JMASM 32: SAS Template for Single-Subject Experimental Designs

Hyewon Chung
Chungnam National University, Deajeon, Korea

Jiseon Kim
University of Washington

Ryoungsun Park
University of Texas

Recommended Citation

JMASM Algorithms and Code
JMASM32: SAS Template for Single-Subject Experimental Designs

Hyewon Chung	Jiseon Kim	Ryoungsun Park
Chungnam National University, Deajeon, Korea	University of Washington, Seattle, WA	University of Texas, Austin, TX

Meta-analysis has been used to synthesize research findings and to evaluate the effectiveness of treatments or the accuracy of diagnostic tools. Although meta-analytic techniques were developed to synthesize the results of several studies, controversy exists as to how to quantify the results from single-subject experimental designs (SSEDs). The most commonly used metrics are reviewed, including non-regression and regression based methods. The application of the SAS template is demonstrated through simulated data sets. The SAS templates can be modified to accommodate a more complex data structure.

Key words: Single-subject experimental designs (SSED), SAS template.

Introduction

Meta-analysis is a "statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings" (Glass, 1976, p. 3). Meta-analyses have been adapted in a variety of settings, such as, special education (Swanson & Sachse-Lee, 2000), counseling (Hagen & Nordahl, 2008) and behavioral intervention research (Filter & Horner, 2009).

Hyewon Chung is an Assistant Professor in the Department of Education at Chungnam National University, Daejeon, Korea. Her research interests are multilevel models, psychometric models and single case design. Email her at: hyewonchung7@gmail.com. Jiseon Kim is a Post-Doctoral researcher in the Department of Rehabilitation Medicine at University of Washington, Seattle, WA, USA. Her research interests are item response theory, computerized adaptive testing, multistage testing and automated test assembly. Email her at: jiseonk@u.washington. edu. Ryoungsun Park is a Ph.D. student in the Department of Educational Psychology at University of Texas, Austin, TX, USA. His research interests are automated test assembly, linear programming, item response theory and computerized adaptive testing. Email him at: rosepark222@gmail.com.

The majority of literature on meta-analyses focuses on group comparison experimental studies; single-subject experimental designs (SSEDs) studies are frequently excluded. Although these meta-analytic techniques effectively summarize and evaluate the results of studies with relatively large sample sizes, on-going controversy exists as to synthesizing results from SSEDs (Faith, Allison & Gorman, 1996). SSED research has grown in popularity in educational and psychological research designed to assess a treatment's effect and different approaches for analyzing single-subject data have been considered over the past decades.

SSEDs have been traditionally evaluated by visual analysis. Bulté and Onghena (2011) suggested that an investigation of treatment effects starts with a visual exploration of the single-case data as opposed to subjective interpretation based on visual analysis; however, evidence-based practice (EBP) emphasizes the importance of more objective outcome measures, especially magnitude of effect indices or effect sizes (Horner, et al., 2005). Methodologists have developed a number of quantitative descriptors as options to describe a treatment's effect in SSEDs. Two main types of statistical summary approaches have been proposed: non-regression based and regression based approaches (Wendt, 2009). Although several metrics have been developed and used

for SSEDs, a question remains regarding what effect size metrics are most appropriate to measure effect size for SSEDs. Thus, this study was designed to develop a SAS procedure to test several metrics to compare their performance.

Regression Based Approaches: Four-Parameter Model

Huitema and McKean (2000) modified the piece-wise regression model by Center, Skiba and Casey (1985-1986) and introduced further regression coefficients that describe change in intercepts as well as in slope from the baseline (A) to treatment (B) phase:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 D_t + \beta_3 \left[T_t - (n_1 + 1) \right] D_t + e_t$$
(1)

Y_t represents outcome score at time t, T_t is the time/session point, D indicates phase either A or B, n_1 is the number of time points in baseline, β_0 represents baseline intercept (i.e., Y at time = 0), β_1 expresses baseline linear trend, β_2 indicates difference in intercept predicted from treatment phase data and that predicted for time = n_1+1 from baseline phase data, and β_3 represents difference in slope between baseline and treatment, $e_t \sim N(0, \sigma^2)$. Thus, β_2 and β_3 provide estimates of a treatment's effect on intercept and on slope, respectively (Wendt, 2009).

