APPENDIX

For the SPCD stage-wise permutation test, we first read in the ADAPT-A dataset (ADAPTA). The variables of interest in the ADAPT-A dataset are madrs_visit3 (stage 1 outcome), madrs_visit6 (stage 2 outcome), treatment_stage1 (0/1 for placebo/drug assignment in stage 1), and treatment_stage2 (0/1 for placebo/drug assignment in stage 2). We run the %NParCov4 macro not adjusting for covariates, %NPACOV4(OUTCOMES = madrs_visit3, TRTGRPS = treatment_stage1, HYPOTH = NULL, ALPHA = 0.05, EXACT = YES, SEED = 44, NREPS = 150000, DSNIN = ADAPTA, DSNOUT = stage1). With this statement, we perform the permutation test on the stage 1 information, at the 0.05 significance, with 50,000 replicates. We can find all the permuted treatment differences in the dataset _STAGE1_BETASAMP, which is produced by the macro. The first row of this dataset shows us the treatment difference for our original dataset. We can run a PROC MEANS to obtain the standard error (0.0061) of these permuted treatment differences for stage 1. Next, we fit %NPACOV4(OUTCOMES = madrs_visit6, TRTGRPS = treatment_stage2, HYPOTH = NULL, ALPHA = 0.05, EXACT = YES, SEED = 44, NREPS = 50000, DSNIN = ADAPTA_STAGE2, DSNOUT = stage2) to the stage 2 ADAPT-A dataset with only placebo non-responders (ADAPTA_STAGE2). This results in a dataset _STAGE2_BETASAMP for the treatment differences for stage 2. Again, we run a PROC MEANS to obtain the standard error (0.0067) of the stage 2 permuted treatment differences. We combine the datasets _STAGE1_BETASAMP and _STAGE2_BETASAMP and create the derived variable, which is the stage-wise SPCD test statistic. This will produce 50,001 SPCD test statistics and the first row of this dataset will have the SPCD test statistic for the original ADAPT-A data. The p-value of this stage-wise permutation test is the rank of the original ADAPT-A SPCD test statistic divided by 50,001. The p-value of the stage-wise permutation test is 0.079. When

adjusting for baseline MADRS score, the p-value of the stage-wise permutation test is 0.17. We see that this is similar to the results produced in Section 5. We did not to employ the %NParCov4 to perform a bootstrap using the HYPOTH = ALT option because the macro does not center the stage 1 and stage 2 data as described in this paper, though modifications of the code could allow for centering of resampled data sets.