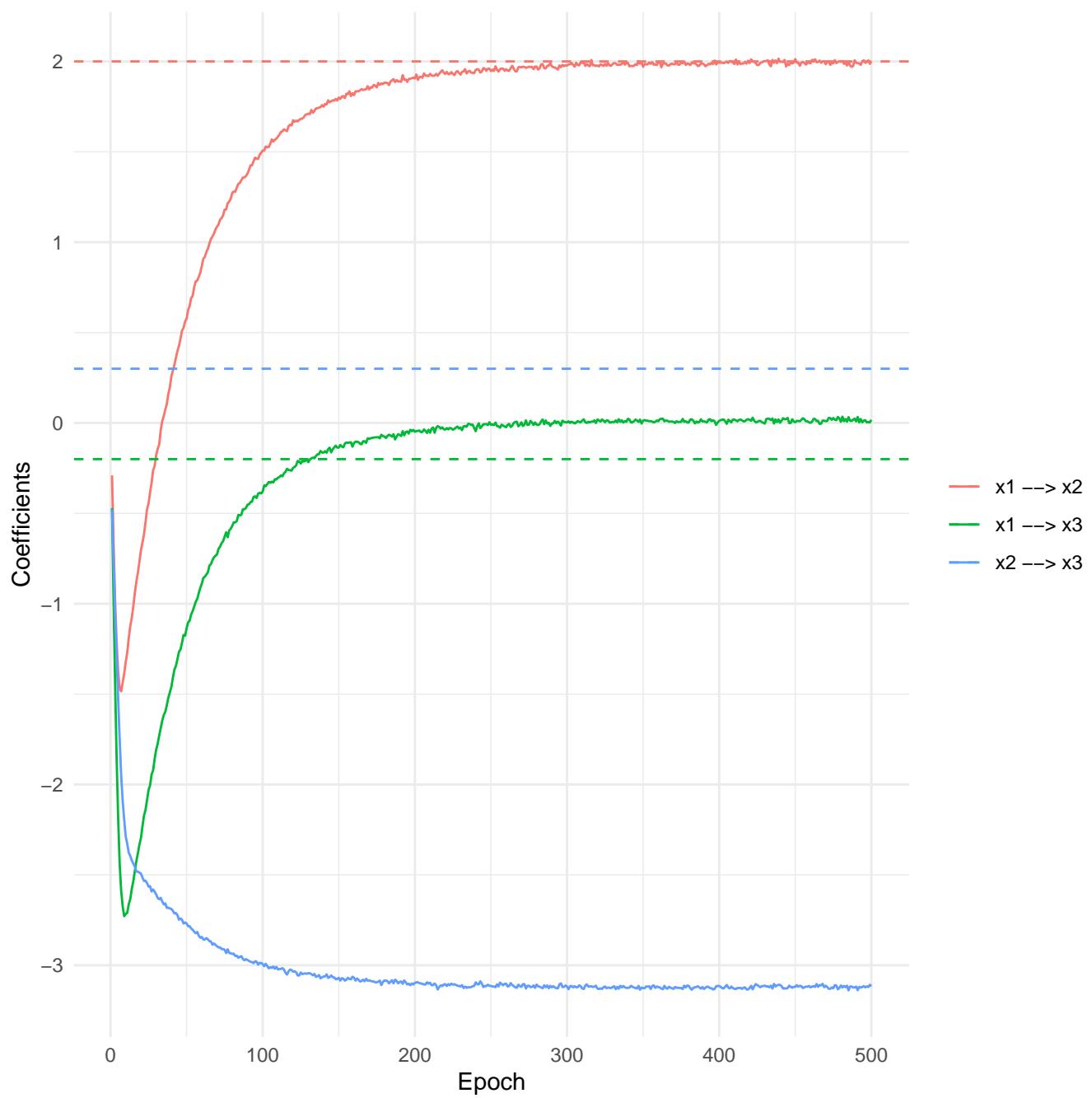


Training (black: train, green: valid)