The four parameter model interprets the results using ΔR^2 approaches, which can be converted to Cohen's d, and uses all data in both baseline and treatment phases (Wendt, 2009). The regression based approaches assume normality, equal variance and serial independence. Because these are not usually met by SSED data, careful interpretation is required depending on whether the data met these assumptions.

Regression Based Approaches: Multilevel Modeling

Researchers have proposed a multilevel modeling approach as an option for combining SSED data (Ferron, et al., 2009; Ferron, Farmer & Owens, 2010; Shadish & Rindskopf, 2007; Van den Noortgarte & Onghena, 2003a, 2003b). Repeated measures are nested within subjects; lowest level measures within subject effect

change from baseline to intervention. The multilevel modeling approaches can explain why some subjects show more change than others; the variability among individuals in the size of treatment effects can be explained by including study and/or characteristics such as an individual's gender or study intervention type in the model. The multilevel modeling approach can resolve many of the limitations of traditional ordinary least square regression approaches (e.g., independent error assumption).

Non-Regression Based Approaches: Percentage of Non-Overlapping Data (PND)

Percentage of Non-Overlapping Data (PND) computes non-overlap between baseline and treatment phases (Scruggs, Mastropieri & Casto, 1987). PND identifies the highest data point in the baseline and calculates the percentage of data points during treatment exceeding this level. If a study includes several experiments, PND scores are aggregated by taking the median rather than mean because scores are usually not distributed normally (Wendt, 2009). The typical interpretation of PND is: The higher the percentage, the more effective the treatment. Readers can find specific criteria for interpreting PND scores in Scruggs, Mastropieri, Cook and Escobar (1986). Although PND is easy to interpret, it ignores all baseline data except one data point and PND cannot detect slope changes (Wendt, 2009).

Non-Regression Based Approaches: Percentage of All Non-Overlapping Data (PAND)

Percentage of All Non-Overlapping Data (PAND) computes the total number of data points that do not overlap between baseline and treatment phases (Parker, Hagan-Burke & Vannest, 2007). PAND determines overlapping data points: the minimum number that would have to be transferred across phases for complete data separation. It calculates % overlap by dividing the number of overlapping points by the total number of points and then subtracts this percent from 100 to obtain PAND (Wendt, 2009). Unlike PND, PAND uses all data points across both phases. It may be translated to *Phi* and *Phi²* to determine an effect size such as Cohen's d (Wendt, 2009).

Non-Regression Based Approaches: Percentage of Data Points Exceeding the Median (PEM)

Percentage of Data Points Exceeding the Median (PEM) computes the percentage of data points exceeding the median of baseline phase (Ma, 2006). It first locates the median point in baseline data and then draws a horizontal middle line passing through the median of the baseline into the treatment phase. PEM calculates the percentage of treatment phase data points above the middle line if behavior increase is anticipated, and below the middle line if behavior decrease is anticipated (Wendt, 2009). It is assumed that – if treatment is ineffective – data points will continually fluctuate around the middle line. The interpretation of PEM is: Values higher than 0.9 reflect highly effective treatment; values between 0.7 and 0.9 indicate moderately effective treatment; values less than 0.7 reflect questionable or ineffective treatment. In the presence of floor or ceiling data points, PEM reflects effect size while PND fails to reflect effect size. However, PEM is insensitive to magnitude of data points above the median, thus, it does not consider trend and variability in data points of treatment phase. It might reflect only partial improvement if an orthogonal slope exists in a baseline treatment pair after the first treatment phase (Wendt, 2009).

Non-Regression Based Approaches: Pairwise Data Overlap (PDO)

Pairwise Data Overlap (PDO) computes the overlap of all possible paired data comparisons between baseline and treatment phases (Parker & Vannest, in press). PDO first compares the baseline data points with all treatment data points, and then it determines the number of overlapping and non-overlapping points. PDO calculates the total number of non-overlapping points divided by the total number of comparison (Wendt, 2009). Compared to other indices, PDO takes slightly longer to calculate and requires that individual data point results be written down and added. The calculation is time consuming for long and crossed data series (Wendt, 2009).

