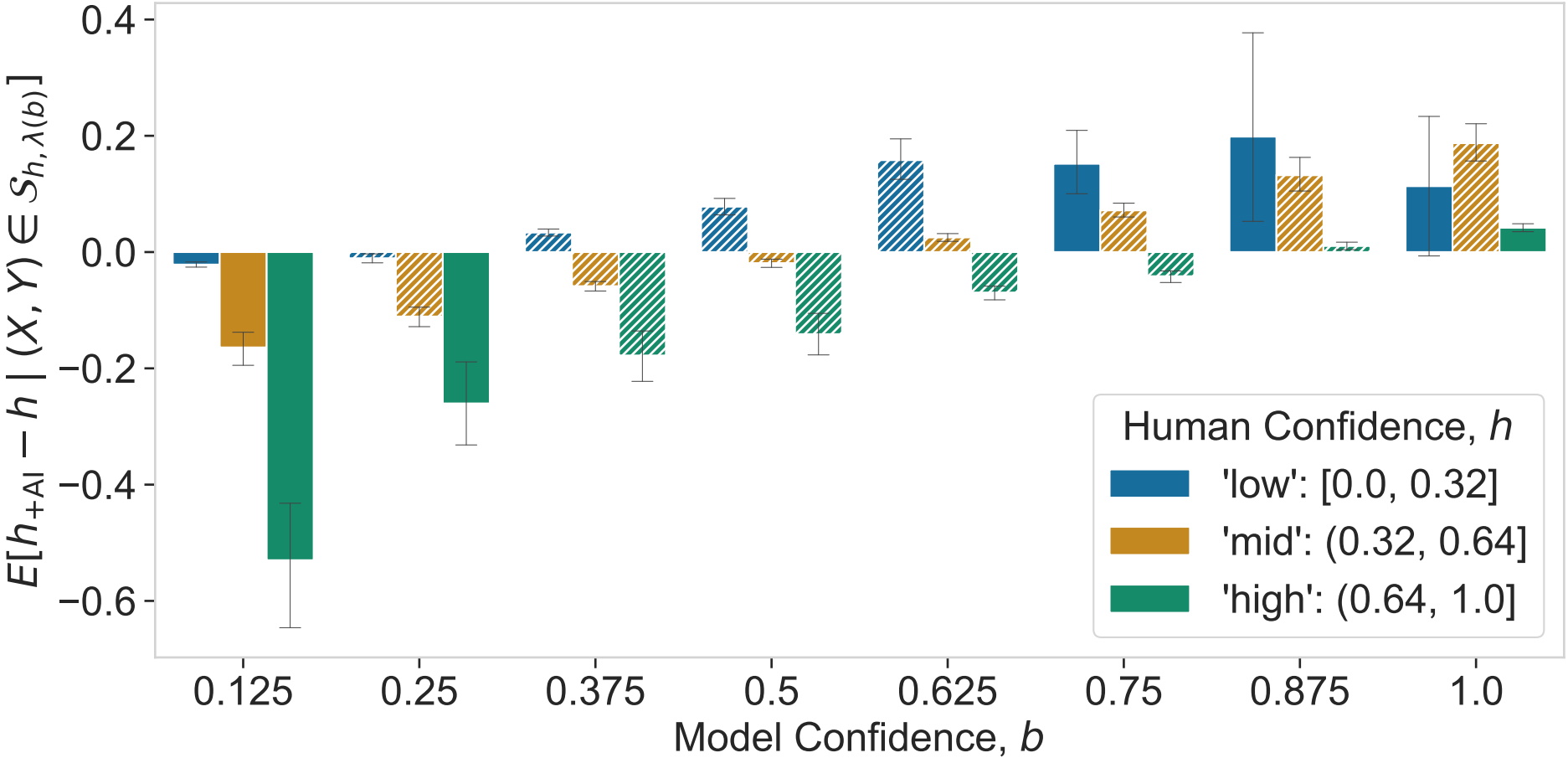
0.4
0.2
0.0
−0.2
−0.4
−0.6
$E[h_{+\mathrm{AI}} - h \mid (X, Y) \in \mathcal{S}_{h,\lambda(b)}]$
0.125
0.25
0.375
0.5
0.625
0.75
0.875
1.0
Model Confidence, b
Human Confidence, h
'low': [0.0, 0.32]
'mid': (0.32, 0.64]
'high': (0.64, 1.0]