AN EVALUATION OF SOME APPROXIMATE F STATISTICS AND THEIR SMALL SAMPLE DISTRIBUTIONS FOR THE MIXED MODEL WITH LINEAR COVARIANCE STRUCTURE

by

Kathleen McCarroll
and Ron Helms

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1838T

October 1987

An Evaluation of Some Approximate F Statistics and Their Small Sample Distributions for the Mixed Model with Linear Covariance Structure

by

Kathleen McCarroll

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics.

Chapel Hill

1987

Approved by:

Ronald W Helms

Advisor

Clarence E Davis

Reader

Keith E. Muller

Reader

KATHLEEN McCARROLL. An Evaluation of Some Approximate F Statistics and Their Small Sample Distributions for the Mixed Model with Linear Covariance Structure. (Under the direction of RONALD W. HELMS).

The purpose of this work was to extend results from the General Linear Univariate Model and the General Linear Multivariate Model to special cases of the mixed model with linear covariance structure. These extensions were then used to motivate approximate F statistics for the mixed model. Three approximate F statistics were proposed; one was based on the canonical form of the mixed model (F_{REML}) and two were based on weighted least squares (F_{WLS}, F_{WLS2}). In a simulation study the three statistics were compared to each other, to the likelihood ratio statistic (LRT) of Andrade and Helms (1984) and to an adjusted ordinary least squares test (F_{Box}).

The test statistics were evaluated for two designs and with missing data, under different hypotheses. The designs that were used were a complete longitudinal design (375 observations), a longitudinal design with 21% missing data (294 observations), a linked cross-sectional design (175 observations) and a linked cross-sectional design with 21% missing data (138 observations).

The F_{REML} statistic produced the most accurate Type I error rates and the closest conformance with the hypothesized distribution. The F_{WLS}, F_{WLS2} and LRT statistics produced inflated Type I error rates. The F_{Box} statistic produced accurate Type I error rates for some cases but the hypothesized distribution did not fit in the noncentral cases.

Table of Contents

Chapter 1

Introduction and Literature Review

The purpose of this work was to develop approximate F statistics that could be used for the mixed model with linear covariance structure. In particular we considered the mixed model as applied in longitudinal designs.

The first section of the literature review will review the use of the mixed model with linear covariance structure in analyzing data from longitudinal studies. Because missing data are a common occurrence in longitudinal studies we will discuss how the mixed model can accomodate this.

The second section will review existing approximations for power when the data are correlated. These include the methods of Geisser and Greenhouse, Huynh and Feldt, and the extension to GLMM of Muller and Barton.

The third section gives a brief review of longitudinal designs. The linked cross-sectional design, a special type of longitudinal design that we will consider, will be discussed.

1.1 Analysis of Incomplete Data from Longitudinal Studies

1.1.1 Introduction and Notation

The analysis of complete, univariate data from a cross-sectional design utilizes standard statistical techniques that have been covered in many texts such as Snedecor and Cochran (1980), Fleiss (1981), Kleinbaum, Kupper and Morgenstern (1982), and Searle (1971). Randomly missing data in such a data set will not make the analysis overly complicated although it may cause a reduction in power due to the reduction in sample size and loss of orthogonality. Missing data may also introduce the possibility of bias and incorrect inferences.

We will use the following notation for the univariate case. In the univariate setting there are N independent observations on one variable. The General Linear Univariate Model (GLUM) is written as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad (1.1.1)$$

where $\mathbf{Y}$ is an Nx1 vector of observations,

$\mathbf{X}$ is an Nxq design matrix of known constants,

$\boldsymbol{\beta}$ is a qx1 vector of parameters to be estimated,

$\boldsymbol{\epsilon}$ is an Nx1 random error vector.

It is assumed that $E(\boldsymbol{\epsilon})=0$ and $Var(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}_n$. Therefore $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ and $Var(\mathbf{Y}) = \sigma^2\mathbf{I}_n$.

The analysis of many types of complete multivariate data from longitudinal designs has been covered in many texts such as Morrison (1976), Timm (1975), Anderson (1984) and Johnson and Wichern

(1982). Also various methods of analysis are summarized and referenced in the review articles by Cook and Ware (1983) and Ware (1985) and in the book by Goldstein (1979).

We will use the following notation for the multivariate case. In the multivariate setting there are N independent vectors of observations. This may represent observations at more than one time point, or observations on more than one variable, or a combination of the two. The General Linear Multivariate Model (GLMM) is written as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1.1.2}$$

where

$\mathbf{Y} = [\ \mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_p\ ]$ is an N x p matrix of observations,

$\mathbf{X}$ is an Nxq design matrix of known constants,

$\boldsymbol{\beta} = [\ \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_p\ ]$ is a qxp matrix of parameters to be estimated,

$\boldsymbol{\epsilon}$ is an Nxp matrix of random error terms.

It is assumed that each $\mathbf{Y}_i$ satisfies the assumptions summarized by GLUM $(\mathbf{Y}_i;\ \mathbf{X}\boldsymbol{\beta},\ \sigma_{ii}\mathbf{I}_n)$ notation and further that

$$\text{Var}[\ \text{Row}_k(\mathbf{Y})\ ] = \boldsymbol{\Sigma}_{pxp}\ ,\ k=1,2,\ldots,N,$$

$$\text{Cov}[\ \text{Row}_i(\mathbf{Y})\ ,\ \text{Row}_j(\mathbf{Y})\ ] = \mathbf{0}_{pxp},\ i \neq j.$$

These assumptions imply that each row of $\mathbf{Y}$ is uncorrelated with every other row and that each row has the same covariance matrix. Also this model implies that each column of $\mathbf{Y}$, i.e. each variable or time point, has the same design matrix $\mathbf{X}$. These can be very restrictive assumptions.

The problem of missing data in longitudinal studies, however, may be more difficult. Many authors have proposed methods to handle this situation. These will be reviewed in detail in this chapter.

A missing data value is said to be missing at random (Rubin, 1976) if both of the following conditions are true:

(1) the probability that a data value is missing is functionally independent of the parameters of a correct model for the complete data,

(2) the event that a data value is missing is stochastically independent of either the value which would have been observed if the value were not missing or the value of any other observation.

Thus data missing at random might include observations missing due to an unrelated illness, laboratory errors, or termination of observations for reasons unrelated to the study, as when subjects move out of town. Data missing for reasons related to the outcomes of interest or to treatment will not be considered in this paper.

Rubin (1976) showed that when the data are missing at random it is valid to "ignore" the process that causes the missing data in the sense that the correct likelihood to be maximized is the marginal likelihood of the complete data.

The analysis of complete longitudinal data may be done using GLMM techniques as described in Davidson (1972). This method assumes that each subject has the same number of observations so that when data are missing special considerations must be made.

One common approach to the analysis of incomplete longitudinal data in both the univariate and the multivariate case has been to delete the entire observation vector (row of $\mathbf{Y}$) on any subject with missing data ("case-wise deletion", see Timm, 1970) and then use traditional complete data methods. This may make the analysis computationally feasible but "wastes" data. If a large proportion of subjects have missing data this is clearly an unreasonable approach.

Another approach is to "impute" missing data values, i.e., to calculate values to "fill in" the data using mean values or regression estimates. These methods are discussed by Wilkinson (1958) and Afifi and Elashoff (1966,1967,1969 a,b). These two "ad hoc" methods should only be used when the proportion of missing data is small. This is often not the case with longitudinal data and so it is necessary to have more precise methods of analysis.

1.1.2 Analysis of Incomplete Longitudinal Data Using the General Linear Multivariate Model

Kleinbaum (1970, 1973a) generalized the above multivariate linear model (GLMM) to four different types of models which include various combinations of missing data and different design matrices for the responses. He calculated BAN (Best Asymptotically Normal) estimators for $\boldsymbol{\beta}$ using weighted least squares. These were then used to construct BAN estimates of secondary parameters and Wald statistics for testing general linear hypotheses. A problem with this method is that the covariance matrix, Σ, is not known and must be estimated. Kleinbaum discussed this problem and proposed some solutions.

Kleinbaum (1970,1973b) also studied weighted least squares (WLS) estimators for the growth curve model of Potthoff and Roy (1964) with missing data and showed that under general conditions the WLS estimators are BAN and corresponding test statistics are Wald tests.

Maximum likelihood techniques have been developed for situations in which the data may be assumed to follow a multivariate normal distribution. An early example of maximum likelihood estimation (MLE) when observations are missing was given by Anderson (1957). Anderson considered the case of a bivariate normal distribution in which n observations were made on the pair (x,y) and N-n observations were made on x alone, so that N-n observations on y are missing. Anderson showed that the bivariate density of x and y can be written as the marginal density of x times the conditional density of y given x. The likelihood formed from this representation of the density can then be factored in such a way that one part is used to derive estimates of E(x) and V(x) and the other part is used to derive estimates of functions of E(y) and V(y). Afifi and Elashoff (1966a) demonstrated how this method could be extended to multivariate normal distributions with more variables.

Hocking and Smith (1968) developed an iterative technique to find estimators for a model in which the complete data are: $[\text{Row}_k(\mathbf{Y})]' \sim \text{NID}_p(\mathbf{O},\boldsymbol{\Sigma})$. Their method consisted of three steps:

1) Data are divided into groups according to which variates are missing. The likelihood function is then partitioned in a similar manner.
2) Initial MLEs of the parameters are obtained from the likelihood associated with the group of observations with no missing variates.
3) Initial estimates are modified by adjoining, optimally, the information in the remaining groups in a sequential manner until all the data are used.

Hocking and Smith assumed that the only parameters of interest were the variances and covariances, and that all means were zero. Although they did not demonstrate that their final estimates were maximum likelihood except in the case of just two groups, they did show that their final estimates had large sample properties similar to MLEs.

Hartley and Hocking (1971) also used the technique of partitioning the data into groups based on the patterns of missing data. They assumed that the complete data would satisfy $[\text{Row}_k(\mathbf{Y})] \sim \text{NID}_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$, $k=1,2,\ldots,N$. In the t-th group the data are assumed to be distributed $N(\boldsymbol{\mu}_t,\boldsymbol{\Sigma}_t)$ where $\boldsymbol{\mu}_t = \mathbf{D}_t\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_t = \mathbf{D}_t\boldsymbol{\Sigma}\mathbf{D}_t'$ and $\mathbf{D}_t$ is a subset of rows of an identity matrix which specifies which observations are recorded. Hartley and Hocking initially assumed no structure on $\boldsymbol{\mu}$ and Σ.

The likelihood equations were derived for each group and then combined to form the total likelihood using the fact that the total likelihood is equal to the product of the group likelihoods. In general, the likelihood equations must be solved using an iterative process. Hartley and Hocking showed how this method could be extended to the case where a structure is imposed on $\boldsymbol{\mu}$, as long as the parameters are estimable within each group.

Woolson, Leeper, and Clarke (1978) also used the technique of partitioning the data into groups based on the patterns of missing data. The data vector is assumed to be distributed $N(\mathbf{X}\boldsymbol{\beta},\Sigma)$. Following the weighted least squares methods of Kleinbaum (1970,1973), the authors showed that the parameters $\boldsymbol{\beta}$ could be estimated and general linear hypotheses about $\boldsymbol{\beta}$ could be tested. They noted that this method is a generalization of the method used by Rao and Rao (1966) for the analysis of the lonked cross-sectional design which is described in Section 1.3.

If it is assumed that the covariance within each group has a definite structure, $\sigma^2\boldsymbol{\Gamma}_j$, where $\boldsymbol{\Gamma}_j$ is known or can be estimated, one can use weighted least squares or, equivalently, a transformation can be made such that the the transformed responses are independently distributed with variance σ^2.

To illustrate these techniques the authors used the LCS design described in Section 1.3. A concern of the authors is the use of a sample estimate for $\boldsymbol{\Gamma}_j$. The small sample behaviour of the test statistic using estimates is not known. Further is is possible that in some cases the estimate of $\mathbf{\Gamma}_j$ may not be positive definite. This problem was also considered by Kleinbaum (1973). Schwertman and Allen (1979) used techniques for finding a positive semidefinite estimate which is 'near' $\mathbf{\Gamma}_j$.