Non-Regression Based Approaches: Percentage Reduction Data (PRD)

Percentage Reduction Data (PRD) computes the reduction of targeted behavior due to treatment (O'Brien & Repp, 1990). PRD is also termed Mean Baseline Reduction (MBR, Campbell, 2003, 2004). PRD determines the mean of the last three data points from the baseline (μA) and of the last three data points from the treatment (μB). It computes the amount of change between the baseline and treatment:

$$[(\mu A - \mu B) \div \mu A] \times 100$$

(2)

Non-Regression Based Approaches: Percentage of Zero Data (PZD)

Percentage of Zero Data (PZD) computes the degree to which treatment completely suppresses targeted behavior (Scotti, et al., 1991). PZD identifies the first data point to reach zero in the treatment phase. It computes the percentage of data points that remain at zero from the first zero point onwards (Wendt, 2009). Interpretation of PZD scores are: PZD lower than 18% reflects ineffectiveness; PZD between 18% - 54% reflects questionable effectiveness; PZD between 55% - 80% reflects fair effectiveness; PZD higher than 80% reflects high effectiveness.

Methodology
SAS Template for SSEDs

The majority of literature on single case data noted that research on SSED effect size metrics requires the use of multiple metrics and the metrics are compared against each other. The comparisons must compare the metrics within the family of non-overlap as well as across regression versus non-regression based approaches. Although researchers are aware of the methodological issues regarding several SSED tools discussed, difficulties are frequently encountered in running and comparing the metrics due to the complexity of each metric involved. This study develops a SAS procedure consisting of several SAS macros to test a number of metrics for SSEDs. Four sets of single case data were generated using SAS macros written by Fan, Felsovalyi, Sivo and Keenan (2002). Table 1 shows one of the four

simulated AB datasets (condition 1). The total number of observations (T_t) is 18, Phase (D_t) is coded with a value of 0 for baseline (A) and 1 for treatment (B). In this dataset, the number of observations (n_1) for A phase is 8 and the number of observations (n_2) for B phase is 10.

Table 1: Simulated AB Dataset

Outcome (Y)	T_t	D_t	$T_t - (n_1 + 1)]D_t$
1.7	1	0	0
1.9	2	0	0
1.6	3	0	0
1.6	4	0	0
2.4	5	0	0
3.9	6	0	0
4.9	7	0	0
4.4	8	0	0
6.5	9	1	0
8.1	10	1	1
9.2	11	1	2
10.1	12	1	3
10.4	13	1	4
10.1	14	1	5
11.1	15	1	6
12.6	16	1	7
13.8	17	1	8
15.4	18	1	10

The SAS procedure in this article provides standardized parameter estimates for Huitema and McKean's model (2000): standardized slope estimates for phase A and standardized intercept and slope estimates for phase B. It also calculates the following non-regression metrics described above: PND, PAND, PEM, PZD and PRD (see Appendix).

Conclusion

It is not surprising that single-case designs are becoming increasingly more popular in diverse fields given the emphasis on establishing an evidence base for practice. This is due to the difficulties encountered in conducting large group experimental studies. Recently, a single case design standard was released by *What Works Clearinghouse* (Kratochwill, Hitchcock, Horner, Levin, Odom, Rindskopf & Shadish, 2010). Methodologists have developed several SSED metrics to assess and compare the efficiency of the design standards. Several methodological concerns have been expressed; controversy exists as to how to quantify the results of SSEDs. In addition, a few studies resulted in user friendly software capable of performing SSED metrics including both non-regression and regression based methods. The use of a SAS template developed in this study will assist researchers to calculate and compare most the commonly used SSED metrics and to synthesize the results from single-case datasets more efficiently.

References

Bulté, I., & Onghena, P. (2011). *R functions to facilitate the visual analysis of single-case data.* Paper presented at the Annual Meeting of the American Educational Research Association, April, New Orleans, LA.

Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities*, *24*(*2*), 120-138.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification*, *28*(*2*), 234-246.

Center, B., Skiba, R., & Casey, A. (1985-1986). A methodology for the quantitative synthesis of intra-subject design research. *The Journal of Special Education*, *19*(*4*), 387-400.

Faith, M. S., Allison, D. B., & Gorman, D. B. (1996). Meta-analysis of single-case research. In *Design and analysis of single-case research*, R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), 245-277. Mahwah, NJ: Lawrence Erlbaum.

Fan, X., Felsovalyi, A., Sivo, S. A., & Keenan, S. C. (2002). *SAS for Monte Carlo Studies: A Guide for Quantitative Researchers*. Cary, NC: SAS Institute Inc.

Ferron, J. M., Bell-Ellison, B. A., Rendina-Gobioff, G., Hibbard, S. T., & Hess, M. R. (2009). Making treatment effect inferences from multiple baseline data: The utility of alternative multilevel modeling approaches. *Behavior Research Methods*, *41*(*2*), 372-384.

