KL Differences Relative to Qwen2.5 0.5B - * Larger Models *

Legend:
- Gemma2 2B (2.0B)
- Qwen2.5 3B (3.0B)
- OPT 6.7B (6.7B)
- GPT2 Large (0.774B)
- OPT 2.7B (2.7B)
- Qwen2.5 0.5B (0.5B)