An alternative maximum likelihood approach was developed by Orchard and Woodbury (1972). This approach is based on their "missing information principle." Let $\mathbf{Y}$ represent an hypothetical complete data matrix. This can be partitioned into two parts: $\mathbf{Y}_p$, the observed data and $\mathbf{Y}_m$, the unobserved (missing) data. The missing information principle states that the missing data, $\mathbf{Y}_m$, are considered to be random variables. Using this principle it can be shown that the likelihood to be maximized is a function of the conditional distribution of the complete data given the observed data.

The fundamental relationships of the missing information principle, as given by Woodbury in a comment on Hartley and Hocking (1971), are the following:

1) the total information is equal to the information contained in the observed data plus the lost information contained in the unobserved data;
2) the score for the observed data may be obtained by taking the conditional expectation of the score for the complete data with the missing values being regarded as random variables.

The advantage of the Orchard and Woodbury approach is that very often the likelihood function for the complete data given the observed data is more readily obtainable then the likelihood for the observed data.

Computationally, the Orchard and Woodbury approach consists of three steps. Assume that the data are a random sample from a $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$ distribution. Some data vectors may be incomplete because of missing observations which are missing at random.

1) Estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using those data vectors that are complete,
2) Use these estimates, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ to estimate the missing data,
3) Recompute the estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using the "completed data", i.e., the observed data complemented with the estimated missing data.

The algorithm cycles between steps 2 and 3 until the estimates converge.

In numerical analysis this technique is considered to be a member of the class of algorithms called "methods of successive substitution". In this case, however, the method arises from statistical, not computational, considerations, and produces MLEs.

Dempster, Laird, and Rubin (1977) generalized the method of Orchard and Woodbury in the EM algorithm. The EM algorithm consists of two steps: an expectation step (E-step) and a maximization step (M-step). The E-step estimates sufficient statistics of the complete data conditional on the observed data. The M-step produces maximum likelihood estimates of the parameters using the sufficient statistics determined in the E-step. The algorithm cycles between these steps until the estimates converge.

Dempster et. al. (1977) attempted to prove that the EM algorithm will converge to a single local maximum of the likelihood function but Wu (1983) and Boyles (1983) noted an error in this "proof". Boyles (1983) corrected the proof that the EM algorithm will converge to a member of closed and bounded set of local maximum of the likelihood function, which may or may not consist of a single point. This proof is based upon the important characteristic that at each step the EM algorithm increases (or at least does not decrease) the log likelihood of the parameters given the observed data.

An attractive aspect of the EM algorithm is that the E-step and M-step are often easy to compute due to the nice form of the complete-data likelihood function (Wu, 1983). Solutions of the M-step often exist in closed form. Also the algorithm often does not require large computer storage space compared, say, to the method of scoring. The EM algorithm does not directly produce an estimate of the covariance matrix of the estimators but for the models considered here the asymptotic covariances are easily computed.

In practice the EM algorithm has been found to require a large number of iterations and step sizes may be particularly small in a relatively large neighborhood about the MLE, sometimes leading one to conclude that the algorithm has converged when in fact it has not (Jennrich and Schluchter, 1985).

1.1.2 Analysis of Incomplete Longitudinal Data Using the Mixed Model

The following notation will be used for the mixed model. The mixed model contains both fixed and random effects. Consider the mixed model given by

$$\mathbf{Y}_i = \mathbf{A}_i\mathbf{\Phi} + \mathbf{B}_i\mathbf{d}_i + \mathbf{e}_i \tag{1.1.3}$$

where $\mathbf{Y}_i$ is an n_ix1 vector of n_i observations on the ith subject,

$\mathbf{\Phi}$ is a px1 vector of unknown, constant, population parameters,

$\mathbf{A}_i$ is an $n_i$xp known, constant design matrix,

$\mathbf{d}_i$ is a qx1 vector of unknown, random individual parameters,

$\mathbf{B}_i$ is an $n_i$xq known, constant, within-subject design matrix corresponding to the random effects $\mathbf{d}_i$,

$\mathbf{e}_i$ is a n_ix1 vector of random error terms.

The following assumptions are made:

$\mathbf{d}_i \sim \mathrm{NID}(\mathbf{O},\mathbf{D})$ independent of $\mathbf{e}_i \sim \mathrm{NID}(\mathbf{O},\sigma^2_w\mathbf{V}_i)$, so that

$$\mathrm{Var}\begin{bmatrix}\mathbf{d}_i\\ \mathbf{e}_i\end{bmatrix} = \begin{bmatrix}\mathbf{D} & \mathbf{O}\\ \mathbf{O} & \sigma^2_w\mathbf{V}_i\end{bmatrix}$$

were $\mathbf{D}$ is qxq positive-definite, symmetric covariance matrix of the random effects, $\mathbf{d}_i$, $\mathbf{V}_i$ is a known, constant positive-definite matrix, and σ^2_w is the scalar within subject variance parameter.

Thus $E(\mathbf{Y}_i) = \mathbf{A}_i\mathbf{\Phi}$ and $Var(\mathbf{Y}_i) = \mathbf{\Sigma}_i = \mathbf{B}_i\mathbf{D}\mathbf{B}_i' + \sigma^2_w\mathbf{V}_i$ where $\mathbf{\Sigma}_i$ is an $n_i x n_i$ positive-definite, symmetric covariance matrix of $\mathbf{Y}_i$. Further it may be assumed that $\mathbf{D}$ has a linear structure given by

$$\mathbf{D} = \sum_{g=1}^{m-1} \tau_g \mathbf{G}_g \qquad (1.1.4)$$

where each $\mathbf{G}_g$ is a known constant matrix. Thus, letting $\tau_m = \sigma^2_w$ and $\mathbf{G}_{im} = \mathbf{V}_i$, $\mathbf{\Sigma}_i$ has the following structure:

$$\begin{aligned}\mathbf{\Sigma}_i &= \mathbf{B}_i\mathbf{D}\mathbf{B}_i' + \tau_m\mathbf{G}_{im}\\ &= \mathbf{B}_i\Big(\sum_{g=1}^{m-1} \tau_g\mathbf{G}_g\Big)\mathbf{B}_i' + \tau_m\mathbf{G}_{im}\\ &= \sum_{g=1}^{m-1} \tau_g(\mathbf{B}_i\mathbf{G}_g\mathbf{B}_i') + \tau_m\mathbf{G}_{im}\\ &= \sum_{g=1}^{m} \tau_g\mathbf{G}_{ig} \qquad (1.1.5)\end{aligned}$$

where $\mathbf{G}_{ig} = \mathbf{B}_i\mathbf{G}_g\mathbf{B}_i'$, $g=1,\ldots,m-1$, and $\mathbf{G}_{im} = \mathbf{V}_i$. It is assumed that $\mathbf{G}_{i1},\ldots,\mathbf{G}_{im}$ are linearly independent and the values of $\tau_1,\ldots,\tau_m$ are such as to make $\mathbf{\Sigma}_i$ positive definite.

An important feature of the mixed model described above is that both $E(\mathbf{Y}_i)$ and $Var(\mathbf{Y}_i)$ are modelled in terms of a "small" number of parameters, $\mathbf{\Phi}$ and $\boldsymbol{\tau} = (\tau_1,\ldots,\tau_m)$; but each subject is allowed to have unique design matrices $\mathbf{A}_i$ and $\mathbf{B}_i$. Because of this the model can accomodate time-dependent covariates and missing and mistimed data. The model also permits different "groups" to have different subsets of the parameters via the use of indicator variables in $\mathbf{A}_i$ and $\mathbf{B}_i$.

The linear covariance structure described here allows the special correlation structure of many types of longitudinal data to be accomodated. Covariance matrices which can be written this way are discussed by Anderson (1970,1973).

An example of the use of the mixed model in analyzing longitudinal data is given by Fairclough and Helms (1984). The data were taken from a longitudinal study of lung function during childhood. The response of interest was Forced Vital Capacity (FVC), the volume of gas expired after inhaling. It is assumed that FVC is linearly related to height over the range of heights in the data, (102-145 cm), but that the slope and intercept vary from child to child. The data set analyzed consisted of 956 measurements from 72 children. The average number of measurements per child was 13 with a minimum of 2 and a maximum of 33. Four demographic groups were studied, black females, black males, white females and white males.

The fixed effects ($\mathbf{\Phi}$) were the intercept and slope of the population "growth curve" (a straight line function of height) and the random effects ($\mathbf{d}_i$) due to each child were increments to the intercept and slope . It was assumed that the covariance matrix had a linear structure as given in (2.5.3). A model was initially fit which allowed for different parameters for each of the four demographic groups. The parameters were estimated using maximum likelihood methods. The details will be discussed later in this section.

A variety of hypotheses were tested using classic likelihood ratio test statistics and asymptotic chi-square approximations (Andrade and Helms, 1984). These included hypotheses concerning the elements of **D** and hypotheses concerning the differences between sex and race groups.

Based on the results of these hypothesis tests reduced models were fit to the data. The final model, accepted as the "best" model for the data, contained different estimates of the fixed effects ($\mathbf{\Phi}$) for the two race groups but common estimates of the between and within-subject variance components (**D** and σ^2).

Laird and Ware (1982) also used the mixed model to analyze longitudinal data. The data were from a longitudinal study of the effects of atmosphere pollutants on pulmonary function. The response of interest was FEV_1, the volume of air exhaled in the first second of a forced exhalation. Approximately 200 school children were examined under normal conditions, then during an air pollution alert and on three successive weeks following the alert. One objective of the analysis was to determine whether FEV_1 was depressed during the alert.

A mixed model was fit in which the fixed effects ($\mathbf{\Phi}$) were the population mean values of FEV_1, on each of the five measurements. The random effect ($\mathbf{d}_i$) was the random deviation in average value of FEV_1 for the i-th child. The model was fit using methods which are discussed later in this section. It was found the FEV_1 declined on and after the alert day and that variances and covariances for the last four measurements were larger than those involving the baseline day.

Another objective of the analysis was to identify sensitive children most severely affected by the pollution episode. This was done by introducing a second random effect to quantify the average decline in FEV_1 for each child, $\mathbf{d}_i = (d_{1i}, d_{2i})$. The estimates of d_{2i} were used to identify those children who showed greatest declines.

The model was then expanded to study the influence of sex, race, location of residence and other individual characteristics on response. Laird and Ware also gave an example in which a "growth curve type" model was fit. This model was similar to those used by Fairclough and Helms (1984) except that Laird and Ware do not impose any structure on $\mathbf{D}$, besides assuming that it is a positive-definite covariance matrix.

1.1.4 Maximum Likelihood Estimation for the Mixed Model

Harville (1977) reviewed the maximum likelihood approach to estimation of parameters in a mixed model. He noted that under general regularity conditions the maximum likelihood estimators (MLE) had the desirable properties of being consistent, asymptotically normal and efficient.

When the assumptions of normality are made the log likelihood function for the model (1.1.3) is

$$\log L(\mathbf{\Phi},\mathbf{\Sigma}_i) = -\tfrac{1}{2}\left[N \log(2\pi) + \sum_{i=1}^{k} \left\{ \log|\mathbf{\Sigma}_i| + \operatorname{tr} \mathbf{\Sigma}_i^{-1}\mathbf{C}_i \right\} \right] \qquad (1.1.6)$$

where $\mathbf{C}_i = \left\{ \mathbf{Y}_i-\mathbf{A}_i\mathbf{\Phi} \right\} \left\{ \mathbf{Y}_i-\mathbf{A}_i\mathbf{\Phi} \right\}'$, n_i observations are made on the i-th subject, $N = \sum_{i=1}^{k} n_i$, and k = total number of subjects.

For all of the cases mentioned above the maximum likelihood estimate of $\mathbf{\Phi}$ is given by

$$\hat{\mathbf{\Phi}} = \left[\sum_{i=1}^{k} \mathbf{A}_i' \hat{\mathbf{\Sigma}}_i^{-1} \mathbf{A}_i \right]^{-1} \sum_{i=1}^{k} \mathbf{A}_i \hat{\mathbf{\Sigma}}_i^{-1} \mathbf{Y}_i \ . \qquad (1.1.7)$$

Laird and Ware (1982) did not discuss the asymptotic behavior of the estimates, but Andrade and Helms (1984) showed that under the assumption of normality, plus mild regularity conditions on the way the n_i tend to infinity, $\hat{\mathbf{\Phi}}$ is asymptotically normal with mean $\mathbf{\Phi}$ and variance

$$\left[\sum_{i=1}^{k} \mathbf{A}_i \mathbf{\Sigma}_i^{-1} \mathbf{A}_i \right]^{-1} .$$

An estimate of the realized value of the random term $\mathbf{d}_i$ is given by Laird and Ware (1982) as:

$$\hat{\mathbf{d}}_i = \hat{D}\, \mathbf{B}_i' \hat{\mathbf{\Sigma}}_i^{-1} (\mathbf{Y}_i-\mathbf{A}_i\hat{\mathbf{\Phi}}) \ . \qquad (1.1.8)$$

This is not maximum likelihood but can be derived using an extension of the Gauss-Markov theorem to cover random effects (Harville, 1976).