Ferron, J. M., Farmer, J. L. & Owens, C. M. (2010). Estimating individual treatment effects from multiple-baseline data: a Monte Carlo study of multilevel-modeling approaches. *Behavior Research Methods*, *42*(*4*), 930-943.

Filter, K. J. & Horner, R. H. (2009). Function-based academic interventions for problem behavior. *Education and Treatment of Children*, *32*(*1*), 1–19.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*(*10*), 3-8.

Hagen, R., & Nordahl, H. M. (2008). Behavioral experiments in the treatment of paranoid schizophrenia: A single case study. *Cognitive Behavioral Practice*, *15*(*3*), 296-305.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(*2*), 165-179.

Huitema, B., & McKean, J., ((2000). Design specification issues in time-series intervention models. *Educational and Psychological Measurement*, *60*(*1*), 38-58.

Kratochwill, T. R., et al. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Ma, H. (2006). An alternative method for quantitative synthesis of single-subject researchers: Percentage of data points exceeding the median. *Behavior Modification*, *30*(*5*), 598-617.

O'Brien, S., & Repp, A. C. (1990). Reinforcement based reductive procedures: A review of 20 years of their use with persons with severe or profound retardation. *The Journal of the Association for Persons with Severe Handicaps*, *15*(*3*), 148-159.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education*, *40*(*4*), 194-204.

Parker, R. I., & Vannest, K. J. (in press). Pairwise data overlap for single case research. *School Psychology Review*.

Scotti, J. R., Evans, I. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal on Mental Retardation*, *96*(*3*), 233-256.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research methodology: Methodology and validation. *Remedial and Special Education*, *8*(*2*), 24-33.

Scruggs, T. E., Mastropieri, M. A., Cook, S. B., & Escobar, C. (1986). Early intervention for children with conduct disorders: A quantitative synthesis of single-subject research. *Behavioral Disorders*, *11*, 260-271.

Shadish, W. R., & Rindskopf, D. M. (2007). Methods for evidence-based practice: Quantitative synthesis of single-subject designs. *New Directions for Evaluation*, *113*, 95-109.

Swanson, H. L., Sachse-Lee, C. (2000). A meta-analysis of single-subject design intervention research for students with LD. *Journal of Learning Disabilities*, *33*(*2*), 114-36.

Van den Noortgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly*, *18*(*3*), 325-346.

Van den Noortgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments and Computers*, *35*(*1*), 1-10.

Wendt, O. (2009). *Calculating effect sizes for single-subject experimental designs: An overview and comparison*. Paper presented at the 7th Annual International Campbell Collaboration Colloquium, May 2011. Oslo, Norway.

Appendix: SAS Procedure

The SAS procedure provides standardized regression parameter estimates as well as PND, PAND, PEM, PZD and PRD.

```
OPTIONS LINESIZE=100 NOSOURCE NOSOURCE2 NONOTES;

LIBNAME AUTOREG 'C:\Users\Desktop\Single Case';
%MACRO AUTOREG_N (ITER,N1,N2,N,RES,AR,ERR,BETA1,BETA2,BETA3,CONDIT);

/* Generate Dataset */;
                DATA TEMP1;

                ARRAY SERIEA{&N} SERIEA1-SERIEA&N;

                SERIEA(1) = RANNOR(-1);

                DO J=2 TO &N;
                        SERIEA(J)=RANNOR(-11)*SQRT(&RES) + SERIEA(J-
1)*SQRT(&AR);
                END;

                KEEP SERIEA1-SERIEA&N;
                OUTPUT;

        PROC TRANSPOSE DATA=TEMP1 OUT=TEMP1;
            DATA TEMP1;
                SET TEMP1;
                ID=_N_;
                SERIESA=COL1;
                OUTPUT;
                DROP COL1 _NAME_;

            DATA TEMP2;
                TIME=1;
        DO ID=1 TO &N;
                        TIME = ID;
                        IF ID <= &N1 THEN D = 0;
                        ELSE D = 1;
            TIMED = (TIME - (&N1 + 1))*D;
                        Y = TIME*SQRT(&BETA1) + D*SQRT(&BETA2) +
TIMED*SQRT(&BETA3);
                OUTPUT;
        END;

            DATA INPUT1;
            MERGE TEMP1 TEMP2;
            BY ID;
```

