IDENTIFICATION AND ESTIMATION OF LOCATION AND DISPERSION EFFECTS IN UNREPLICATED 2K-P DESIGNS USING GENERALIZED LINEAR MODELS

Rainier M. Sabangan

A Thesis

Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2010

Committee:

Richard N. McGrath, Advisor

Jane Chang

James Albert
Unreplicated fractional factorial experiments are sometimes used as screening experiments whenever replication is not practical due to cost or time constraints. It is then important to be prudent when choosing which factors are to be included in the model. This study suggests a method in identifying significant location and dispersion effects using generalized linear modeling approach. A critical values table was generated and the difference in the minus twice the log-likelihood of two different joint models is compared against the critical values to determine significance.
I would like to extend my heartfelt thanks to my committee members Dr. Jane Chang and Dr. James Albert for providing me the support and encouragement. And most especially, to my thesis advisor, Dr Herb McGrath for his unwavering belief in me and my ability to finish what I set out to do despite the odds.

I would also like to thank my family back in the Philippines. I know that I am always in their prayers and my success is not ever possible without their unconditional love and support.
# TABLE OF CONTENTS

1. INTRODUCTION .......................................................................................................................... 1

2. METHODOLOGY ....................................................................................................................... 4

3. ALGORITHMS ............................................................................................................................. 8

4. WHICH IS THE BETTER MODEL ............................................................................................... 13

5. SIMULATION RUNS .................................................................................................................. 17

   5.1 Using the Three Algorithms............................................................................................... 17

   5.2 One True Model With Known Coefficients......................................................................... 18

   5.3 Different True Models .......................................................................................................... 25

   5.4 Simulation For Models With ‘Best” Location Terms.......................................................... 30

6. DISCUSSION ............................................................................................................................... 37

   6.1 Model Building ...................................................................................................................... 38

   6.2 Model Comparison.................................................................................................................. 46

7. MULTI-MODAL PATTERN FOR 3 DISPERSION TERMS OR MORE.................................... 50

8. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE STUDY.............................. 52

REFERENCES .................................................................................................................................. 54

APPENDIX A. R Codes .................................................................................................................. 56

APPENDIX B. Tables and Figures.................................................................................................... 75
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>A typical design matrix of a $2^4$ factorial design</td>
<td>5</td>
</tr>
<tr>
<td>Table 2</td>
<td>Recommended design matrix for a fractional $2^{5-1}$</td>
<td>6</td>
</tr>
<tr>
<td>Table 3</td>
<td>Welding experiment data</td>
<td>11</td>
</tr>
<tr>
<td>Table 4</td>
<td>Comparison of estimates from the 3 algorithms</td>
<td>12</td>
</tr>
<tr>
<td>Table 5</td>
<td>Mean of coefficient estimates from a simulation</td>
<td>24</td>
</tr>
<tr>
<td>Table 6</td>
<td>95th quantile, mean and standard deviation of -2L difference</td>
<td>29</td>
</tr>
<tr>
<td>Table 7</td>
<td>95th quantile, mean and standard deviation of chi-square distribution</td>
<td>24</td>
</tr>
<tr>
<td>Table 8</td>
<td>Partial critical values table</td>
<td>37</td>
</tr>
<tr>
<td>Table 9</td>
<td>-2L values for model building using Pan and Taam (2002) data</td>
<td>39</td>
</tr>
<tr>
<td>Table 10</td>
<td>Injection molding experiment data</td>
<td>42</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R code for -2L and -2R computation</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Histograms of -2L differences when terms were added or removed</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Histograms of -2R differences when terms were added or removed</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>Plot of -2L and -2R</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>Regression analysis for -2L and -2R</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>Histograms of -2L differences for models with additional terms</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>Histograms of 10000 random samples from a chi-square distribution</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>Half-normal plot for the welding data</td>
<td>31</td>
</tr>
<tr>
<td>9</td>
<td>Histograms of -2L differences for models with 3 dispersion terms</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>Half-normal plot for the injection molding data</td>
<td>43</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

When designing experiments, it is always ideal to provide replication. However, some experiments are either too expensive, time consuming or dangerous to replicate. Factor screening is used to initially study several factors with a small amount of data in order to find the few more important ones. Often, un-replicated two level fractional factorial designs are used for this purpose. The usual goal is to identify the factors (and interactions) that impact the mean response. A more ambitious experimenter will also wish to identify those that impact the variance of the response, hence the importance of identifying both location and dispersion effects. Factors with significant location effects will be used to determine the desired setting that will lead to the desired mean, while determining the factors with dispersion effects will provide an understanding how to keep variability to its minimum. Factors that are determined to have no effects will be set to their most economical levels (Bergman and Hynen 1997).

Analysis of un-replicated experiments is difficult due to lack of an error term. Most suggested procedures are somewhat subjective. Perhaps the most widely used was advocated by Daniel (1959), employing the use of half-normal plots. Daniel’s method plots the absolute values of effect estimates against half-normal scores, and the significant effects will look like outliers in the plot. But, identification of location and dispersion effects is somewhat subjective. If the terms were ‘correctly’ identified, then the procedure is a reasonable approach to estimate the effects. However, McGrath and Lin (2001) showed that dispersion effect estimates are confounded to some
extent with location effect estimates. So, correctly identifying the location model is essential for studying dispersion.

Several authors have studied location and dispersion effects in unreplicated factorial experiments. Nelder and Lee (1998a) estimated the location and dispersion effects using two interlinked generalized linear models and discussed the advantages of using the GLM approach. Lee and Nelder (1998) used GLM to analyze data from quality-improvement experiments. Pan and Taam (2002) analyzed the severe impact of unidentified location effects to the dispersion model using the GLM method. A few more authors have championed the use of GLM in estimating the location and dispersion effects. However, their use of GLM is confined in estimating the coefficients when the location and dispersion model is already known.

The main focus of this paper is to find an approach on how to identify the active location and dispersion effects using GLM, not just estimation of the coefficients. Using this process, simulation was done to come up with a critical values table that can be used to determine if adding a location or a dispersion term is significantly helpful to the model. All of the program codes are written in R.

During the course of this study, several computer simulations were done with different methods to determine a feasible approach to come up with the critical values table, as well as ensure that the process execution time is reasonable. For each of these simulations, sampling from different true models was done which served as observations for a ‘hypothetical’ experiment. Then,
location and dispersion terms were added to the true model. For each joint location and dispersion models, a numerical value was computed for each. It was hypothesized that a function of the value of the two models that we want to compare asymptotically follows chi-square distribution. Knowing that, histograms of the differences were generated. These simulations and the results for each will be discussed in the succeeding sections of this paper. The critical values table will be presented and two methods on how to use the table will be discussed with several examples. Later in the paper, an unexpected finding will be reported for models with 3 or more dispersion. Articles related to this finding will be discussed and recommendations for future study will be suggested.
2. METHODOLOGY

For this study, the location model is defined as

\[ Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_p X_{pi} + \varepsilon_i \]

where \( \varepsilon_i \sim N(0, \sigma_i^2) \)

and the dispersion model (with a log-link) is defined as

\[ \log(\sigma_i^2) = \gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + ... + \gamma_q Z_{qi} \rightarrow \sigma_i^2 = \exp(\gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + ... + \gamma_q Z_{qi}). \]

In other words,

\[ E(Y_i) = \mu = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_p X_{pi} \] and

\[ \text{Var}(Y_i) = \sigma_i^2 = \exp(\gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + ... + \gamma_q Z_{qi}). \]

The main objective of this study is to identify the active location and dispersion effects

where \( \beta_j \neq 0 \) and \( \gamma_j \neq 0 \ j = 1, 2, ..., n-1 \) on a \( 2^{k-p} \) fractional factorial experiment where \( k-p=4 \) (i.e. \( n = 16 \) observations). This is a common design among unreplicated experiments. A typical design of a \( 2^4 \) full factorial experiment with high and low levels coded 1 and -1 is shown below.
If a $2^{k-p}$ fractional design is run, each column represents a string of $2^p$ effects. A sample $2^{k-p}$ fractional design would be $2^{5-1}$ where the design is shown below. Note that the design is similar to a $2^4$ full factorial except that the fifth factor (E) is in column ABCD, and each column represent two terms. The advantage of fractional factorial designs is that the number of runs is less than the runs required for full factorials by a factor of 2. A full factorial $2^5$ with 5 factors would need 32 runs, as compared to 16 runs for a $2^{5-1}$. The disadvantage of fractional factorials is that some of the effects may be confounded with other terms. For the design in Table 2 below, the main factors would be confounded with fourth-order interactions. Therefore, there is a need to separate the main factor effects from their confounded terms. However, the effects sparsity principle states only a few main factors and second order interactions will be significant. For further discussion on fractional factorials, please refer to Montogomery (2001).

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>AB</th>
<th>AC</th>
<th>AD</th>
<th>BD</th>
<th>CD</th>
<th>ABC</th>
<th>ABD</th>
<th>ACD</th>
<th>BCD</th>
<th>ABCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1. A typical design matrix of a $2^4$ factorial design with 4 main factors and all their interactions*
There are a total of 16 observations in the design and therefore, only 15 terms at most can be estimated (including intercept) for a $2^k-p$ unreplicated design where $k-p = 4$. However, there are a total of 15 location terms and 15 dispersion terms (plus two intercepts) or a total of 32 coefficients that need to be estimated for a saturated model. Therefore, there is a need to be prudent in identifying the terms that significantly affect the mean and variance among the responses.

It has been decided that generalized linear model (GLM) is the approach to be used for modeling in this study. One advantage of GLM is that observations need not be forced to be normal, but instead a distribution that more naturally models the data can be used without transformation. Distributions that are members of the exponential family can be analyzed using GLM. These distributions include normal, binomial, poisson, exponential and gamma distributions. One other advantage of using GLM is that since transformation is implicitly done, no de-
transformation of the data is needed since they are retained in their original state. For a more
detailed discussion of GLM, see McGrath (2007) for reference.
3. ALGORITHMS

For this study, three different algorithms were analyzed to estimate the location and dispersion effects and determine the coefficients. All of these algorithms pre-suppose that the terms in both location and dispersion models have been identified by some other method. These methods include the use of normal or half-normal plot as suggested by Daniel (1959).

The first algorithm uses the iterative least squares procedure discussed in Myers, Montgomery and Vining (2002, p.296-297) and Engel, J. and Huele, A (1996). This procedure essentially suggested the following iterative method (or IRLS which stands for Iterative Regression Least Squares) to estimate both location and dispersion effects.

0. Use OLS to obtain the estimated location coefficients.
1. Fit a model and calculate residuals e,
2. Fit the variance model using log(e$_i^2$) as a response to estimate the dispersion effects
3. Use resulting active dispersion coefficients (from 2) to calculate the fitted variances
4. Do regression with weights as 1/fits computed in #3 and do #1 again, i.e. use weighted least squares (WLS)
5. Continue until convergence$^1$.

$^1$ The authors did not specify what criterion is to be used for convergence. However for the R code developed, we used the -2L criterion. By using -2L as a criterion, the two succeeding sets of coefficients estimates will be computed with their -2L value. If the absolute value of their difference is almost zero (i.e. 0.01), then convergence is achieved.
Shown below is the R code to implement the IRLS algorithm. The comments in the program match the steps mentioned above for the procedure.

```r
# STEP 0 - DONE MANUALLY USING SUBJECTIVE APPROACH
# STEP 1
locmodel=glm(y~mat[,bcols])
mlcoefs=locmodel$coef
ressq=locmodel$residuals^2
# STEP 2
dispmodel=glm(ressq~mat[,dcols],family=Gamma(link = "log"))
mdcoefs=dispmodel$coef
# STEP 3
varfit=dispmodel$fitted.values
newml=MLVAL(c(mlcoefs,mdcoefs))
oldml=99999
j=1
while (abs(oldml-newml) > 0.001) { #STEP 5 Loop
  locmodel=glm(y~mat[,bcols],weights=1/varfit) #STEP 4
  mlcoefs=locmodel$coefficients
  ressq=locmodel$residuals^2
  dispmodel=glm(ressq~mat[,dcols],family=Gamma(link = "log"))
  mdcoefs=dispmodel$coef
  varfit=dispmodel$fitted.values
  oldml=newml
  newml=MLVAL(c(mlcoefs,mdcoefs))
  j=j+1
}
list(mlcoefs=mlcoefs,mdcoefs=mdcoefs,neg2logl=newml)
```

Note: MLVAL function is shown on the next page

The other two algorithms use two linked GLMs, i.e. the location and dispersion models. The distribution used for location model is normal while the dispersion model has a gamma distribution with log-link. The difference between these two algorithms is one uses the dglm package available in R and the other uses an algorithm developed by the author using R optim function.
A function that computes the minus twice the log-likelihood (−2L) or minus-twice the restricted log-likelihood (−2R) is written in R. Writing this function is necessary as optim feature in R optimizes the value returned by a function. This function is called MLVAL and is shown in Fig. 1 below.

```r
# This function computes the -2L or -2R given the location and dispersion model coefficients
MLVAL=function(bhats)
{
  beta_cnt=length(bcols)+1
  gamma_cnt=length(dcols)+1
  beta=bhats[1:beta_cnt]
  a=beta_cnt+1
  b=beta_cnt+gamma_cnt
  gamma=bhats[a:b]
  terms=c(1,bcols)
  mui=beta[1]
  if (beta_cnt>1)
    {for (i in 2:beta_cnt)
      {mui=mui+(beta[i]*mat[,terms[i]])}
    }
  terms=c(1,dcols)
  phi=exp(gamma[1])
  if (gamma_cnt>1)
    {for (i in 2:gamma_cnt)
      {phi=phi*exp(gamma[i]*mat[,terms[i]])}
    }
  neg2logl= sum(((y-mui)^2)/phi +log(2*pi*phi))
  if (method=="ml") {neg2logl}
  Else
    {phiimat=diag(phi)
     effectmat=cbind(rep(1,16),mat[,bcols])
     term=(t(effectmat) %*% solve(phiimat) %*% effectmat) / (2*pi)
     neg2logr=neg2logl + log(det(term))
     neg2logr}
}
```

**Fig 1.** R code for the computation of minus twice the negative log-likelihood (−2L) and minus twice the negative restricted log-likelihood (−2R).

The idea of the method using optim is that the process will optimize the value returned by MLVAL (−2L or -2R). As a result, optim will provide the location and dispersion coefficients that minimizes -2L or -2R, which in turn maximizes log-likelihood L or restricted log-likelihood R. The R code below shows how optim was used in the study.
# USING OPTIMIZER

```r
outoptim = optim(rep(0, 7), MLVAL)
out$value  # minimized -2L/-2R value
out$par    # set loc and disp coefficients/parameters that minimized -2L/-2R
```

For the other approach using dglm package in R, the R code is shown below.

```r
# USING DGLM
outdglm = dglm(y ~ mat[, bcols], ~ mat[, dcols], dlink="log", method='ml'))
outdglm$coefficients  # coefficients for the location model
outdglm$dispersion.fit$coefficients # coefficients for the dispersion model
outdglm$m2loglik      # minimized -2L
```

The complete code for each of these algorithms can be found in Table A.1 in Appendix A.

The welding data discussed in Pan and Taam (2002) is used and the resulting output is compared with the coefficient estimates that Nelder and Lee (1998) found. The welding data can be seen in Table 3 below.

<table>
<thead>
<tr>
<th>yi</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>AB</th>
<th>AC</th>
<th>AD</th>
<th>BC</th>
<th>BD</th>
<th>CD</th>
<th>ABC</th>
<th>ABD</th>
<th>ACD</th>
<th>BCD</th>
<th>ABCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43.7</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>40.2</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>42.4</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>44.7</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>42.4</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>45.9</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>42.2</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>40.6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>9</td>
<td>42.4</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>10</td>
<td>45.5</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>43.6</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>12</td>
<td>40.6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>44.0</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>14</td>
<td>40.2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>15</td>
<td>42.5</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>46.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. The Welding Experiment data from Pan and Taam (2002)
A comparison of the output from the R code for each of the three different algorithms and the results from Nelder and Lee (1998) shows similar results. Take note that the coding used by Nelder and Lee for high and low levels is 1 and 0, respectively. The comparison is shown on Table 4 below. Since all of the resulting values are similar, we can therefore use any of these algorithms for this data set. However, convergence issues may arise and for different data sets, these three algorithms may not always come up with the same results.

<table>
<thead>
<tr>
<th></th>
<th>Location Coefficients</th>
<th>Dispersion Coefficients</th>
<th>Log Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$</td>
<td>$\beta_{ACD}$</td>
<td>$\beta_{BCD}$</td>
</tr>
<tr>
<td>ML</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRLS</td>
<td>43.8317032</td>
<td>-0.2354314</td>
<td>1.8642036</td>
</tr>
<tr>
<td>Optim</td>
<td>43.8317838</td>
<td>-0.2354546</td>
<td>1.8641320</td>
</tr>
<tr>
<td>Dglm</td>
<td>43.8317973</td>
<td>-0.2354318</td>
<td>1.8641305</td>
</tr>
<tr>
<td>REML</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelder &amp; Lee</td>
<td>43.825</td>
<td>-0.239</td>
<td>1.869</td>
</tr>
<tr>
<td>Optim</td>
<td>43.8254780</td>
<td>-0.2387095</td>
<td>1.8693695</td>
</tr>
<tr>
<td>Dglm</td>
<td>43.82698945</td>
<td>-0.2380058</td>
<td>1.8681624</td>
</tr>
</tbody>
</table>

Table 4. Comparison of estimates using 3 different GLM algorithm and the results from Nelder and Lee (1998)

¹ Nelder and Lee (1998) came up with +0.235 which we think is a typographical error.
² Dglm only reports the first component of the formula for -2R.
4. WHICH IS THE BETTER MODEL

As mentioned at the beginning of this section, the starting point of each of these three algorithms is that the active location and dispersion effects are already pre-determined. And determining which terms would be included in the location model is somewhat subjective. The choice of active terms in the location model greatly affects the terms that would eventually be identified as significant in the dispersion model. These algorithms and their end results, which are the estimated coefficients, are only as good as the starting location and dispersion model. Therefore, there should be a way to identify what is the ‘best’ combination of active location and dispersion effects, i.e. the ‘best’ joint model.

To measure the advantage of using one model over the other, a quantitative yardstick needs to be used. Several tools for model selection have been developed through the years. Some of the more popular ones include AIC, Mallow’s Cp and stepwise regression. However, all these methods are developed for location models only, and do not take into consideration the contribution of the dispersion effects into the model.

