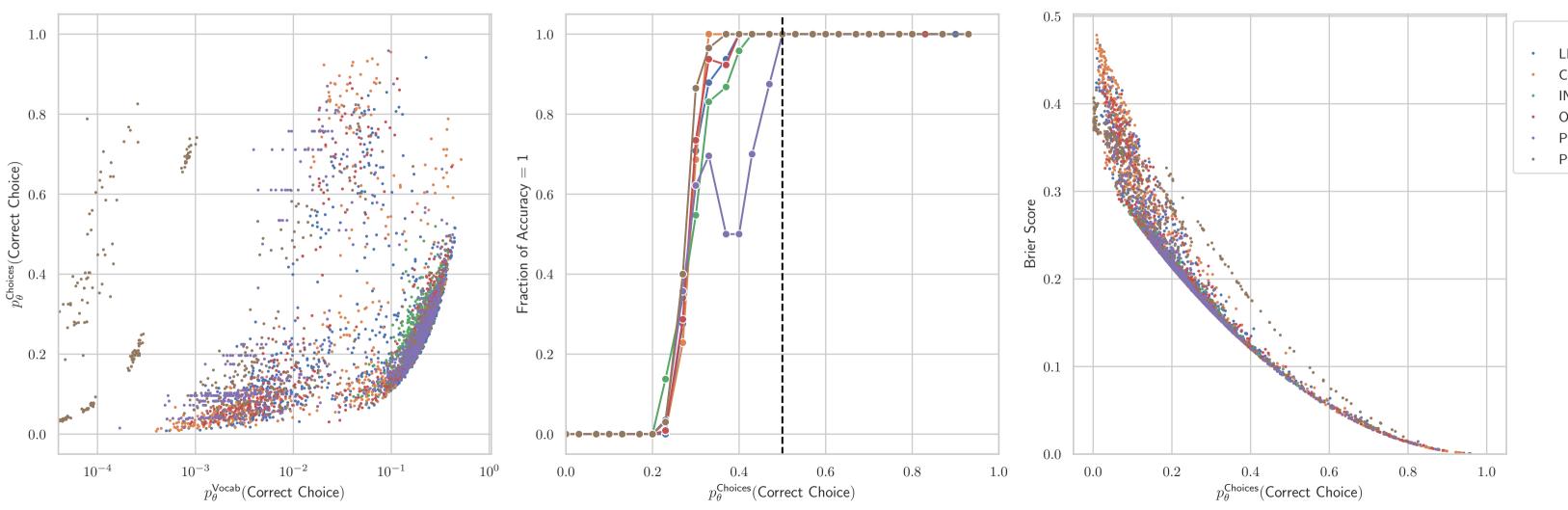# MMLU Public Relations

Legend — Model Family:
- LLM360 Amber 7B Tokens (Param Scaling)
- Cerebras (Param. and Data Scaling)
- INCITE 7B Param. (Data Scaling)
- OLMo 7B Param. (Data Scaling)
- Pythia 300B Tokens (Param. Scaling)
- Pythia 12B Param. (Data Scaling)

Left panel: x-axis $p_\theta^{\text{Vocab}}(\text{Correct Choice})$, y-axis $p_\theta^{\text{Choices}}(\text{Correct Choice})$

Middle panel: x-axis $p_\theta^{\text{Choices}}(\text{Correct Choice})$, y-axis Fraction of Accuracy = 1

Right panel: x-axis $p_\theta^{\text{Choices}}(\text{Correct Choice})$, y-axis Brier Score