Comparing the Top 5 of 30 of pretraining amounts 0.800 0.795 0.790 0.785 0.780 0.775 0.770 0.765 0.760 0 1000 2000 5000 8000 Batches used for pretraining per layer