Meta Analyses of Multiple Baseline Time Series Design Intervention Models for Dependent and Independent Series

Oluwagbohunmi Adetunji Awosoga,
Advisors: Joseph W. McKean (Dept: Statistics)
&
Bradley E. Huitema (Dept: Psychology)

Western Michigan University, Kalamazoo, Michigan
&
Faculty of Health Sciences,
University of Lethbridge, Alberta, Canada

IMS New Researchers Conference Presentation
July 28th - July 31st, 2009

Introduction

This work evaluates a proposed meta-analytic procedure for the analysis of multiple baseline series (Huitema, 2004) and compares that pvalue with a robust analog. Diagnostic procedures for the analysis of these designs are also developed.

This presentation focuses on new approaches for the analyses of multiple baseline (series) AB-type designs. AB type designs are frequently used in Psychology, Economics, and Education.

Purpose Statement

This work studies the behavior of new parametric and robust approaches being introduced with specific focus on the following scenarios:

1. Case of independence within series, and between two or more baselines;

Purpose Statement

This work studies the behavior of new parametric and robust approaches being introduced with specific focus on the following scenarios:

1. Case of independence within series, and between two or more baselines;
2. Case of dependence within series, but independence between two or more baselines;

Purpose Statement

This work studies the behavior of new parametric and robust approaches being introduced with specific focus on the following scenarios:

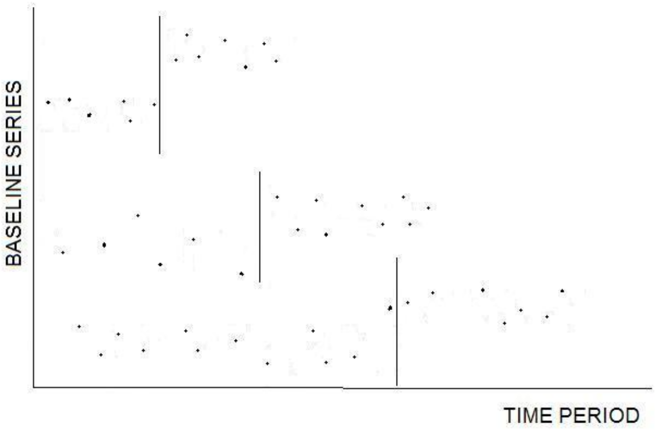
1. Case of independence within series, and between two or more baselines;
2. Case of dependence within series, but independence between two or more baselines;
3. Case of independence within series, but dependence between two or more baselines;

Purpose Statement

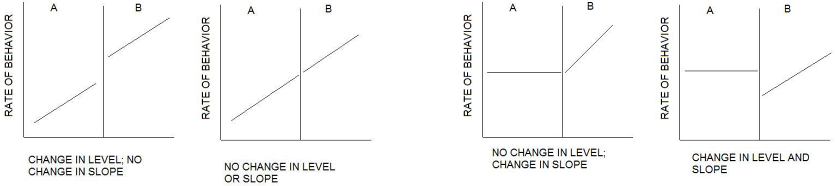
This work studies the behavior of new parametric and robust approaches being introduced with specific focus on the following scenarios:

1. Case of independence within series, and between two or more baselines;
2. Case of dependence within series, but independence between two or more baselines;
3. Case of independence within series, but dependence between two or more baselines;
4. Case of dependence within series, and between two or more baselines.

Three Multiple Baseline Time Series Design with staggered intervention lines



Graphs showing change in level / slope



Four Parameter Model for an AB Design

As recommended by Huitema and McKean (1998), we adopt the following model:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 D_t + \beta_3 [T_t - (n_1 + 1)] D_t + \epsilon_t \tag{1}$$

where
Y_t is the dependent variable score at time t,
T_t is the value of the measurement occasion variable T at time t,
D_t is the value of the level change dummy variable D indicating the treatment phase (0 for the first phase and 1 for the second phase) at time t,
$[T_t - (n_1 + 1)] D_t$ is the value of the slope change variable at time t,
β_0 is the regression intercept,
β_1 through β_3 are the process partial regression coefficients, and
ϵ_t is the process error of the model at time t.

Autocorrelation Test

Once we fit the model, then check for dependence (autocorrelation) within series. Three methods were considered:

1. The Durbin-Watson (DW) test statistic is used to detect the presence of autocorrelation in the residuals from a given regression analysis and is defined as:

$$d = \frac{\sum_{t=2}^{T} (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{T} \hat{e}_t^2} \tag{2}$$

where $\hat{e}_t$ is the residual of the fitted model at time t, and T is the total number of observations in the series.

