Sensitivity Curves for b1=0.95, b2=0.95, ViT trained on Tiny ImageNet 0.450.40 Test Accuracy 0.35 0.30 0.25 BC: False — Warmup BC: True **-**■**-** Constant 0.20  $10^{-4}$  $10^{-3}$ Learning Rate