For the covariance structure used by Laird and Ware the maximum likelihood estimators of σ^2 and $\mathbf{D}$ are given by

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{k} \operatorname{tr} \hat{\Sigma}_i^{-1} (\hat{\mathbf{C}}_i - \mathbf{B}_i \hat{\mathbf{D}} \mathbf{B}_i') \hat{\Sigma}_i^{-1}}{\sum_{i=1}^{k} \operatorname{tr} \hat{\Sigma}_i^{-1} \hat{\Sigma}_i^{-1}} , \tag{1.1.9}$$

$$\hat{\mathbf{D}} = \left(\sum_{i=1}^{k} \mathbf{B}_i' \hat{\Sigma}_i^{-1} \mathbf{B}_i\right)^{-1} \left(\sum_{i=1}^{k} \mathbf{B}_i' \hat{\Sigma}_i^{-1} (\hat{\mathbf{C}}_i - \hat{\sigma}^2 \mathbf{I}) \hat{\Sigma}_i^{-1} \mathbf{B}_i\right) \left(\sum_{i=1}^{k} \mathbf{B}_i' \hat{\Sigma}_i^{-1} \mathbf{B}_i\right)^{-1}, \tag{1.1.10}$$

$$\hat{\Sigma}_i = \mathbf{B}_i \hat{\mathbf{D}} \mathbf{B}_i' + \hat{\sigma}^2 \mathbf{I} . \tag{1.1.11}$$

Laird and Ware did not discuss the asymptotic behavior of the estimators.

For the case where a linear structure is imposed on **D**, Andrade and Helms (1984) and Fairclough and Helms (1984) showed that the maximum likelihood estimator for $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ based on Anderson's work (1970,1973) is given by

$$\hat{\boldsymbol{\tau}} = \left[\langle \sum_{i=1}^{k} \mathrm{tr}\, \hat{\Sigma}_i^{-1} G_{ig} \hat{\Sigma}_i^{-1} G_{ih} \rangle_{gh} \right]^{-1} \left[\langle \sum_{i=1}^{k} \mathrm{tr}\, \hat{\Sigma}_i^{-1} G_{ig} \hat{\Sigma}_i^{-1} \hat{C}_i \rangle_{g} \right] \tag{1.1.12}$$

where $\hat{C}_i = \left\{ \mathbf{Y}_i - \mathbf{A}_i \hat{\boldsymbol{\Phi}} \right\} \left\{ \mathbf{Y}_i - \mathbf{A}_i \hat{\boldsymbol{\Phi}} \right\}'$.

Here $[\langle\ \rangle_{gh}]$ is an mxm matrix whose (g,h)-th element is the expression within $\langle\ \rangle$ and $[\langle\ \rangle_{g}]$ is an mx1 vector whose g-th element is the expression within $\langle\ \rangle$.

Andrade and Helms (1984) showed that under the assumptions above, plus mild regularity conditions on the way the n_i tend to infinity, $\hat{\boldsymbol{\tau}}$ asymptotically follows a normal distribution with mean $\boldsymbol{\tau}$ and covariance matrix:

$$2\left[\langle \sum_{i=1}^{k} \mathrm{tr}\, \Sigma_i^{-1} G_{ig} \Sigma_i^{-1} G_{ih} \rangle_{gh} \right]^{-1}.$$

They also showed that $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\tau}}$ are asymptotically independent.

Fairclough and Helms (1984) used the following expressions, based on the results cited above, as asymptotic estimates of variance matrices for the model with linear covariance structure:

$$\text{Asm. } \hat{\text{Var}}\,(\hat{\boldsymbol{\Phi}}) = \left[\sum_{i=1}^{k} \mathbf{A}_i \hat{\Sigma}_i^{-1} \mathbf{A}_i \right]^{-1} \tag{1.1.13}$$

$$\text{Asm. } \hat{\text{Var}}\,(\hat{\tau}) = 2\left[\ \langle \sum_{i=1}^{k} \text{tr}\,(\ \hat{\Sigma}_i^{-1} G_{ig} \hat{\Sigma}_i^{-1} G_{ih}\ ) \rangle_{gh}\ \right]^{-1}. \qquad (1.1.14)$$

For the general covariance structure considered by Jennrich and Schluchter (1985) the maximum likelihood estimate of Σ will depend on the specific structure assumed but, in general, is obtained by solving the following likelihood equations

$$\sum_{i=1}^{k} \text{tr}\,(\ \Sigma_i^{-1} (\ C_i - \Sigma_i)\ \Sigma_i^{-1})\ \frac{\delta \Sigma_i}{\delta \theta_r} = 0 \qquad (1.1.15)$$

In this case the elements of Σ are considered to be known functions of q unknown covariance parameters contained in the vector $\boldsymbol{\theta}$.

Laird and Ware (1982) discussed the use of Restricted Maximum Likelihood (REML) estimators. They noted that in "balanced ANOVA models, ML estimates of variance components fail to take into account the degrees of freedom lost in estimating the fixed effects and are thus biased downwards". In some models REML estimates are unbiased. Laird and Ware showed that a Bayesian formulation of the mixed model (1.1.3) yields REML estimates for the variance components. Further discussion of REML estimators can be found in Harville (1977). REML estimators will not be considered in the research proposed here.

The likelihood equations for each of the models mentioned above are sets of simultaneous nonlinear equations, one equation per parameter, which require iterative techniques for solution. Judge, et al. (1980, Chap. 17) presented a general summary of algorithms for simultaneous nonlinear equations.

To solve the equations Laird and Ware (1982) used the EM algorithm, Fairclough and Helms (1984,1985) used the EM algorithm and the Method of Scoring, and Jennrich and Schluchter (1985) used the EM algorithm, the Method of Scoring and the Newton Raphson method.

The Newton Raphson method and the Method of Scoring are both gradient algorithms based on second order derivatives. The forms for the r-th iterate in terms of the (r-1)-th iterate are:

Newton Raphson: $$\boldsymbol{\theta}^{(r)} = \boldsymbol{\theta}^{(r-1)} + \left[-\frac{\delta^2 \log L}{\delta\theta_g \delta\theta_h} \right]^{-1} \frac{\delta \log L}{\delta\theta_h} \qquad (1.1.16)$$

Method of Scoring: $$\boldsymbol{\theta}^{(r)} = \boldsymbol{\theta}^{(r-1)} + \left[E \left[-\frac{\delta^2 \log L}{\delta\theta_g \delta\theta_h} \right] \right]^{-1} \frac{\delta \log L}{\delta\theta_h} \qquad (1.1.17)$$

Because of the asymptotic independence of $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\tau}}$ the information matrix used in the Method of Scoring will have large zero submatrices and the inverse can be computed piecewise. The corresponding matrix in the Newton Raphson method will generally have nonzero elements in those positions and will require the inverse of a much larger matrix at at each iteration. Thus, in general, the Method of Scoring will require less computer time per iteration.

Harville (1977) reviewed these two algorithms and discussed two problems which may arise: 1) the algorithms are sensitive to poor starting values and may converge to local maxima, and 2) the procedures may converge to maxima outside the boundaries constraining $\boldsymbol{\Sigma}_i$ to be positive definite.

The EM algorithm has been described previously. To use the EM algorithm to fit the mixed model both Laird and Ware (1982) and Fairclough and Helms (1984) treated the $\mathbf{d}_i$ as if they were observable but missing. Jennrich and Schlucter, however, used the EM algorithm to estimate missing data and then computed sufficient statistics from the completed data. The difference between these two approaches can be considerable. Using the first approach, at each iteration the sufficient statistics are computed while using the second approach the missing data is estimated at each iteration. Thus the computer time for each iteration may be much larger using the second method.

Laird and Ware (1982) noted that in their use of the EM algorithm it is often slow to converge. They suggested that the problem of slow convergence may be especially severe if the maximum likelihood occurs on or near a boundary of the parameter space.

Fairclough and Helms (1985) also found that in many cases the EM algorithm failed to converge within a reasonable number of iterations. They found the Method of Scoring to be computationally more efficient. However, when the covariance matrix of the random effects, **D**, was nearly singular the Method of Scoring sometimes produced non-positive definite estimates of **D** while the EM estimates of **D** were structurally positive definite. In all other cases the two methods converged to the same set of parameter estimates.

Jennrich and Schluchter (1985) summarized their use of the three algorithms as follows:

"Direct comparison of the Newton, scoring, and generalized EM algorithms in terms of required computation are difficult, because this depends to a large degree on how efficiently the algorithms are coded. The Newton algorithm, with a quadratic convergence rate, generally converges in a small number of iterations, with a sometimes high cost per iteration. On the other extreme, generalized EM has the lowest cost per iteration but at times requires a large number of iterations. The scoring algorithm is intermediate in terms of cost per iteration and required number of iterations. However, the cost per iteration of scoring is often not much less than that of the Newton algorithm, whereas the scoring algorithm sometimes requires a considerably higher number of iterations than does the Newton algorithm. Because the scoring algorithm is often more robust to poor starting values than is the Newton algorithm, a good compromise is to start with scoring, and switch to the Newton algorithm after the first several steps.

"When the number of covariance parameters q is not large, any of the algorithms will generally be satisfactory, but we prefer the Newton algorithm because of its generally clean and fast convergence. With large q, as when fitting a large unstructured covariance matrix ($t > 10$ say), the GEM algorithm is the only feasible algorithm because the cost per iteration of the Newton and scoring algorithms becomes excessive.

"Another consideration is that the Newton and scoring algorithms produce standard errors based on the empirical and Fisher information matrices as a by-product. The generalized EM algorithm gives no standard errors, but these can be obtained by taking a single Newton or scoring step after convergence."

1.2 Power Calculations

1.2.1 Approximations for the Multivariate Case

Because the exact distribution of many of the test statistics used in the univariate case are known, power calculations for this case are relatively straightforward. This topic has been discussed in many texts including Cochran and Cox (1957), Cox (1958), Scheffe (1959) and Searle (1971).

In the multivariate case the exact noncentral distributions of the usual test statistics are generally too complicated for practical power calculations, except for special cases. Several strategies for approximating the exact distributions or for using other distributions to estimate power have been examined in the literature.

Box (1954a,b), Greenhouse and Geisser (1959), Huynh and Feldt (1970 and 1976) and Huynh (1978) examined the distribution of the F statistics in univariate ANOVA when the data are correlated, as in data from a repeated measures design. Using the results from Box, Greenhouse and Geisser showed that the F-statistic for testing the significance of factor A has an approximate F distribution, where A is the factor over which repeated measurements are made. The degrees of freedom for this approximate distribution are adjusted downward by a correction factor, ϵ, which is a function of the elements of the population variance-covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$. The formula is

$$\epsilon = p^2 \,(\sigma_{ii} - \sigma_{..})^2 \,/\, (p-1)\left(\sum_i\sum_j \sigma^2_{ij} - 2p\sum_j \sigma^2_{j.} + p^2\sigma^2_{..}\right) \qquad (1.2.1)$$

where p is the number of levels of factor A (i.e. the number of repeated measures), σ_{ij} are the elements of Σ, σ_{ii} is the mean of the diagonal terms, $\sigma_{i.}$ is the mean of the ith row (or ith column), and $\sigma_{..}$ is the grand mean.

Since the population variance-covariance matrix is almost never known, ϵ must be estimated from the sample variance-covariance matrix. The effect of using an estimated ϵ on the approximate F distributions is unknown. Therefore, unless Σ is estimated with a large number of degrees of freedom, Geisser and Greenhouse suggested the use of a conservative test. The conservative test is based on the minimum possible value of ϵ, which does not depend on the elements of the variance matrix. The minimum value is $(p-1)^{-1}$ where p is, as above, the number of repeated measures. This test is conservative since the minimum value of ϵ gives the maximum reduction in degrees of freedom.