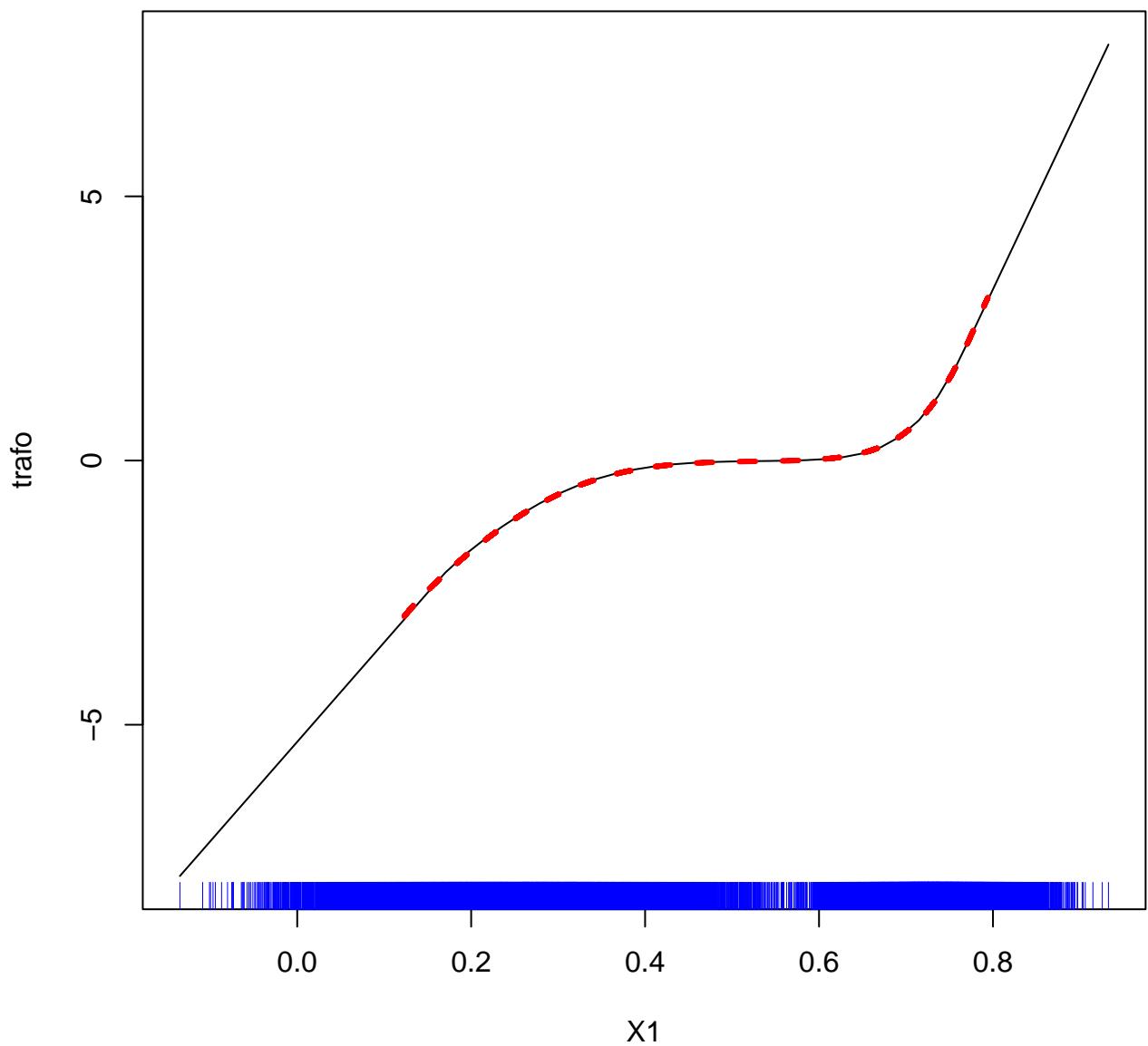


Last 50 epochs

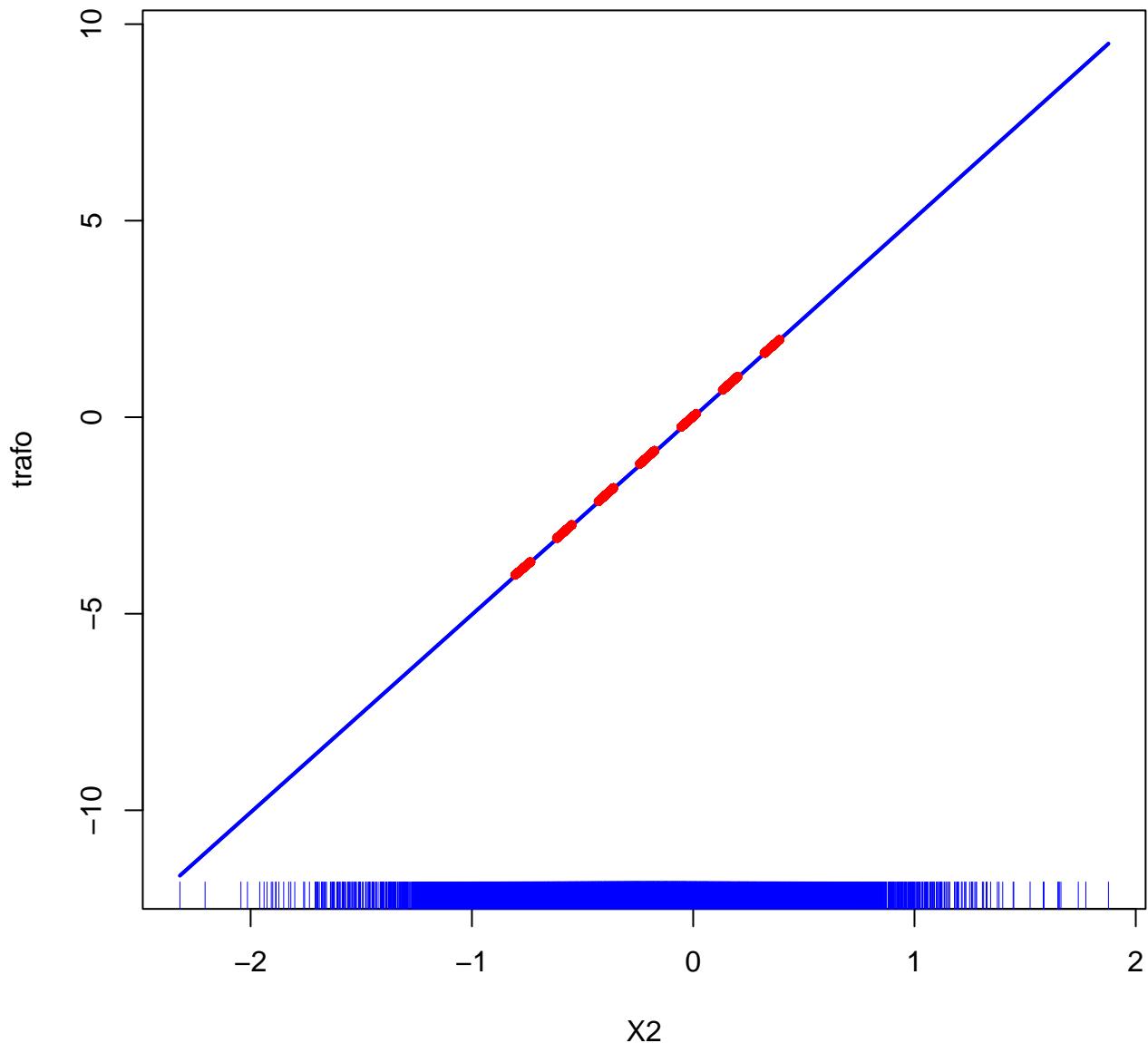




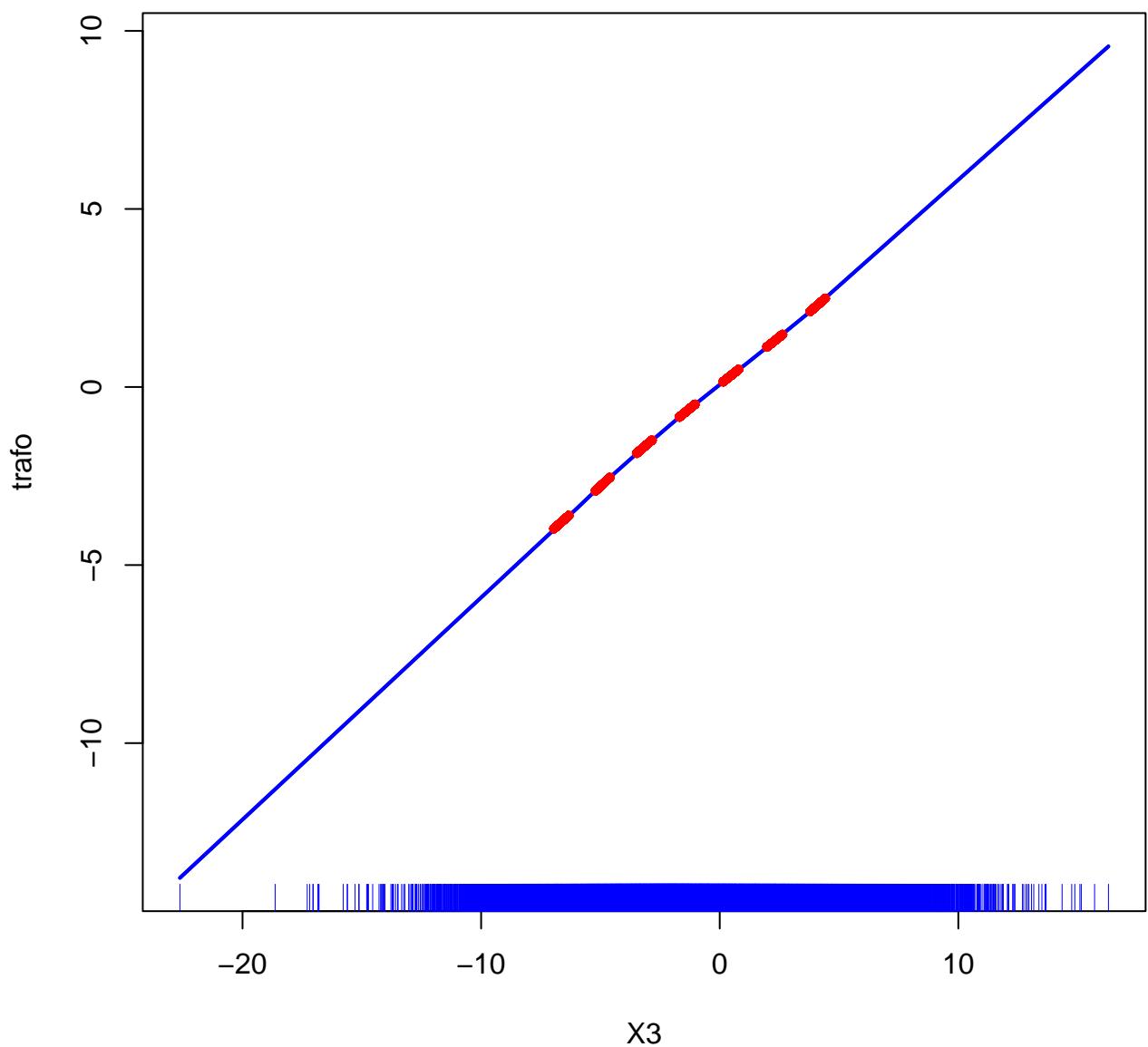
Black: COLR, Red: Our Model