Another approach to test if one model is better than the other is by using the likelihood ratio test. For nested models, likelihood ratio test is usually conducted to determine if a model with fewer terms is significantly different from a model with more terms. Suppose we are considering two nested models $\omega_1$ and $\omega_2$ such that $\omega_1$ is a subset of $\omega_2$. In other words, $\omega_1$ can be derived by setting some of the coefficients of the terms in $\omega_2$ to 0. The main idea of the likelihood ratio test is
to compare the maximum likelihood of the two models. Therefore, the maximized likelihood for the simple model $\omega_1$ is

$$\max_{\theta \in \omega_1} L(\theta, y) = L(\hat{\theta}_{\omega_1}, y)$$

where $\hat{\theta}_{\omega_1}$ is the MLE of $\theta$ under $\omega_1$.

Similarly, the maximized likelihood for model $\omega_2$ is

$$\max_{\theta \in \omega_2} L(\theta, y) = L(\hat{\theta}_{\omega_2}, y)$$

where $\hat{\theta}_{\omega_2}$ is the MLE of $\theta$ under $\omega_2$.

Minus twice the log of the ratio of these two quantities

$$-2\log(\lambda) = -2 \log \left( \frac{L(\hat{\theta}_{\omega_1}, y)}{L(\hat{\theta}_{\omega_2}, y)} \right)$$

can be rewritten as

$$-2\log(L(\hat{\theta}_{\omega_1}, y)) + 2\log(L(\hat{\theta}_{\omega_2}, y))$$

Minus twice the log of the likelihood ratio converges to chi-square distribution as $v \to \infty$ with $v$ degrees of freedom which is the difference in the number of terms between the two models. Thus for large sample size,

$$-2\log \lambda = 2 \log L(\hat{\theta}_{\omega_2}, y) - 2 \log L(\hat{\theta}_{\omega_1}, y) \sim \chi^2_v$$

where the degrees of freedom $v = \text{dim}(\omega_2) - \text{dim}(\omega_1)$.

In other words, the degrees of freedom is the number of parameters in the larger model $\omega_2$ minus the number of parameters in the simpler model $\omega_1$. 
However, since we deal with small sample size of 16, asymptotic assumption will not hold and the chi-square table cannot be used. Also, the likelihood ratio test only works for nested models. The ‘active’ dispersion effects depend on what terms were chosen in the location model. Therefore, the two models that are being compared may not necessarily be nested. For example, a location model with A and B as active model may have D as an active dispersion effect. However, if we add C to the location model, the resulting active dispersion effect may be AC. Therefore the original joint location and dispersion model \( y \sim A + B \) and \( \sim D \) is not nested with joint location and dispersion model \( y \sim A + B + C \) and \( \sim AC \).

For the purpose of this study, we would like to introduce a new approach to model selection. Let’s denote L (or log-likelihood) as the log\([L(\theta_{o,i}, y)]\) of model \( \omega_i \). Pan and Taam (2002) provided the formula to compute the minus twice the log-likelihood (-2L) of a particular model as:

\[
-2L = \sum_{i=1}^{n} \left( \frac{(y_i - \mu_i)^2}{\sigma_i^2} \right) + \ln (2\pi \sigma_i^2)
\]

(1)

where \( y_i \) = ith response observation

\( \mu_i \) = location model fit of the ith observation

\( \sigma_i^2 \) = dispersion model fit of the ith observation

The set of location and dispersion effect coefficient estimates that minimizes -2L (i.e. maximizes L) are called the maximum-likelihood estimates (ML).
We know that the maximum likelihood estimator for the variance is biased. To adjust for this bias, another form of likelihood estimation called restricted maximum likelihood (REML) can be used. Let's denote R as the restricted log-likelihood of a particular model. The formula to compute for twice the negative restricted log-likelihood (-2R) is also given in Pan and Taam (2002) and is defined as:

\[-2R = \sum \left( \frac{(y_i - \mu)^2}{\sigma_i^2} \right) + \ln \left( 2\pi \sigma_i^2 \right) + \ln(\det(\frac{\chi^T \phi^{-1} \chi}{2\pi})) \]  

(2)

where \( \chi \) = location effects matrix
\( \Phi = \text{diag}(\sigma_i^2) \)

The MLVAL function is used to compute for the -2L and -2R of a model given the coefficients in the terms of the location and dispersion effect.

Using the MLVAL function, it is then possible to compute the -2L or -2R of model \( \omega_1 \) and compare it with the computed -2L or -2R of model \( \omega_2 \) by getting their difference, and determine if that difference is statistically different. But how do we determine if they are statistically different? Similar to other statistical tests, we suggest with coming up with a critical values table for the -2L or -2R difference. The next section will discuss the process in coming up with critical values table for experiments with sample size equal to 16.
5. SIMULATION RUNS

5.1 USING THE THREE ALGORITHMS

Initial simulation of 10,000 runs was executed using the three algorithms to compute for both -2L and -2R yielded with the following observations:

1. The IRLS method takes the longest execution time to process.
2. The dglm method encounters non-convergence warning message for some datasets.
   Various fatal error messages were also encountered that stop the execution of the program. An attempt has been made to contact individuals who may have an answer to the problem, but a reply has not been received to date.
3. Using optim function is the most stable of the three methods.

Since using dglm always ends up in the program stopping due to errors, another simulation of 10000 runs was done using a combination of IRLS and optim. Even though IRLS is slow, we need to find starting values when a call to optim is made. Instead of initial estimates of all zeroes, it may be faster to run IRLS initially to find reasonable starting values for optim. By providing reasonable starting coefficients, we expect that the slow processing time of the IRLS will be offset by making optim run faster. In this round of simulation, we would also monitor the execution time so we can decide if it is feasible to use this approach.


5.2 ONE TRUE MODEL WITH KNOWN COEFFICIENTS

In reality, a true model can never be determined. However, for the purpose of this study, a true model can be known and assumed, and observations would be sampled from that model. The true model used for the initial runs is given below:

Location model \[ y \sim 43 - 2A + 3B - 5C \]

Dispersion model \[ \sim -3 - 4A + 2C \]

Knowing the active terms in the location and dispersion model as well as their coefficients, random samples can then be generated from the model. The R code to generate 16 random samples from the above true model is shown below:

\[
sd = \sqrt{\exp(-3 - 4A + 2C)}
\]

\[
y = 43 - (2A) + (3B) - (5C) + \text{rnorm}(16,0,sd)
\]

For each set of \( y \) with 16 observations, the \(-2L\) and \(-2R\) is computed and the corresponding coefficients estimates were determined for the following scenario:

1. The active location and dispersion terms are the same as the true model (i.e. A, B & C for location model, A and C for dispersion model).
2. D is ‘mistakenly’ added to the location model.
3. B is ‘mistakenly’ added to the dispersion model.
4. D and B are ‘mistakenly’ added to location and dispersion model, respectively.
5. C is ‘mistakenly’ removed from the location model.

6. A is ‘mistakenly’ removed from the dispersion model.

7. C and A are ‘mistakenly’ removed from the location and dispersion model, respectively.

The complete R code for this round of simulation can be found in Table A.2 in Appendix A.

When terms were added to the location and/or dispersion model, the shape of the histogram for the \(-2L\) difference resemble a chi-square distribution, except for the negative values. However, when terms were removed from the true model, it does not resemble a chi-square distribution anymore. Please note that we still adhere to the formula

\[
\text{Difference} = -2L(\text{model with lesser terms}) \\
\text{minus} -2L(\text{model with more terms}).
\]

When terms were added, the true model has lesser terms. But when terms were removed, the true model has more terms. Please refer to Fig 2 below.
Fig 2. Histograms of -2L differences between the true model and the model where terms are added or removed.

Histograms for -2R differences are shown in Fig 3. Similar to -2L differences, resemblance to a chi-square distribution is observed when terms are added to the true location or dispersion model. And which is not the case when terms are removed from the true model. It looks more like
bell-shaped curve though than the -2L difference. The significance of these findings can be the subject for future research.

Fig 3. Histograms of -2R differences between the true model and the model where terms are added or removed.

For a less complicated program code and to cut down on program execution time, we have to choose whether we want to use -2L or -2R for the critical values table. The plot in Fig 4 below
shows that for each model, the -2L and -2R are strongly correlated. The solid line in each plot is the (0,1) line. A regression analysis for all of the models confirms that observation, with a p-value of almost 0. Fig 5 below shows the regression analysis done for the true model.

Fig 4. Plot of -2L and -2R for each of the models shows a strong linear relationship.
> summary(lm(reml~ml))
Call:
  lm(formula = reml ~ ml)
Residuals:
        Min          1Q    Median          3Q         Max
-6.818671 -0.790570 -0.001307  0.778087  5.211992
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 13.136861   0.046562   282.1   <2e-16 ***
mlmata[1, ]  0.704131   0.001606   438.5   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 1.229 on 9998 degrees of freedom
Multiple R-squared: 0.9506,    Adjusted R-squared: 0.9506
F-statistic: 1.923e+05 on 1 and 9998 DF,  p-value: < 2.2e-16

**Fig 5.** Regression analysis for -2L and -2R

Analysis of the estimate for the coefficients generated was also done. The estimated coefficients for the location model are close to their real value whether the ML or REML method is used. However, estimates of dispersion coefficients are closer using REML method when terms are added to the model. When terms are removed from the model, the coefficient estimates for dispersion effects are off their true value, very much more so when an active location term is removed from the model. Please refer to Table 5 below.
This approach may not work because removing terms from the true model does not in any way asymptotically converge to a chi-square distribution. The approach should be to start from a basic true model, and just keep adding terms into the model. The next round of simulation makes use of that concept. 

The algorithm that is a combination of IRLS and optim takes too long to execute. Keeping in mind that eventually, we might be running for more than 1 million simulations, it is imperative to find a way on how to cut down on program execution time as well. If IRLS is dropped and optim function is to be run by itself, convergence problem is being encountered. A solution that was arrived at is to call optim a number of times, starting with initial all zero coefficients. If optim does not converge at first pass, then optim will be called again. Instead of starting with 0 coefficients, the coefficients passed from the previous call will be the starting coefficients for the next optim call. This will be done in a loop until it converges. The R code of this algorithm is shown below:

<table>
<thead>
<tr>
<th>Model</th>
<th>Location Model Coefficients</th>
<th>Dispersion Model Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$ (=43)</td>
<td>$\beta_A$ (=−2)</td>
</tr>
<tr>
<td>True</td>
<td>ML 42.999528</td>
<td>-1.994949</td>
</tr>
<tr>
<td></td>
<td>REML 42.99998</td>
<td>-1.99919</td>
</tr>
<tr>
<td>+D to Loc</td>
<td>ML 43.000294</td>
<td>-2.00064</td>
</tr>
<tr>
<td></td>
<td>REML 43.00113</td>
<td>-2.001304</td>
</tr>
<tr>
<td>+B to Disp</td>
<td>ML 43.001189</td>
<td>-2.00074</td>
</tr>
<tr>
<td>+D to Loc &amp; +B to Disp</td>
<td>ML 43.001774</td>
<td>-2.001304</td>
</tr>
<tr>
<td>-C from Loc</td>
<td>ML 42.720142</td>
<td>-1.998966</td>
</tr>
<tr>
<td></td>
<td>REML 42.71008</td>
<td>-1.99865</td>
</tr>
<tr>
<td>-A from Disp</td>
<td>ML 43.000072</td>
<td>-1.998496</td>
</tr>
<tr>
<td></td>
<td>REML 43.00007</td>
<td>-1.998386</td>
</tr>
</tbody>
</table>

Table 5. Mean of estimated coefficients computed from a simulation of 10,000 (with their true value enclosed in parenthesis).
A mini-simulation that was conducted that this confirms that this algorithm works much faster than the combination of IRLS and optim and comes up with same results for both -2L and -2R values, and similar results up to 4th decimal place for the estimate of the coefficients.

## 5.3 DIFFERENT TRUE MODELS

A different model design is used for this round of simulation. Instead of taking samples from a single model with known coefficients, several true models were used. The following are the true models and their notation:

- **(0,0)** – Model with no location and no dispersion effect
- **(1,0)** – Model with 1 location effect and no dispersion effect
- **(0,1)** – Model with no location effect and 1 dispersion effect
- **(1,1)** – Model with 1 location effect and 1 dispersion effect
- **(2,0)** – Model with 2 location effects and no dispersion effect
- **(2,1)** – Model with 2 location effects and 1 dispersion effect
- **(2,2)** – Model with 2 location effects and 2 dispersion effects
- **(1,2)** – Model with 1 location effect and 2 dispersion effects
The observations samples from each of the true models mentioned above will be sample from a population with intercepts of 0 and the active terms have coefficients of 1 for both the location and dispersion model. For example, if the assigned active location terms for true model (2,2) are C and D, and the active dispersion terms are A and B, then the location model is $y \sim C + D$ and dispersion model is $\sim A + B$. We assume normal distribution and mean 0, and the R code to generate 16 samples from (2,2) is:

\[
\begin{align*}
    \text{sd} & = \sqrt{\exp(A+B)} \\
    y & = C + D + \text{rnorm}(16,0,\text{sd})
\end{align*}
\]

The computation of -2R was dropped since it has been determined from the previous simulation that -2R is strongly linearly correlated with -2L.

For each of the true models mentioned above, randomly assigned additional terms were then added to the location and/or dispersion model and the -2L value is computed, the difference from the true model is computed and histograms were generated. The R code for the true model (2,2) can be found in Table A.3 in Appendix A. The histogram of the differences in -2L for true models (0,0) and (2,2) is shown below in Fig 6 below. To see the histograms for all the true models, please refer to Fig B.1 in Appendix B. From the histograms, we see more resemblance to chi-square distribution now, but the asymptotics are still not very helpful.
Fig 6. Histogram of -2L differences between the true models and the model with additional randomly selected terms.
The histograms now look more like the characteristic shape of chi-square distribution. From the previous discussion on likelihood ratio test, it was mentioned that the difference in the number of terms between the two models is the degrees of freedom. From the histograms above, we can see that for different degrees of freedom, the peak of the curve somehow flattens and spreads to the right away from 0 as degrees of freedom increases. This is the same behavior that can be seen for samples from chi-square distributions, as shown in Fig 7 below.

![Histograms of chi-square distributions with different degrees of freedom](image)

**Fig 7.** Histogram of 10000 random samples from a chi-square distribution with 1, 2, 3 and 4 degrees of freedom

The histogram of models with no dispersion closely resemble the chi-square distribution for the same degrees of freedom. This is not true, however, for models with dispersion terms. Even more so, as the number of dispersion terms increases, the more it does not resemble a chi-square distribution for the same degrees of freedom. To get a deeper understanding of the effect of the
number of dispersion terms in the distribution, a summary of the 95\textsuperscript{th} quantile, mean and standard deviation of the differences for each model is summarized in Table 6 below for true models (0,0) and (2,2). The complete summary for all the models can be seen in Table B.3 in Appendix B. We can see from Table 6 that for the same degrees of freedom, different statistics are obtained depending on the number of location and dispersion terms are in the model. And if we compare these values to the expected 95\textsuperscript{th} quantile, mean and standard deviation if they are actually chi-square distributed (as shown in Table 7), that would give us further confirmation that we cannot use the existing chi-square table.

<table>
<thead>
<tr>
<th>df</th>
<th>(0,0)</th>
<th>95Q</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>(2,2)</th>
<th>95Q</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1,0)</td>
<td>4.700857</td>
<td>1.206331</td>
<td>1.701453</td>
<td>(3,2)</td>
<td>8.309203</td>
<td>2.193903</td>
<td>3.072102</td>
</tr>
<tr>
<td>1</td>
<td>(0,1)</td>
<td>4.7658</td>
<td>1.248415</td>
<td>1.749845</td>
<td>(2,3)</td>
<td>6.897766</td>
<td>1.850837</td>
<td>2.526546</td>
</tr>
<tr>
<td>2</td>
<td>(1,1)</td>
<td>7.327649</td>
<td>2.433343</td>
<td>2.434533</td>
<td>(3,3)</td>
<td>14.0293</td>
<td>4.89368</td>
<td>4.627996</td>
</tr>
<tr>
<td>2</td>
<td>(2,0)</td>
<td>7.347385</td>
<td>2.489031</td>
<td>2.477084</td>
<td>(4,2)</td>
<td>12.8323</td>
<td>4.296251</td>
<td>4.278364</td>
</tr>
<tr>
<td>2</td>
<td>(0,2)</td>
<td>7.46947</td>
<td>2.500361</td>
<td>2.487133</td>
<td>(2,4)</td>
<td>11.0015</td>
<td>3.791013</td>
<td>3.623804</td>
</tr>
<tr>
<td>3</td>
<td>(1,2)</td>
<td>11.03515</td>
<td>4.308915</td>
<td>3.469508</td>
<td>(3,4)</td>
<td>20.29339</td>
<td>8.328025</td>
<td>6.152605</td>
</tr>
<tr>
<td>3</td>
<td>(3,0)</td>
<td>9.996214</td>
<td>3.915147</td>
<td>3.189729</td>
<td>(5,2)</td>
<td>14.53544</td>
<td>5.425751</td>
<td>4.663842</td>
</tr>
<tr>
<td>4</td>
<td>(2,2)</td>
<td>13.00016</td>
<td>5.461071</td>
<td>3.887711</td>
<td>(4,4)</td>
<td>35.862</td>
<td>16.60115</td>
<td>10.28244</td>
</tr>
</tbody>
</table>

Table 6. 95\textsuperscript{th} quantile, mean and standard deviation of -2L differences between models

<table>
<thead>
<tr>
<th>df</th>
<th>95Q</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.841459</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5.991465</td>
<td>2</td>
<td>1.414214</td>
</tr>
<tr>
<td>3</td>
<td>7.814728</td>
<td>3</td>
<td>1.732051</td>
</tr>
<tr>
<td>4</td>
<td>9.487729</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 7. 95\textsuperscript{th} quantile, mean and standard deviation of chi-square distribution

With the information gained from the previous simulation, we can now design and fine-tune the approach for the next round of simulation.
5.4 SIMULATION FOR MODELS WITH ‘BEST’ LOCATION TERMS

One of the changes from the previous approach is to re-group the true models by the number of dispersion terms, and keep adding location terms for each grouping. That way, we can keep track the behavior or pattern by the number of dispersion terms in the model. From previous simulations, we observed that as dispersion terms in the model increases, the processing execution time increases as well. Also, we do not need to monitor the estimated coefficients anymore as the previous two rounds of simulations suggest that the coefficients that come up with the minimum -2L are close to their true values.