Appendix: SAS Procedure (Continued)

```
                  TIMEAR=TIME; DAR=D; TIMEDAR=TIMED;
                        Y = Y + SERIESA*SQRT(&ERR);
                  ITERATION = &ITER;
                        CONDITION = &CONDIT;
            RUN;
      PROC APPEND BASE = INPUT DATA = INPUT1 FORCE; RUN;


/* Four Parameter Regression */

PROC REG DATA=INPUT1 OUTEST=RES_REG;
        MODEL Y = TIME D TIMED/OUTSTB OUTSEB NOPRINT;
            RUN;


/* outstb (1st line in data generate) shows standardized parameter
estimates*/
data REG_EST; set RES_REG;if _n_=1; keep time d timed;
output;
/* outseb (2nd line in data generate) shows standardized error of
parameter estimates*/
data REG_SE; set RES_REG; if _n_=2; keep time d timed SE_time SE_d
SE_timed;
   SE_time = time;
      SE_d = d;
      SE_timed =timed;
      output;
data REG_SE;set REG_SE; drop time d timed;

data RESULTS;
      merge REG_EST REG_SE ;

      BETA1=&BETA1;
      BETA2=&BETA2;
      BETA3=&BETA3;
      AR=&AR;
      ITERATION=&ITER;
      CONDITION=&CONDIT;

      KEEP TIME D TIMED SE_time SE_d SE_timed
       BETA1 BETA2 BETA3 AR ITERATION CONDITION;
RUN;

PROC APPEND BASE=Four_Par_RESULT (CNTLLEV=MEMBER) DATA=RESULTS FORCE;
/* Four-Paremeter_Regression result*/
RUN;
```

Appendix: SAS Procedure (Continued)

```
/*PND*/
DATA BASE;
        SET INPUT1;       IF D=0;
        KEEP ID D Y ITERATION CONDITION;

RUN;

DATA TTT;
        SET INPUT1;
        IF D=1;
        KEEP ID D Y ITERATION CONDITION;

RUN;

PROC MEANS NOPRINT DATA=BASE;
        VAR Y;
        OUTPUT OUT=RES MAX=MAXIMUM;
RUN;

DATA RES1;
        SET RES;
    ARRAY MAX{&N2} MAX1-MAX&N2;
        DO J=1 TO &N2;
        MAX(J)=MAXIMUM;
        END;
        DROP MAXIMUM _FREQ_ _TYPE_ J;
RUN;

PROC TRANSPOSE DATA=RES1 OUT=RES2;
RUN;

DATA BASE_TTT;
        MERGE TTT RES2;
        IF Y-COL1>0 THEN CHECK=1; ELSE CHECK=0;
        RUN;

PROC MEANS DATA=BASE_TTT NOPRINT;
    CLASS CONDITION;
    VAR CHECK;
        OUTPUT OUT=BASEPND ;
        RUN;

DATA PND; SET BASEPND;
        IF _STAT_='MEAN' AND _TYPE_=1;
        KEEP CHECK PND ITERATION CONDITION;
        PND=CHECK*100;
        ITERATION=&ITER;
        OUTPUT;
        FORMAT PND 10.3;
RUN;
```

Appendix: SAS Procedure (Continued)

```
/*PAND*/
proc means data = input1 noprint;
      var y;
      by d;
      output out = out1 min=min_y;
run;

data out2;
      set out1;
      keep min_y;
      if d=1;
run;

DATA out3;
            SET out2;
      ARRAY MIN{&N1} MIN1-MIN&N1;
DO J=1 TO &N1;
            MIN(J)=min_y;
            END;
            drop min_y J;
RUN;

data base;
            set input1;
            if D=0;
            KEEP ID D Y ITERATION CONDITION;
run;

PROC TRANSPOSE DATA=OUT3 OUT=RES;
RUN;

DATA BASE_TTT;
            MERGE BASE RES;
            IF Y-COL1>=0 THEN CHECK=1; ELSE CHECK=0;
            RUN;

PROC MEANS DATA=BASE_TTT NOPRINT;
    CLASS CONDITION;
    VAR CHECK;
            OUTPUT OUT=BASE_PAND ;
      RUN;

data PAND; set BASE_PAND;
      IF _STAT_='MEAN' AND _TYPE_=1;
      PAND=(1-CHECK)*100;
      output;
      keep CHECK PAND ITERATION CONDITION;
      FORMAT PAND 10.3;
RUN;
```

