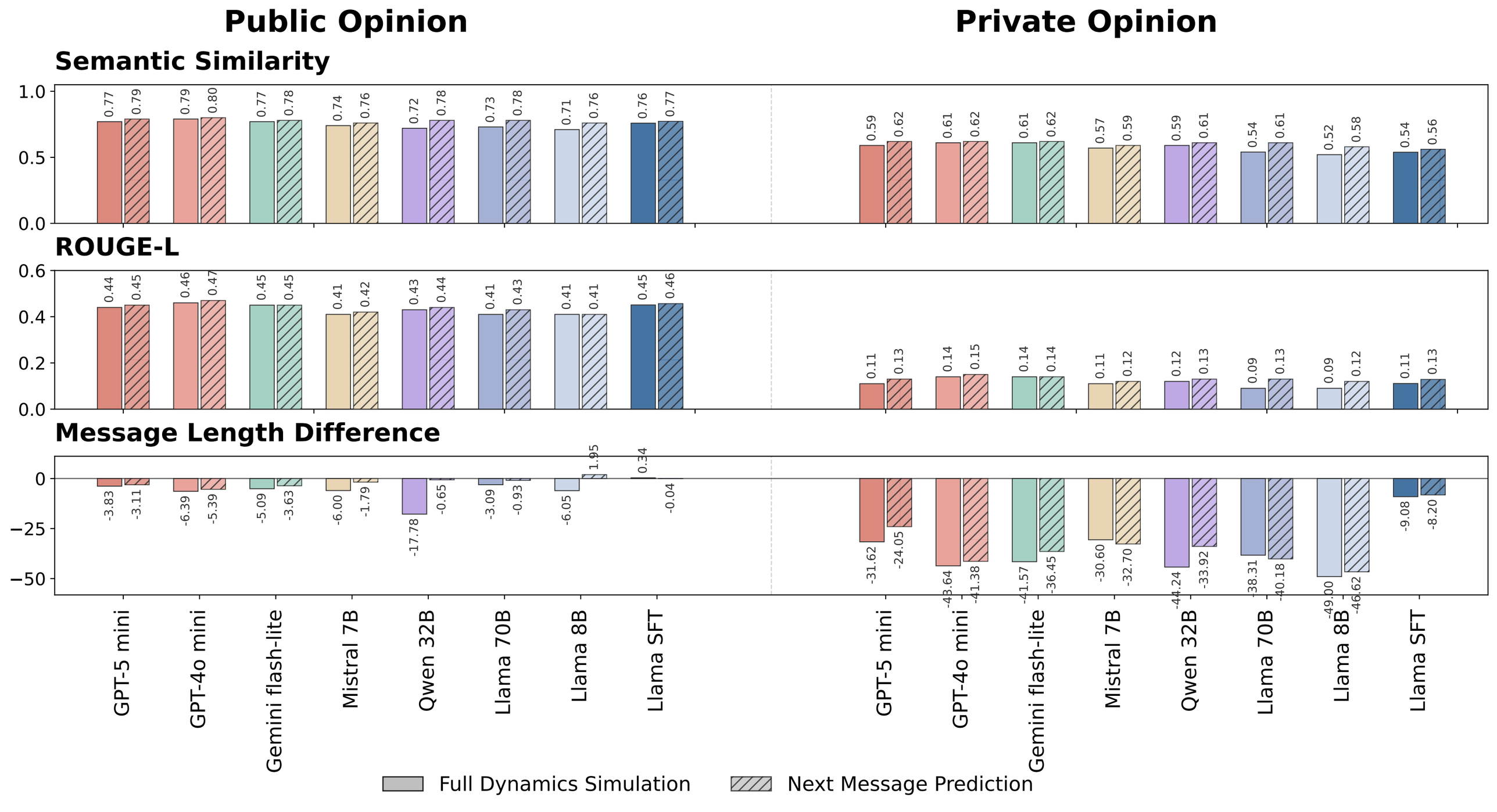

Public Opinion
Private Opinion
Semantic Similarity
ROUGE-L
Message Length Difference
1.0
0.5
0.0
0.6
0.4
0.2
0.0
0
−25
−50
0.77 0.79
0.79 0.80
0.77 0.78
0.74 0.76
0.72 0.78
0.73 0.78
0.71 0.76
0.76 0.77
0.59 0.62
0.61 0.62
0.61 0.62
0.57 0.59
0.59 0.61
0.54 0.61
0.52 0.58
0.54 0.56
0.44 0.45
0.46 0.47
0.45 0.45
0.41 0.42
0.43 0.44
0.41 0.43
0.41 0.41
0.45 0.46
0.11 0.13
0.14 0.15
0.14 0.14
0.11 0.12
0.12 0.13
0.09 0.13
0.09 0.12
0.11 0.13
-3.83 -3.11
-6.39 -5.39
-5.09 -3.63
-6.00 -1.79
-17.78 -0.65
-3.09 -0.93
-6.05 1.95
0.34 -0.04
-31.62 -24.05
-43.64 -41.38
-41.57 -36.45
-30.60 -32.70
-44.24 -33.92
-38.31 -40.18
-49.00 -46.62
-9.08 -8.20
GPT-5 mini
GPT-4o mini
Gemini flash-lite
Mistral 7B
Qwen 32B
Llama 70B
Llama 8B
Llama SFT
GPT-5 mini
GPT-4o mini
Gemini flash-lite
Mistral 7B
Qwen 32B
Llama 70B
Llama 8B
Llama SFT
Full Dynamics Simulation
Next Message Prediction