Greenhouse and Geisser (1959) presented the F-test "corrections" as tests of convenience to avoid the laborious computations of a MANOVA analysis. They did not compare the power of their approximate tests to the power of the exact MANOVA test statistics.

Huynh and Feldt (1970) examined the conditions under which the test statistics from ANOVA have exact F distributions when the data are correlated. They found that a necessary and sufficient condition for the relevant test statistics to have exact F distributions is that the elements of the covariance matrix, $\Sigma = (\sigma_{ij})$, have the following linear structure:

$$\sigma_{ij} = \alpha_i + \alpha_j + \lambda\delta_{ij}\,, \quad \text{where } \lambda>0 \text{ and } \delta_{ij} = 1 \text{ if } i=j, \\ = 0 \text{ if } i\neq j.$$

This condition implies that all possible differences have equal variance and that the correction factor, ϵ, is unity. Huynh and Feldt gave a test for this condition based on Mauchley's Sphericity Criterion, W (Mauchley, 1940). The necessary and sufficient condition given above and the test for this condition are given initially for the randomized block design and then extended to the split-plot design.

Huynh and Feldt (1976) also derived an estimator, $\tilde{\epsilon}$, for the Geisser-Greenhouse correction factor, ϵ. They compared it to the Geisser-Greenhouse estimator, $\hat{\epsilon}$,based on the sample variance-covariance matrix, $S = (s_{ij})$. The Huynh-Feldt estimator is

$$\tilde{\epsilon} = \frac{nA - 2B}{(p-1)\ [(n-1)B - A]} \tag{1.2.2}$$

where p = number of treatments,

n = number of blocks

$$A = p^2(s_{ii} - s_{..})^2 \tag{1.2.3}$$

$$\text{and } B = \sum_i\sum_j s^2_{ij} - 2k\sum_i s^2_{i.} + k^2 s^2_{..} \tag{1.2.4}$$

This can be written in terms of $\hat{\epsilon}$ as:

$$\tilde{\epsilon} = \frac{n(p-1)\hat{\epsilon} - 2}{(p-1)[n-1-(p-1)\hat{\epsilon}\,]} \tag{1.2.5}$$

Here $\tilde{\epsilon} \leq \hat{\epsilon}$,with equality holding when $\hat{\epsilon} = 1/(p-1)$, the minimum possible value. If $\tilde{\epsilon} > 1$ the estimator is set equal to 1.0. The difference between $\tilde{\epsilon}$ and $\hat{\epsilon}$ decreases as the number of blocks (n) increases. The two estimators were compared using a Monte Carlo study. It was found that $\tilde{\epsilon}$ is less biased then $\hat{\epsilon}$ in the range $0.75 < \epsilon < 1.00$. The opposite is true when $\epsilon < 0.50$.

Muller and Barton (1987) gave the following summary of the performance of the univariate F tests using the various estimates of ϵ. "No consensus exists as to which test to use. Simulation results indicate that usually the uncorrected test [F] is very liberal, giving far too many type I errors. Usually the Huynh-Feldt test [$\tilde{\epsilon}$] is somewhat liberal, with type I error rate as high as .10, often .05 - .07, for a nominal α of .05. In contrast, the Geisser-Greenhouse test [$\hat{\epsilon}$] usually results in a type I error rate equal to or less than .05 (for $\alpha = .05$). Finally, the Box test [Geisser-Greenhouse conservative test] is overly conservative, rarely allowing the type I error rate to be near α."

Muller and Barton also gave approximate power functions for these four tests. They expressed the tests within the framework of the

GLMM (2.3.2). The general linear hypothesis is H_o: $\mathbf{\Theta} = \mathbf{C}\boldsymbol{\beta}\mathbf{U} - \mathbf{\Theta}_o = \mathbf{0}$, where **C** is an axq matrix of known constants, **U** is a pxb matrix of known constants and $\mathbf{\Theta}_o$ is an axb matrix of known constants. The test statistic is

$$F = \frac{tr(\mathbf{H})/ab}{tr(\mathbf{E})/b(N-r)} \qquad (1.2.6)$$

where $\mathbf{H} = \hat{\mathbf{\Theta}}'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}')^{-1}\hat{\mathbf{\Theta}}$ and $\mathbf{E} = \mathbf{U}'\hat{\mathbf{\Sigma}}\mathbf{U}(N-r)$. When **U** is an orthonormal matrix and $\mathbf{U}'\mathbf{\Sigma}\mathbf{U} = \sigma^2\mathbf{I}$ the test given above is the uncorrected univariate F test and the test statistic follows an exact F distribution with ab and b(N-r) degrees of freedom.

When $\mathbf{U}'\mathbf{\Sigma}\mathbf{U} \neq \sigma^2\mathbf{I}$ the test statistic follows an approximate F distribution with $ab\epsilon$ and $b(N-r)\epsilon$ degrees of freedom, where ϵ is the correction factor given previously. In terms of GLMM notation ϵ can be written as:

$$\epsilon = \frac{tr^2(\mathbf{U}'\mathbf{\Sigma}\mathbf{U})}{b\,tr(\mathbf{U}'\mathbf{\Sigma}\mathbf{U}^2)} = \frac{(\sum_{k=1}^{b}\lambda_k)^2}{b(\sum_{k=1}^{b}\lambda_k^2)} \qquad (1.2.7)$$

where λ_k, k=1,2,...,b, are the ordered eigenvalues of $\mathbf{U}'\mathbf{\Sigma}\mathbf{U}$. Since $\mathbf{\Sigma}$, and thus ϵ, is not known the estimates discussed previously must be used.

If $\epsilon = 1$ then F exactly follows a noncentral F distribution under the alternative. For the case when $\epsilon < 1$ Muller and Barton proved that under the alternative hypothesis F follows approximately a noncentral F distribution with $ab\epsilon$ and $b(N-r)\epsilon$ degrees of freedom. The noncentrality parameter is ϵabF_a where F_a is the value of F computed using the population values of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.

An approximation for the power of the conservative test is gotten by using $\epsilon = 1/b$. Approximations for the Geisser-Greenhouse and Huynh-Feldt tests are gotten by substituting the expected values of $\hat{\epsilon}$ and $\tilde{\epsilon}$ for ϵ. The expected values of $\hat{\epsilon}$ and $\tilde{\epsilon}$ are not known exactly but Muller and Barton developed approximations for them based on asymptotic expansions.

Muller and Peterson (1984) provided approximations for the null and non-null distributions of three multivariate test statistics. These are Wilk's Lambda, Hotelling-Lawley trace and Pillai-Bartlett trace. The approximations are based on noncentral F distributions and generalize existing F approximations for the central case. Muller and Peterson compared the power approximations to existing approximations and exact cases and found the new approximations to be accurate to nearly two decimal places. They claimed that the new F approximations are "simpler to implement, simpler to understand and less expensive to compute."

Unfortunately none of the finite sample results given in this section are applicable to the mixed model with linear covariance structure. In the next section asymptotic results which do apply to the mixed model will be given.

1.2.2 Asymptotic Power Calculations for the Mixed Model with Linear Covariance Structure

Andrade and Helms (1984) presented tests for general linear hypotheses about the parameters of the mixed model with linear covariance structure (1.1.3-5). For each hypothesis of interest a likelihood ratio test and large sample properties were given.

The general linear hypothesis about the fixed effects can be written as H_o: $\mathbf{L'\Phi} = \mathbf{\Theta}_o$ vs. H_o: $\mathbf{L'\Phi} \neq \mathbf{\Theta}_o$, where $\mathbf{L'}$ is an lxp matrix with rank $l \leq p$ and $\mathbf{\Theta}_o$ is a known lx1 vector. The likelihood ratio test (Theorem 4.7) is given by

$$\lambda = \prod_{i=1}^{k} \left[\frac{|\hat{\mathbf{\Sigma}}_i|}{|\hat{\mathbf{\Sigma}}_{oi}|} \right] \tag{1.2.8}$$

where $\hat{\mathbf{\Sigma}}_i$ and $\hat{\mathbf{\Sigma}}_{oi}$ are the MLEs of $\mathbf{\Sigma}_i$ under the null and alternative hypotheses. The null hypothesis is rejected if λ is too small (Theorem 4.8).

Andrade and Helms (1984) showed that under suitable regularity conditions the asymptotic distribution of $-2\log\lambda$ is χ^2_l when H_o is true. Under a sequence of "local" alternatives (Theorem 4.9) H_a: $\mathbf{L'\Phi} - \mathbf{\Theta}_o = N^{-\frac{1}{2}}\mathbf{\Delta}$ where $\mathbf{\Delta}$ is an lx1 fixed vector, the asymptotic distribution of $-2\log\lambda$ is $\chi^2_l(a)$, a noncentral χ^2 with l degrees of freedom and noncentrality parameter

$$a = \boldsymbol{\Delta}'(\mathbf{L}'(\sum_{i=1}^{k}(1/N)\mathbf{A}_i'\Sigma_i^{-1}\mathbf{A}_i)^{-1}\mathbf{L})^{-1}\boldsymbol{\Delta}\,, \qquad (1.2.9)$$

where k is the total number of subjects.

The general linear hypothesis about the random effects can be written as H_o: $\mathbf{S}'\boldsymbol{\tau} = \boldsymbol{\gamma}_o$ vs. H_a: $\mathbf{S}'\boldsymbol{\tau} \neq \boldsymbol{\gamma}_o$ where $\mathbf{S}'$ is a sxm matrix of rank s≤m and $\boldsymbol{\gamma}_o$ is a sx1 known vector. The likelihood ratio test is given by (Theorem 4.10)

$$\lambda = \left[\prod_{i=1}^{k}\frac{|\hat{\Sigma}_i|}{|\hat{\Sigma}_{oi}|}\right]\exp(\hat{\mathbf{V}}) \qquad (1.2.10)$$

where $\hat{\mathbf{V}} = \boldsymbol{\gamma}_o'(\,\mathbf{S}'\hat{\mathbf{Z}}^{-1}\mathbf{S}\,)^{-1}(\mathbf{S}'\hat{\mathbf{Z}}^{-1}\hat{\mathbf{z}} - \boldsymbol{\gamma}_o)$

$$\hat{\mathbf{Z}} = \left[\,\langle\sum_{i=1}^{k}\mathrm{tr}\ \hat{\Sigma}_i^{-1}G_{ig}\hat{\Sigma}_i^{-1}G_{ih}\rangle_{gh}\,\right]$$

$$\hat{\mathbf{z}} = \left[\,\langle\sum_{i=1}^{k}\mathrm{tr}\hat{\Sigma}_i^{-1}G_g\hat{\Sigma}_i^{-1}\hat{C}_i\rangle_g\,\right]$$

and $\hat{\Sigma}_i, \hat{\Sigma}_{io}$ are the MLEs of Σ_i under the null and alternative hypotheses. The null hypothesis is rejected if λ is too small (Theorem 4.11).

Andrade and Helms (1984) showed that under suitable regularity conditions the asymptotic distribution of $-2\log\lambda$ is χ^2_s when H_o is true. Under a sequence of local alternatives (Theorem 4.12) H_a: $\mathbf{S}'\boldsymbol{\tau} - \boldsymbol{\gamma}_o = N^{-\frac{1}{2}}\boldsymbol{\Delta}$, where $\boldsymbol{\Delta}$ is a fixed sx1 vector, the asymptotic distribution of $-2\log\lambda$ is $\chi^2_s(a)$, a noncentral χ^2 with s degrees of freedom and noncentrality parameter parameter

$$a = 2\boldsymbol{\Delta}'\left\{ \mathbf{S}'\left[\langle\sum_{i=1}^{k} (1/N)\ \mathrm{tr}\boldsymbol{\Sigma}_i^{-1}\mathbf{G}_{ig}\boldsymbol{\Sigma}_i^{-1}\mathbf{G}_{ih}\rangle_{gh} \right]^{-1}\mathbf{S}\right\}^{-1}\boldsymbol{\Delta}\ . \quad (1.2.11)$$

To test the composite hypothesis H_o: $\mathbf{L}'\boldsymbol{\beta} = \boldsymbol{\Theta}_o$ and $\mathbf{S}'\boldsymbol{\tau} = \boldsymbol{\gamma}_o$ vs. H_a: $\mathbf{L}'\boldsymbol{\beta} \neq \boldsymbol{\Theta}_o$ and/or $\mathbf{S}'\boldsymbol{\tau} \neq \boldsymbol{\gamma}_o$ Andrade and Helms stated that "the LR statistic and its asymptotic null and non-null distributions can easily be obtained from the results" presented in their paper.

