Stability of Measure Comparisons for wig+100 on nfcorpus (p=0.1 **Count Delta** 0.5 600 1200 0.4 1800 Swap Probability 2400 0.3 3000 0.2 0.1 0.0 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00

Average Performance Delta