2. The Huitema-McKean (HMZ) test statistic is given by

$$z_{HM} = \frac{r_1 + \frac{p}{N}}{\sqrt{\frac{(N-2)^2}{(N-1)N^2}}} \tag{3}$$

where
$r_1 = \frac{\sum_{t=2}^{N} (\hat{e}_t)(\hat{e}_{t-1})}{\sum_{t=1}^{N} \hat{e}_t^2}$ is the lag 1 autocorrelation coefficient, e_t is the residual of the fitted time series regression model, measured at time t; p is the number of parameters in the time series regression model and N is the total number of residuals in the observed time series.

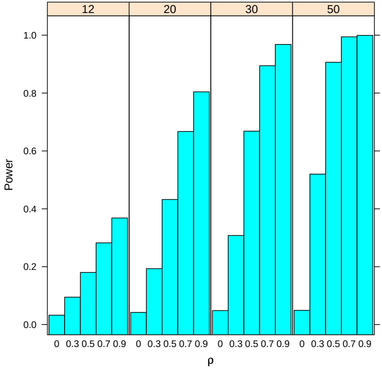
The Robust analog to the Huitema-McKean test

3. A nonparametric analog for the test of autocorrelated errors is given by

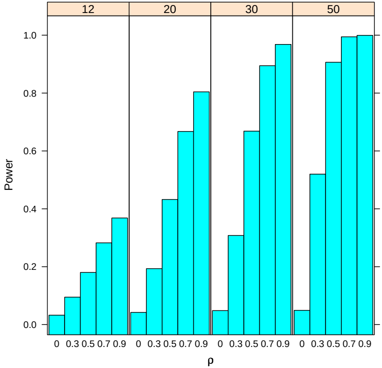
$$z_{HMN} = \frac{\gamma_{ww} + \frac{p}{N}}{\sqrt{\frac{(N-2)^2}{(N-1)N^2}}} \tag{4}$$

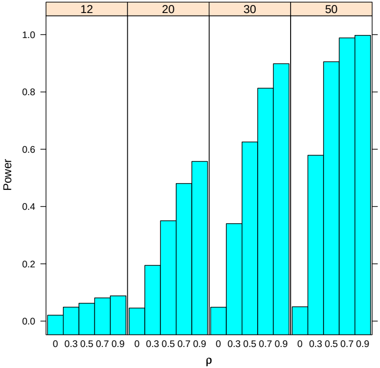
where γ_{ww} is the robust analog of the lag-1 autocorrelation coefficient obtained from the weighted Wilcoxon (WW) estimate of the slope, based on the R residuals formed after estimating $\boldsymbol{\beta}$; a power study and empirical alpha level test confirmed it to be equally competitive with the parametric tests such as the D-W test and HMZ test.

Graphs of HMZ & RHMZ with Normal Error

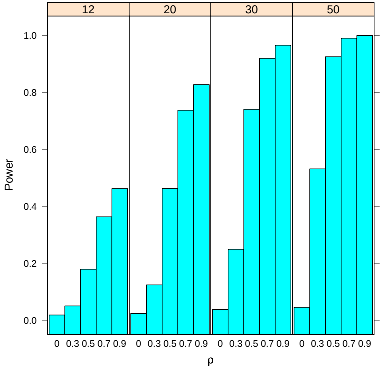


Graphs of HMZ & RHMZ with Normal Error

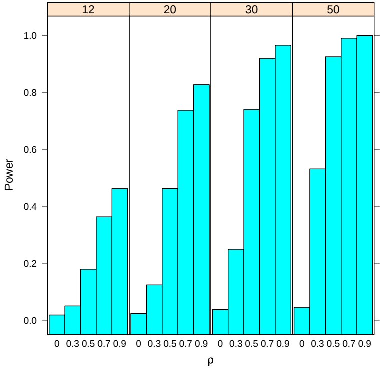


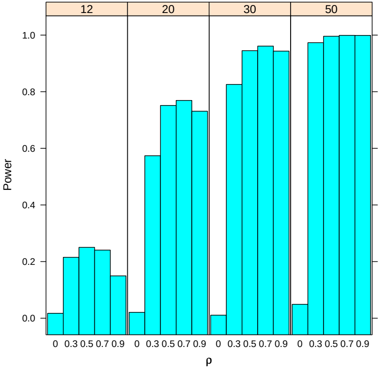


Graphs of HMZ & RHMZ with Contaminated Normal Error



Graphs of HMZ & RHMZ with Contaminated Normal Error





Methodology Algorithm for Scenario 1

1. Fit the H-M four parameter model to each of the series using OLS.
2. Test for autocorrelated errors using H-M or D-W; if autocorrelation is present then fit the model using the double bootstrap (McKnight et al. (2000)), else, continue with OLS. This is the purpose of our diagnostic procedure.
3. If the slope change coefficient is non-significant, refit the model using the level-change specification (i.e. H-M two-parameter model); do this for each series.
4. Re-compute H-M or D-W test for the possibly simplified models and re-confirm step 2.