1.3 Designs for Longitudinal Studies

The design of an epidemiologic or clinical study is critical, not only in determining the quality of the data collected but also in determining how the data will be analyzed and what inferences may be drawn from the study. It is essential that the design selected for a study allow the effects of interest to be measured efficiently and with a minimum of time, effort and cost.

This chapter of the literature review will discuss the advantages and disadvantages of the longitudinal design as compared to the cross-sectional design. The linked cross-sectional design (LCS) design will be described and some of its properties will be discussed.

1.3.1 Comparison of Cross-Sectional and Longitudinal Designs

Two general classes of designs which have historically been used in epidemiologic and clinical research are the cross-sectional design and the longitudinal design. In this paper cross-sectional designs will be defined as those in which subjects are observed at only one time point. Longitudinal designs will be defined as those in which subjects are observed at more than one time point.

A cross-sectional study and a longitudinal study are compared pictorially in Figures 1.3.1 and 1.3.2, which illustrate a study of the growth of school children in grades 5 through 11. Figure 1.3.1 illustrates a crosssectional study of all seven grades at one time and

Figure 1.3.2 illustrates a traditional, single-cohort longitudinal study. In these Figures, cohort, defined by year of birth, is given on the vertical axis; age, represented by school year, is given on the upper horizontal axis; and period, or time-of-measurement, is indicated in the lower margin. As can be seen, the cross-sectional study contains seven cohorts observed at one time point while the longitudinal study follows one cohort over seven measurement periods.

Some of the advantages and disadvantages of the two designs are immediately apparent. The cross-sectional design in Figure 1.3.1 can be completed in one year, or one measurement period, while the longitudinal study in Figure 1.3.2 will take 7 years to complete. The longer time required for the longitudinal study implies, other factors being equal, that the study will probably be more expensive and more more difficult to implement and monitor. Also the likelihood of subjects dropping out or missing appointments will increase. Baltes (1968) discussed some of the potential problems of longitudinal designs including selective survival and drop-out and testing effects.

In spite of these problems longitudinal designs are scientifically important because (1) patterns of change or growth over time can be characterized and (2) certain types of "treatment comparisons" can be made on a within-subject basis, thus improving precision by eliminating intersubject variations.

The first property is very important in the study of growth and development. In a cross-sectional design growth rates cannot be estimated without bias because age effects are confounded with cohort effects. (Two effects are said to be confounded when they are not separately estimable.) This can be seen in Figure 1.3.1 by noting that each cohort encompasses a different age group. To estimate an age effect in this situation it would be necessary to assume that the cohort effects and secular trends are negligible, which is not always true. In a longitudinal without period or secular effects , study cohort and age are not confounded and thus growth rates may be estimated without assuming that cohort effects and secular trends are negligible.

The second property, that comparisons are made on a within-subject basis, is important in many longitudinal designs, including crossover designs. A crossover design is one in which each subject is studied for several "periods" and is subjected to different treatments in different periods. Since subjects are measured on more than one occasion the crossover design is a longitudinal design. Because the crossover design allows comparisons to be made on a within subject basis, the resulting treatment comparisons may be more precise than comparisons made from two independent groups.

There is a rich literature on crossover designs (see, e.g., Brown (1980) for references). The focus in this paper, however, will be on general longitudinal designs. Specific issues which arise from crossover designs will be addressed at appropriate places in subsequent chapters.

It is important to note that the defining characteristic of a longitudinal design is that each subject is observed on two or more occasions. This typically means that multiple observations from the same subject are correlated. It is this correlation that allows comparisons to be made with more precision. A correct statistical analysis of data from a longitudinal study should take this correlation into account. This important point will be elaborated on further in a subsequent chapter. These and other aspects of general longitudinal designs are discussed in more detail in Cook and Ware (1983) and Goldstein (1979).

1.3.2 Linked Cross-Sectional Designs

Rao and Rao (1966) proposed the class of Linked Cross-Sectional (LCS) designs for use in studies of growth and development. The LCS design can be thought of as a compromise between the longitudinal design and the cross-sectional design which takes advantage of the good properties of longitudinal designs while reducing the associated problems. The LCS design is, in effect, a longitudinal study of several cohorts, each over a different age range but designed so that age ranges overlap, as illustrated in Figures 1.3.3 and 1.3.4. The number of measurement periods which overlap will be referred to as the degree of overlap. Thus these examples represent LCS designs with overlap of degree 1 and 2, respectively. The advantage of a lower degree LCS is that it requires less total time to complete but may still cover a wide age range. The disadvantage is that certain types of higher order secular (or time) trends cannot be evaluated.

A LCS design could also be constructed that is a combination of different degree LCS designs, as illustrated in Figure 1.3.5. This would permit a greater emphasis on certain degrees of expected trends with the capability of detecting higher order trends, if present.

The estimates provided by cross-sectional, longitudinal and LCS designs are qualitatively different unless, as noted by Rao and Rao (1966), the external factors affecting the growth of variables under study are reasonably stable over a long period (and cohort,period and secular effects are negligible); then the three studies provide estimates of the same parameters.

The LCS design is an example of a purposefully incomplete longitudinal design. It is a longitudinal design because subjects are measured on more than one occasion and it is purposefully incomplete because each cohort is measured on a different subset of the possible measurement periods.

An alternative to the LCS design is the purposefully incomplete full-period longitudinal design illustrated in Figure 1.3.6. This design is constructed by planning, in advance, to omit measurements at specified follow-up times from a full longitudinal design. Each subject, however, is always measured at the initial and final time points. Each group of subjects with the same measurement schedule forms a cohort.

The LCS design was first described and termed as such by Rao and Rao (1966) in their studies on the growth of Indian boys. In the field of psychology, however, Schaie (1965) proposed "sequential research designs" which are, in fact, LCS designs. These designs arose from a general developmental model which stated that growth, or response, was a function of three variables: age, cohort and time-of-measurement, where cohort refers to birth cohorts. The "sequential research designs" were designed to estimate the effect of these three variables.

Baltes (1968) argued that only two variables are necessary in the general developmental model since any two will determine the third. Baltes described two designs: a longitudinal sequence and a cross-sectional sequence. The longitudinal sequence consists of measuring several cohorts longitudinally. The cross-sectional sequence consists of several cross-sectional studies done at different time points using different groups of subjects. These designs, however, are not LCS designs.

A LCS design was also described by van't Hoff, et al. (1977) for a study of the growth and development of normal children. The study design was termed a "mixed longitudinal design". Data was collected during 18 measurement periods for six cohorts covering the age range of 4.0 to 13.5 years. The measurement periods were divided into three groups determined by the individual taking the measurements. Periods 1 and 2 formed one group; periods 3-8 formed a second group; and periods 9-18 formed the third group. A quadratic function was used to approximate the growth curve. Six cohort and three time-of-measurement effects were also included in the model. The time-of-measurement effects corresponded to possible inter-observer differences in the measurement process.

The method of least squares was used to fit the model. The authors recognized the limitations of this method of analysis in that the data are not independent and thus the usual properties of least squares estimators do not necessarily hold. They used the estimated parameters to calculate predicted values. The average deviation was then computed as a measure of goodness-of-fit.

Rao and Rao (1966) estimated norms and growth rates of school boys in India using data from a LCS study. They assumed an equal-correlation structure for multiple measurements from the same subject and estimated norms using the method of weighted least squares. The efficiency of these estimates was compared to estimates which might have been obtained from a cross-sectional study. Those from the LCS study were found to have standard errors that are nearly half of those in the comparable cross-sectional study.

Rao and Rao (1966) also considered the optimal design for a mixed cross-sectional/LCS study for three cases: 1) estimating norms, 2) estimating growth rates, and 3) estimating differential growth rate. For the first two cases it is assumed that we have a two time point design involving a total of 2n measurements, n at each time period. πn subjects are measured at the first time point only, πn subjects are measured at the second time point only and $(1-\pi)n$ subjects are measured at both time points. When $\pi=1$ the design is cross-sectional and when $\pi=0$ the design is pure longitudinal.

For estimating norms the optimum value of π is found by minimizing the variance of the estimate of the norm. The resulting optimum value is

$$\pi_{opt} = \frac{1}{[\ 1+(1-\rho^2)^{\frac{1}{2}}\ ]}$$

where ρ is the correlation between the measurements at the two ages. For values of ρ up to 0.60, the number of subjects to be measured twice is about 50% of the subjects who are measured just once; for higher values of ρ the optimum proportion measured at both times decreases.

For the estimation of growth rates π=0 is optimal, indicating that all subjects should be measured on both occasions. For the case where it is of interest to estimate the differential growth rate, i.e., the difference in growth in successive years, two alternative designs are considered. These are:

a) A pure longitudinal design where each of n subjects is measured at three successive ages (say 6,7 and 8);

b) A LCS design with n subjects measured at ages 6 and 7 and n different subjects measured at ages 7 and 8.

It was found that design a) is better than b) when

$$2-2\rho_1 + \rho_2 < 1 \quad \text{or} \quad 2\rho_1 > 1 + \rho_2$$

where ρ_1 is the correlation between measurements at age 6 and 7 and ρ_2 is the correlation between measurements at age 7 and age 8. Rao and Rao concluded that the LCS design is the optimum design for studying norms, growth rates and differential growth.

Woolson, Leeper, and Clarke (1978) also found the LCS design (mixed longitudinal study in their terminology) to be more efficient than the crosssectional design. The LCS design they used consisted of three measurements, done in 1971, 1973 and 1975, on children in grades 1 through 12. During each of the three measurement years children were measured with respect to various characteristics such as height, weight and blood pressure. Each child potentially had three sets of measurements but due to a variety of reasons many children were not present during each measurement year and thus had only one or two sets of measurements.

Each variable (height, weight, etc.) was analyzed separately. It was assumed that each variable had a particular covariance structure, $\sigma^2\boldsymbol{\Gamma}_i$, where $\boldsymbol{\Gamma}_i$ referred to the correlation matrix for a particular variable and σ^2 was unspecified. Further it was assumed that $\boldsymbol{\Gamma}_i$ was known and given by the estimated correlation matrix. The data were assumed to be distributed $N(\mathbf{X}\boldsymbol{\beta},\sigma^2\boldsymbol{\Gamma}_i)$ and the method of weighted least squares was used to estimate $\boldsymbol{\beta}$. In both the analysis of means and the analysis of growth rates substantial reductions in the standard errors were found for the LCS study.

Machin (1975) compared the LCS design to the pure longitudinal design for the following special case. The total number of observations (N) was held constant. The pure longitudinal design consisted of n subjects measured on p occasions, (N=np). The LCS designs consisted of m(>n) subjects measured on k(<p) occasions, (N=np=mk). The data were assumed to follow a multivariate normal distribution with an auto

regressive covariance structure of order 1. A straight line growth model was assumed. Using the generalized variance as a measure of efficiency, Machin found that, in general, as ρ increases the efficiency of the LCS design increases, where ρ is the correlation between any two adjacent observations.

Given the previous discussion it would seem that the LCS design would be the design of choice in many research situations, yet a review of the literature finds only the examples given above. LCS designs are mentioned only briefly in the book by Goldstein (1979) and the review article by Cook and Ware (1983). This may be attributed, in part, to the fact that the analysis of data from such a design may be very complex. Aside from certain special cases, such as equal correlation structure, appropriate methods for the analysis of this type of data ,such as those discussed in Section 1.1, have only recently become feasible.

Chapter 2

Exact and Approximate F Tests for the General Linear Multivariate Model and the Mixed Model

In this chapter we first review the notation that we will use. We then discuss those cases of the general linear univariate and multivariate models in which an exact F statistic can be calculated. Essentially these consist of cases in which the data are uncorrelated or a linear transformation of the data produces uncorrelated data. We also discuss some approximate results that apply when the data are correlated.

We then extend these results to special cases of the mixed model with linear covariance structure. These extensions are used to motivate approximate F statistics for the mixed model. Three approximate F statistics are proposed; one is based on the canonical form of the mixed model and two others are based on weighted least squares.

