

Income

Mobility

Employment

TravelTime

PublicCoverage

GPT 4o mini (it)

Mixtral 8x22B (it)

Llama 3 70B (it)

Mixtral 8x7B (it)

Yi 34B (it)

Llama 3 8B (it)

Mistral 7B (it)

Gemma 7B (it)

Gemma 2B (it)

Prompting scheme



Multiple-choice



Numeric

0.0

0.25

0.50

0.0

0.25

0.50

0.0

0.25

0.50

0.0

0.25

0.50

0.0

0.25

0.5

Expected Calibration Error (ECE)