5. Compute the overall level-change test statistic using our proposed parametric method or robust approach in case there are outliers in the data set.
6. Compute the weighted overall level-change using our formula.
7. Compute the individual standardized level-change coefficients.
8. Compute the overall standardized effect size using our formula.
9. Compute the individual pair-wise test on the homogeneity of effects across baselines.
10. Obtain a table of summary for all the results of the final models.

Methodology Algorithm for Scenario 2

In step 2, evidence of autocorrelation of lag-1 in the series indicates series of scenario 2. In this case, we use the double bootstrap method (McKnight et al. (2000)). Steps to be taken are as follows:

1. Fit an appropriate time series model to each of the series using the double bootstrap method.
2. Compute the residuals for each series.
3. Compute the inter-correlation among the series residuals (these cross correlation should be at lag zero unless there is reason to believe that one series is leading or lagging the others).
4. If the slope change coefficient is non-significant, refit the model using the level-change specification (i.e. H-M two-parameter model); do this for each series.
5. Follow through steps 6 to 10 as shown under the independent case.

General Form of Summary Table for LC (/ SC / RLC /RSC)

Series	LC	SE(LC)	TS(LC)	Pvallc	z(LC)	Wt	ES
1	LC_1	SE_1	$t(LC_1)$	$P(LC_1)$	$z(LC_1)$	W_1	ES_1
2	LC_2	SE_2	$t(LC_2)$	$P(LC_2)$	$z(LC_2)$	W_2	ES_2
3	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
J	LC_J	SE_J	$t(LC_J)$	$P(LC_J)$	$z(LC_J)$	W_J	ES_J

The Meta Analytic Components

In addition to the preliminary ouput summary table discussed above, we provide the user with the meta analytic components needed for an overall decision making. This include the following computations (each gives a weighted average of all statistics):

Overall-test statistic for the LC

$$z_{overall} = \frac{\sum_{j=1}^{J} z_j}{\sqrt{J}} \tag{5}$$

The Meta Analytic Components

In addition to the preliminary ouput summary table discussed above, we provide the user with the meta analytic components needed for an overall decision making. This include the following computations (each gives a weighted average of all statistics):

Overall-test statistic for the LC

$$z_{overall} = \frac{\sum_{j=1}^{J} z_j}{\sqrt{J}} \tag{5}$$

The weighted overall LC statistic

$$Overallwted = \frac{\sum_{j=1}^{J} [\frac{1}{\hat{\sigma}^2_{e_j}}(LCCoeff_j)]}{\sum_{j=1}^{J} \frac{1}{\hat{\sigma}^2_{e_j}}} \tag{6}$$

The overall standardized effect size

$$Overall - effect = \frac{\sum_{j=1}^{J} \left[\frac{LCCoeff_j}{\sqrt{MSResid_j}} \right]}{J} \tag{7}$$

The overall standardized effect size

$$Overall - effect = \frac{\sum_{j=1}^{J}\left[\frac{LCCoeff_j}{\sqrt{MSResid_j}}\right]}{J} \tag{7}$$

Test on the homogeneity of effects across units

$$Pair - wise \;\; test = \frac{z_j - z_i}{\sqrt{2}} \tag{8}$$

Motivational Example

Suppose we are required to compute the LC table (using the new parametric approach) given the following data set from the fictitious classroom setting,

Table: Data Input format

Series Number	Phase	Observations
1	1	2.5, 3.0, 4.5, 2.0, 3.0
1	2	3.0, 4.9, 3.0, 5.0, 3.0
2	1	2.5, 1.0, 3.0, 2.0, 4.0, 1.0, 2.5
2	2	5.5, 4.0, 4.0
3	1	1.0
3	2	2.0, 4.0, 2.5, 5.0, 3.0, 3.5, 2.5, 5.0, 4.0