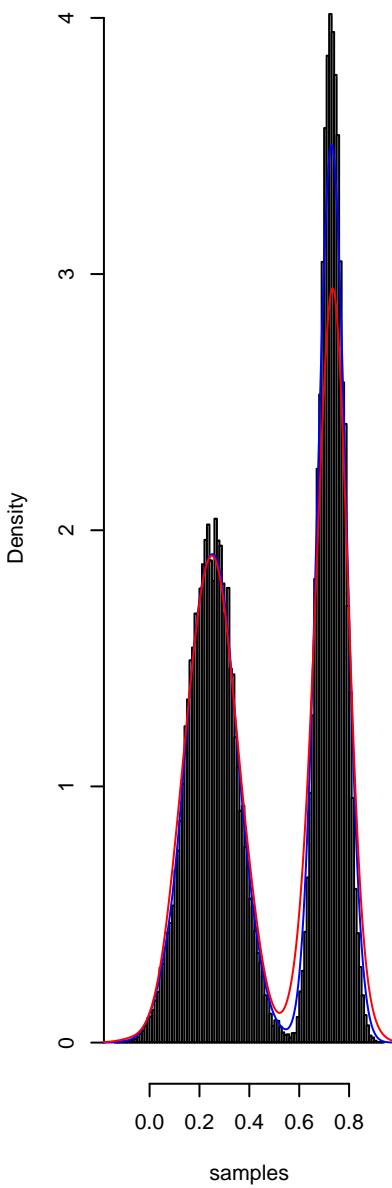
$h_I(X_2)$ Black: COLR, Red: Our Model



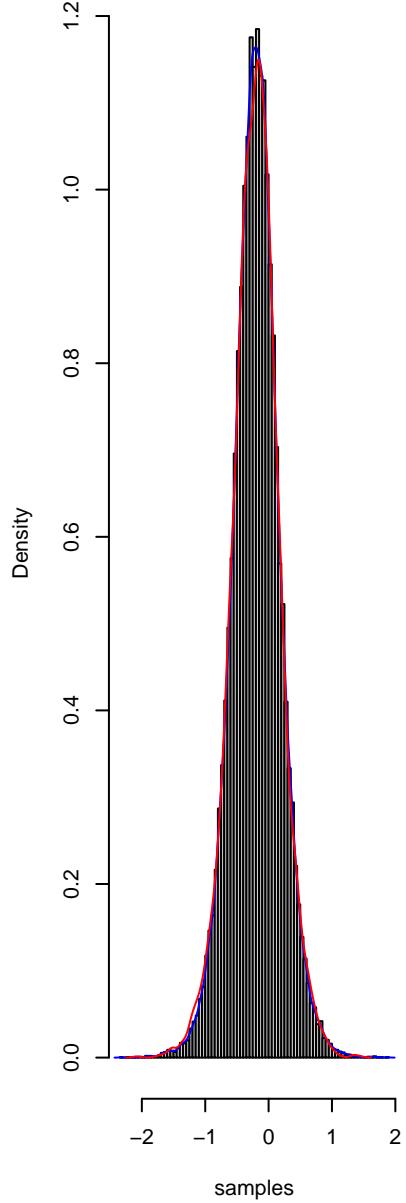
h_I(X3) Colr and Our Model