2.1 General Linear Univariate Model (GLUM)

The notation "GLUM($\mathbf{y};\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I}_N$) with normality" is an abbreviation for the following:

$\mathbf{y}$ is an Nx1 vector of observations,

$\mathbf{X}$ is an Nxq design matrix of known constants with rank($\mathbf{X}$) = r,

$\boldsymbol{\beta}$ is a qx1 vector of parameters to be estimated,

and $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I}_N)$.

If $\mathbf{X}$ has full rank, i.e. r=q, then the notation GLUM-FR will be used.

The ordinary least squares estimator of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y} \tag{2.1.1}$$

An unbiased estimator for σ^2 is given by

$$\begin{aligned}\hat{\sigma}^2 &= s/(N-\text{rank}(\mathbf{X})) \\ &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})/(N-\text{rank}(\mathbf{X})).\end{aligned} \tag{2.1.2}$$

Secondary parameters may be written as

$$\boldsymbol{\theta} = \mathbf{L}\boldsymbol{\beta} \tag{2.1.3}$$

where $\mathbf{L}$ is an axq matrix of fixed, known constants. We shall assume that $\boldsymbol{\theta}$ is estimable. The minimum variance unbiased estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}} = \mathbf{L}\hat{\boldsymbol{\beta}}. \tag{2.1.4}$$

We assume that the hypothesis H_0: $\boldsymbol{\theta} = \mathbf{0}$ is testable, i.e. $\mathbf{L}(\mathbf{X}'\mathbf{X})^{-}\mathbf{L}'$ is nonsingular. The likelihood ratio test is given by the test statistic

$$F = \frac{\hat{\boldsymbol{\theta}}'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-}\mathbf{L}')^{-1}\hat{\boldsymbol{\theta}}/a}{s/(N-r)} \quad . \tag{2.1.5}$$

This statistic follows a noncentral F distribution with a and N-r degrees of freedom and noncentrality parameter $\boldsymbol{\theta}'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-}\mathbf{L}')^{-1}\boldsymbol{\theta}/\sigma^2$.

2.2 General Linear Multivariate Model (GLMM)

The notation "GLMM ($\mathbf{Y}$; $\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma}$) with normality" is an abbreviation for the following:

- $\mathbf{Y}$ is an Nxp matrix of observations on dependent variables,
- $\text{Vec}(\mathbf{Y}')$ is an Npx1 vector containing the elements of $\mathbf{Y}$ "rolled out by rows",
- $\mathbf{X}$ is an Nxq design matrix of known constants with rank($\mathbf{X}$)=r,
- $\boldsymbol{\beta}$ is a qxp matrix of parameters to be estimated,
- $\text{Vec}(\boldsymbol{\beta}')$ is a qpx1 vector containing the elements of $\boldsymbol{\beta}$ "rolled out by rows" (Searle 1982, p. 332),

and

$$\text{Vec}(\mathbf{Y}') \sim N_{Np}((\mathbf{X} \otimes \mathbf{I}_p)\text{Vec}(\boldsymbol{\beta}'),\ \mathbf{I}_N \otimes \boldsymbol{\Sigma}),$$

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}, \tag{2.2.1}$$

$$E(\text{Vec}(\mathbf{Y}')) = (\mathbf{X} \otimes I_p)\text{Vec}(\boldsymbol{\beta}'), \tag{2.2.2}$$

$$\text{Cov}(\text{ Row}_i(\mathbf{Y})', \text{Row}_j(\mathbf{Y})') = \begin{cases} \mathbf{0}_{pxp} & i \neq j \\ \Sigma & i=j \end{cases} \tag{2.2.3}$$

$$\text{Var}(\text{Vec}(\mathbf{Y}')) = I_N \otimes \Sigma . \tag{2.2.4}$$

If the design matrix has full rank then the notation GLMM-FR will be used.

The secondary parameters to be considered for this model are assumed to be estimable and may be expressed as

$$\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\beta}\mathbf{U} \tag{2.2.5}$$

or $$\boldsymbol{\gamma} = \text{L}\text{Vec}(\boldsymbol{\beta}') \tag{2.2.6}$$

where $\mathbf{C}$ is an axq known constant matrix with rank = a $\leq$ q,

$\mathbf{U}$ is a pxb known constant matrix with rank = b $\leq$ p,

L is a cxqp known constant matrix with rank = c $\leq$ qp.

If $\mathbf{L}$ is of the form $\mathbf{L} = (\mathbf{C} \otimes \mathbf{U}')$ then $\boldsymbol{\gamma} = \text{Vec}(\boldsymbol{\theta}')$.

Maximum likelihood estimators are given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y} \tag{2.2.7}$$

$$\text{Vec}(\hat{\boldsymbol{\beta}}') = ((\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' \otimes I_p)\text{Vec}(\mathbf{Y}') \tag{2.2.8}$$

$$\hat{\boldsymbol{\theta}} = \mathbf{C}\hat{\boldsymbol{\beta}}\mathbf{U} \tag{2.2.9}$$

$$\hat{\boldsymbol{\gamma}} = \text{L}\text{Vec}(\hat{\boldsymbol{\beta}}'). \tag{2.2.10}$$

An unbiased estimate of Σ is given by

$$\hat{\Sigma} = \text{S}/(\text{N-r}) = (\mathbf{Y}\text{-}\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}\text{-}\mathbf{X}\hat{\boldsymbol{\beta}})/(\text{N-r}). \tag{2.2.11}$$

Multivariate test statistics such as Hotelling-Lawley trace, Roy's largest root, Wilk's Lambda, or Pillai-Bartlett trace, may be used to test hypotheses of the form H_0: $\boldsymbol{\theta} = \mathbf{0}$. Unfortunately, closed form, finite series expressions generally do not exist for the distribution functions of these test statistics. In practice asymptotic distributions, based on F or chi-square distributions, are used for the null case. These approximations and those for the non-null case are discussed by Muller and Peterson (1984).

Because we do not have exact distributions for these test statistics and they are restricted to GLMM($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma}$) with complete data, they will not be considered in this work. Instead we will look for test statistics for which we can discover exact distributions and which can be extended to include missing data and different covariance matrices.

2.3 Alternative F tests for the General Linear Multivariate Model

In this section we will consider some approximate and exact F tests for the GLMM. We begin by considering a naive test for the GLMM. If we "roll out" the Nxp data matrix $\mathbf{Y}$ into an Npx1 vector we could naively calculate a test statistic for H_0: $\boldsymbol{\gamma} = \mathbf{0}$ using univariate theory. That test statistic would be

$$F_\gamma = \frac{\hat{\boldsymbol{\gamma}}'(\mathbf{L}((\mathbf{X}\otimes\mathbf{I}_p)'(\mathbf{X}\otimes\mathbf{I}_p))^{-}\mathbf{L}')^{-1}\hat{\boldsymbol{\gamma}}/c}{S/(Np-rp)} \tag{2.3.1}$$

where $S = (\text{Vec}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})')'(\text{Vec}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})')$ (2.3.2)

and $\hat{\boldsymbol{\gamma}}$ is given in the previous section.

In general F_γ will not follow an exact F distribution. The following theorem presents one situation in which the statistic does follow an exact F distribution. In essence, this theorem states that if the data are uncorrelated the F statistic will have an exact F distribution.

<u>Theorem 2.3.1</u>

Given GLMM-FR$(\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma})$ with normality, if $\boldsymbol{\Sigma} = \lambda\mathbf{I}_p$ then the test statistic (2.3.1) follows an exact F distribution with c and Np−rp degrees of freedom and noncentrality parameter cF_a, where F_a is the value of the test statistic computed using the population values of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.

<u>Proof</u> If $\boldsymbol{\Sigma} = \lambda\mathbf{I}_p$ then $\text{Vec}(\mathbf{Y}') \sim N_{Np}((\mathbf{X}\otimes\mathbf{I}_p)\text{Vec}(\boldsymbol{\beta}'),\lambda\mathbf{I}_{Np})$ and we have the univariate case as described in section 2.1. The test statistic above (2.3.1) is the univariate test statistic. □

Other situations in which the test statistic (2.3.1) follows an exact F distribution have been discussed by Geisser and Greenhouse (1959), Huynh (1978) and Huynh and Feldt (1970,1976). These have been discussed in the literature review.

We will next consider a test statistic based on a transformation of the data. Let **H** be a px(p−1) column orthonormal matrix such that

$$\mathbf{H}'\mathbf{H} = \mathbf{I}_{p-1} \quad \text{and} \quad \mathbf{H}'\mathbf{1}_p = \mathbf{0}. \qquad (2.3.3)$$

where $\mathbf{1}_p = (1,1,\ldots,1)'$. The columns of **H** form an orthonormal basis for the contrast space. The nonconstant columns of a Helmert matrix are an example of a family of matrices with this property.

The following theorem deals with the case where $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$, where **H** is a px(p−1) matrix which satisfies (2.3.3). This theorem states that if the data can be transformed such that the resulting data is uncorrelated then there is an exact F statistic for the test H_o: $\boldsymbol{\gamma} = \mathbf{0}$.

<u>Theorem 2.3.2</u>

Given GLMM-FR($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma} : \mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$) with normality, where **H** is a px(p−1) matrix satisfying (2.3.3), let

$$\boldsymbol{\gamma} = \mathrm{Vec}(\boldsymbol{\theta}') = (\mathbf{C}\otimes(\mathbf{HU})')\mathrm{Vec}(\boldsymbol{\beta}') = \mathbf{L}\mathrm{Vec}(\boldsymbol{\beta}') \qquad (2.3.4)$$

where **C** is an axq known constant matrix with rank = a ≤ q,

U is a (p-1)xb known constant matrix with rank = b ≤ p-1,

L is an abxqp known constant matrix with rank = ab ≤ qp.

Assume that $\boldsymbol{\gamma}$ is testable . The likelihood ratio test statistic for testing H_o: $\boldsymbol{\gamma} = \mathbf{0}$ is given by

$$F_\gamma = \frac{\hat{\boldsymbol{\gamma}}'(\mathbf{L}\,[(\mathbf{X}\otimes\mathbf{I}_p)'(\mathbf{X}\otimes\mathbf{I}_p)]^{-1}\mathbf{L}')^{-1}\hat{\boldsymbol{\gamma}}\,/ab}{S_H/(N-r)(p-1)} \qquad (2.3.5)$$

where $S_H = [\mathrm{Vec}((\mathbf{YH}-\mathbf{X}\hat{\boldsymbol{\beta}}\mathbf{H})')]'[\mathrm{Vec}((\mathbf{YH}-\mathbf{X}\hat{\boldsymbol{\beta}}\mathbf{H})')]$, (2.3.6)
$\hat{\boldsymbol{\gamma}} = \mathrm{L}\mathrm{Vec}(\hat{\boldsymbol{\beta}}')$ and $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.
This test statistic (2.3.5) follows an exact F distribution with ab and (p-1)(N-r) degrees of freedom and noncentrality parameter abF_a, where F_a is the value of the test statistic computed using the population values of $\boldsymbol{\beta}$ and Σ.

Proof Consider the transformation $\mathbf{Z} = \mathbf{YH}$. The transformed data matrix $\mathbf{Z}$ satisfies $\mathrm{GLMM}(\mathbf{Z};\mathbf{X}\boldsymbol{\beta}_z,\lambda\mathbf{I}_{p-1})$ with normality, where $\boldsymbol{\beta}_z = \boldsymbol{\beta}\mathbf{H}$.

Next consider "rolling out" the data matrix $\mathbf{Z}$ into an N(p-1)x1 vector, $\mathbf{z}$, i.e., $\mathbf{z} = \mathrm{Vec}(\mathbf{Z}') = \mathrm{Vec}((\mathbf{YH})')$. Then
$\mathbf{z} \sim N_{N(p-1)}[(\mathbf{X}\otimes\mathbf{I}_{p-1})\mathrm{Vec}(\boldsymbol{\beta}_z'),\lambda\mathbf{I}_{N(p-1)}]$.
Let $\boldsymbol{\theta}_z = \mathbf{C}_z\boldsymbol{\beta}_z\mathbf{U}_z$ and $\boldsymbol{\gamma}_z = \mathrm{Vec}(\boldsymbol{\theta}_z') = (\mathbf{C}_z\otimes\mathbf{U}_z')\mathrm{Vec}(\boldsymbol{\beta}_z') = \mathrm{L}_z\mathrm{Vec}(\boldsymbol{\beta}_z')$
where $\mathbf{C}_z$ is an axq known constant matrix with rank = a ≤ q,

$\mathbf{U}_z$ is a (p-1)xb known constant matrix with rank = b ≤ (p-1),

L_z is an abxq(p-1) known constant matrix with rank= ab ≤ q(p-1).