Results from the new parametric approach

Summary Table for the Output

Series	LC	SE(LC)	TS(LC)	Pvallc	z(LC)	Wt	ES
1	0.60	0.579	1.037	0.165	0.9737	2.99	0.66
2	2.22	0.708	3.126	0.007	2.4548	1.99	2.16
3	2.50	1.149	2.176	0.031	1.8720	0.76	2.30

Results from the new parametric approach

Summary Table for the Output

Series	LC	SE(LC)	TS(LC)	Pvallc	z(LC)	Wt	ES
1	0.60	0.579	1.037	0.165	0.9737	2.99	0.66
2	2.22	0.708	3.126	0.007	2.4548	1.99	2.16
3	2.50	1.149	2.176	0.031	1.8720	0.76	2.30

Meta Analysis table

Overall test	Overall Wted	Overall ES
3.060241	1.411972	1.702348

Results from the new parametric approach

Summary Table for the Output

Series	LC	SE(LC)	TS(LC)	Pvallc	z(LC)	Wt	ES
1	0.60	0.579	1.037	0.165	0.9737	2.99	0.66
2	2.22	0.708	3.126	0.007	2.4548	1.99	2.16
3	2.50	1.149	2.176	0.031	1.8720	0.76	2.30

Meta Analysis table

Overall test	Overall Wted	Overall ES
3.060241	1.411972	1.702348

Test of Homogeneity

Individual z test	Pairwise z test
0.9736745	−1.0473046
2.4547869	−0.6352345
1.8720318	0.4120701

Results from the new robust approach

Summary Table for the Output

Series	LC	SE(LC)	TS(LC)	TPvallc	z(LC)	Wt	ES
1	0.5	0.6610	0.7565	0.2355	0.7208	2.2889	0.4784
2	2.0	0.8294	2.4115	0.0212	2.0295	1.4538	1.6641
3	2.0	2.2964	0.8709	0.2046	0.8254	0.1896	0.6885

Results from the new robust approach

Summary Table for the Output

Series	LC	SE(LC)	TS(LC)	TPvallc	z(LC)	Wt	ES
1	0.5	0.6610	0.7565	0.2355	0.7208	2.2889	0.4784
2	2.0	0.8294	2.4115	0.0212	2.0295	1.4538	1.6641
3	2.0	2.2964	0.8709	0.2046	0.8254	0.1896	0.6885

Meta Analysis table

Overall test	Overall Wted	Overall ES
2.064376	1.126895	0.9436843

Results from the new robust approach

Summary Table for the Output

Series	LC	SE(LC)	TS(LC)	TPvallc	z(LC)	Wt	ES
1	0.5	0.6610	0.7565	0.2355	0.7208	2.2889	0.4784
2	2.0	0.8294	2.4115	0.0212	2.0295	1.4538	1.6641
3	2.0	2.2964	0.8709	0.2046	0.8254	0.1896	0.6885

Meta Analysis table

Overall test	Overall Wted	Overall ES
2.064376	1.126895	0.9436843

Test of Homogeneity

Individual z test	Pairwise z test
0.7208	-0.9254
2.0295	-0.074
0.8254	0.8514

Comparison of Parametric & Robust method for Examples with / without Outliers

	Series Without Outliers			Series With Outliers	
Example	Series	Parametric LC	Robust LC	Parametric LC	Robust LC
2	1	−4.75	−5	−1253.5	−5
2	2	−3.222	−3	70.111	−2
2	3	−3.771	−3	6344.8	−2
4	1	1.615	1.7	390.12	1.9
4	2	1.8897	2.0	−5931.63	1.8
4	3	1.7235	1.9	607664.95	2.0
4	4	0.9795	0.8	2455863.48	1.5
5	1	−41.54	−44	−2278.54	−55
5	2	−28.56	−27	−925945.23	−27
5	3	−16.10	−12	1723.9	−10
6	1	13.90	14	38899.0	14
6	2	7.125	7	−66647.88	3
6	3	9.667	12	743.0	12
7	1	4.60	5	1.11e−4	6
7	2	5.5833	5	2.53e−2	6
7	3	5.667	6	5.67	6
7	4	3.300	3	4.44e−6	4
8	1	4.0714	4.0	3972.79	4.5
8	2	8.0606	8.0	1614.42	9.0
8	3	2.4306	2.5	2.4306	2.5
8	4	8.8571	9.0	−950.0	8.0
8	5	3.6690	3.5	7073.67	3.5

Scenario 3 (Dependence between series but independence within series)

We investigate two procedures:

1. The classical procedure based on normal likelihood theory and
2. A robust R procedure recently proposed by Kloke et al.(2009).