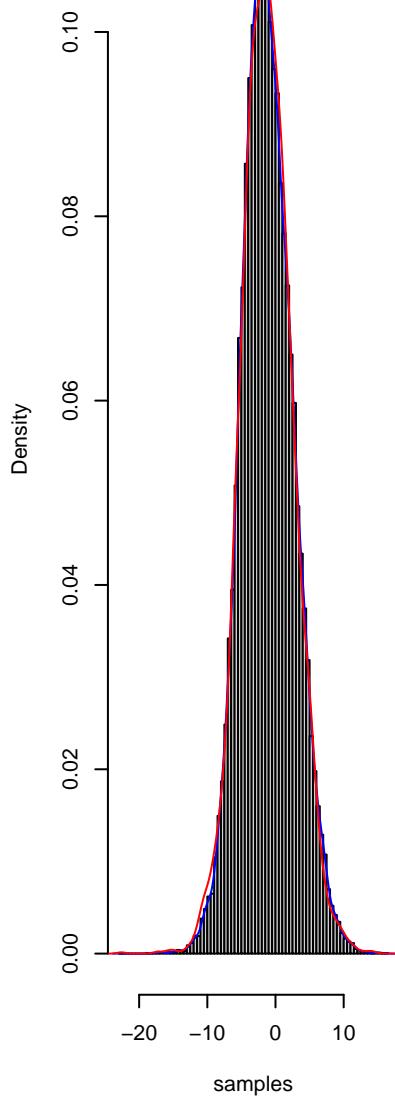
X1 red: ours, black: data



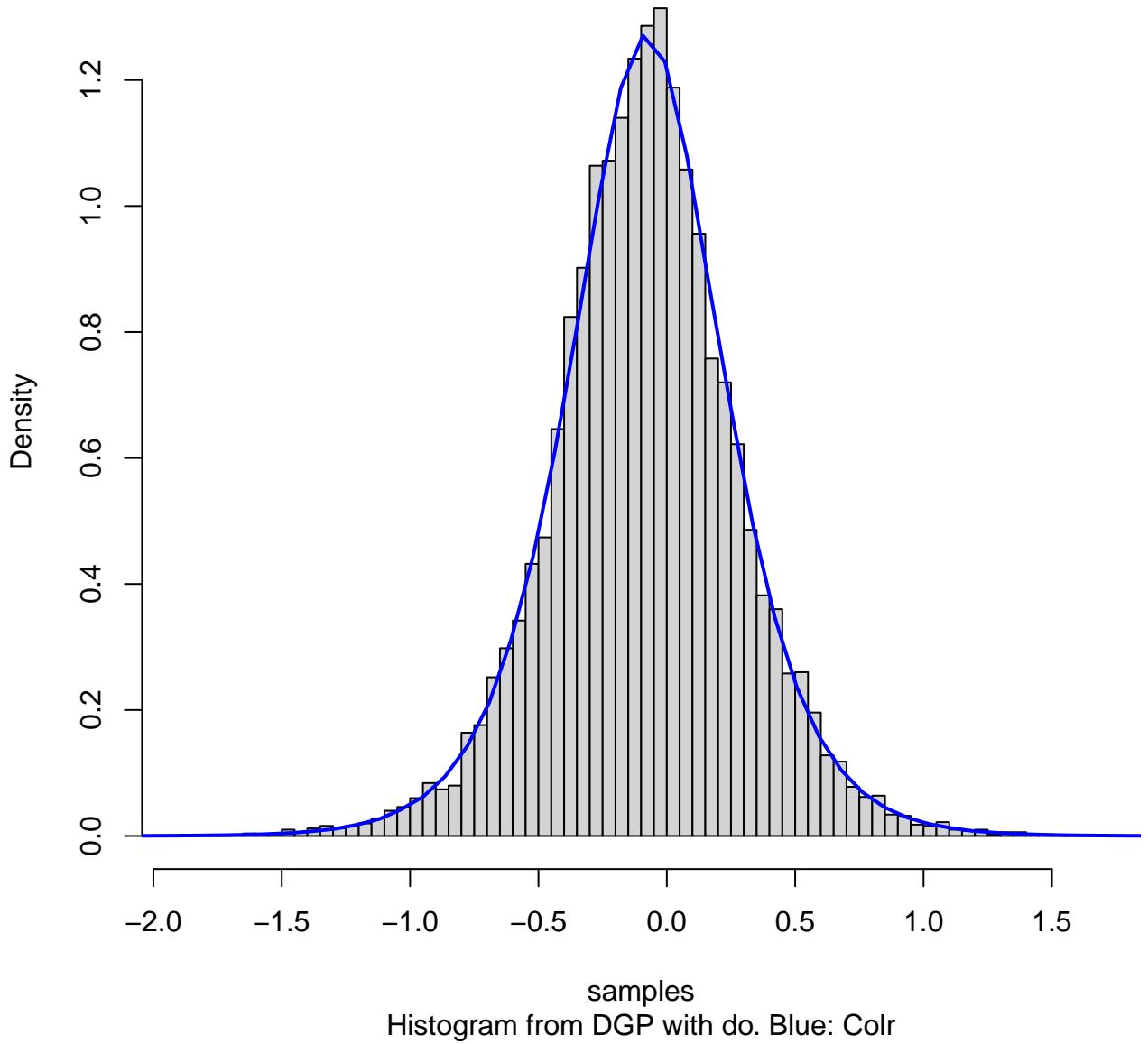
X2 red: ours, black: data



