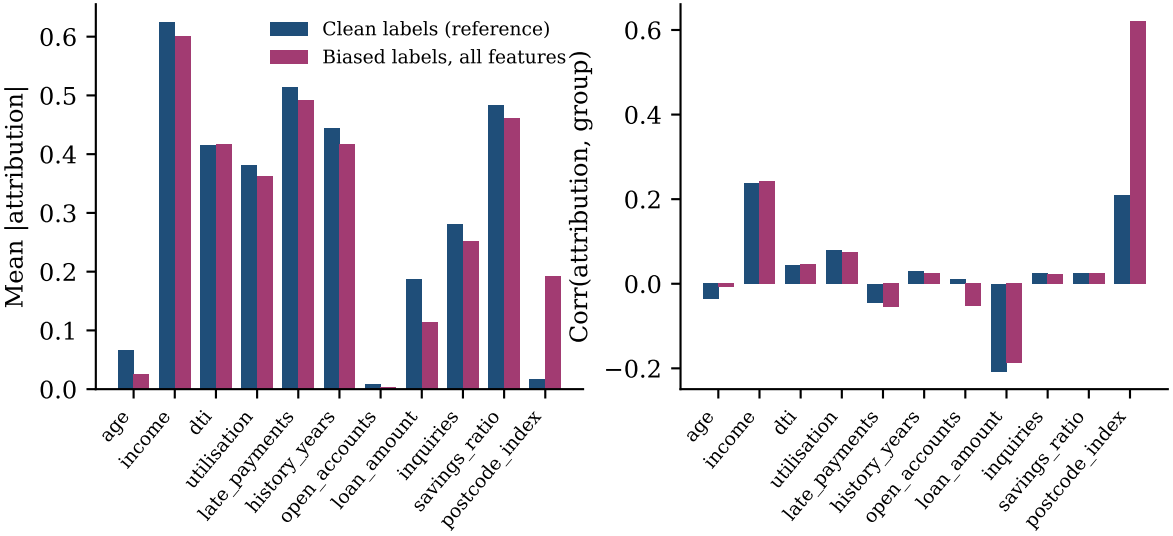
Clean labels (reference)
Biased labels, all features
Mean |attribution|
Corr(attribution, group)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
−0.2
0.0
0.2
0.4
0.6
age
income
dti
utilisation
late_payments
history_years
open_accounts
loan_amount
inquiries
savings_ratio
postcode_index