Assume H_o: $\boldsymbol{\gamma}_z = \mathbf{0}$ is testable; the likelihood ratio test statistic is given by

$$F_{\gamma z} = \frac{\hat{\boldsymbol{\gamma}}_z'(\mathbf{L}_z[(\mathbf{X}\otimes\mathbf{I}_{p-1})'(\mathbf{X}\otimes\mathbf{I}_{p-1})]^{-1}\mathbf{L}_z')^{-1}\hat{\boldsymbol{\gamma}}_z\ /ab}{S_z/(N-r)(p-1)} \tag{2.3.7}$$

where $S_z = (\mathbf{z}-(\mathbf{X}\otimes\mathbf{I}_{p-1})\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z'))'(\mathbf{z}-(\mathbf{X}\otimes\mathbf{I}_{p-1})\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z'))$. (2.3.8)
Since $\mathbf{z}$ has a normal distribution and $V(\mathbf{z}) = \lambda\mathbf{I}$, $F_{\gamma z}$ follows an exact F distribution with ab and (N-r)(p-1) degrees of freedom and noncentrality parameter abF_a.

The strategy in this proof will be to show that $F_\gamma = F_{\gamma z}$ when the secondary parameters are equal, and thus that F_γ follows the same exact F distribution as $F_{\gamma z}$. Assume that $\mathbf{C} = \mathbf{C}_z$ and $\mathbf{U} = \mathbf{U}_z$ and thus $\boldsymbol{\gamma} = \boldsymbol{\gamma}_z$. Substituting $\mathbf{L}_z = (\mathbf{C}_z\otimes\mathbf{U}_z')$, the numerator of $F_{\gamma z}$ is

$$\hat{\boldsymbol{\gamma}}_z'[(\mathbf{C}_z\otimes\mathbf{U}_z')[(\mathbf{X}\otimes\mathbf{I}_{p-1})'(\mathbf{X}\otimes\mathbf{I}_{p-1})]^{-1}(\mathbf{C}_z\otimes\mathbf{U}_z')]^{-1}\hat{\boldsymbol{\gamma}}_z\ /ab$$

$$= \hat{\boldsymbol{\gamma}}_z'(\mathbf{C}_z(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}_z' \otimes (\mathbf{U}_z'\mathbf{U}_z))^{-1}\hat{\boldsymbol{\gamma}}_z/ab. \tag{2.3.9}$$

In comparison the numerator of F_γ is

$$\hat{\boldsymbol{\gamma}}'[(\mathbf{C}\otimes(\mathbf{HU})')[(\mathbf{X}\otimes\mathbf{I}_p)'(\mathbf{X}\otimes\mathbf{I}_p)]^{-1}(\mathbf{C}\otimes(\mathbf{HU})')']^{-1}\hat{\boldsymbol{\gamma}}/ab$$
$$= \hat{\boldsymbol{\gamma}}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' \otimes \mathbf{U}'\mathbf{H}'\mathbf{H}\mathbf{U}]^{-1}\hat{\boldsymbol{\gamma}}/ab$$
$$= \hat{\boldsymbol{\gamma}}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C} \otimes \mathbf{U}'\mathbf{U}]^{-1}\hat{\boldsymbol{\gamma}}/ab. \tag{2.3.10}$$

Since $\mathbf{C}_z = \mathbf{C}$ and $\mathbf{U}_z = \mathbf{U}$ by assumption, and $\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_z$ as shown below, the numerators (2.3.9) and (2.3.10) are equal.

To show that $\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_z$ we write $\hat{\boldsymbol{\gamma}}$ as follows,

$$\hat{\boldsymbol{\gamma}} = (\mathbf{C}\otimes(\mathbf{HU})')\mathrm{Vec}(\hat{\boldsymbol{\beta}}') = (\mathbf{C}\otimes\mathbf{U}')\mathrm{Vec}((\hat{\boldsymbol{\beta}}\mathbf{H})')$$
$$= (\mathbf{C}\otimes\mathbf{U}')\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z') = (\mathbf{C}_z\otimes\mathbf{U}_z')\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z') = \hat{\boldsymbol{\gamma}}_z.$$

The denominator of $F_{\gamma z}$ (without the degrees of freedom) is

$$S_z = (\mathbf{z}-(\mathbf{X}\otimes\mathbf{I}_{p-1})\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z'))'(\mathbf{z}-(\mathbf{X}\otimes\mathbf{I}_{p-1})\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z')).$$

Substituting $\mathbf{z} = \mathrm{Vec}(\mathbf{Z}') = \mathrm{Vec}((\mathbf{YH})')$ and $(\mathbf{X}\otimes\mathbf{I}_{p-1})\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z') =$ $\mathrm{Vec}(\hat{\boldsymbol{\beta}}_z'\mathbf{X}') = \mathrm{Vec}((\hat{\boldsymbol{\beta}}\mathbf{H})'\mathbf{X}') = \mathrm{Vec}(\mathbf{H}'\hat{\boldsymbol{\beta}}'\mathbf{X}')$ into this expression we get

$$S_z = (\mathrm{Vec}((\mathbf{YH})'-\mathrm{Vec}((\hat{\boldsymbol{\beta}}\mathbf{H})'\mathbf{X}'))'(\mathrm{Vec}((\mathbf{YH})'-\mathrm{Vec}((\hat{\boldsymbol{\beta}}\mathbf{H})'\mathbf{X}'))$$
$$= [\mathrm{Vec}((\mathbf{YH}-\mathbf{X}\hat{\boldsymbol{\beta}}\mathbf{H})')]'[\mathrm{Vec}((\mathbf{YH}-\mathbf{X}\hat{\boldsymbol{\beta}}\mathbf{H})')]$$
$$= S_H,$$

which is the denominator of F_γ (2.3.5). Therefore the test statistics are equal and they follow the same distribution. □

The following corollaries establish the structure that $\boldsymbol{\Sigma}$ must have to insure that $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$. Esentially we show that if $\boldsymbol{\Sigma}$ has the Huynh-Feldt structure described in the literature review then $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$ and the results of Theorem 2.3.2 will hold. This is also true when $\boldsymbol{\Sigma}$ has equicorrelation structure, a special case of the Huynh-Feldt structure.

Corollary 2.3.1

Given GLMM-FR($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma}$) with normality, if $\mathbf{H}$ is a px(p-1) matrix which spans the contrast space then $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$ if and only if $\boldsymbol{\Sigma}$ has the Huynh-Feldt linear structure.

Proof (Huynh and Feldt, 1970)

1) Assume that $\boldsymbol{\Sigma} = [\sigma_{ij}]$ has the Huynh-Feldt linear structure, i.e.,

$$\sigma_{ij} = \alpha_i + \alpha_j + \lambda\delta_{ij} \qquad \text{where } \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases} \tag{2.3.11}$$

The Huynh-Feldt structure can also be written in matrix form as

$$\boldsymbol{\Sigma} = (\boldsymbol{\alpha} \otimes \mathbf{1}') + (\boldsymbol{\alpha} \otimes \mathbf{1}')' + \lambda\mathbf{I}_p \tag{2.3.12}$$

where $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_p)'$ and $\mathbf{1}$ is a px1 vector of one's.

The elements of the ith row of the matrix $(\boldsymbol{\alpha} \otimes \mathbf{1}')$ are all equal to α_i.

Then

$$\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \mathbf{H}'(\boldsymbol{\alpha} \otimes \mathbf{1}')\mathrm{H} + \mathbf{H}'(\boldsymbol{\alpha} \otimes \mathbf{1}')'\mathrm{H} + \lambda\mathbf{H}'\mathbf{H}. \tag{2.3.13}$$

Since $\mathbf{H}$ is a contrast matrix and the rows of $(\boldsymbol{\alpha} \otimes \mathbf{1}')$ are vectors of equal value, $(\boldsymbol{\alpha} \otimes \mathbf{1}')\mathbf{H} = \mathbf{0}$. Thus the first two elements on the right hand side of equation (2.3.13) are equal to zero. From (2.3.4) we get that $\mathbf{H}'\mathbf{H} = \mathbf{I}_{p-1}$ and thus (2.3.13) reduces to $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$.

2) Assume that $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$. If we premultiply by $\mathbf{H}$ and postmultiply by $\mathbf{H}'$ we have $\mathbf{H}\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H}\mathbf{H}' = \lambda\mathbf{H}\mathbf{H}'$. Substituting $\mathbf{H}\mathbf{H}' = (\mathbf{I}-\mathbf{1}\mathbf{1}'/k)$ into this expression we get

$$(\mathbf{I}-\mathbf{1}\mathbf{1}'/k)\boldsymbol{\Sigma}(\mathbf{I}-\mathbf{1}\mathbf{1}'/k) = \lambda(\mathbf{I}-\mathbf{1}\mathbf{1}'/k). \qquad (2.3.14)$$

The general elements of the matrix $(\mathbf{I}-\mathbf{1}\mathbf{1}'/k)$ and $\boldsymbol{\Sigma}(\mathbf{I}-\mathbf{1}\mathbf{1}'/k)$ are, respectively, $\delta_{ij}-(1/k)$ and $\sigma_{ij}-\sigma_{i.}$, where the "." in the subscript indicates that we have averaged over elements in the row and/or column. Hence from (2.3.14) we get

$$\sigma_{ij} = \sigma_{.j} + \sigma_{i.} - \sigma_{..} + \lambda(\delta_{ij}-(1/k)), \text{ for all i and j.} \qquad (2.3.15)$$

Substituting the following

$$\alpha_i = \sigma_{i.} - (\sigma_{..}/2) - \lambda/2k$$

$$\alpha_j = \sigma_{.j} - (\sigma_{..}/2) - \lambda/2k$$

into (2.3.15) we get the Huynh-Feldt structure as described in (2.3.11). □

Corollary 2.3.2

Given GLMM-FR($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma}$) with normality, if $\boldsymbol{\Sigma}$ has an equicorrelation structure,

$$\boldsymbol{\Sigma} = \sigma^2 \left[(1-\rho)\mathbf{I}_p + \rho \mathbf{1}\mathbf{1}' \right] \tag{2.3.16}$$

then $\boldsymbol{\Sigma}$ has the Huynh-Feldt linear structure.

Proof Writing the Huynh-Feldt linear structure in matrix form (2.3.12) we get that

$$\boldsymbol{\Sigma} = (\boldsymbol{\alpha} \otimes \mathbf{1}') + (\boldsymbol{\alpha} \otimes \mathbf{1}')' + \lambda \mathbf{I}_p$$

Let $(\boldsymbol{\alpha} \otimes \mathbf{1}') = (\rho/2)\mathbf{1}\mathbf{1}'$ and $\lambda = (1-\rho)$. Substituting these into the matrix expression for $\boldsymbol{\Sigma}$ we get

$$\boldsymbol{\Sigma} = \rho \mathbf{1}\mathbf{1}' + (1-\rho)\mathbf{I}_p$$

the desired result. □

The following theorem considers testing hypotheses of the form H_o: $\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\beta}\mathbf{U}$ and gives an exact result for the case where $\boldsymbol{\Sigma}$ has Huynh-Feldt linear structure and an approximate result for when $\boldsymbol{\Sigma}$ does not have this structure. Our previous results applies only to those cases where the data, or the transformed data was uncorrelated. In the following theorem we give exact and approximate results for the case where the data are correlated.