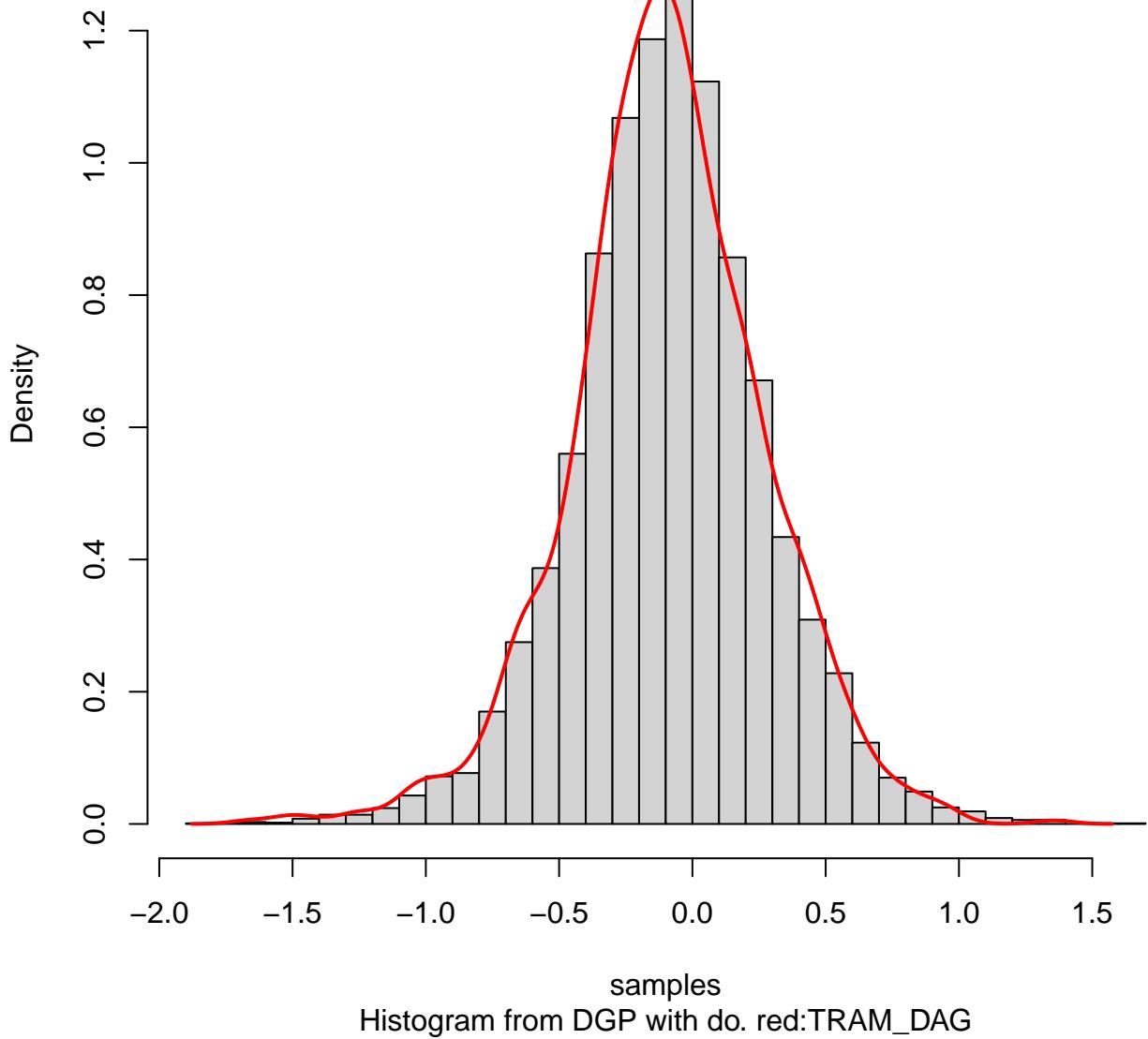
X3 red: ours, black: data



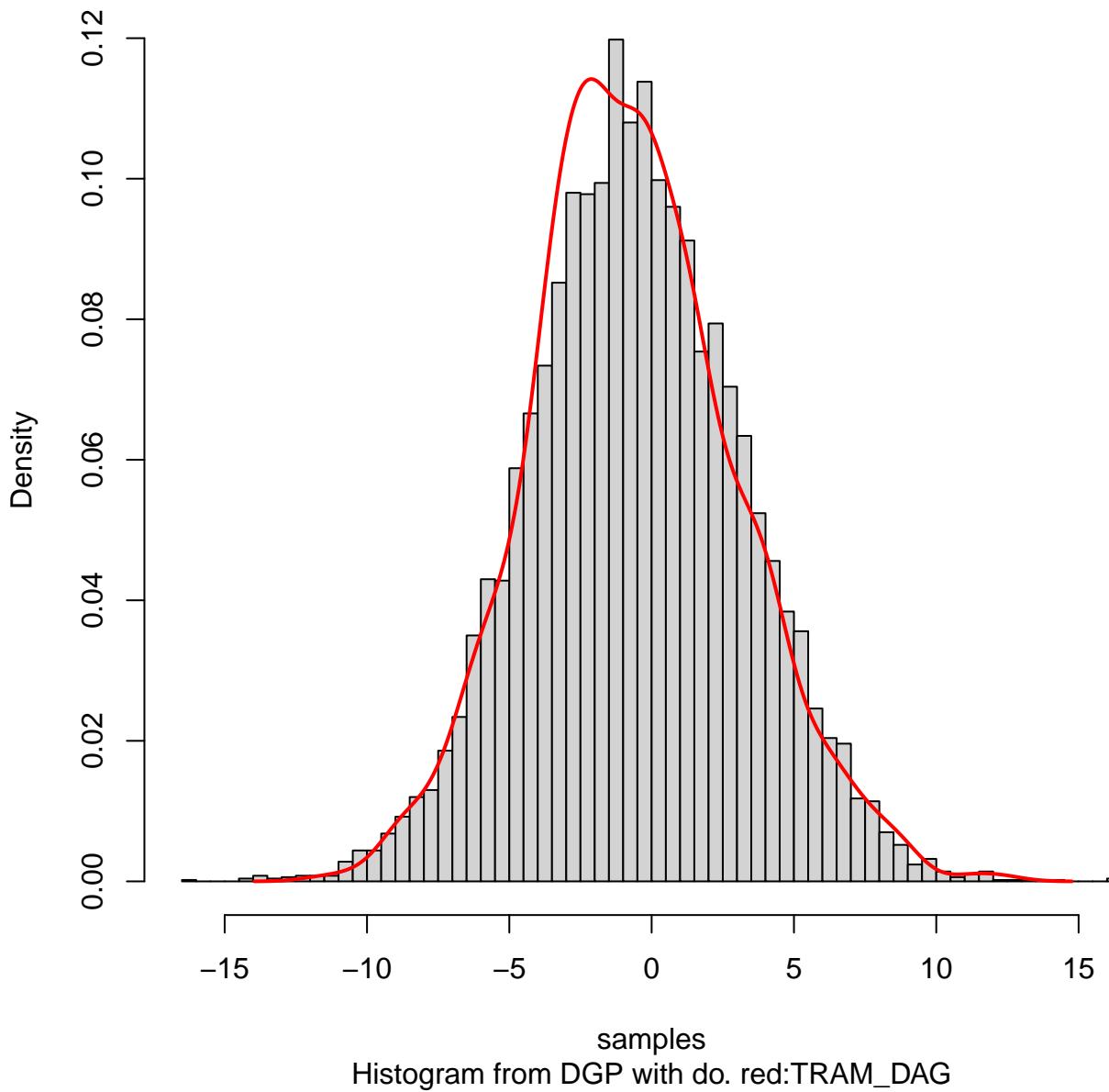
Do(X1=0.2) X2



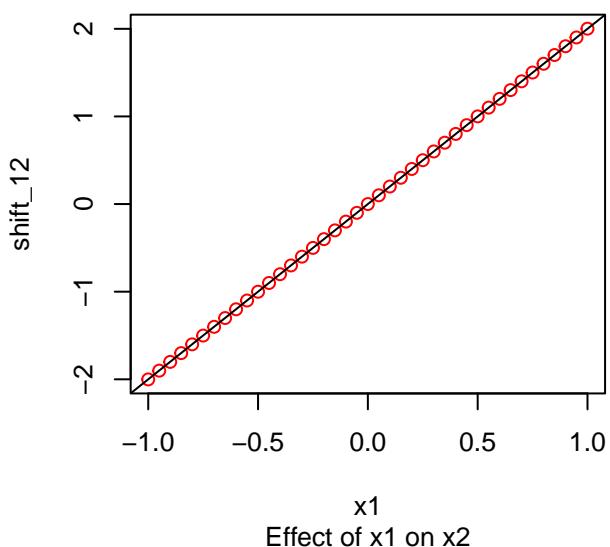