The asympototic theory of the estimators is similar to that of R estimators for linear model where all random errors are independent of one another.
Four methods namely CT, LME, JR & WW were compared under scenario 3.

1. CT Method

Initiated by Huitema (2007), it is a parametric method which uses the least squares procedure to generate the correlation test statistic for LC (or SC). It is computed as:

$$t = \frac{J^T t_i}{\sqrt{J^T R J}} \tag{9}$$

where **R** is the pooled within phase correlation matrix of all series
J is an m by 1 vector of ones
$\mathbf{t_i}$ is the corresponding test statistic for each baseline series
$df = N - 2 \times$ (Number of baseline series)

2. Mixed-Effects Model

A model with both fixed effects and random effects is called a mixed-effects model. For between series, we have $b_i \sim N(0, \sigma_b^2)$ and for within series, $\epsilon_{ij} \sim N(0, \sigma^2)$. The covariance between observations in the same series is σ_b^2 corresponding to a correlation of $\frac{\sigma_b^2}{\sigma^2+\sigma_b^2}$ assuming multivariate normality.

The LME function from the nlme library for R is used for fitting the linear mixed-effects model using, maximum likelihood (ML).

3. **JR Method**

An extension of the R estimators based on the joint ranks (JR) of all residuals developed by Kloke et al.(2009) to estimate the fixed effects in a linear model with cluster correlated continuous error distributions for general score functions is adopted here. The test statistic is given by:

$$t = \frac{1'\hat{\psi}}{\sqrt{3}} \tag{10}$$

where $\hat{\psi} = \hat{\sum}_{\eta}^{-\frac{1}{2}} \hat{\eta}$, $V(\hat{\eta}) = \hat{\sum}_{\eta}$, and $\hat{\eta}$ is the parameter estimate. A special design matrix was formulated for this purpose.

Empirical Alpha level, Type I error & Power

Extensive computer simulation studies were carried out (using R software and FORTRAN codes) over a variety of sample sizes and error distributions. Results were compared under each scenario stated above. We reported the results obtained for power and empirical type I error rate using a nominal alpha level of 0.05, 0.10. An AB design with 3-series specific and sample sizes $N = 12, 20, 30, 50, 100$ respectively were used where the simulation size was 15000 for each situation.

For the null model, we chose the following distributions: normal; contaminated normal (with $P = 0.10, 0.20$ and $\sigma_c = 10$ & 25 respectively); Cauchy; and contaminated Skewed Cauchy distributions (with $\mu = 30, P = 0.10, 0.20$ and $\sigma_c = 10$ & 25 respectively).

Distributions for random errors

1. Normal: The error structure associated with our model is estimated by
$$\hat{\epsilon}_t = \rho\hat{\epsilon}_{t-1} + a_t$$
which was used as the first-order autoregressive process error.
2. CN: A contaminated normal random variable is of the form
$$H = (1 - I_\epsilon)Z + I_\epsilon \sigma_c Z$$
where I_ϵ and Z are independent random variables, Z has a standard normal distribution, and I_ϵ is either 1 or 0 with probability ϵ or $1 - \epsilon$, respectively.
3. Cauchy: A contaminated skewed Cauchy random variable has the form
$$G = (1 - I_\epsilon)W + I_\epsilon[\sigma_c W + \mu],$$
where I_ϵ is a binomial random variable which takes on values 0 or 1 with probability $1 - \epsilon$ or ϵ, respectively, and W is a Cauchy random variable which is independent of $I_{1-\epsilon}$.

Conclusion

- Our new parametric method and its robust analog are recommended for senario 1, respectively, when series are without and with outliers.
- The double bootstrap method discussed by McKnight et al. (2000) in conjuction with our meta analysis summary table should be used for the case of independence between series but dependence within series (senario 2).
- For senario 3 (i.e. the case of dependence between series and independence within series), we recommend the correlation test (CT) or JR as discussed in Awosoga's thesis.
- Finally, dependence between series and dependence within series is handled using a combination of paired difference and double bootstrap method. More study still need be done in scenario 4, and this will be the focus of our future study.

Future Study

As an extension of this work, we will develop a R package, which can generate appropriate design matrices for individual users, combined with current program codes for multiple baseline time-series design intervention models. We plan to extend this work to more than two phases multiple baseline designs i.e. reversal designs such as ABA, ABAB, and ABABA designs. More work still need to be done on the scenario of dependence within series and dependence between series in terms of determining the best approach to handle different cases.

THANK YOU ALL!!!