Stability of Measure Comparisons for scq on msmarco-passage (p=0.5) Count Delta 0.5 0.4 Swap Probability 5.0 8.0 0.1 0.0 Average Performance Delta