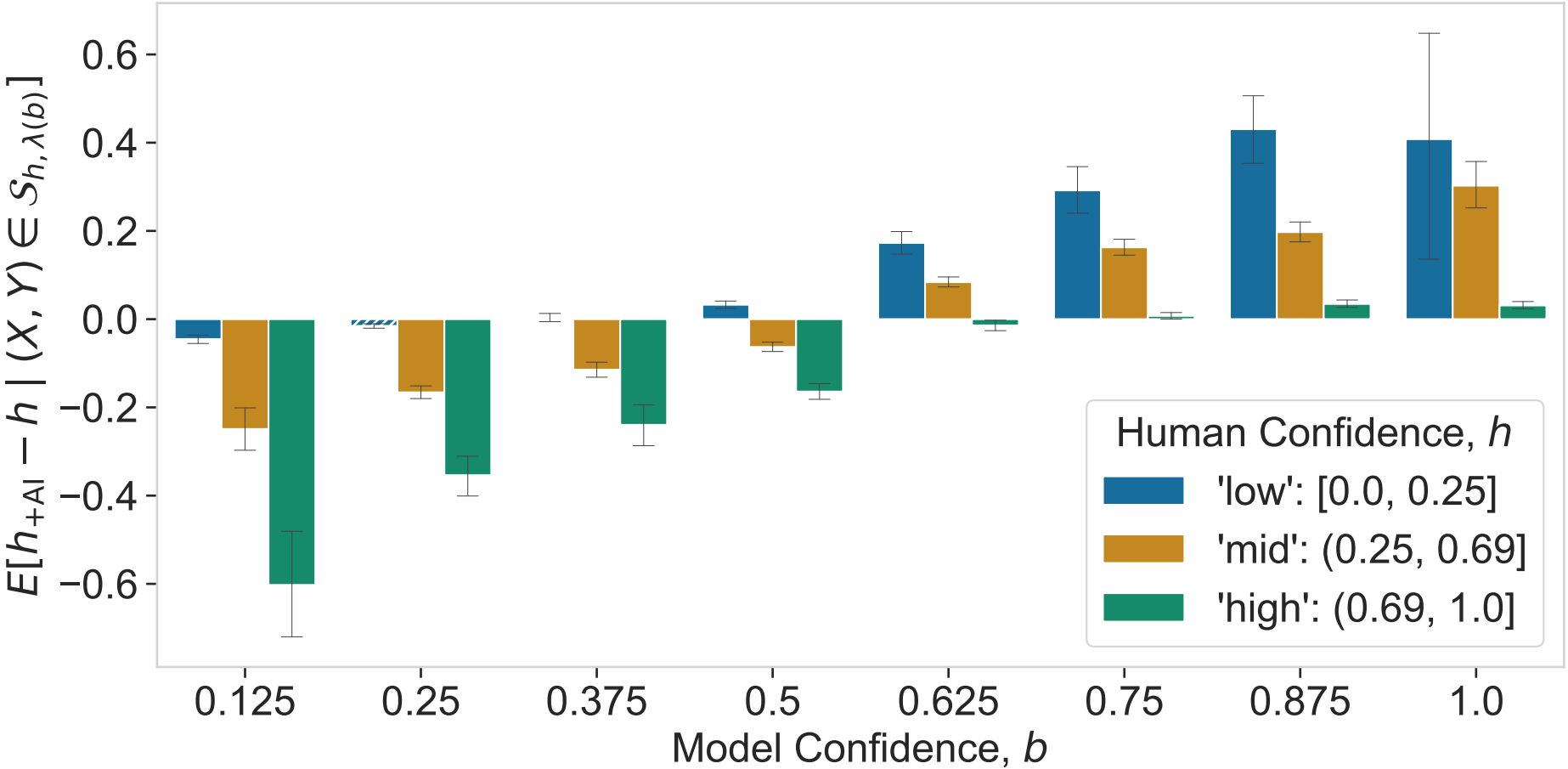
$E[h_{+AI} - h \mid (X, Y) \in \mathcal{S}_{h,\lambda(b)}]$
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
0.125
0.25
0.375
0.5
0.625
0.75
0.875
1.0
Model Confidence, b
Human Confidence, h
'low': [0.0, 0.25]
'mid': (0.25, 0.69]
'high': (0.69, 1.0]