Another major change to the previous approach is the selection of terms added in the location and dispersion model. Whereas previously, terms are added randomly and are fixed throughout the entire simulation, this is not the practice in real life situation. In practice, the terms that are included in the model is determined by an empirical method. One of the most commonly used is the half-normal plot suggested by Daniel (1959). Using this method, the absolute value of the coefficient estimates of all the terms in the saturated location model are first plotted. The points that appear to be outliers or separated are the ones that can be thought of as significant. This method is subjective and there is no numerical measure as to whether adding an additional term would greatly improve the model. The papers from Box and Meyer (1986), Wang (1989), Fuller and Bisgaard (1995), Bergman and Hynen (1997), and Blomkvist, Hynen, and Bergman (1997) espoused this idea. They also discussed that after significant location terms were known, only then can identification of dispersion effects proceed.
To illustrate how to determine the significant terms using half-normal plot, a saturated model is first fitted and the absolute value of the coefficient estimates (excluding intercept) are sorted. These values are then plotted on a normal plot. Those values that stick out from the rest of the coefficients can be considered significant. Below is the sorted order of the coefficient estimates for the Pan and Taam (2002) data:

<table>
<thead>
<tr>
<th>BC</th>
<th>AD</th>
<th>BD</th>
<th>CD</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>AB</th>
<th>ACD</th>
<th>ABC</th>
<th>AC</th>
<th>D</th>
<th>BD</th>
<th>BCD</th>
<th>ABCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>0.050</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.150</td>
<td>0.150</td>
<td>0.300</td>
<td>0.375</td>
<td>0.375</td>
<td>0.400</td>
<td>0.400</td>
<td>0.425</td>
<td>2.150</td>
<td>3.100</td>
</tr>
</tbody>
</table>

From the half-normal (QQ) plot below, there are 2 points that separate from the rest. Looking at the sorted coefficient estimates, we can tell that these are estimates for BCD and ABCD.

![Normal Q-Q Plot](image)

**Fig. 8.** Half normal plot for the Welding Data from Pan and Taam (2002) of coefficient estimates for the saturated model.
So, a joint model with $l$ location effects and $d$ dispersion effects (plus intercepts) will always include the largest (in absolute values) $l$ location terms and the largest $d$ dispersion terms. But these $d$ terms depend on which $l$ terms were selected.

In programming terms, the coefficients of the location terms for the saturated model will first be estimated and ranked based on the absolute values. The terms with the highest ranks gets priority in being included in the model. For example, with $l = 3$, the 3 largest location terms will be in the model. The location model with $l$ terms will then be fit and the resulting squared residuals become the response observations to model for the dispersion. This process would result in the estimates of dispersion coefficients, which would also be ranked and their ranking would determine the order when they get to be added in the dispersion model.

The process is similar to the IRLS, except the iterative part.

This approach will be incorporated in the next round of simulation, and the R code to implement this new approach can be seen in Table A.4 in Appendix A for models with no dispersion and Table A.5 in Appendix A for models with 1 dispersion effects. The R code for 1 dispersion effect can be expanded to other model with more dispersion effects by simply changing a parameter in the program.

An initial simulation of 10,000 for 0 to 7 dispersion terms in the model with 0 to 7 additional location terms was executed. For a sample size of 16, a maximum of 14 terms can be estimated (excluding intercept) for both the location and dispersion models. Two extreme cases
would be 14 location terms and no dispersion terms or no location terms and 14 dispersion terms. Following the notation of the previous simulation, these would be (14,0) and (0,14), respectively. To simplify matters, we decided to start at (0,0) and stop at (7,7).

As expected, the more dispersion terms in the model, the longer it took for the software to execute. When the execution for the 5 dispersion model finished after more than 24 hours, we decided to not proceed with the simulation for the 6 and 7 dispersion terms model.

The histograms generated by this round of simulation can be seen in Table B.4 of Appendix B. For the models with 2 or less dispersion terms, the histogram looks as expected. First of all, the histograms still resemble chi-square distributions. As more location terms are added to the model, the histogram of the -2L differences goes further away from 0. However, for models with dispersion terms 3 or more, a surprising trend shows in the histograms. There appears to be a second mode in the right tail, as shown in Fig. 9 below.
This trend starts to appear when there are 3 location and 3 dispersion terms. Further analysis was done with and will be discussed in the later section of this paper. Due to this, it was decided that for the critical values table that will be generated by this study, we will only proceed for 0 to 2 dispersion terms in the model.

The final simulation would be the basis for the critical values table. Basically, the information that will be shown in the critical values table would be the 99th, 95th, 90th and 80th quantile for each group of number of dispersion terms. And for each group denoted by the number of dispersion terms, a sub-grouping will be created by the number of location terms added.
Two versions of R code were written for the final simulation. The first version is for models with no dispersion terms and the other version is for models with dispersion terms. The codes can be seen on Table A.4 and Table A.5 in Appendix A. The version of the R code for models with dispersion terms can be used for any number of dispersion terms with just a change in the value of a parameter in the code.

Another change in the R code from the previous simulation is the manner in which simulations are conducted. Instead of executing all the simulations at once (i.e. one main loop executed n times where n is the total number of simulations), the design was changed by grouping the simulation into cycles, and each cycle has 50,000 simulations each. This change is necessary to enable the executions to run in parallel using multiple PCs. A pilot run of 10 cycles (or 500,000 simulations) for the no dispersion version was executed to estimate the execution timing as well as gather data to determine the total number of cycles that needs to be executed. Since $\alpha = 0.05$ is most commonly used, we decided to use the 0.95 quantile information to compute for total cycles. Also, it was observed that the -2L difference of the model with 0 location and 7 location terms provides the biggest variance. For the mentioned reasons, we decided to use the 0.95 quantile of the -2L difference of the no location terms and the 7 location terms model with no dispersion terms as our samples to determine the total number of cycles we need to execute.

From central limit theorem, $V(\bar{w}) = \frac{\sigma^2}{n}$ where $\bar{w}$ is the mean of the samples. Let $\bar{w}$ be the mean of 0.95 quantile of the first 10 cycles of 50,000 simulations. Since we don’t know $\sigma^2$, we can
estimate it by using the variance of the 10 samples. Let’s denote this estimate of \( \sigma^2 \) as \( s^2 \). We want a standard error of 0.05, and therefore \( 2 \sqrt{\text{variance}} = 0.05 \).

\[
2 \sqrt{\text{variance}} = 0.05 \rightarrow 2 \sqrt{\frac{s^2}{n}} = 0.05 \rightarrow n = \frac{4s^2}{0.05^2}
\]

The computed \( s^2 \) is 0.01480387, and using the formula above, \( n = 24 \) cycles.

The initial 10 cycles took 7 hours to execute. Therefore, with 24 cycles, it is expected that the duration of the execution time for no dispersion model is (7hrs * 2.4) or 17 hours. But that is only for models with no dispersion effects. It is expected that models with dispersion will take much longer to execute. A request was made to use the PC’s located the BA building computer labs for the weekend. When the request was approved, and with the availability of more PC’s to execute the programs, we decided to go beyond the 24 cycles needed so that we can accurately estimate the critical values by having more cycles. During the allotted weekend when the PC’s were reserved for us, the programs were started and were running starting Friday evening until morning of the next Monday.

A total of 70 cycles were executed for models with no dispersion terms (or 3.5M simulations), 33 cycles for the models with 1 dispersion terms (or 1.65M simulations) and 56 cycles for the models with 2 dispersion terms (or 2.8M simulations). Partial results for each cycle can be found in Tables B.5.1, B.5.2 and B.5.3 in Appendix B, grouped by the number of dispersion of terms in the model. The data was then tabulated and summarized.
6. DISCUSSION

The complete critical values table can be found in Table B.7 in Appendix B. Table 8 below only shows the critical values for selected models at 95th quantile. The number in each cell indicates the critical value for the -2L difference of a pair of model where one model has lesser terms than the other. The row corresponds to the model with lesser terms. For example, if we would like to find the critical value of the -2L difference of the model with 2 location and no dispersion terms (2,0) over a model with 3 location and no dispersion terms (3,0), then we should be looking at the cell intersecting row (2,0) and column (3,0). For the 95th quantile, that value is 7.94.

There are two ways that this table can be used: one is for model building, and the other is by direct comparison of two different joint models.

<table>
<thead>
<tr>
<th></th>
<th>0.95</th>
<th>(1,0)</th>
<th>(2,0)</th>
<th>(3,0)</th>
<th>(4,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
<th>(2,1)</th>
<th>(3,1)</th>
<th>(0,2)</th>
<th>(1,2)</th>
<th>(2,2)</th>
<th>(3,2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td></td>
<td>10.18</td>
<td>17.24</td>
<td>23.83</td>
<td>30.41</td>
<td>9.46</td>
<td>17.97</td>
<td>25.83</td>
<td>33.82</td>
<td>13.70</td>
<td>22.39</td>
<td>30.76</td>
<td>39.73</td>
</tr>
<tr>
<td>(1,0)</td>
<td></td>
<td>0.00</td>
<td>8.34</td>
<td>15.49</td>
<td>22.46</td>
<td>4.69</td>
<td>11.23</td>
<td>18.64</td>
<td>26.67</td>
<td>8.86</td>
<td>15.81</td>
<td>23.79</td>
<td>32.60</td>
</tr>
<tr>
<td>(2,0)</td>
<td></td>
<td>NA</td>
<td>0.00</td>
<td>7.94</td>
<td>15.31</td>
<td>NA</td>
<td>6.83</td>
<td>13.26</td>
<td>20.84</td>
<td>4.60</td>
<td>11.36</td>
<td>18.48</td>
<td>27.12</td>
</tr>
<tr>
<td>(3,0)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>0.00</td>
<td>8.03</td>
<td>NA</td>
<td>8.95</td>
<td>15.80</td>
<td>NA</td>
<td>7.16</td>
<td>14.12</td>
<td>22.14</td>
<td></td>
</tr>
<tr>
<td>(4,0)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.00</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>9.84</td>
<td>17.73</td>
<td></td>
</tr>
<tr>
<td>(0,1)</td>
<td></td>
<td>7.77</td>
<td>14.38</td>
<td>20.80</td>
<td>27.25</td>
<td>0.00</td>
<td>14.81</td>
<td>22.52</td>
<td>30.44</td>
<td>7.28</td>
<td>19.05</td>
<td>27.30</td>
<td>36.23</td>
</tr>
<tr>
<td>(1,1)</td>
<td></td>
<td>NA</td>
<td>5.85</td>
<td>12.58</td>
<td>19.24</td>
<td>NA</td>
<td>0.00</td>
<td>14.76</td>
<td>22.87</td>
<td>4.60</td>
<td>8.35</td>
<td>19.66</td>
<td>28.76</td>
</tr>
<tr>
<td>(2,1)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>5.29</td>
<td>12.12</td>
<td>NA</td>
<td>NA</td>
<td>0.00</td>
<td>16.31</td>
<td>NA</td>
<td>6.38</td>
<td>9.85</td>
<td>22.28</td>
</tr>
<tr>
<td>(3,1)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>5.09</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>8.32</td>
<td>12.14</td>
</tr>
<tr>
<td>(4,1)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>10.94</td>
<td></td>
</tr>
<tr>
<td>(0,2)</td>
<td></td>
<td>NA</td>
<td>12.00</td>
<td>18.36</td>
<td>24.78</td>
<td>NA</td>
<td>12.36</td>
<td>19.79</td>
<td>27.79</td>
<td>0.00</td>
<td>16.53</td>
<td>24.54</td>
<td>33.53</td>
</tr>
<tr>
<td>(1,2)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>9.96</td>
<td>16.55</td>
<td>NA</td>
<td>NA</td>
<td>11.93</td>
<td>19.82</td>
<td>NA</td>
<td>0.00</td>
<td>16.73</td>
<td>25.65</td>
</tr>
<tr>
<td>(2,2)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>9.19</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.00</td>
<td>18.85</td>
</tr>
<tr>
<td>(3,2)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.00</td>
<td>NA</td>
</tr>
<tr>
<td>(4,2)</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 8. Partial Critical Values Table
6.1 MODEL BUILDING

This method is to be used when the objective is to identify the significant location and dispersion terms. Examples will be provided below:

Example 1. To demonstrate the use of this table for model building, let’s use the Pan and Taam (2002) data. After fitting for the saturated model

\[ y \sim A + B + C + D + AB + AC + AD + BC + BD + CD + ABC + ABD + ACD + BCD + ABCD, \]

the following are the location effects coefficient estimates and their ranks:

<table>
<thead>
<tr>
<th>Term</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>AB</th>
<th>AC</th>
<th>AD</th>
<th>BC</th>
<th>BD</th>
<th>CD</th>
<th>ABC</th>
<th>ABD</th>
<th>ACD</th>
<th>BCD</th>
<th>ABCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff. Est.</td>
<td>0.125</td>
<td>-0.150</td>
<td>0.150</td>
<td>0.400</td>
<td>-0.400</td>
<td>0.050</td>
<td>-0.025</td>
<td>0.425</td>
<td>0.125</td>
<td>-0.375</td>
<td>0.125</td>
<td>-0.375</td>
<td>2.150</td>
<td>-3.100</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>12</td>
<td>8</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td>13</td>
<td>4</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

The next step is to successively add location terms to the model in the order of their ranks. For each location model, model the squared residuals to the design matrix and rank their dispersion coefficients. Similar to the ranks of the location terms, the ranks of the dispersion terms will determine the order when adding the terms into the dispersion model. For each location model – dispersion model combination, the -2L was computed.
The table below summarizes the -2L values.

<table>
<thead>
<tr>
<th>Location Terms</th>
<th>Dispersion Terms</th>
<th>Model Notation</th>
<th>-2L</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>None</td>
<td>(0,0)</td>
<td>66.76539</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>(0,1)</td>
<td>55.04781</td>
</tr>
<tr>
<td></td>
<td>+ABCD</td>
<td>(0,2)</td>
<td>49.19817</td>
</tr>
<tr>
<td>ABCD</td>
<td>None</td>
<td>(1,0)</td>
<td>50.7952</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>(1,1)</td>
<td>50.13683</td>
</tr>
<tr>
<td></td>
<td>+BC</td>
<td>(1,2)</td>
<td>48.74896</td>
</tr>
<tr>
<td>+BCD</td>
<td>None</td>
<td>(2,0)</td>
<td>22.68635</td>
</tr>
<tr>
<td></td>
<td>ABCD</td>
<td>(2,1)</td>
<td>8.430096</td>
</tr>
<tr>
<td></td>
<td>+AB</td>
<td>(2,2)</td>
<td>7.35689</td>
</tr>
<tr>
<td>+BD</td>
<td>None</td>
<td>(3,0)</td>
<td>19.37764</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>(3,1)</td>
<td>15.35374</td>
</tr>
<tr>
<td></td>
<td>+AC</td>
<td>(3,2)</td>
<td>7.119332</td>
</tr>
</tbody>
</table>

Table 9. -2L values for model building using Pan and Team (2002) data

The following are the steps to determine the ‘best’ model for this data (at 95th quantile):

1. Starting with (0,0), add the location effect with the highest rank (i.e. ABCD). After adding 1 location effect, we are now at (1,0).

2. Compute the -2L difference of model (0,0) and model (1,0), which is computed as:

\[
66.76539 - 50.7952 = 15.97019
\]

3. Determine if difference is significant by looking at the critical values table and compare it with is compared with the cell that intersects with row (0,0) and column (1,0). This value in the Table 9 is 10.18.
3. If the computed -2L value is greater than the critical value, then we can say that the model with 1 location term is significantly better than no location term. Since the computed difference (15.97) is greater than the critical value, proceed by adding more location terms.

4. By adding 1 more location term with the next highest rank (i.e. BCD), we are now at model (2,0). The -2L difference between (1,0) and (2,0) is 50.7952-22.68635=28.10885.

5. The critical value from (1,0) to (2,0) is 8.34. The computed -2L difference (28.1085) is greater than this critical value. So we can say that adding BCD in the model is justified.

6. Add 1 more location term with the next highest rank (i.e. BD) and we are now at model (3,0). The -2L difference between (2,0) and (3,0) is 3.30871.

7. The critical value from (2,0) to (3,0) is 7.94. Since the computed -2L difference (3.30871) is less than the critical value, we can say that adding a third location term does not significantly make a better model. We now stop at adding location terms at (2,0) and location model is \( y \sim ABCD + BCD \).

9. Fit the squared residuals of the location model to come up with coefficient estimates for the dispersion model, and rank these coefficients. Priority in adding the dispersion term would depend on the rank. The order of the dispersion terms based on their coefficient estimate is shown below (from lowest to highest):
CD,AC,ABC,BC,AD,C,B,BD,ACD,D,A,BCD,ABD,AB,ABCD

10. We then add the highest ranked dispersion effect (i.e. ABCD) to (2,0) and we are now at (2,1). The -2L difference from (2,0) to (2,1) is 22.68635-8.430096=14.25625. The critical value for the comparison of this model pair is 13.26. Since the computed difference is greater than the critical value, then we proceed by adding 1 more dispersion term.

11. The next dispersion model in the ordered list is AB, and we are now at model (2,2). The -2L difference between (2,1) and (2,2) is 8.430096 –7.35689=1.073206 and the critical value is 9.85. Since the -2L difference is not greater than the critical value, then we stop.

12. Using the table, the final suggested joint model is:

Location model \( y \sim ABCD + BCD \) and
Dispersion model \( \sim ABCD \)

This model is the same as the model suggested by Box and Meyer (1986). In that paper, test statistics and graphical tools were used in combination to arrive at the model that we came up with.

**Example 2.** Let’s look at another example using the injection molding experiment data discussed in McGrath (2003). The experiment is a \( 2^{7-4} \) fractional factorial with design generators
E=ABC, G=ACD and F=BCD. Therefore, there are $2^4 = 16$ terms confounded with each other. The data set, coefficient estimates and their ranks is shown in Table 10 below, while the half-normal plot is shown in Fig 10.