$X2 | \text{Do}(X1=0.2)$



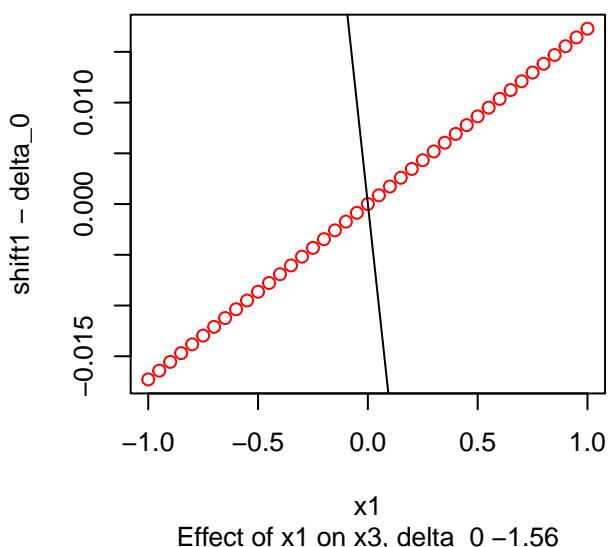
$X3 | \text{Do}(X1=0.2)$



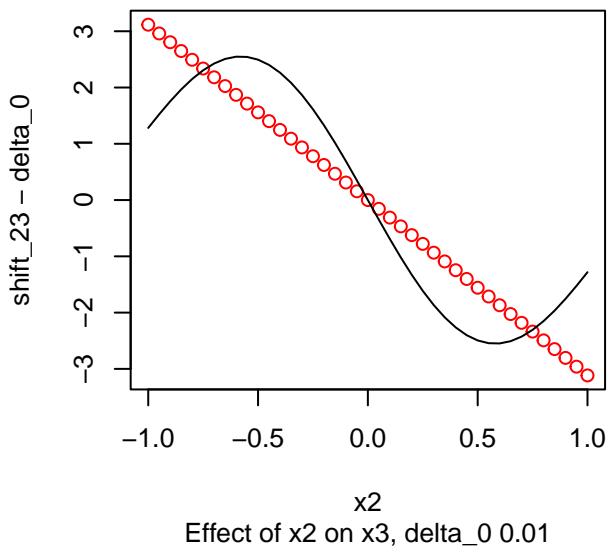
LS-Term (black DGP, red Ours)

