

### H3 Summary — MATH-500, Qwen2.5-1.5B

#### Bucket-wise Aha Margins ( $\Delta \text{Acc}$ ):

Variant	Bucket	N	Share Aha	$\Delta \text{Acc}$
Formal	(0.0558, 0.459]	24800	0.000	-0.193
Gpt	(0.459, 0.748]	24799	0.001	-0.393
Formal	(0.748, 1.302]	24800	0.001	+0.178
Gpt	(1.302, 2.26]	24799	0.001	-0.111
Formal	(2.26, 5.509]	24800	0.000	-0.165
Gpt	(0.0558, 0.459]	24800	0.001	-0.319
Gpt	(0.459, 0.748]	24799	0.004	-0.198
Gpt	(0.748, 1.302]	24800	0.006	-0.022
Gpt	(1.302, 2.26]	24799	0.004	+0.048
Words	(2.26, 5.509]	24800	0.001	-0.122
Words	(0.0558, 0.459]	24800	0.010	-0.017
Words	(0.459, 0.748]	24799	0.015	-0.105
Words	(0.748, 1.302]	24800	0.017	-0.033
Words	(1.302, 2.26]	24799	0.013	+0.017
Words	(2.26, 5.509]	24800	0.005	-0.034