<table>
<thead>
<tr>
<th>yi</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>AB</th>
<th>AC</th>
<th>AD</th>
<th>BC</th>
<th>BD</th>
<th>CD</th>
<th>E=ABC</th>
<th>ABD</th>
<th>G=ACD</th>
<th>F=BCD</th>
<th>ABCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>26</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>34</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>12</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>16</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>15</td>
<td>37</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>16</td>
<td>52</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

$\hat{\beta}_j$ | 6.94 | 17.81 | -0.44 | .69  | 5.93 | -0.81 | -2.69 | -0.94 | -0.06 | -0.06 | .19  | .06  | -2.44 | .19  | .31  |

**Table 10.** The Injection Molding data from McGrath (2003).
1. Add the highest ranked location term (i.e. B) in the location model. For location model $y^{\sim B}$ with no dispersion, the $-2L$ is 118.92. This is our model $(0,0)$.

2. Add the next highest ranked location term (A). For location model $y^{\sim A+B}$ with no dispersion, the $-2L$ is 108.2526. This is our model $(1,0)$. The difference is $118.92 - 108.2526 = 10.6674$. At 95th quantile, this difference is greater than the critical value 10.18.
3. Add AB to the location model. For location model $y \sim A + B + AB$ with no dispersion, the $-2\text{L}$ is 89.30779. This model is (3,0), and its $-2\text{L}$ difference from (2,0) is $108.2526 - 89.30779 = 18.94481$. This is greater than the critical value 7.94, therefore we keep this term in the model.

4. Add AD to the location model. For location model $y \sim A + B + AB + AD$ with no dispersion, the $-2\text{L}$ is 79.31274. This model is (4,0), and its $-2\text{L}$ difference from (3,0) is $89.30779 - 79.31274 = 9.99505$. This is greater than the critical value 8.03, therefore we keep this term in the model.

5. Add G to the location model. For location model $y \sim A + B + AB + AD + G$ with no dispersion, the $-2\text{L}$ is 79.31274. This model is (5,0), and its $-2\text{L}$ difference from (4,0) is $79.31274 - 59.29854 = 20.0142$. This is greater than the critical value 8.41, therefore we keep this term in the model.

6. Add BC to the location model. For location model $y \sim A + B + AB + AD + G + BC$ with no dispersion, the $-2\text{L}$ is 51.93509. This model is (6,0), and its $-2\text{L}$ difference from (5,0) is $59.29854 - 51.93509 = 7.36345$. This is now less than the critical value 9.06, therefore we do not keep this term in the model and stop at $y \sim A + B + AB + AD + G$. 
7. Our final location model is $y \sim A + B + AB + AD + G$. With this location model, below are the order of the absolute value of coefficient estimates for the dispersion terms (from lowest to highest):

$$G, C, E, BD, B, D, BC, AD, ABD, AC, A, CD, ABCD, F, AB$$

8. Adding $AB$ to the dispersion model, the joint model is now $y \sim A + B + AB + AD + G$ for location and $\sim AB$ for dispersion or $(5,1)$ in our model notation. The -2L value for this joint model is 55.72745, and the -2L difference of this model from $(5,0)$ is 59.29854–55.72745= 3.57109. This is less than the critical value 23.57. Therefore, the ‘best’ model that is significantly better than other models with this data set is

$$y \sim A + B + AB + AD + G$$ with no dispersion.

This final model is similar to the model suggested by McGrath(2003) where a dispersion-conditioned location effect test is employed.
6.2 MODEL COMPARISON

This approach is to be used to compare two different joint models.

**Example 1.** To demonstrate the use of this method, let's use the the Pan and Taam (2002) welding data in our earlier example by comparing the this paper’s findings with Nelder and Lee (1998) recommendation.

1. This paper’s recommended model:
   - Location model: \( y \sim ABCD + BCD \) and
   - Dispersion model: \( \sim ABCD \)
   - -2L value: 8.430096
   - Model notation: (2,1)

   - Location model: \( y \sim ABCD + BCD + ACD \)
   - Dispersion model: \( \sim ACD + ABCD \)
   - -2L value: -9.648669
   - Model notation: (3,2)

We will then subtract the -2L difference between this paper’s model (#1 above) and Nelder and Lee’s (1998) recommended model.

\[-2L\text{ difference} = 8.430096 - (-9.648689) = 18.07878\]
The critical value for models (2,1) and (3,2) is 22.28. Since the computed -2L difference is less than the critical value, we can say that the Nelder and Lee’s (1998) recommended model, with more terms, does not significantly reduce -2L.

Bergman and Hynen (1997) used the same welding data to suggest a model. To compare our model with their suggested joint model of y~BCD+ABCD and ~B+ACD+ABCD (or (2,3) in our model notation) would also be interesting. However, their model has 3 dispersion terms, and since the table we generated only has up to 2 dispersion terms, we cannot compare our model with Bergman and Hynen’s (1997) findings.

**Example 2.** Let’s compare our suggested model using the injection molding data in Table 11 with the findings suggested by Montgomery(1990).

1. This paper’s recommended model:
   Location model  \( y \sim \text{A+B+AB+AD+G} \)
   Dispersion model  none
   -2L value  :  59.29854
   Model notation:  (5,0)

   Location model  \( y \sim \text{A + B + AB} \)
Dispersioon model  ~ C

-2L value : 71.24946

Model notation: (3,1)

The Montgomery (1990) model has lesser terms, so to get the -2L difference, we should subtract (5,0) from (3,1).

-2L difference = 71.24946 – 59.29854 = 11.95092

The critical value for the comparison of these two models is 12.18. Since the -2L difference is less than the critical value, then we can say that our model, with more terms, does not significantly reduce -2L than Montgomery’s (1990) model at 95th quantile. It is interesting to note that the values we are comparing are close. At the 90th quantile, the critical value is at 10.36, and thus the conclusion would be reversed.

The interaction of the two extra terms in our model (AD and G=ACD) is C, which is the reported dispersion term in the Montgomery (1990) model. McGrath and Lin (2001) discussed that the failure to include a pair of location effect s (i.e. terms AD and ACD in our example) creates a spurious dispersion effect (i.e. term C in Montgomery’s (1990) model). The paper also discussed the reverse, that failure to account for a dispersion model (i.e term C in Montgomery (1990) model) would create two location effects, an example of which is our model where the terms AD and G=ACD were reported as location effects in the absence of a dispersion effect.
If we go back to the stepwise method for the injection molding data, we can see that when adding the fourth location term AD, the -2L difference computed is close to the critical value. Let’s re-do the stepwise method and stop at the third location term (ie. y^~A+B+AB) and proceed in determining the dispersion effects. The order (from lowest to highest) of the dispersion coefficient estimates is the following:

$$CD, B, G, A, BC, AD, AB, F, BD, AC, ABD, D, E, ABCD, C$$

1. We add C to the dispersion model, and our model is now at (3,1). The -2L value for this model is 71.24946. The -2L difference from (3,0) is

$$-2L \text{ difference} = 89.30779 - 71.24946 = 18.05833$$

This value is greater than the critical value of 15.80 so we can say keep this term in the dispersion model.

2. We add ABCD in the dispersion model, and our model is now at (3,2). The -2L value for this model is 65.67648 and the -2L difference is

$$-2L \text{ difference} = 71.24946 - 65.67648 = 5.57298.$$ 

This value is less than the critical value of 12.14, therefore we stop at 1 dispersion effect, which is C.

After re-doing the stepwise method to stop after the third location term, we came up with the same model recommended by Montgomery (1990).
7. MULTI-MODAL PATTERN FOR 3 DISPERSION TERMS OR MORE

The histogram generated for a model with 3 or more dispersion effects shows an unusual pattern. This starts when 3 location terms are in the model. As shown in Fig. 8, large differences start to accumulate in the right tail of the histogram for the -2L difference between no location terms and 3 location terms. This trend continues when more location terms and dispersion terms are added. For the model with 5 dispersion terms, small differences also start to accumulate on the left tail of the histogram. Please refer to Table B.4 of Appendix B to view the -2L difference histograms for models with 0 to 5 dispersion terms.

Pan (1999) studied the effects of unidentified location terms in the determination of dispersion terms. His paper tried to answer the question “How severe is the impact of unidentified location effects on dispersion-effects identification?”, and provides a mathematical explanation of this impact. In the paper, it is stated that the main problem lies in the fact that the empirical method of determining significant location effect leaves a lot of room for leaving out potentially significant terms. McGrath and Lin (2001) also discussed that mistakes can be made in identifying dispersion effects when active location effects are excluded. McGrath (2002) mentioned that the interaction of a pair of excluded active location effects can be identified as an active dispersion effect. An unpublished paper also discussed the possibility of multimodalities when there are more than 2 dispersion effects.

Analysis were done on the observations found on the right tail of the histogram of the -2L difference between (0,3) and (3,3). We filtered out those observations were the -2L difference is
more than 110, and there were 182 out of 10,000 that were in this area. The location terms, the corresponding dispersion terms and the estimates of the coefficient were listed to see if there is a pattern that emerge. Some of the common characteristics found are the following:

1. The interaction of two dispersion terms is the third dispersion term for all.
2. For all of the observations, either the interaction of 2 location terms is the third location term or the interaction of 3 location terms is one of the dispersion terms.
3. There are an unusually high absolute dispersion coefficient estimates. On the average, the absolute value of the dispersion coefficient estimates for these observations is about 17, considering that the expected estimate is 0.

Whether these observations are in line with the findings of the papers mentioned above needs further study. However, due to the unreliability of the data for the 3 or more dispersion terms case, we decided to stop generating the critical values table up to 2 dispersion terms only.
8. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE STUDY

In principle, a good design is always the design with the least terms in the model. Unreplicated fractional factorial experiments have been very useful as a tool to weed out factors that are not contributing significantly to the response. However, to date, there has not been a definitive method to determine the ‘best’ joint location and dispersion model. This study came up with a procedure on how to identify the best combination of location and dispersion terms that aims to replace the empirical way that was the most common method currently used. The critical values table that we came up with can be used to compare the -2L difference of two models. If the -2L difference is greater than the value in the table (depending on the number of location and dispersion terms and quantile), then we can say that the model with more terms is justified because it significantly reduces -2L. This is also true for the reverse. The objective is to make sure that when adding terms in the model, we can statistically be confident that the addition of the terms significantly improves the model.

Using several examples and comparison of our suggested model with the models recommended from different papers, we have illustrated the usefulness and validity of the results using the critical values table, not to mention ease of use.

As mentioned earlier in the paper, several authors have cautioned the consequence of either not including a location effect that is potentially significant, or adding an insignificant location term in the determination of the dispersion effects. By using the critical values table, there is now a tool that we can use to test statistically, if adding a location or dispersion term makes a
model better than another. The real advantage of using this table is more evident on factor screening experiments, especially experiments that are time consuming or expensive to conduct.

Minor problems have been encountered in coming up with the critical values table. The dglm package available in R, which is the package developed for double generalized linear modeling, faces convergence issues and some errors that was beyond the author’s knowledge on how to rectify. The processing time for the simulations may be faster and the R codes could have been much more readable if dglm worked that way it should.

Dispersion effects are generally considered to be secondary to location effects. In fact, homoscedasticity (i.e. equal variance) is often assumed in modeling. Generally, if there are active dispersion effects, they are few in number. So, though it may be uncommon to have 3 or more dispersion effects, it would be better if the critical values table generated by this study be expanded for more dispersion effects.

The unusual pattern of the resulting dispersion effects for 3 or more dispersion and location effects in a model is worth looking into and further analysis is recommended to be made. Several papers have theorized and explained in general terms the unusual pattern that was observed. However, so far, there is no study that states in specific terms the kind of pattern as what was mentioned in this study. It is not a far-fetched idea then to predict that the current method of identifying dispersion effects needs to be revised. Still, it is worth looking at.
REFERENCES


Appendix A – R Codes

Table A.1 R code for the three different algorithms to estimate coefficients and compute -2L and -2R

```r
# R PROGRAM : MethodsCode.txt
# DESCRIPTION : This program demonstrates 3 different methods
# of generalized linear modelling available
# 1. Iterative Regression Least Squares method
# 2. Using R optim function
# 3. Using R Dglm package
#
# PROGRAMMER : Rainier M. Sabangan
# DATE : February 5, 2010
#
# 2^4 design
A=rep(c(-1,1),8)
B=rep(c(-1,-1,1,1),4)
C=rep(c(-1,-1,-1,-1,1,1,1,1),2)
D=c(rep(-1,8),rep(1,8))
#A=rep(c(0,1),8)
#B=rep(c(0,0,1,1),4)
#C=rep(c(0,0,0,0,1,1,1,1),2)
#D=c(rep(0,8),rep(1,8))
AB=A*B; AC=A*C; BC=B*C; ABC=A*B*C; AD=A*D; BD=B*D; CD=C*D;
ABD=A*BD; ACD=A*CD; BCD=B*CD; ABCD=A*B*CD;
ABCD=-ABCD;ABD=-ABD;AD=-AD;ABC=-ABC;AC=-AC
mat=cbind(A,B,C,D,AB,AC,AD,BC,BD,CD,ABC,ABD,ACD,BCD,ABCD)
# A=1    B=2    C=3     D=4    AB=5   AC=6    AD=7    BC=8
# BD=9   CD=10  ABC=11  ABD=12 ACD=13 BCD=14  ABCD=15
# Convert -1's to 0
for(i in 1:16)
  { for(j in 1:15)
    { if (mat[i,j]==-1)
        { mat[i,j]=0 }
    }
  }
#
# WELDING EXPERIMENT DATA FROM PAN,G & TAAM, W (2002)
y=c(43.7,40.2,42.4,44.7,42.4,45.9,42.2,40.6,42.4,45.5,43.6,40.6,44,40.2,42.5,46.5)
#
# DEFINE THE LOCATION TERMS (bcols) & DISPERSION TERMS (dcols)
bcols=c(13,14,15)  # ACD, BCD & ABCD
dcols=c(13,15)     # ACD & ABCD

###############################################################
# GLM 1 - Iterative method #
###############################################################
IRLS=function()
#
# DETERMINE INITIAL LOCATION COEFFS
locmodel=glm(y~mat[,bcols])
mlcoeffs=locmodel$coeff
ressq=locmodel$residuals^2
#
# DETERMINE INITIAL DISPERSION COEFFS
dismodel=glm(ressq~mat[,dcols],family=Gamma(link = "log"))
mdcoeffs=dismodel$coeff
varfit=dismodel$fitted.values
newml=MLVAL(c(mlcoeffs,mdcoeffs))
oldml=999999
j=1
while (abs(oldml-newml) > 0.001)
  { locmodel=glm(y~mat[,bcols],weights=1/varfit)
    mlcoeffs=locmodel$coefficients
    oldml=newml
    newml=MLVAL(c(mlcoeffs,mdcoeffs))
  }
```

```r
# GLM 2 - Using optim function
#
# ML METHOD
method="ml"
# Set all initial coefficients to 0
out=optim(rep(0,7),MLVAL)
while (out$convergence==1)
  {out=optim(out$par,MLVAL)}
out

# Set all initial coefficients to 0
out=optim(rep(0,7),MLVAL)
while (out$convergence==1)
  {out=optim(out$par,MLVAL)}
out

# GLM3 - Using DGLM
#
library(dglm)
# ML Method
summary(dglm(y~mat[,bcols], ~mat[,dcols], dlink="log",method='ml'))
# REML Method
summary(dglm(y~mat[,bcols], ~mat[,dcols], dlink="log",method='reml'))
```
Table A.2 Program to compute for the -2L and -2R when random samples are obtained from a known true model with 3 location terms and 2 dispersion terms

```r
# R PROGRAM : TrueModel.txt
# DESCRIPTION : This iteratively does the following
# 1. Generate 16 random sample from
#    a population with
#    Loc Model : y ~ 43 + 2A + 3B + 5C and
#    Disp Model: ~ -3 - 4A + 2C
# 2. Computes the -2L and -2R when active terms from the
#    true model are retained, when a loc or disp term is added or #
#    removed
# 3 Determines the coefficients that arrived at the -2L or -2R value
# 4. Plots and histogram the information
#
# PROGRAMMER : Rainier M. Sabangan
# DATE : February 10, 2010
#

date()
stdt=date()
MLVAL=function(hats)
{ num_bcoefs=length(bcols)+1
  num_dcoefs=length(dcols)+1
  terms=c(1,bcols,1,dcols)
  mui=hats[1]
  start_indx=2
  end_indx=num_bcoefs
  for(i in start_indx:end_indx)
  |
    mui=mui+(hats[i]*mat[,terms[i]])
  phi=exp(hats[num_bcoefs+1])
  start_indx=num_bcoefs+2
  end_indx=num_bcoefs+num_dcoefs
  for(i in start_indx:end_indx)
  |
    phi=phi*exp(hats[i]*mat[,terms[i]])
  neg2logl= sum( ((y-mui)^2)/phi +log(2*pi*phi))
  neg2logl }

REMLVAL=function(hats)
{ num_bcoefs=length(bcols)+1
  num_dcoefs=length(dcols)+1
  terms=c(1,bcols,1,dcols)
  effectmat=cbind(rep(1,16),mat[,bcols])
  mui=hats[1]
  start_indx=2
  end_indx=num_bcoefs
  for(i in start_indx:end_indx)
  |
    mui=mui+(hats[i]*mat[,terms[i]])
  phi=exp(hats[num_bcoefs+1])
  start_indx=num_bcoefs+2
  end_indx=num_bcoefs+num_dcoefs
  for(i in start_indx:end_indx)
  |
    phi=phi*exp(hats[i]*mat[,terms[i]])
  neg2logl= sum( ((y-mui)^2)/phi +log(2*pi*phi))
  phiimat=diag(phi)
  term=(t(effectmat) %% solve(phiimat) %% effectmat) / (2*pi)
  neg2logr=neg2logl + log(det(term))
  neg2logr }

IRLS=function(y,bcols,dcols,cdiff,maxit)
# --- This is the IRLS function. Parameters are:
#  y    - response observations
#  bcols - factors in the location model
#  dcols - factors in the dispersion model
#  cdiff - maximum difference between coefficients to converge
```
# maxit - maximum number of iterations
#
{ # ---- Initial location model coefficients ---#
    locmodel=glm(y~mat[,bcols])
    mlcoefs=locmodel$coef
    res=locmodel$residuals
    ressq=res^2
    # determine initial dispersion model coefficients
    dispmodel=glm(ressq~mat[,dcols],family=quasi(link = "log")
                  ,control=list(epsilon=0.00001,maxit=200,trace = F))
    mdcoefs=dispmodel$coef
    ressqfit=dispmodel$fitted.values
    varfit=ressqfit
    weights=1/varfit
    coeffs1=c(mlcoefs,mdcoefs)
    coeffs2=c(mlcoefs,mdcoefs)
    end=0
    i=0
    while (end==0)
    {
        i=i+1
        locmodel=glm(y~mat[,bcols],weights=weights)
        mlcoefs=locmodel$coefficients
        res=locmodel$residuals
        ressq=res^2
        dispmodel=glm(ressq~mat[,dcols],family=quasi(link = "log")
                      ,control=list(epsilon=0.00001,maxit=200,trace = F))
        mdcoefs=dispmodel$coef
        coeffs1=coeffs2
        coeffs2=c(mlcoefs,mdcoefs)
        ressqfit=dispmodel$fitted.values
        varfit=ressqfit
        weights=1/varfit
        end=1
        maxdiff=max(abs(coeffs1-coeffs2))
        if (maxdiff>cdiff)  { end=0 }
        if (i==maxit)       { end=1 }
        c(coeffs2,i)
    }
}
# 2^4 design
A=rep(c(0,1),8)
B=rep(c(0,0,1,1),4)
C=rep(c(0,0,0,0,1,1,1,1),2)
D=c(rep(0,8),rep(1,8))
mat=cbind(A,B,C,D)
numit=1
# True Model
Loc Model: y ~ 43 - 2A + 3B - 5C
Disp: ~ -3 - 4A + 2C
sd = sqrt(exp(-3 - (4*mat[,1]) + (2*mat[,3])))
y = 43 - (2*mat[,1]) + (3*mat[,2]) - (5*mat[,3])
y = matrix(0,16,numit)
for (i in 1:numit)
  { y[i] = y + rnorm(16,0,sd) }
# 2^4 design
A=rep(c(0,1),8)
B=rep(c(0,0,1,1),4)
C=rep(c(0,0,0,0,1,1,1,1),2)
D=c(rep(0,8),rep(1,8))
mat=cbind(A,B,C,D)
# True Model
Loc Model: y ~ 43 - 2A + 3B - 5C
Disp: ~ -3 - 4A + 2C
numit=10000