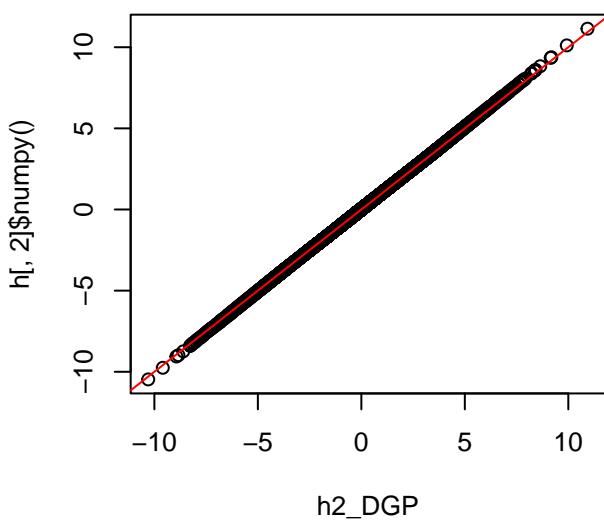
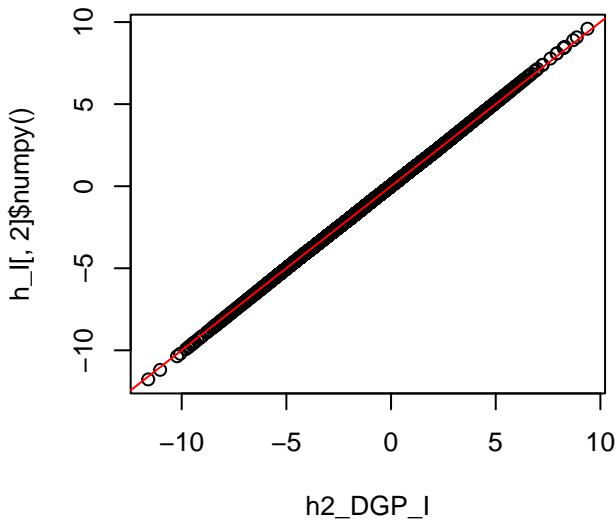
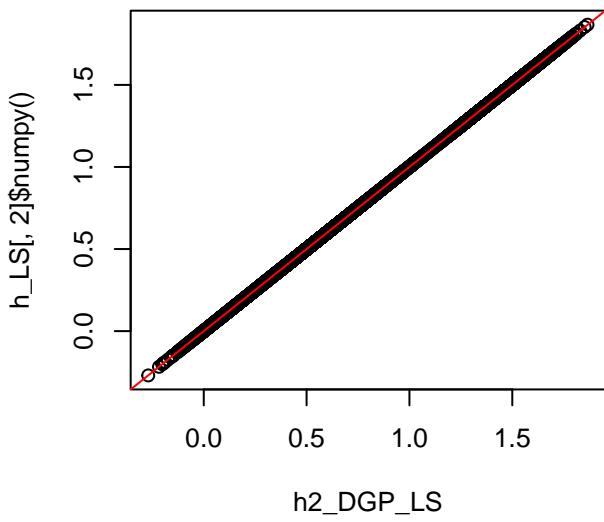
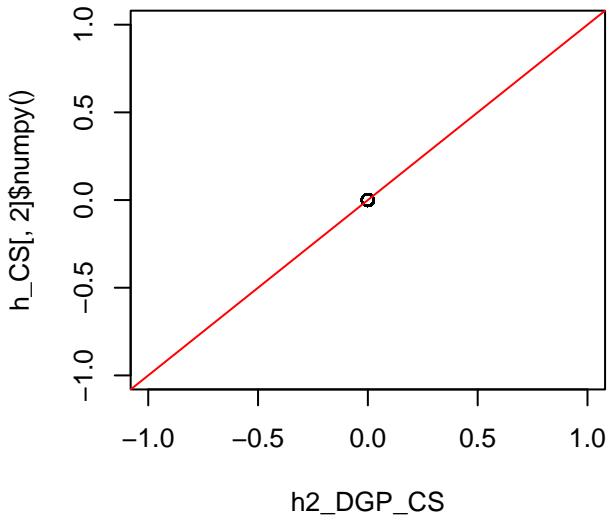


