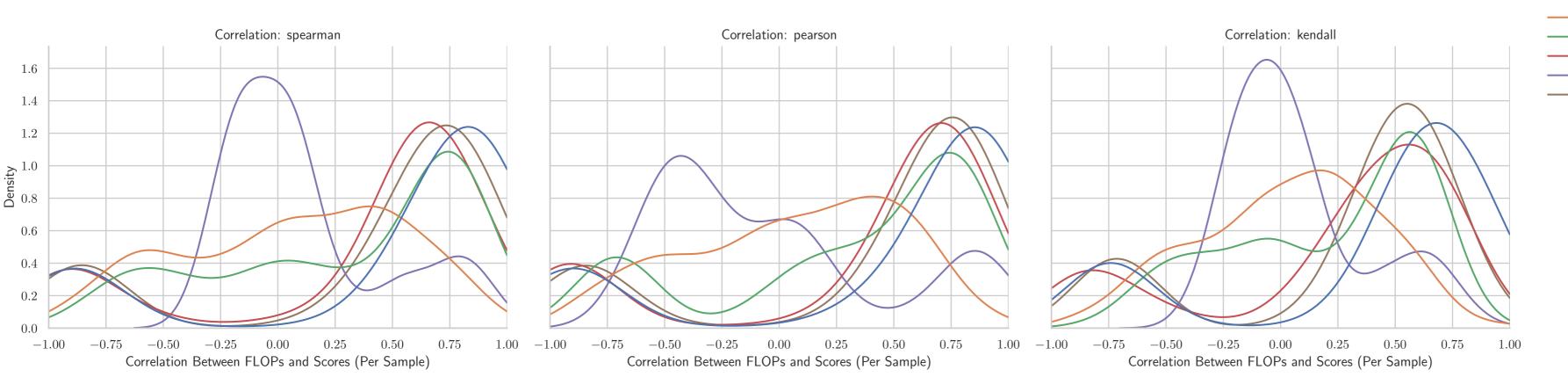
$\begin{array}{c} \text{MMLU HS Macroeconomics} \\ p_{\theta}^{\text{Choices}}(\text{Correct Choice}) \end{array}$ 



Model FamilyCerebras (Param. and Data Scaling)

— INCITE 7B Param. (Data Scaling)

INCITE 7B Param. (Data Scaling)LLM360 Amber 7B Tokens (Param Scaling)

OLMo 7B Param. (Data Scaling)

— Pythia 12B Param. (Data Scaling)

—— Pythia 300B Tokens (Param. Scaling)