# 1 Correct Model
bcols=c(1,2,3)
dcols=c(1,3)
bhats1b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats1m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats1r = matrix(0, length(bcols) + length(dcols) + 2, numit)

#---------------------------
# 2 Add D to location model
#---------------------------
bcols = c(1, 2, 3, 4)
dcols = c(1, 3)
bhats2b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats2m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats2r = matrix(0, length(bcols) + length(dcols) + 2, numit)

#----------------------------
# 3 Add B to dispersion model
#----------------------------
bcols = c(1, 2, 3)
dcols = c(1, 2, 3)
bhats3b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats3m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats3r = matrix(0, length(bcols) + length(dcols) + 2, numit)

#--------------------------------------------------------
# 4 Add D to location model and add B to dispersion model
#--------------------------------------------------------
bcols = c(1, 2, 3, 4)
dcols = c(1, 2, 3)
bhats4b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats4m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats4r = matrix(0, length(bcols) + length(dcols) + 2, numit)

#-------------------------------
# 5 Remove C from location model
#-------------------------------
bcols = c(1, 2)
dcols = c(1, 3)
bhats5b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats5m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats5r = matrix(0, length(bcols) + length(dcols) + 2, numit)

#---------------------------------
# 6 Remove A from dispersion model
#---------------------------------
bcols = c(1, 2, 3)
dcols = c(3)
bhats6b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats6m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats6r = matrix(0, length(bcols) + length(dcols) + 2, numit)

#-----------------------------------------------------------
# 7 Remove C from location model and A from dispersion model
#-----------------------------------------------------------
bcols = c(1, 2)
dcols = c(3)
bhats7b = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats7m = matrix(0, length(bcols) + length(dcols) + 2, numit)
bhats7r = matrix(0, length(bcols) + length(dcols) + 2, numit)

mlmatb = matrix(0, 7, numit)
mlmata = matrix(0, 7, numit)
remlmatb = matrix(0, 7, numit)
remlmata = matrix(0, 7, numit)
conv = matrix(0, 7, numit)

beg = 1

# True Model     Loc Model: y ~ 43 - 2A + 3B - 5C   Disp: ~ -3 - 4A + 2C     #
# 1 Correct Model

date()
for (sim in beg:numit)
{
  y=ys[,sim]
bcols=c(1,2,3)
dcols=c(1,3)

dataout=IRLS(y,bcols,dcols,0.001,100)
bhats1b[,sim]=dataout[1:length(dataout)-1]
conv[1,sim]=dataout[length(dataout)]

mlmatb[1,sim]=MLVAL(bhats1b[,sim])
mlout=optim(bhats1b[,sim],MLVAL)
mlmata[1,sim]=mlout$value
bhats1m[,sim]=mlout$par

remlmatb[1,sim]=REMLVAL(bhats1b[,sim])
remlout=optim(bhats1b[,sim],REMLVAL)
remlmata[1,sim]=remlout$value
bhats1r[,sim]=remlout$par
}
date()

# 2 Add D to location model

date()
for (sim in beg:numit)
{
  y=ys[,sim]
bcols=c(1,2,3,4)
dcols=c(1,3)

dataout=IRLS(y,bcols,dcols,0.001,100)
bhats2b[,sim]=dataout[1:length(dataout)-1]
conv[2,sim]=dataout[length(dataout)]

mlmatb[2,sim]=MLVAL(bhats2b[,sim])
mlout=optim(bhats2b[,sim],MLVAL)
mlmata[2,sim]=mlout$value
bhats2m[,sim]=mlout$par

remlmatb[2,sim]=REMLVAL(bhats2b[,sim])
remlout=optim(bhats2b[,sim],REMLVAL)
remlmata[2,sim]=remlout$value
bhats2r[,sim]=remlout$par
}
date()

# 3 Add B to dispersion model

date()
for (sim in beg:numit)
{
  y=ys[,sim]
bcols=c(1,2,3)
dcols=c(1,2,3)

dataout=IRLS(y,bcols,dcols,0.001,100)
bhats3b[,sim]=dataout[1:length(dataout)-1]
conv[3,sim]=dataout[length(dataout)]

mlmatb[3,sim]=MLVAL(bhats3b[,sim])
mlout=optim(bhats3b[,sim],MLVAL)
mlmata[3,sim]=mlout$value
bhats3m[,sim]=mlout$par

remlmatb[3,sim]=REMLVAL(bhats3b[,sim])
remlout=optim(bhats3b[,sim],REMLVAL)
remlmata[3,sim]=remlout$value
bhats3r[,sim]=remlout$par
date()
#--------------------------------------------------------
# 4 Add D to location model and add B to dispersion model
#--------------------------------------------------------
for (sim in beg:numit) {
  y=ys[,sim]
bcols=c(1,2,3,4)
dcols=c(1,2,3)

  dataout=IRLS(y,bcols,dcols,0.001,100)
  bhats4b[,sim]=dataout[1:length(dataout)-1]
  conv[4,sim]=dataout[length(dataout)]
  mlmatb[4,sim]=MLVAL(bhats4b[,sim])
  mlout=optim(bhats4b[,sim],MLVAL)
  mlmata[4,sim]=mlout$value
  bhats4m[,sim]=mlout$par
  remlmatb[4,sim]=REMLVAL(bhats4b[,sim])
  remlout=optim(bhats4b[,sim],REMLVAL)
  remlmata[4,sim]=remlout$value
  bhats4r[,sim]=remlout$par
}
date()
#-------------------------------
# 5 Remove C from location model
#-------------------------------
for (sim in beg:numit) {
  y=ys[,sim]
bcols=c(1,2)
dcols=c(1,3)

  dataout=IRLS(y,bcols,dcols,0.001,100)
  bhats5b[,sim]=dataout[1:length(dataout)-1]
  conv[5,sim]=dataout[length(dataout)]
  mlmatb[5,sim]=MLVAL(bhats5b[,sim])
  mlout=optim(bhats5b[,sim],MLVAL)
  mlmata[5,sim]=mlout$value
  bhats5m[,sim]=mlout$par
  remlmatb[5,sim]=REMLVAL(bhats5b[,sim])
  remlout=optim(bhats5b[,sim],REMLVAL)
  remlmata[5,sim]=remlout$value
  bhats5r[,sim]=remlout$par
}
date()
#---------------------------------
# 6 Remove A from dispersion model
#---------------------------------
for (sim in beg:numit) {
  y=ys[,sim]
bcols=c(1,2,3)
dcols=c(3)

  dataout=IRLS(y,bcols,dcols,0.001,100)
  bhats6b[,sim]=dataout[1:length(dataout)-1]
  conv[6,sim]=dataout[length(dataout)]
  mlmatb[6,sim]=MLVAL(bhats6b[,sim])
  mlout=optim(bhats6b[,sim],MLVAL)
  mlmata[6,sim]=mlout$value
  bhats6m[,sim]=mlout$par
  remlmatb[6,sim]=REMLVAL(bhats6b[,sim])
  remlout=optim(bhats6b[,sim],REMLVAL)
  remlmata[6,sim]=remlout$value
bhats6r[,sim]=remlout$par
}
date()
#-----------------------------------------------------------
# 7 Remove C from location model and A from dispersion model
#-----------------------------------------------------------
for (sim in beg:numit)
{
y=ys[,sim]
bcols=c(1,2)
dcols=c(3)
dataout=IRLS(y,bcols,dcols,0.001,100)
bhats7b[,sim]=dataout[1:length(dataout)-1]
conv[7,sim]=dataout[length(dataout)]
mlmatb[7,sim]=MLVAL(bhats7b[,sim])
mlout=optim(bhats7b[,sim],MLVAL)
mlmata[7,sim]=mlout$value
bhats7m[,sim]=mlout$par
remlmatb[7,sim]=REMLVAL(bhats7b[,sim])
remlout=optim(bhats7b[,sim],REMLVAL)
remlmata[7,sim]=remlout$value
bhats7r[,sim]=remlout$par
}
date()
enddt=date()

# HISTOGRAM OF -2L DIFFERENCES FROM THE TRUE MODEL
par(mfrow=c(2,3))
hist(mlmata[1,]-mlmata[2,],main="+D Loc",xlab="")
hist(mlmata[1,]-mlmata[3,],main="+B Disp",xlab="")
hist(mlmata[1,]-mlmata[4,],main="+D Loc,+B Disp",xlab="")
hist(mlmata[5,]-mlmata[1,],main="-C Loc",xlab="")
hist(mlmata[5,]-mlmata[1,],main="-A Disp",xlab="")
hist(mlmata[5,]-mlmata[1,],main="-C Loc,-A Disp",xlab="")

# HISTOGRAM OF -2R DIFFERENCES FROM THE TRUE MODEL
par(mfrow=c(2,3))
hist(remlmata[1,]-remlmata[2,],main="+D Loc")
hist(remlmata[1,]-remlmata[3,],main="+B Disp")
hist(remlmata[1,]-remlmata[4,],main="+D Loc,+B Disp")
hist(remlmata[5,]-remlmata[1,],main="-C Loc")
hist(remlmata[5,]-remlmata[1,],main="-A Disp")
hist(remlmata[5,]-remlmata[1,],main="-C Loc & -A Disp")

# PLOT OF COMPUTED -2L AND -2R
par(mfrow=c(2,4))
plot(mlmata[1,],remlmata[1,],main="True Model")
abline(1,0)
plot(mlmata[2,],remlmata[2,],main="+D Loc")
abline(1,0)
plot(mlmata[3,],remlmata[3,],main="+B Disp")
abline(1,0)
plot(mlmata[4,],remlmata[4,],main="+D Loc,+B Disp")
abline(1,0)
plot(mlmata[5,],remlmata[5,],main="-C Loc")
abline(1,0)
plot(mlmata[6,],remlmata[6,],main="-A Disp")
abline(1,0)
plot(mlmata[7,],remlmata[7,],main="-C Loc,-A Disp")
abline(1,0)

# MEAN OF COEFFICIENTS USING ML METHOD
round(apply(bhats1m,1,mean),6)
round(apply(bhats2m,1,mean),6)
round(apply(bhats3m,1,mean),6)
round(apply(bhats4m,1,mean),6)
round(apply(bhats5m,1,mean),6)
round(apply(bhats6m,1,mean),6)
round(apply(bhats7m,1,mean),6)

# MEAN OF COEFFICIENTS USING REML METHOD
round(apply(bhats1r,1,mean),6)
round(apply(bhats2r,1,mean),6)
round(apply(bhats3r,1,mean),6)
round(apply(bhats4r,1,mean),6)
round(apply(bhats5r,1,mean),6)
round(apply(bhats6r,1,mean),6)
round(apply(bhats7r,1,mean),6)
Table A.3 Program code to compute the -2L from a true model of (2,2) and models where terms are added to the location and dispersion model.

```r
R PROGRAM : True2-2.txt
DESCRIPTION : This program does the following:
1. Generate 16 random sample from a population (2,2)
   (i.e) no loc and disp effects
   Loc Model : y ~ C+D
   Disp Model: ~ A+B
2. Computes the -2L
3. Adds terms to the model and compute the -2L for each
   model with added terms
4. Histogram the information
5. Display means, sd and quantile for each pair of the
   -2L differences
PROGRAMMER : Rainier M. Sabangan
DATE : February 10, 2010

date()
stdt=date()
MLVAL=function(hats)
{ num_bcoefs=length(bcols)+1
  num_dcoefs=length(dcols)+1
  terms=c(1,bcols,1,dcols)
  mui=hats[1]
  start_indx=2
  end_indx=num_bcoefs
  if (start_indx<=end_indx)
  {for (i in start_indx:end_indx)
   { mui=mui+(hats[i]*mat[,terms[i]]) }}
  phi=exp(hats[num_bcoefs+1])
  start_indx=num_bcoefs+2
  end_indx=num_bcoefs+num_dcoefs
  if (start_indx<=end_indx)
  {for (i in start_indx:end_indx)
   {phi=phi*exp(hats[i]*mat[,terms[i]])}}
  neg2logl= sum( ((y-mui)^2)/phi +log(2*pi*phi))
  neg2logl}

MLVALO=function(hats)
{ num_bcoefs=length(bcols)+1
  terms=c(1,bcols,1)
  mui=hats[1]
  start_indx=2
  end_indx=num_bcoefs
  if (start_indx<=end_indx)
  {for (i in start_indx:end_indx)
   { mui=mui+(hats[i]*mat[,terms[i]]) }}
  phi=rep(sigmasq,numobs)
  neg2logl= sum( ((y-mui)^2)/phi +log(2*pi*phi))
  neg2logl}

#2^4 design
numobs=16
A=rep(c(-1,1),8)
B=rep(c(-1,-1,1,1),4)
C=rep(c(-1,-1,-1,-1,1,1,1,1),2)
D=rep(c(-1,1,-1,1,1,1,1,1),2)
AB=A*B; AC=A*C; BC=B*C; ABC=A*B*C
AD=A*D; BD=B*D; CD=C*D; ABD=A*BD; ACD=A*C; BCD=B*CD; ABCD=A*B*CD
mat=cbind(A,B,C,D,AB,AC,AD,BC,BD,CD,ABC,ABD,ACD,BCD,ABCD)
dim_mat=dim(mat)
for(i in 1:dim_mat[1])
  for(j in 1:dim_mat[2])
    if(mat[i,j]==-1)
      { mat[i,j]=0 } } }
```
numobs=16
numit=10000
ys=matrix(0,numobs,numit)
# True Model       Loc Model: y ~ C + D  Disp: ~ A + B
#--------------------------------------------------------------------------------#
# True Model       Loc Model: y ~ C + D  Disp: ~ A + B
#--------------------------------------------------------------------------------#
bcols=c(3,4)
dcols=c(1,2)
sd = sqrt(exp(mat[,1]+mat[,2]))
for (i in 1:numit)
{ ys[,i] = mat[,3] + mat[,4] + rnorm(numobs,0,sd) }
mlmata  =matrix(0,20,numit)
conv    =matrix(0,20,numit)
beg=1
# True Model       Loc Model: y ~ C + D  Disp: ~ A + B
#--------------------------------------------------------------------------------#
# True Model       Loc Model: y ~ C + D  Disp: ~ A + B
#--------------------------------------------------------------------------------#
date()
#-------------------------------
#1. (2,2)
#-------------------------------
bcols=c(3,4)
dcols=c(1,2)
p=length(bcols)+length(dcols)+2
bhats1a =matrix(0,length(bcols)+length(dcols)+2,numit)
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  mlmata[1,sim]=mlout$value
  bhats1a[,sim]=mlout$par
  }
}
date()
#-------------------------------
#2. (3,2)
#-------------------------------
bcols=c(2,3,4)
dcols=c(1,2)
p=length(bcols)+length(dcols)+2
bhats2a =matrix(0,length(bcols)+length(dcols)+2,numit)
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  mlmata[2,sim]=mlout$value
  bhats2a[,sim]=mlout$par
  }
}
date()
#-------------------------------
#3. (2,3)
#-------------------------------
bcols=c(3,4)
dcols=c(1,2,3)
bhats3a =matrix(0,length(bcols)+length(dcols)+2,numit)
p=length(bcols)+length(dcols)+2
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  mlmata[3,sim]=mlout$value
  bhats3a[,sim]=mlout$par
  }
}
while (!mlout$convergence==0)
    {j=j+1
     mlout=optim(mlout$par,MLVAL,method="BFGS")
     mlmata[3,sim]=mlout$value
     bhats3a[,sim]=mlout$par
    }
}
date()
#-------------------------------
#4. (3,3)
#-------------------------------
bcols=c(2,3,4)
dcols=c(1,2,3)
bhats4a =matrix(0,length(bcols)+length(dcols)+2,numit)
p=length(bcols)+length(dcols)+2
for (sim in beg:numit)
    { y=ys[,sim]
      mlout=optim(rep(0,p),MLVAL,method="BFGS")
      j=1
      while (!mlout$convergence==0)
          {j=j+1
           mlout=optim(mlout$par,MLVAL,method="BFGS")
           mlmata[4,sim]=mlout$value
           bhats4a[,sim]=mlout$par
          }
    }