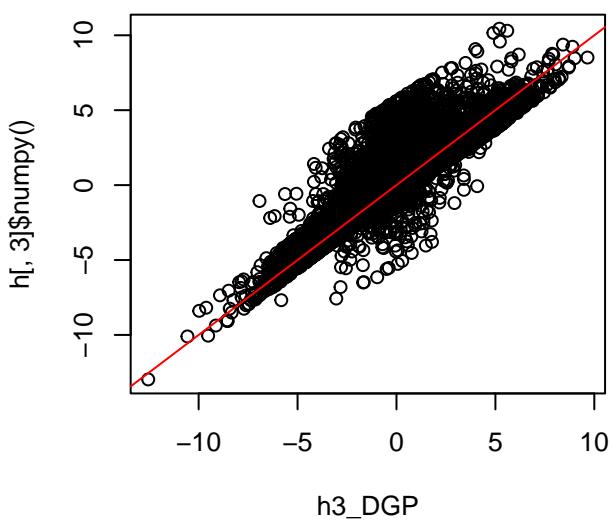
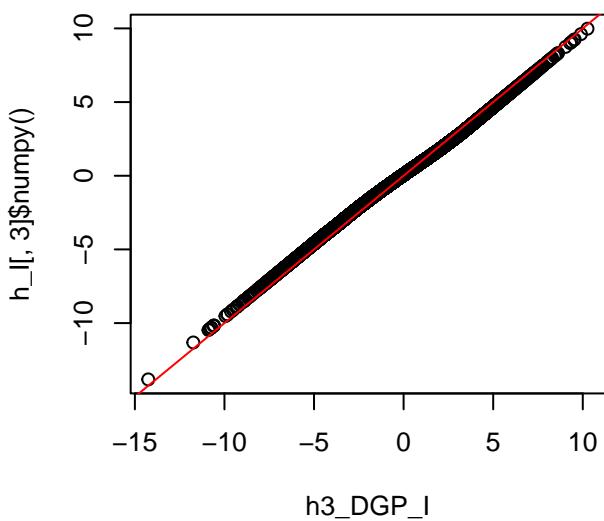
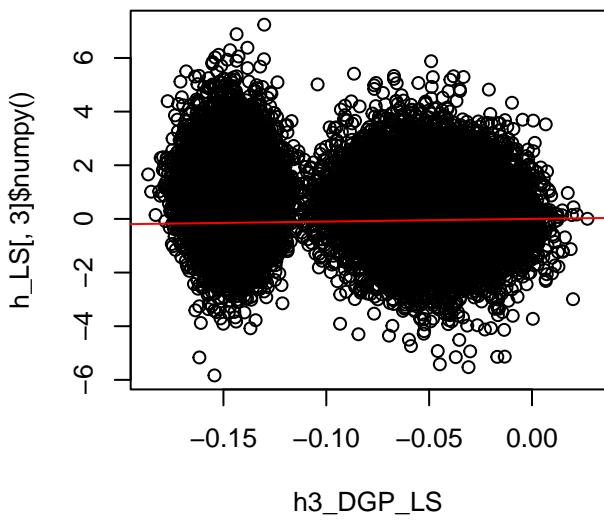
LS-Term (black DGP, red Ours)



LS-Term (black DGP, red Ours)



h2**h2_I****h2_LS****h2_CS**

h3**h3_I****h3_LS****h3_CS**