Appendix: SAS Procedure (Continued)

```
PROC APPEND BASE=PAND_RESULT (CNTLLEV=MEMBER) DATA=PAND; /* PAND
result*/
RUN;

/*PEM*/
proc means data = input1 noprint;
     var y;
     by d;
     output out = out1 median=med_y;
run;

data out2;
     set out1;
     keep med_y;
     if d=0;
run;

DATA out3;
          SET out2;
          ARRAY med{&N1} med1-med&N1;
DO J=1 TO &N1;
          med(J)=med_y;
          END;
          drop med_y J;
RUN;

PROC TRANSPOSE DATA=OUT3 OUT=RES;
RUN;

data ttt;
          set input1;
          if D=1;
          KEEP ID D Y ITERATION CONDITION;
run;

DATA BASE_TTT;
          MERGE ttt res;
          IF Y-COL1>=0 THEN CHECK=1; ELSE CHECK=0;
                              RUN;

PROC MEANS DATA=BASE_TTT NOPRINT;
    CLASS CONDITION;
    VAR CHECK;
          OUTPUT OUT=BASE_pem ;
     RUN;
```

Appendix: SAS Procedure (Continued)

```
data pem; set BASE_pem;
     IF _STAT_='MEAN' AND _TYPE_=1;
     pem=CHECK*100;
     ITERATION=&ITER;        output;
     keep CHECK pem ITERATION CONDITION;
     FORMAT pem 10.3;
RUN;

PROC APPEND BASE=PEM_RESULT (CNTLLEV=MEMBER) DATA=PEM; /* PEM result*/
RUN;

/*PZD*/
DATA TTT;
     SET INPUT1;
     IF D = 1;
     IF Y = 0;
RUN;

DATA TTT2;
     SET TTT;
     IF _N_ = 1;
RUN;

DATA TTT3;
     SET TTT2;
     DO I = 1 TO &N1;
       OUTPUT;
     END;
     KEEP ID;
     RENAME ID=START;
RUN;

DATA TTT4;
     SET INPUT1;
     IF D = 1;
RUN;

DATA RES;
     MERGE TTT3 TTT4;
     IF TIME >= START;
     IF Y = 0 THEN CHECK_Z = 1; ELSE CHECK_Z = 0;
     KEEP CHECK_Z;
RUN;

PROC MEANS DATA = RES NOPRINT;
     OUTPUT OUT = RES2
          MEAN(CHECK_Z) = PZD;
RUN;
```

Appendix: SAS Procedure (Continued)

```
DATA PZD;
     SET RES2;
     PZD = PZD * 100;
     KEEP PZD;
RUN;

PROC APPEND BASE=PZD_RESULT (CNTLLEV=MEMBER) DATA=PZD; /* PZD result*/
RUN;

/*PRD*/
DATA BASE;
     SET INPUT1;
     IF D = 0;
     CNT = &N1;
  IF _N_ > CNT - 3;
     KEEP Y;
     RENAME Y = BASE_Y;
RUN;

DATA TTT;
     SET INPUT1;
     IF D = 1;
RUN;

DATA TTT; SET TTT;
     CNT = &N1;
     IF _N_ > CNT - 3;
     KEEP Y;
     RENAME Y = TTT_Y;
RUN;

DATA BASE_TTT;
     MERGE BASE TTT;
RUN;

PROC MEANS DATA = BASE_TTT NOPRINT;
     OUTPUT OUT = RES (KEEP=MEAN_B MEAN_T)
          MEAN(BASE_Y) = MEAN_B MEAN(TTT_Y) = MEAN_T;
RUN;

DATA PRD;
     SET RES;
     PRD = (MEAN_B - MEAN_T) / MEAN_B * 100;
RUN;

PROC APPEND BASE=PRD_RESULT (CNTLLEV=MEMBER) DATA=PRD; /* PRD result*/
RUN;

%MEND AUTOREG_N;
%AUTOREG_N (1, 8, 10, 18, 0.5, 0.5, 0.4, 0.2, 1, 0.4, 1);
%AUTOREG_N (1, 10, 20, 30, 0.2, 0.8, 0.1, 0.3, 0.3, 0.3, 2);
%AUTOREG_N (1, 16, 16, 32, 1, 0, 0.1, 0.6, 0.3, 0, 3);
%AUTOREG_N (1, 8, 24, 32, 0.8, 0.2, 0.4, 0.0, 0.0, 0.6, 4);
```