date()
#-------------------------------
#5. (4,2)
#-------------------------------
bcols=c(1,2,3,4)
dcols=c(1,2)
p=length(bcols)+length(dcols)+2
bhats5a =matrix(0,length(bcols)+length(dcols)+2,numit)
for (sim in beg:numit)
    { y=ys[,sim]
      mlout=optim(rep(0,p),MLVAL,method="BFGS")
      j=1
      while (!mlout$convergence==0)
          {j=j+1
           mlout=optim(mlout$par,MLVAL,method="BFGS")
           mlmata[5,sim]=mlout$value
           bhats5a[,sim]=mlout$par
          }
    }

date()
#-------------------------------
#6. (4,3)
#-------------------------------
bcols=c(1,2,3,4)
dcols=c(1,2,3)
p=length(bcols)+length(dcols)+2
bhats6a =matrix(0,length(bcols)+length(dcols)+2,numit)
for (sim in beg:numit)
    { y=ys[,sim]
      mlout=optim(rep(0,p),MLVAL,method="BFGS")
      j=1
      while (!mlout$convergence==0)
          {j=j+1
           mlout=optim(mlout$par,MLVAL,method="BFGS")
           mlmata[6,sim]=mlout$value
           bhats6a[,sim]=mlout$par
          }
    }

date()
#-------------------------------
#7. (4,4)
#-------------------------------
conv7=matrix(0,numit)
bcols=c(1,2,3,4)
dcols=c(1,2,3,4)
bhats7a =matrix(0,length(bcols)+length(dcols)+2,numit)
p=length(bcols)+length(dcols)+2
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  }
  conv7[sim]=j
  mlmata[7,sim]=mlout$value
  bhats7a[,sim]=mlout$par
}
date()
#-------------------------------
#8. (3,4)
#-------------------------------
bcols=c(2,3,4)
dcols=c(1,2,3,4)
bhats8a =matrix(0,length(bcols)+length(dcols)+2,numit)
p=length(bcols)+length(dcols)+2
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  }
  mlmata[8,sim]=mlout$value
  bhats8a[,sim]=mlout$par
}
date()
#-------------------------------
#9. (2,4)
#-------------------------------
bcols=c(3,4)
dcols=c(1,2,3,4)
bhats9a =matrix(0,length(bcols)+length(dcols)+2,numit)
p=length(bcols)+length(dcols)+2
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  }
  mlmata[9,sim]=mlout$value
  bhats9a[,sim]=mlout$par
}
date()
#-------------------------------
#10. (5,2)
#-------------------------------
bcols=c(1,2,3,4,5)
dcols=c(1,2)
bhats10a =matrix(0,length(bcols)+length(dcols)+2,numit)
p=length(bcols)+length(dcols)+2
for (sim in beg:numit)
{ y=ys[,sim]
  mlout=optim(rep(0,p),MLVAL,method="BFGS")
  j=1
  while (!mlout$convergence==0)
  {j=j+1
   mlout=optim(mlout$par,MLVAL,method="BFGS")
  }
  mlmata[10,sim]=mlout$value
  bhats10a[,sim]=mlout$par
}
date()
par(mfrow=c(3,3))
hist(mlmata[1,]-mlmata[2,],main="(2,2) vs (3,2)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[3,],main="(2,2) vs (2,3)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[4,],main="(2,2) vs (3,3)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[5,],main="(2,2) vs (4,2)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[6,],main="(2,2) vs (4,3)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[7,],main="(2,2) vs (4,4)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[8,],main="(2,2) vs (3,4)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[9,],main="(2,2) vs (2,4)",xlab="","ylab="");
hist(mlmata[1,]-mlmata[10,],main="(2,2) vs (5,2)",xlab="","ylab="");

quants=c(quantile(mlmata[1,]-mlmata[2,],0.95),quantile(mlmata[1,]-mlmata[3,],0.95),
         quantile(mlmata[1,]-mlmata[4,],0.95),quantile(mlmata[1,]-mlmata[5,],0.95),
         quantile(mlmata[1,]-mlmata[6,],0.95),quantile(mlmata[1,]-mlmata[7,],0.95),
         quantile(mlmata[1,]-mlmata[8,],0.95),quantile(mlmata[1,]-mlmata[9,],0.95),
         quantile(mlmata[1,]-mlmata[10,],0.95))
means=c(mean(mlmata[1,]-mlmata[2,]),mean(mlmata[1,]-mlmata[3,]),
         mean(mlmata[1,]-mlmata[4,]),mean(mlmata[1,]-mlmata[5,]),
         mean(mlmata[1,]-mlmata[6,]),mean(mlmata[1,]-mlmata[7,]),
         mean(mlmata[1,]-mlmata[8,]),mean(mlmata[1,]-mlmata[9,]),
         mean(mlmata[1,]-mlmata[10,]))
sds=c(sd(mlmata[1,]-mlmata[2,]),sd(mlmata[1,]-mlmata[3,]),
     sd(mlmata[1,]-mlmata[4,]),sd(mlmata[1,]-mlmata[5,]),
     sd(mlmata[1,]-mlmata[6,]),sd(mlmata[1,]-mlmata[7,]),
     sd(mlmata[1,]-mlmata[8,]),sd(mlmata[1,]-mlmata[9,]),
     sd(mlmata[1,]-mlmata[10,]))
cat(rbind(quants,means,sds),sep="","}
Table A.4 Program that computes for the -2L of a true model with no dispersion effect and the choice of additional location terms are determined by the rank of their estimated coefficients

```r
# DESCRIPTION: This program does the following:
# 1. Generates 16 random samples from std normal and will be considered
#    as the response observations \( y \)
# 2. Models the observations from the pre-defined design matrix &
#    ranks the coefficients
# 3. Iteratively determines the terms for the model with no loc, 1 loc,
#    2 loc, etc using the highest ranking determined at #2 and computes -2L
# PROGRAMMER: Rainier M. Sabangan
# DATE: February 5, 2010

# FUNCTION THAT RETURNS THE -2L
MLVAL0=function(y,bcols)
{ if (length(bcols)==0) {locmodel=lm(y~1)}
else                  {locmodel=lm(y~mat[,bcols])}
mlcoefs=locmodel$coefficients
sigmasq=summary(locmodel)$sigma^2 * (n-(length(bcols)+1))
num_bcoefs=length(bcols)+1
terms=c(1,bcols,1)
mui=mlcoefs
if (num_bcoefs>1)
  {for (i in 2:num_bcoefs)
   { mui=mui+(mlcoefs[i]*mat[,terms[i]]) }}
phii=rep(sigmasq,n)
neg2logl= sum( ((y-mui)^2)/phii +log(2*pi*phii))
neg2logl }

#CODE UNITS IN DESIGN MATRIX TO 0,1
for(i in 1:dim_mat[1])
  { for(j in 1:dim_mat[2])
    { if(mat[i,j]==-1)
      { mat[i,j]=0 } }}

library(dglm)
n=16                        #sample size
col=1:15                    #column number of the terms in the model
sz=50000                    #number of simulations per batch
btch=70                     #number of batches
quants=c(0.99,0.95,0.90,0.80)

quant.out=array(data=NA,dim=c(4,13,btch),dimnames=list(Quantiles=quants,
  Diff=c("0-1","0-2","0-3","0-4","0-5","0-6","0-7","1-2","2-3","3-4","4-5","5-6","6-7")))

# FUNCTION THAT RETURNS THE QUANTILES
Q=function(sims,quants)
{ mlmat0=matrix(NA,8,sims)
  #GEN"RATE y's
  #LOOP FOR 0 DISPERSION EFFECT
  for (i in 1:sims)
    { y=rnorm(16)
```

### MLVAL0 Function Explanation

```r
MLVAL0=function(y,bcols)
# If no location terms are specified (bcols == 0), model the response y
# against a constant term. Otherwise, model y against the design matrix.

# Extract coefficients and estimate sigma squared
mlcoefs = coef(locmodel)
sigmasq = summary(locmodel)$sigma^2 * (n - (length(bcols) + 1))

# Determine the terms to include in the model
num_bcoefs = length(bcols) + 1
terms = c(1, bcols, 1)

# Calculate mui (expected value)
mui = mlcoefs
if (num_bcoefs > 1)
  for (i in 2:num_bcoefs)
    mui = mui + (mlcoefs[i] * mat[, terms[i]])

# Calculate phii (dispersion parameter)
phii = rep(sigmasq, n)

# Calculate the negative log-likelihood
neg2logl = sum(((y - mui)^2)/phii + log(2*pi*phii))
```

### Code Units in Design Matrix

```r
# Convert code units in the design matrix to 0,1
for (i in 1:dim_mat[1])
  for (j in 1:dim_mat[2])
    if (mat[i,j] == -1)
      mat[i,j] = 0
```

### Loading dglm Package

```r
library(dglm)
```

### Simulation Parameters

```r
n = 16                        # Sample size
col = 1:15                    # Column number of terms in the model
sz = 50000                    # Number of simulations per batch
btch = 70                     # Number of batches
quants = c(0.99, 0.95, 0.90, 0.80)
```

### Calculating Quantiles

```r
quant.out = array(data = NA, dim = c(4, 13, btch),
  dimnames = list(Quantiles = quants,
    Diff = c("0-1", "0-2", "0-3", "0-4", "0-5", "0-6", "0-7", "1-2", "2-3", "3-4", "4-5", "5-6", "6-7")))
```

### Generating y's

```r
# Generate y's for simulations
for (i in 1:sims)
  y = rnorm(16)
```
lcoef=lm(y~mat)$coefficients[2:n]
lr=rank(abs(lcoef))
# (0,0)
bcols=c()
mlmat0[1,i]=MLVAL0(y,bcols)
# (1,0)
bcols=col[lr>=n-1]
mlmat0[2,i]=MLVAL0(y,bcols)
# (2,0)
bcols=col[lr>=n-2]
mlmat0[3,i]=MLVAL0(y,bcols)
# (3,0)
bcols=col[lr>=n-3]
mlmat0[4,i]=MLVAL0(y,bcols)
# (4,0)
bcols=col[lr>=n-4]
mlmat0[5,i]=MLVAL0(y,bcols)
# (5,0)
bcols=col[lr>=n-5]
mlmat0[6,i]=MLVAL0(y,bcols)
# (6,0)
bcols=col[lr>=n-6]
mlmat0[7,i]=MLVAL0(y,bcols)
# (7,0)
bcols=col[lr>=n-7]
mlmat0[8,i]=MLVAL0(y,bcols)

a01=quantile(mlmat0[1,]-mlmat0[2,],quants)
a02=quantile(mlmat0[1,]-mlmat0[3,],quants)
a03=quantile(mlmat0[1,]-mlmat0[4,],quants)
a04=quantile(mlmat0[1,]-mlmat0[5,],quants)
a05=quantile(mlmat0[1,]-mlmat0[6,],quants)
a06=quantile(mlmat0[1,]-mlmat0[7,],quants)
a07=quantile(mlmat0[1,]-mlmat0[8,],quants)
a12=quantile(mlmat0[2,]-mlmat0[3,],quants)
a23=quantile(mlmat0[3,]-mlmat0[4,],quants)
a34=quantile(mlmat0[4,]-mlmat0[5,],quants)
a45=quantile(mlmat0[5,]-mlmat0[6,],quants)
a56=quantile(mlmat0[6,]-mlmat0[7,],quants)
a67=quantile(mlmat0[7,]-mlmat0[8,],quants)
c(a01,a02,a03,a04,a05,a06,a07,a12,a23,a34,a45,a56,a67)

for (i in 1:btch)
(quant.out[,i]=Q(sz,quants))
Table A.5 Program that computes for the -2L of a true model with dispersion effects and the choice of additional location terms are determined by the rank of their estimated coefficients

<table>
<thead>
<tr>
<th>Function THAT RETURNS THE -2L</th>
</tr>
</thead>
<tbody>
<tr>
<td>GETDCOLS=function(noofdisp)</td>
</tr>
<tr>
<td>{ if (length(bcols)==0)  {res=lm(y~1)$residuals}</td>
</tr>
<tr>
<td>else                   {res=lm(y~mat[,bcols])$residuals}</td>
</tr>
<tr>
<td>dcoef=lm(log(res^2)-mat)$coefficients[2:16]</td>
</tr>
<tr>
<td>dr=rank(abs(dcoef))</td>
</tr>
<tr>
<td>col[dr&gt;=n-noofdisp]   }</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FUNCTION THAT RETURNS THE -2L</th>
</tr>
</thead>
<tbody>
<tr>
<td>GETML=function()</td>
</tr>
<tr>
<td>{p=length(bcols)+length(dcols)+2</td>
</tr>
<tr>
<td>mlout=optim(rep(0,p),MLVAL,method=&quot;BFGS&quot;)</td>
</tr>
<tr>
<td>while (!mlout$convergence==0)</td>
</tr>
<tr>
<td>{mlout=optim(mlout$par,MLVAL,method=&quot;BFGS&quot;)}</td>
</tr>
<tr>
<td>mlout$value }</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MLVAL=function(hats)</th>
</tr>
</thead>
<tbody>
<tr>
<td>{ num_bcoefs=length(bcols)+1</td>
</tr>
<tr>
<td>num_dcoefs=length(dcols)+1</td>
</tr>
<tr>
<td>#COMPUTE MUI</td>
</tr>
<tr>
<td>terms=c(1,bcols,1,dcols)</td>
</tr>
<tr>
<td>mui=hats[1]</td>
</tr>
<tr>
<td>start_indx=2</td>
</tr>
<tr>
<td>end_indx=num_bcoefs</td>
</tr>
<tr>
<td>if (start_indx==end_indx)</td>
</tr>
<tr>
<td>{for (i in start_indx:end_indx)</td>
</tr>
<tr>
<td>{mui=mui+(hats[i]*mat[,terms[i]]) }</td>
</tr>
<tr>
<td>#COMPUTE PHII</td>
</tr>
<tr>
<td>phii=exp(hats[num_bcoefs+1])</td>
</tr>
<tr>
<td>start_indx=num_bcoefs+2</td>
</tr>
<tr>
<td>end_indx=num_bcoefs+num_dcoefs</td>
</tr>
<tr>
<td>if (start_indx==end_indx)</td>
</tr>
<tr>
<td>{for (i in start_indx:end_indx)</td>
</tr>
<tr>
<td>{phii=phii*exp(hats[i]*mat[,terms[i]])}</td>
</tr>
<tr>
<td>neg2logl= sum( ((y-mui)^2)/phii +log(2<em>pi</em>phii) }</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q1=function(mlmat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>{ c(quantile(mlmat[1,]-mlmat[2,],quants),quantile(mlmat[1,]-mlmat[3,],quants),</td>
</tr>
<tr>
<td>quantile(mlmat[1,]-mlmat[4,],quants),quantile(mlmat[1,]-mlmat[5,],quants),</td>
</tr>
<tr>
<td>quantile(mlmat[1,]-mlmat[6,],quants),quantile(mlmat[1,]-mlmat[7,],quants),</td>
</tr>
<tr>
<td>quantile(mlmat[1,]-mlmat[8,],quants),quantile(mlmat[2,]-mlmat[3,],quants),</td>
</tr>
<tr>
<td>quantile(mlmat[3,]-mlmat[4,],quants),quantile(mlmat[4,]-mlmat[5,],quants),</td>
</tr>
<tr>
<td>quantile(mlmat[5,]-mlmat[6,],quants),quantile(mlmat[6,]-mlmat[7,],quants),</td>
</tr>
<tr>
<td>quantile(mlmat[7,]-mlmat[8,],quants)} }</td>
</tr>
</tbody>
</table>

A=rep(c(-1,1),8) |
B=rep(c(-1,-1,1,1),4) |
C=rep(c(-1,-1,-1,1,1,1,1,1),2) |
D = c(rep(-1,8),rep(1,8))
AB = A*B; AC = A*C; BC = B*C; ABC = A*B*C; AD = A*D; BD = B*D;
CD = C*D; ABD = A*B*D; ACD = A*C*D; BCD = B*C*D; ABCD = A*B*C*D
mat = cbind(A, B, C, D, AB, AC, AD, BC, BD, CD, ABD, ACD, BCD, ABCD)
dim_mat = dim(mat)
for(i in 1:dim_mat[1])
  for(j in 1:dim_mat[2])
    if(mat[i,j] == -1)
      mat[i,j] = 0

# DEFINE OBJECTS HERE
sims = 50000  # number of simulations
batn = 10      # number of batches
n = 16         # sample size
mlmat = matrix(NA, 8, sims)  # ml value matrix
col = 1:15     # column number of the terms in the model
quants = c(0.99, 0.95, 0.90, 0.80)  # quantiles we are interested with
        "BCD", "ABCD")

# MAIN ARRAY TO STORE THE QUANTILES FOR 1 DISPERSION MODELS
quant.out1 = array(data = NA, dim = c(4, 13, batn),
                     dimnames = list(Quantiles = c(0.99, 0.95, 0.90, 0.80),
                                     Diff = c("0-1", "0-2", "0-3", "0-4", "0-5", "0-6", "0-7", "1-2",
                                              "2-3", "3-4", "4-5", "5-6", "6-7")))

# LOOP FOR 1 DISPERSION EFFECT
noofdisp = 1
for( k in 1:batn)
  {for (i in 1:sims)
   {y = rnorm(16)
    lcoef = lm(y ~ mat)$coefficients[2:n]
    lr = rank(abs(lcoef))
    # (0,1)
    bcols = c()
    dcols = GETDCOLS(noofdisp)
    mlmat[1,i] = GETML()
    # (1,1)
    bcols = col[lr >= n-1]
    dcols = GETDCOLS(noofdisp)
    mlmat[2,i] = GETML()
    # (2,1)
    bcols = col[lr >= n-2]
    dcols = GETDCOLS(noofdisp)
    mlmat[3,i] = GETML()
    # (3,1)
    bcols = col[lr >= n-3]
    dcols = GETDCOLS(noofdisp)
    mlmat[4,i] = GETML()
    # (4,1)
    bcols = col[lr >= n-4]
    dcols = GETDCOLS(noofdisp)
    mlmat[5,i] = GETML()
    # (5,1)
    bcols = col[lr >= n-5]
    dcols = GETDCOLS(noofdisp)
    mlmat[6,i] = GETML()
    # (6,1)
    bcols = col[lr >= n-6]
    dcols = GETDCOLS(noofdisp)
    mlmat[7,i] = GETML()
    # (7,1)
    bcols = col[lr >= n-7]
    dcols = GETDCOLS(noofdisp)
    mlmat[8,i] = GETML()
  }
  quant.out1[,k] = Q1(mlmat)
  print(paste(k, date()))
}
#MAIN ARRAY TO STORE THE QUANTILES FOR 2 DISPERSION MODELS
quant.out2=array(data=NA,dim=c(4,13,batn),
dimnames=list(Quantiles=c(0.99,0.95,0.90,0.80),
Diff=c("0-1","0-2","0-3","0-4","0-5","0-6","0-7","1-2",
      "2-3","3-4","4-5","5-6","6-7")))

#LOOP FOR 2 DISPERSION EFFECTS
noofdisp=2
for (k in 1:batn)
  {for (i in 1:sims)
    {y=rnorm(16)
lcoef=lm(y~mat)$coefficients[2:n]
    #(0,1)
bcols=c()
dcols=GETDCOLS(noofdisp)
    mmat[1,i]=GETML()
    #(1,1)
bcols=col[lr>=n-1]
dcols=GETDCOLS(noofdisp)
    mmat[2,i]=GETML()
    #(2,1)
bcols=col[lr>=n-2]
dcols=GETDCOLS(noofdisp)
    mmat[3,i]=GETML()
    #(3,1)
bcols=col[lr>=n-3]
dcols=GETDCOLS(noofdisp)
    mmat[4,i]=GETML()
    #(4,1)
bcols=col[lr>=n-4]
dcols=GETDCOLS(noofdisp)
    mmat[5,i]=GETML()
    #(5,1)
bcols=col[lr>=n-5]
dcols=GETDCOLS(noofdisp)
    mmat[6,i]=GETML()
    #(6,1)
bcols=col[lr>=n-6]
dcols=GETDCOLS(noofdisp)
    mmat[7,i]=GETML()
    #(7,1)
bcols=col[lr>=n-7]
dcols=GETDCOLS(noofdisp)
    mmat[8,i]=GETML()
    }
    quant.out2[,k]=Q1(mmat)
    print(paste(k,date()))
  }

date()}
Appendix B – Tables and Figures

Fig. B.1 Resulting histogram of -2L difference for different true models

(0.0) vs (1.0)  (0.0) vs (0.1)  (0.0) vs (1.1)
(0.0) vs (2.0)  (0.0) vs (2.1)  (0.0) vs (2.2)
(0.0) vs (1.2)  (0.0) vs (0.2)  (0.0) vs (3.0)
(1.0) vs (2.0)  (1.0) vs (1.1)  (1.0) vs (2.1)
(1.0) vs (3.0)  (1.0) vs (3.1)  (1.0) vs (3.2)
(1.0) vs (2.2)  (1.0) vs (1.2)  (1.0) vs (4.0)
Table B.2 Estimates of coefficients for different true models and models with added terms

<table>
<thead>
<tr>
<th>Model</th>
<th>Location</th>
<th>Dispersion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$</td>
<td>$\beta_1$</td>
</tr>
<tr>
<td>(0.0)</td>
<td>-0.0027</td>
<td></td>
</tr>
<tr>
<td>(1.0)</td>
<td>-0.0046</td>
<td>0.00866</td>
</tr>
<tr>
<td>(0.1)</td>
<td>-0.0168</td>
<td></td>
</tr>
<tr>
<td>(1.1)</td>
<td>-0.0046</td>
<td>0.00866</td>
</tr>
<tr>
<td>(2.0)</td>
<td>-0.0032</td>
<td>0.00866</td>
</tr>
<tr>
<td>(2.1)</td>
<td>-0.0064</td>
<td>0.01131</td>
</tr>
<tr>
<td>(2.2)</td>
<td>-0.0043</td>
<td>0.01071</td>
</tr>
<tr>
<td>(1.0)</td>
<td>-0.0079</td>
<td>0.01583</td>
</tr>
<tr>
<td>(2.0)</td>
<td>-0.0059</td>
<td></td>
</tr>
<tr>
<td>(3.0)</td>
<td>-0.0063</td>
<td>0.00966</td>
</tr>
<tr>
<td>(1.0)</td>
<td>-0.0011</td>
<td>0.99999</td>
</tr>
<tr>
<td>(2.0)</td>
<td>0.0111</td>
<td>0.99999</td>
</tr>
<tr>
<td>(1.1)</td>
<td>-0.001</td>
<td>0.99905</td>
</tr>
<tr>
<td>(2.1)</td>
<td>0.0011</td>
<td>0.99981</td>
</tr>
<tr>
<td>(3.0)</td>
<td>-0.0012</td>
<td>0.99999</td>
</tr>
<tr>
<td>(3.1)</td>
<td>0.0065</td>
<td>0.98841</td>
</tr>
<tr>
<td>(3.2)</td>
<td>-0.0087</td>
<td>0.92297</td>
</tr>
<tr>
<td>(2.2)</td>
<td>0.0081</td>
<td>0.98748</td>
</tr>
<tr>
<td>(1.2)</td>
<td>0.09035</td>
<td></td>
</tr>
<tr>
<td>(4.0)</td>
<td>-0.00751</td>
<td>0.99999</td>
</tr>
<tr>
<td>(0.1)</td>
<td>0.00156</td>
<td></td>
</tr>
<tr>
<td>(1.0)</td>
<td>0.0026</td>
<td>-0.0027</td>
</tr>
<tr>
<td>(2.0)</td>
<td>-0.00137</td>
<td></td>
</tr>
<tr>
<td>(1.2)</td>
<td>0.00107</td>
<td>-0.00166</td>
</tr>
<tr>
<td>(2.1)</td>
<td>0.00721</td>
<td>-0.00384</td>
</tr>
<tr>
<td>(2.2)</td>
<td>0.0071</td>
<td>-0.00498</td>
</tr>
<tr>
<td>(3.1)</td>
<td>-0.0083</td>
<td>-0.01083</td>
</tr>
<tr>
<td>(1.3)</td>
<td>0.00478</td>
<td>-0.00025</td>
</tr>
<tr>
<td>(0.3)</td>
<td>0.00066</td>
<td></td>
</tr>
<tr>
<td>(3.1)</td>
<td>0.00929</td>
<td>-0.00517</td>
</tr>
<tr>
<td>(1.1)</td>
<td>-0.00478</td>
<td>1.00743</td>
</tr>
<tr>
<td>(2.1)</td>
<td>-0.00765</td>
<td>1.00611</td>
</tr>
<tr>
<td>(1.2)</td>
<td>-0.00496</td>
<td>1.00718</td>
</tr>
<tr>
<td>(2.2)</td>
<td>-0.00112</td>
<td>1.00032</td>
</tr>
<tr>
<td>(3.1)</td>
<td>0.00161</td>
<td>1.00042</td>
</tr>
<tr>
<td>(3.2)</td>
<td>0.00496</td>
<td>0.99928</td>
</tr>
<tr>
<td>(3.3)</td>
<td>0.01057</td>
<td>0.9928</td>
</tr>
<tr>
<td>(2.3)</td>
<td>0.00114</td>
<td>1.00024</td>
</tr>
<tr>
<td>(1.3)</td>
<td>-0.00306</td>
<td>1.00252</td>
</tr>
<tr>
<td>(4.1)</td>
<td>0.00326</td>
<td>1.00259</td>
</tr>
<tr>
<td>(2.0)</td>
<td>-0.00479</td>
<td>0.99421</td>
</tr>
<tr>
<td>(3.0)</td>
<td>-0.0056</td>
<td>0.99421</td>
</tr>
<tr>
<td>(2.1)</td>
<td>-0.00556</td>
<td>0.99421</td>
</tr>
<tr>
<td>(3.1)</td>
<td>-0.00682</td>
<td>0.9942</td>
</tr>
<tr>
<td>(4.0)</td>
<td>-0.00733</td>
<td>0.99421</td>
</tr>
<tr>
<td>(4.1)</td>
<td>-0.00302</td>
<td>0.99423</td>
</tr>
<tr>
<td>(4.2)</td>
<td>0.01264</td>
<td>0.98059</td>
</tr>
<tr>
<td>(3.2)</td>
<td>0.0071</td>
<td>0.98097</td>
</tr>
<tr>
<td>(2.2)</td>
<td>-0.00383</td>
<td>0.99409</td>
</tr>
</tbody>
</table>

- $\beta_0$: Coefficient for the intercept
- $\beta_1$: Coefficient for the first true model
- $\beta_2$: Coefficient for the second true model
- $\beta_3$: Coefficient for the third true model
- $\beta_4$: Coefficient for the fourth true model
- $\beta_5$: Coefficient for the fifth true model
- $\gamma_0$: Coefficient for the dispersion parameter in the first true model
- $\gamma_1$: Coefficient for the dispersion parameter in the second true model
- $\gamma_2$: Coefficient for the dispersion parameter in the third true model
- $\gamma_3$: Coefficient for the dispersion parameter in the fourth true model
- $\gamma_4$: Coefficient for the dispersion parameter in the fifth true model
<table>
<thead>
<tr>
<th>(5,0)</th>
<th>-0.00714</th>
<th>0.99421</th>
<th>1.01107</th>
<th>0.00345</th>
<th>0.00164</th>
<th>-0.0037</th>
<th>-0.56973</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2,1)</td>
<td>-0.00355</td>
<td>1.0075</td>
<td>0.99629</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,1)</td>
<td>0.00802</td>
<td>0.99532</td>
<td>0.98609</td>
<td>-0.00772</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2,2)</td>
<td>0.00444</td>
<td>0.99808</td>
<td>0.98451</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,2)</td>
<td>0.00951</td>
<td>0.9927</td>
<td>0.96682</td>
<td>0.01479</td>
<td>-0.01266</td>
<td></td>
<td>-0.48951</td>
</tr>
<tr>
<td>(3,3)</td>
<td>0.01373</td>
<td>0.99311</td>
<td>0.98157</td>
<td>-0.01298</td>
<td></td>
<td></td>
<td>-0.25846</td>
</tr>
<tr>
<td>(5,1)</td>
<td>-0.00163</td>
<td>1.00105</td>
<td>0.99514</td>
<td>0.01559</td>
<td>-0.01265</td>
<td>0.00729</td>
<td>-0.11946</td>
</tr>
<tr>
<td>(2,3)</td>
<td>0.00222</td>
<td>1.00067</td>
<td>0.98009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,1)</td>
<td>0.00802</td>
<td>0.99532</td>
<td>0.98609</td>
<td>-0.00772</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4,0)</td>
<td>0</td>
<td>1.00361</td>
<td>0.99222</td>
<td>1.00283</td>
<td>0.00422</td>
<td></td>
<td>-0.47446</td>
</tr>
<tr>
<td>(4,1)</td>
<td>-0.0035</td>
<td>1.00361</td>
<td>0.99043</td>
<td>1.00762</td>
<td></td>
<td></td>
<td>-0.49923</td>
</tr>
<tr>
<td>(5,0)</td>
<td>-0.00355</td>
<td>1.00361</td>
<td>0.99043</td>
<td>1.00762</td>
<td></td>
<td></td>
<td>-0.49923</td>
</tr>
<tr>
<td>(5,1)</td>
<td>-0.00351</td>
<td>1.00361</td>
<td>0.99043</td>
<td>1.00762</td>
<td></td>
<td></td>
<td>-0.49923</td>
</tr>
<tr>
<td>(5,2)</td>
<td>-0.00306</td>
<td>1.00361</td>
<td>0.99043</td>
<td>1.00762</td>
<td></td>
<td></td>
<td>-0.49923</td>
</tr>
<tr>
<td>(6,0)</td>
<td>-0.00183</td>
<td>1.00361</td>
<td>0.99222</td>
<td>1.00283</td>
<td>0.00422</td>
<td>0.00726</td>
<td>-0.00341</td>
</tr>
<tr>
<td>(0)</td>
<td>.95Q</td>
<td>Mean</td>
<td>Std. Dev</td>
<td>(0)</td>
<td>.95Q</td>
<td>Mean</td>
<td>Std. Dev</td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>1</td>
<td>(1,0)</td>
<td>4.70857</td>
<td>1.206331</td>
<td>(1,0)</td>
<td>5.897428</td>
<td>1.294636</td>
<td>1.790121</td>
</tr>
<tr>
<td>1</td>
<td>(0,1)</td>
<td>4.7658</td>
<td>1.248415</td>
<td>(1,1)</td>
<td>5.462189</td>
<td>1.499456</td>
<td>2.087033</td>
</tr>
<tr>
<td>2</td>
<td>(1,1)</td>
<td>7.327649</td>
<td>2.433343</td>
<td>2.434533</td>
<td>(2,1)</td>
<td>8.047634</td>
<td>2.745854</td>
</tr>
<tr>
<td>2</td>
<td>(2,0)</td>
<td>7.347386</td>
<td>2.480381</td>
<td>2.477084</td>
<td>(3,0)</td>
<td>7.83304</td>
<td>2.655121</td>
</tr>
<tr>
<td>3</td>
<td>(0,2)</td>
<td>4.49647</td>
<td>1.503081</td>
<td>2.487133</td>
<td>(4,0)</td>
<td>9.173391</td>
<td>3.097151</td>
</tr>
<tr>
<td>3</td>
<td>(2,2)</td>
<td>11.46299</td>
<td>4.497468</td>
<td>3.86025</td>
<td>(3,1)</td>
<td>13.90188</td>
<td>5.439927</td>
</tr>
<tr>
<td>3</td>
<td>(1,2)</td>
<td>11.03515</td>
<td>3.698515</td>
<td>3.49508</td>
<td>(2,2)</td>
<td>13.89604</td>
<td>5.461055</td>
</tr>
<tr>
<td>3</td>
<td>(3,0)</td>
<td>9.996214</td>
<td>3.915147</td>
<td>3.89729</td>
<td>(4,0)</td>
<td>10.89652</td>
<td>4.767459</td>
</tr>
<tr>
<td>4</td>
<td>(2,2)</td>
<td>13.00016</td>
<td>5.461071</td>
<td>3.887711</td>
<td>(2,1)</td>
<td>15.36467</td>
<td>6.531887</td>
</tr>
</tbody>
</table>

**Table B.3** Summary of relevant statistics for different true models and models with added terms
Table B.4 Histogram of -2L differences grouped by number of dispersion terms in the model
<table>
<thead>
<tr>
<th>Cycle</th>
<th>Quantile</th>
<th>Model 1 (more terms)</th>
<th>Model 2 (less terms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05</td>
<td>10.24</td>
<td>17.77</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>8.716</td>
<td>15.26</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>7.244</td>
<td>13.26</td>
</tr>
<tr>
<td>2</td>
<td>0.99</td>
<td>13.69</td>
<td>21.62</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>10.18</td>
<td>17.22</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>8.683</td>
<td>15.26</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>7.233</td>
<td>13.28</td>
</tr>
<tr>
<td>3</td>
<td>0.99</td>
<td>16.88</td>
<td>21.55</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>15.17</td>
<td>21.76</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>7.055</td>
<td>15.24</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>7.216</td>
<td>13.23</td>
</tr>
<tr>
<td>4</td>
<td>0.99</td>
<td>17.72</td>
<td>21.56</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>17.08</td>
<td>21.73</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>8.705</td>
<td>15.24</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>7.245</td>
<td>13.29</td>
</tr>
<tr>
<td>5</td>
<td>0.99</td>
<td>16.81</td>
<td>21.57</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>15.31</td>
<td>21.76</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>8.705</td>
<td>15.32</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>7.245</td>
<td>13.39</td>
</tr>
<tr>
<td>6</td>
<td>0.99</td>
<td>17.65</td>
<td>21.56</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.97</td>
<td>21.76</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>8.895</td>
<td>15.33</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>7.245</td>
<td>13.39</td>
</tr>
<tr>
<td>7</td>
<td>0.99</td>
<td>16.72</td>
<td>21.62</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.19</td>
<td>21.76</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>8.905</td>
<td>15.33</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>7.245</td>
<td>13.39</td>
</tr>
</tbody>
</table>

Table B.5.2 Results of the simulation for models with 1 dispersion effects

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Quantiles</th>
<th>Model 1 (less location terms)</th>
<th>Model 2 (more location terms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.90</td>
<td>19.70</td>
<td>27.90</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>14.74</td>
<td>22.45</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>12.48</td>
<td>19.74</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>10.03</td>
<td>16.84</td>
</tr>
<tr>
<td>2</td>
<td>0.99</td>
<td>19.98</td>
<td>28.33</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>14.83</td>
<td>22.45</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>12.56</td>
<td>19.81</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>10.08</td>
<td>16.83</td>
</tr>
<tr>
<td>3</td>
<td>0.99</td>
<td>19.98</td>
<td>28.33</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>14.83</td>
<td>22.45</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>12.56</td>
<td>19.81</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>10.08</td>
<td>16.83</td>
</tr>
<tr>
<td>Cycle</td>
<td>Quantities</td>
<td>Model 1 (lessor location terms)</td>
<td>Model 2 (more location terms)</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>--------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>3</td>
<td>0.99 12.51 19.98 27.35 35.34 44.06 54.57 67.91 12.43 13.46 15.16 17.74 21.42 28.21</td>
<td>0.80 10.08 16.96 23.76 31.01 38.94 47.96 59.05 9.75 10.36 11.49 13.16 15.27 19.15</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.99 19.91 28.34 37.06 47.48 60.50 76.47 105.76 20.26 22.43 26.64 33.94 44.62 67.44</td>
<td>0.95 14.79 22.50 30.34 39.08 48.89 60.67 76.98 14.81 16.20 18.75 22.35 27.93 37.53</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.99 12.52 19.83 27.35 35.26 44.07 54.41 67.93 12.41 13.36 15.19 17.75 21.31 27.76</td>
<td>0.80 10.03 16.83 23.68 30.91 38.89 48.01 58.73 9.76 10.39 11.46 13.15 15.33 18.79</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.99 20.02 28.27 37.62 47.45 59.85 75.76 104.89 20.13 22.71 26.77 33.58 44.46 66.47</td>
<td>0.95 14.88 22.51 30.48 39.12 48.98 60.98 76.79 14.67 16.34 18.64 22.40 28.07 37.46</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.99 12.54 19.83 27.36 35.26 44.11 54.56 67.64 12.29 13.46 15.19 17.65 21.48 27.71</td>
<td>0.90 12.53 19.78 27.22 35.12 43.99 54.70 67.72 12.32 13.49 15.17 17.73 21.91 28.17</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.99 19.86 28.29 37.12 48.03 60.21 76.87 106.58 20.15 22.52 26.70 33.81 45.26 67.87</td>
<td>0.95 14.80 22.35 30.43 39.06 49.07 61.14 77.05 14.76 16.29 18.63 22.40 28.47 38.08</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.90 12.51 19.84 27.19 35.19 43.96 54.56 67.50 12.37 13.30 15.18 17.72 21.61 27.64</td>
<td>0.80 10.06 16.87 23.69 30.96 38.89 48.01 58.86 9.76 10.34 11.54 13.17 15.45 18.74</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.99 19.58 28.37 36.99 47.68 60.03 77.45 106.78 20.06 22.50 27.05 34.18 45.82 67.71</td>
<td>0.95 14.71 22.50 30.39 39.24 48.90 60.82 77.24 14.74 16.25 18.92 22.48 27.93 38.30</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.90 12.49 19.79 27.20 35.19 44.11 54.34 67.81 12.35 13.38 15.34 17.77 21.52 28.03</td>
<td>0.80 10.05 16.86 23.71 31.04 38.81 47.91 58.86 9.76 10.36 11.52 13.08 15.35 18.91</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.99 19.74 28.45 37.14 47.49 60.54 76.94 105.71 20.19 22.74 27.16 33.95 45.18 68.03</td>
<td>0.95 14.80 22.61 30.35 39.10 49.18 60.80 77.12 14.80 16.23 18.73 22.63 28.18 38.09</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.90 12.51 19.85 27.22 35.16 44.04 54.60 67.76 12.35 13.33 15.15 17.80 21.53 28.03</td>
<td>0.80 10.04 16.87 23.75 30.89 39.84 48.09 58.83 9.76 10.32 11.50 13.14 15.32 19.01</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.99 19.74 28.39 37.36 47.32 59.82 78.48 107.82 20.15 22.77 26.88 33.78 46.65 69.70</td>
<td>0.95 14.93 22.56 30.52 38.96 48.99 61.04 77.68 14.83 16.22 18.72 22.41 28.11 38.36</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.90 12.60 19.85 27.21 35.24 44.07 54.59 68.05 12.35 13.36 15.11 17.68 21.61 27.96</td>
<td>0.80 10.06 16.86 23.74 31.05 38.99 48.01 59.04 9.69 10.38 11.56 13.11 15.38 19.17</td>
<td></td>
</tr>
</tbody>
</table>

Table B.5.3. Results of the simulation for models with 2 dispersion effects
<p>| | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>0.99</td>
<td>22.05</td>
<td>30.62</td>
<td>41.27</td>
<td>53.17</td>
<td>68.88</td>
<td>89.78</td>
<td>152.82</td>
<td>22.55</td>
<td>26.07</td>
<td>31.44</td>
<td>40.38</td>
<td>55.23</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.57</td>
<td>24.58</td>
<td>33.67</td>
<td>43.76</td>
<td>56.33</td>
<td>71.71</td>
<td>92.98</td>
<td>16.72</td>
<td>18.79</td>
<td>22.38</td>
<td>27.94</td>
<td>35.92</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>14.00</td>
<td>21.63</td>
<td>30.05</td>
<td>39.27</td>
<td>50.54</td>
<td>64.00</td>
<td>80.93</td>
<td>13.90</td>
<td>15.55</td>
<td>18.24</td>
<td>22.47</td>
<td>28.18</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>11.15</td>
<td>18.35</td>
<td>26.04</td>
<td>34.49</td>
<td>44.41</td>
<td>55.85</td>
<td>69.48</td>
<td>10.91</td>
<td>12.05</td>
<td>13.84</td>
<td>16.69</td>
<td>20.14</td>
</tr>
<tr>
<td>53</td>
<td>0.99</td>
<td>22.07</td>
<td>30.72</td>
<td>41.22</td>
<td>53.30</td>
<td>68.08</td>
<td>90.60</td>
<td>153.10</td>
<td>22.70</td>
<td>26.26</td>
<td>31.57</td>
<td>39.83</td>
<td>55.04</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.58</td>
<td>24.50</td>
<td>33.62</td>
<td>43.76</td>
<td>56.00</td>
<td>71.66</td>
<td>93.22</td>
<td>16.66</td>
<td>18.96</td>
<td>22.40</td>
<td>27.68</td>
<td>36.11</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>13.94</td>
<td>21.57</td>
<td>30.01</td>
<td>39.47</td>
<td>50.51</td>
<td>63.93</td>
<td>80.90</td>
<td>13.92</td>
<td>15.61</td>
<td>18.32</td>
<td>22.39</td>
<td>28.26</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>11.15</td>
<td>18.29</td>
<td>26.04</td>
<td>34.55</td>
<td>44.22</td>
<td>55.88</td>
<td>69.74</td>
<td>10.87</td>
<td>12.03</td>
<td>13.87</td>
<td>16.49</td>
<td>20.54</td>
</tr>
<tr>
<td>54</td>
<td>0.99</td>
<td>21.97</td>
<td>30.84</td>
<td>41.18</td>
<td>52.88</td>
<td>68.42</td>
<td>90.50</td>
<td>154.34</td>
<td>22.87</td>
<td>26.10</td>
<td>31.18</td>
<td>40.38</td>
<td>55.53</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.48</td>
<td>24.56</td>
<td>33.61</td>
<td>43.68</td>
<td>56.13</td>
<td>71.54</td>
<td>93.68</td>
<td>16.68</td>
<td>18.94</td>
<td>22.19</td>
<td>27.98</td>
<td>36.28</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>13.96</td>
<td>21.60</td>
<td>30.02</td>
<td>39.35</td>
<td>50.47</td>
<td>64.03</td>
<td>81.55</td>
<td>14.04</td>
<td>15.50</td>
<td>18.17</td>
<td>22.38</td>
<td>28.26</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>11.13</td>
<td>18.35</td>
<td>26.00</td>
<td>34.52</td>
<td>44.37</td>
<td>56.08</td>
<td>69.90</td>
<td>10.96</td>
<td>11.92</td>
<td>13.89</td>
<td>16.54</td>
<td>20.43</td>
</tr>
<tr>
<td>55</td>
<td>0.99</td>
<td>22.04</td>
<td>30.86</td>
<td>41.29</td>
<td>53.53</td>
<td>68.72</td>
<td>90.47</td>
<td>163.55</td>
<td>22.98</td>
<td>26.14</td>
<td>31.73</td>
<td>40.38</td>
<td>55.52</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.55</td>
<td>24.72</td>
<td>33.58</td>
<td>43.84</td>
<td>56.28</td>
<td>71.81</td>
<td>93.62</td>
<td>16.81</td>
<td>18.93</td>
<td>22.59</td>
<td>27.90</td>
<td>36.22</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>13.97</td>
<td>21.75</td>
<td>30.07</td>
<td>39.50</td>
<td>50.50</td>
<td>64.37</td>
<td>81.58</td>
<td>14.11</td>
<td>15.63</td>
<td>18.38</td>
<td>22.47</td>
<td>28.49</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>11.12</td>
<td>18.43</td>
<td>26.06</td>
<td>34.64</td>
<td>44.47</td>
<td>56.15</td>
<td>70.07</td>
<td>11.01</td>
<td>12.02</td>
<td>13.91</td>
<td>16.76</td>
<td>20.54</td>
</tr>
<tr>
<td>56</td>
<td>0.99</td>
<td>22.07</td>
<td>30.76</td>
<td>40.88</td>
<td>53.37</td>
<td>68.70</td>
<td>89.81</td>
<td>148.07</td>
<td>22.68</td>
<td>25.96</td>
<td>31.29</td>
<td>40.45</td>
<td>55.21</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16.59</td>
<td>24.47</td>
<td>33.57</td>
<td>43.80</td>
<td>56.23</td>
<td>71.42</td>
<td>92.93</td>
<td>16.80</td>
<td>18.85</td>
<td>22.53</td>
<td>27.84</td>
<td>36.37</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>13.96</td>
<td>21.53</td>
<td>30.02</td>
<td>39.41</td>
<td>50.48</td>
<td>63.89</td>
<td>81.05</td>
<td>14.00</td>
<td>15.54</td>
<td>18.35</td>
<td>22.29</td>
<td>28.19</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>11.06</td>
<td>18.33</td>
<td>26.00</td>
<td>34.49</td>
<td>44.25</td>
<td>55.73</td>
<td>69.72</td>
<td>10.98</td>
<td>12.00</td>
<td>13.83</td>
<td>16.61</td>
<td>20.32</td>
</tr>
</tbody>
</table>
### Table B.6. Summarized findings for the observations with big -2L difference for models (0,3) and (3,3)

<table>
<thead>
<tr>
<th>Disp Model Terms</th>
<th>Disp Term Rank</th>
<th>Disp Model Coefficients</th>
<th>Sum of Abs</th>
<th>ML Diff</th>
<th>Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>C AC BC D ABC ABCD</td>
<td>10 11 4</td>
<td>-17.6936 -15.1786 17.0041 17.16004</td>
<td>67.03633</td>
<td>131.5036</td>
<td>2 3</td>
</tr>
<tr>
<td>B D AD AB ACD BCD</td>
<td>11 2 10</td>
<td>-16.8968 15.5277 15.84655 -14.8557</td>
<td>63.12676</td>
<td>127.8562</td>
<td>2 3</td>
</tr>
<tr>
<td>AD ABC BCD B ACD</td>
<td>10 7 8</td>
<td>-18.8564 18.67696 16.85743 -16.7247</td>
<td>71.11554</td>
<td>146.5315</td>
<td>1 3</td>
</tr>
<tr>
<td>D ABC ABD B C BC</td>
<td>2 7 1</td>
<td>-18.67 15.74163 -15.7485 16.26171</td>
<td>64.62182</td>
<td>138.3268</td>
<td>1 3</td>
</tr>
<tr>
<td>AD BC BD AC ACD</td>
<td>7 3 4</td>
<td>-17.1393 15.56964 16.32981 -18.2582</td>
<td>65.32336</td>
<td>128.8352</td>
<td>1 3</td>
</tr>
<tr>
<td>BC BCD ACD C D CD</td>
<td>4 5 1</td>
<td>-18.1914 17.45836 15.34544 -15.621</td>
<td>72.01546</td>
<td>139.6801</td>
<td>2 3</td>
</tr>
<tr>
<td>B C D BC ABC ABCD</td>
<td>9 1 8</td>
<td>15.34544 -15.6214 -17.4061 -16.1117</td>
<td>64.48426</td>
<td>137.3691</td>
<td>2 3</td>
</tr>
<tr>
<td>D ABC ACD BC ABD</td>
<td>5 4 10</td>
<td>-16.8968 15.5277 15.84655 -14.8557</td>
<td>63.12676</td>
<td>127.8562</td>
<td>2 3</td>
</tr>
<tr>
<td>A BCD BCD B ACD</td>
<td>10 7 8</td>
<td>-18.2648 17.3452 15.0309 16.66973</td>
<td>67.30894</td>
<td>131.9457</td>
<td>2 3</td>
</tr>
<tr>
<td>D ABC ACD BC ABD</td>
<td>6 2 11</td>
<td>-19.1745 17.5903 -17.7319 16.76312</td>
<td>66.42852</td>
<td>141.3096</td>
<td>1 3</td>
</tr>
<tr>
<td>A BCD BCD B ACD</td>
<td>10 7 8</td>
<td>-18.2648 17.3452 15.0309 16.66973</td>
<td>67.30894</td>
<td>131.9457</td>
<td>2 3</td>
</tr>
<tr>
<td>A BCD BCD B ACD</td>
<td>10 7 8</td>
<td>-18.2648 17.3452 15.0309 16.66973</td>
<td>67.30894</td>
<td>131.9457</td>
<td>2 3</td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>AB</td>
<td>AD</td>
<td>ABC</td>
<td>BC</td>
<td>ABD</td>
<td>AC</td>
</tr>
<tr>
<td>D</td>
<td>BD</td>
<td>BCD</td>
<td>AB</td>
<td>CD</td>
<td>ABCD</td>
</tr>
<tr>
<td>AC</td>
<td>BD</td>
<td>ABC</td>
<td>C</td>
<td>D</td>
<td>CD</td>
</tr>
<tr>
<td>ABD</td>
<td>ACD</td>
<td>ABCD</td>
<td>AD</td>
<td>ABC</td>
<td>BCD</td>
</tr>
<tr>
<td>ACD</td>
<td>BCD</td>
<td>ABCD</td>
<td>BC</td>
<td>BD</td>
<td>CD</td>
</tr>
<tr>
<td>AB</td>
<td>AD</td>
<td>BD</td>
<td>D</td>
<td>AC</td>
<td>ACD</td>
</tr>
<tr>
<td>B</td>
<td>BD</td>
<td>ABC</td>
<td>AB</td>
<td>CD</td>
<td>ABCD</td>
</tr>
<tr>
<td>D</td>
<td>ABC</td>
<td>ABCD</td>
<td>AB</td>
<td>ACD</td>
<td>BCD</td>
</tr>
</tbody>
</table>

Tags: 1 = interaction of 2 location terms is the third location term  
2 = interaction of all 3 location terms is one of the dispersion term  
3 = interaction of 2 dispersion terms is the third dispersion term
<table>
<thead>
<tr>
<th>Quantile</th>
<th>0.0 (0)</th>
<th>0.01 (1)</th>
<th>0.02 (2)</th>
<th>0.03 (3)</th>
<th>0.04 (4)</th>
<th>0.05 (5)</th>
<th>0.06 (6)</th>
<th>0.07 (7)</th>
<th>0.08 (8)</th>
<th>0.09 (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.0)</td>
<td>NA</td>
<td>0.00</td>
<td>11.04</td>
<td>19.38</td>
<td>27.37</td>
<td>35.42</td>
<td>43.97</td>
<td>51.57</td>
<td>6.77</td>
<td>15.78</td>
</tr>
<tr>
<td>(2.0)</td>
<td>NA</td>
<td>0.00</td>
<td>10.44</td>
<td>19.21</td>
<td>27.85</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(3.0)</td>
<td>NA</td>
<td>0.00</td>
<td>9.65</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(4.0)</td>
<td>NA</td>
<td>0.00</td>
<td>8.97</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(5.0)</td>
<td>NA</td>
<td>0.00</td>
<td>8.38</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(6.0)</td>
<td>NA</td>
<td>0.00</td>
<td>7.86</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(7.0)</td>
<td>NA</td>
<td>0.00</td>
<td>7.42</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(8.0)</td>
<td>NA</td>
<td>0.00</td>
<td>7.04</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(9.0)</td>
<td>NA</td>
<td>0.00</td>
<td>6.72</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
<tr>
<td>(10.0)</td>
<td>NA</td>
<td>0.00</td>
<td>6.44</td>
<td>19.09</td>
<td>27.67</td>
<td>36.02</td>
<td>45.93</td>
<td>53.79</td>
<td>37.59</td>
<td>68.83</td>
</tr>
</tbody>
</table>

Table B.7. Critical Values Table
<table>
<thead>
<tr>
<th>0.90</th>
<th>0.91</th>
<th>0.92</th>
<th>0.93</th>
<th>0.94</th>
<th>0.95</th>
<th>0.96</th>
<th>0.97</th>
<th>0.98</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>NA 0.00</td>
<td>7.24</td>
<td>13.79</td>
<td>20.25</td>
<td>26.88</td>
<td>33.89</td>
<td>41.40</td>
<td>2.94</td>
<td>9.18</td>
</tr>
<tr>
<td>(0,1)</td>
<td>NA 0.00</td>
<td>8.94</td>
<td>13.64</td>
<td>20.47</td>
<td>27.65</td>
<td>35.38</td>
<td>44.00</td>
<td>3.49</td>
<td>10.66</td>
</tr>
<tr>
<td>(0,2)</td>
<td>NA 0.00</td>
<td>6.90</td>
<td>12.55</td>
<td>19.83</td>
<td>26.57</td>
<td>33.89</td>
<td>41.40</td>
<td>2.94</td>
<td>9.18</td>
</tr>
<tr>
<td>(0,3)</td>
<td>NA 0.00</td>
<td>6.90</td>
<td>12.55</td>
<td>19.83</td>
<td>26.57</td>
<td>33.89</td>
<td>41.40</td>
<td>2.94</td>
<td>9.18</td>
</tr>
</tbody>
</table>

**Table B.7. Critical Values Table - Continued**

<table>
<thead>
<tr>
<th>0.90</th>
<th>0.91</th>
<th>0.92</th>
<th>0.93</th>
<th>0.94</th>
<th>0.95</th>
<th>0.96</th>
<th>0.97</th>
<th>0.98</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td>NA 0.10</td>
<td>5.71</td>
<td>11.51</td>
<td>18.19</td>
<td>24.86</td>
<td>31.53</td>
<td>38.20</td>
<td>31.78</td>
<td>41.40</td>
</tr>
<tr>
<td>(1,2)</td>
<td>NA 0.10</td>
<td>5.71</td>
<td>11.51</td>
<td>18.19</td>
<td>24.86</td>
<td>31.53</td>
<td>38.20</td>
<td>31.78</td>
<td>41.40</td>
</tr>
<tr>
<td>(1,3)</td>
<td>NA 0.10</td>
<td>5.71</td>
<td>11.51</td>
<td>18.19</td>
<td>24.86</td>
<td>31.53</td>
<td>38.20</td>
<td>31.78</td>
<td>41.40</td>
</tr>
</tbody>
</table>

**Table B.7. Critical Values Table - Continued**

<table>
<thead>
<tr>
<th>0.90</th>
<th>0.91</th>
<th>0.92</th>
<th>0.93</th>
<th>0.94</th>
<th>0.95</th>
<th>0.96</th>
<th>0.97</th>
<th>0.98</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2,1)</td>
<td>NA 0.10</td>
<td>5.71</td>
<td>11.51</td>
<td>18.19</td>
<td>24.86</td>
<td>31.53</td>
<td>38.20</td>
<td>31.78</td>
<td>41.40</td>
</tr>
<tr>
<td>(2,2)</td>
<td>NA 0.10</td>
<td>5.71</td>
<td>11.51</td>
<td>18.19</td>
<td>24.86</td>
<td>31.53</td>
<td>38.20</td>
<td>31.78</td>
<td>41.40</td>
</tr>
<tr>
<td>(2,3)</td>
<td>NA 0.10</td>
<td>5.71</td>
<td>11.51</td>
<td>18.19</td>
<td>24.86</td>
<td>31.53</td>
<td>38.20</td>
<td>31.78</td>
<td>41.40</td>
</tr>
</tbody>
</table>