Theorem 2.3.3

Assume GLMM-FR$(\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma})$ with normality. Let $\tilde{\boldsymbol{\theta}} = \tilde{\mathbf{C}}\boldsymbol{\beta}\tilde{\mathbf{U}}$ be estimable with $a = \text{rank}(\tilde{\mathbf{C}}) \leq q$ and $b = \text{rank}(\tilde{\mathbf{U}}) \leq p$. (The tilde (˜) over the matrices is used to indicate that these quantities are not necessarily the same as those given in Theorem 2.3.2.) A test statistic for testing H_0: $\tilde{\boldsymbol{\theta}} = \mathbf{0}$ is given by

$$F_\theta = \frac{\text{tr}(\hat{\tilde{\boldsymbol{\theta}}}'(\tilde{\mathbf{C}}(\mathbf{X}'\mathbf{X})^{-1}\tilde{\mathbf{C}}')^{-1}\hat{\tilde{\boldsymbol{\theta}}})/ab}{\text{tr}(\tilde{\mathbf{U}}'\mathbf{S}_y\tilde{\mathbf{U}})/b(N-r)} \tag{2.3.17}$$

where $\mathbf{S}_y = (\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})$, (2.3.18)

$\hat{\tilde{\boldsymbol{\theta}}} = \tilde{\mathbf{C}}\hat{\boldsymbol{\beta}}\tilde{\mathbf{U}}$ and $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

$$\text{Let } \epsilon = \frac{\left(\sum_{i=1}^{b} \lambda_i\right)^2}{b\left(\sum_{i=1}^{b} \lambda_i^2\right)} = \frac{[\text{tr}(\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}})]^2}{b(\text{tr}\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}}^2)} \tag{2.3.19}$$

where λ_i, $i=1,\ldots,b$ are the eigenvalues of $\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}}$. If $\epsilon = 1$ then F_θ follows an exact F distribution with ab and $b(N-r)$ degrees of freedom and noncentrality parameter abF_a, where F_a is the value of F_θ calculated using the true values of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.

Proof The proof is given by Muller and Barton(1987).

The case where $\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}} = \lambda\mathbf{I}$ is usually referred to as "sphericity" since it corresponds to the transformed scores having a spherical normal distribution. It is an important case.

Corollary 2.3.3

If $\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}} = \lambda\mathbf{I}$ then $\epsilon = 1$.

Proof Substituting $\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}} = \lambda\mathbf{I}$ into (2.3.19) gives the following

$$\epsilon = \frac{[\mathrm{tr}(\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}})]^2}{b[\mathrm{tr}(\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}})^2]} = \frac{[\mathrm{tr}(\lambda\mathbf{I})]^2}{b[\mathrm{tr}\lambda^2\mathbf{I}]} = \frac{(b\lambda)^2}{b^2\lambda^2} = 1. \qquad \square$$

If $\epsilon \neq 1$, i.e. $\tilde{\mathbf{U}}'\boldsymbol{\Sigma}\tilde{\mathbf{U}} \neq \lambda\mathbf{I}$, then Muller and Barton (1987) proved that the test statistic (2.3.17) approximately follows a noncentral F distribution with $ab\epsilon$ and $b(N-r)\epsilon$ degrees of freedom and noncentrality parameter $abF_a\epsilon$.

Theorem 2.3.2 gives an exact result for testing H_o: $\boldsymbol{\gamma} = \mathrm{LVec}(\boldsymbol{\beta}') = \mathbf{0}$ when $\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} = \lambda\mathbf{I}_{p-1}$, that is when $\boldsymbol{\Sigma}$ has Huynh-Feldt structure. However, no results are given for the case where $\boldsymbol{\Sigma}$ does not have this special structure.

Theorem 2.3.3 gives an approximate result for testing H_0: $\mathbf{C\beta U} = \mathbf{0}$ in the case where $\mathbf{\Sigma}$ does not have Huynh-Feldt structure. We would like to be able to apply the approximate results of Theorem 2.3.3 to the situation described in Theorem 2.3.2. The following theorem gives conditions under which the situations are equivalent.

<u>Theorem 2.3.4</u>

Assume GLMM-FR$(\mathbf{Y};\mathbf{X\beta},\mathbf{\Sigma})$ with normality, as given in Theorems 2.3.2 and 2.3.4. Further, assume $\tilde{\mathbf{C}} = \mathbf{C}$ and $\tilde{\mathbf{U}} = \mathbf{HU}$ so that the secondary parameters for the two cases are equal, i.e., $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}$ and $\tilde{\boldsymbol{\gamma}} = \boldsymbol{\gamma}$. Then the test statistics F_γ (2.3.5) and F_θ (2.3.17) are equal if $\mathbf{U}$ is a $(p-1)\times(p-1)$ orthogonal matrix such that $\mathbf{U'U} = \mathbf{UU'} = \mathbf{I}_{p-1}$.

<u>Proof</u> We will show that $F_\gamma = F_\theta$ by first showing that the denominators of the test statistics are equal and then that the numerators are equal. Substituting $\tilde{\mathbf{U}} = \mathbf{HU}$ into the denominator of F_θ we get

$$\begin{aligned} tr(\tilde{\mathbf{U}}'\mathbf{S}_y\tilde{\mathbf{U}}) &= tr((\mathbf{HU})'\mathbf{S}_y(\mathbf{HU})) \\ &= tr(\mathbf{U'H'S}_y\mathbf{HU}). \\ &= tr(\mathbf{H'S}_y\mathbf{HUU'}). \end{aligned}$$

This last step is due to the permutation property of the trace operator. Then if $\mathbf{UU'} = \mathbf{I}_{p-1}$ we get

$$tr(\mathbf{U'S}_y\mathbf{U}) = tr(\mathbf{H'S}_y\mathbf{H})$$

and substituting (2.3.18), $\mathbf{S}_y = (\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})$, we get

$$tr(\mathbf{U'S}_y\mathbf{U}) = tr(\mathbf{H'}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})\mathbf{H}).$$

Then using Searle(1982), Theorem 2, page 333 we get

$$\mathrm{tr}(\mathbf{U}'\mathbf{S}_y\mathbf{U}) = [\mathrm{Vec}((\mathbf{Y}-\mathbf{X}\hat{\boldsymbol\beta})\mathbf{H}')'[\mathrm{Vec}((\mathbf{Y}-\mathbf{X}\hat{\boldsymbol\beta})\mathbf{H})'] = S_H$$

and thus the denominators are equal.

Substituting $\mathbf{L} = (\mathbf{C}\otimes(\mathbf{HU})')$ into the numerator of F_γ we get

$$\hat{\boldsymbol\gamma}'((\mathbf{C}\otimes(\mathbf{HU})')[(\mathbf{X}\otimes\mathbf{I}_p)'(\mathbf{X}\otimes\mathbf{I}_p)]^{-1}(\mathbf{C}\otimes(\mathbf{HU})')')^{-1}\hat{\boldsymbol\gamma}$$

$$= \hat{\boldsymbol\gamma}'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' \otimes \mathbf{U}'\mathbf{H}'\mathbf{H}\mathbf{U})^{-1}\hat{\boldsymbol\gamma}$$

$$= \hat{\boldsymbol\gamma}'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' \otimes \mathbf{U}'\mathbf{U})^{-1}\hat{\boldsymbol\gamma} \text{ since } \mathbf{H}'\mathbf{H} = \mathbf{I}_{p-1}.$$

Then if $\mathbf{U}'\mathbf{U} = \mathbf{I}_{p-1}$, and substituting $\hat{\boldsymbol\gamma} = \mathrm{Vec}(\hat{\boldsymbol\theta}') = \mathrm{Vec}(\hat{\tilde{\boldsymbol\theta}}')$ we get that the denominator is

$$\mathrm{Vec}(\hat{\tilde{\boldsymbol\theta}}')'[(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1} \otimes \mathbf{I}_{p-1}]\mathrm{Vec}(\hat{\tilde{\boldsymbol\theta}}').$$

Then using Searle(1982), Theorem 3, page 333 we get that the denominator is

$$\mathrm{tr}((\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}\hat{\tilde{\boldsymbol\theta}}\hat{\tilde{\boldsymbol\theta}}')$$

which can be written as

$$\mathrm{tr}(\hat{\tilde{\boldsymbol\theta}}'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}\hat{\tilde{\boldsymbol\theta}}),$$

which is the quadratic form in the numerator of F_θ. Therefore the test statistics are equal. □

2.4 F Tests for GLMM with Missing Data

Up to this point we have considered GLMM $(\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma})$ with complete data. Now we would like to allow for missing data. We will assume that data are missing at random, using the definition given in the literature review, or that data are missing by design. Data missing for reasons related to the outcomes of interest or to treatment will not be considered in this paper.

Let $\mathbf{M}_i$ be a p_ixp matrix $[m_{jk}]$ $j = 1,\ldots,p_i$ and $k = 1,\ldots,p$, where

$$m_{jk} = \begin{cases} 1 \text{ if } j=k \text{ and the } k^{th} \text{ observation is present} \\ 0 \text{ otherwise.} \end{cases}$$

Let $\mathbf{M}$ be a block diagonal matrix with $\mathbf{M}_i$ along the diagonal, i.e.,

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & & & \\ & \mathbf{M}_2 & & \\ & & \cdots & \\ & & & \mathbf{M}_N \end{bmatrix} \tag{2.4.1}$$

Thus $\mathbf{M}$ is a $p.$ x Np matrix where $p. = \sum_{i=1}^{N} p_i$.

The data matrix with missing values may then be written as a function of the complete data matrix as follows: $\mathbf{y}^{\star} = \mathbf{My} = \mathbf{M}\text{Vec}(\mathbf{Y}')$. An F statistic for testing hypotheses concerning the parameters $\boldsymbol{\beta}$ is given in the following theorem.

Theorem 2.4.1

Given GLMM($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma}$) with normality and the $\mathbf{M}$ matrix as defined above (2.4.1), assume that the secondary parameter $\boldsymbol{\gamma} = \mathrm{L}\mathrm{Vec}(\boldsymbol{\beta}')$ is testable, (i.e., that the inverse in (2.4.2) exists), where $\mathbf{L}$ is a cxqp known constant matrix with rank(L) = c ≤ qp. A statistic for testing H_o: $\boldsymbol{\gamma} = \mathbf{0}$ is given by

$$F^{\star}_{\gamma} = \frac{\hat{\boldsymbol{\gamma}}'[\mathrm{L}((\mathbf{X}\otimes\mathbf{I}_p)'\mathbf{M}'\mathbf{M}(\mathbf{X}\otimes\mathbf{I}_p))^{-}\mathrm{L}']^{-1}\hat{\boldsymbol{\gamma}}/c}{S^{\star}/(p.-rp)} \qquad (2.4.2)$$

where
$$\begin{aligned} S^{\star} &= (\mathbf{y}^{\star}-\mathbf{M}(\mathbf{X}\otimes\mathbf{I}_p)\mathrm{Vec}(\hat{\boldsymbol{\beta}}'))'(\mathbf{y}^{\star}-\mathbf{M}(\mathbf{X}\otimes\mathbf{I}_p)\mathrm{Vec}(\hat{\boldsymbol{\beta}}')) \\ &= (\mathbf{y}-(\mathbf{X}\otimes\mathbf{I}_p)\mathrm{Vec}(\hat{\boldsymbol{\beta}}'))'\mathbf{M}'\mathbf{M}(\mathbf{y}-(\mathbf{X}\otimes\mathbf{I}_p)\mathrm{Vec}(\hat{\boldsymbol{\beta}}')) \qquad (2.4.3) \end{aligned}$$
and $\mathrm{Vec}(\boldsymbol{\beta}')$ is the ordinary least sqares estimator from GLUM($\mathbf{M}\mathrm{Vec}(\mathbf{Y}'),(\mathbf{X}\otimes\mathbf{I}_p)\mathrm{Vec}(\boldsymbol{\beta}'),\lambda\mathbf{I}_p$).

If $\mathbf{M}(\mathrm{I}_N\otimes\boldsymbol{\Sigma})\mathbf{M}' = \lambda\mathbf{I}_p$ this test statistic has an exact F distribution with c and p.-rp degrees of freedom and noncentrality parameter cF_a.

Proof If we make the transformation $\mathbf{y}^{\star} = \mathbf{M}\mathrm{Vec}(\mathbf{Y}')$ then $\mathbf{y}^{\star} \sim N(\mathbf{M}(\mathbf{X}\otimes\mathbf{I}_p)\mathrm{Vec}(\boldsymbol{\beta}'), \mathbf{M}(\mathbf{I}_N\otimes\Sigma)\mathbf{M}')$. If $\mathbf{M}(\mathbf{I}_N\otimes\Sigma)\mathbf{M}' = \lambda\mathbf{I}_{p.}$ we have the univariate case as described in section 2.1 and a univariate test statistic as given above (2.4.2). This test statistic has an exact F distribution with c and (p.-rp) degrees of freedom and noncentrality parameter cF_a. □

Schwertman (1978) has shown that for the split plot design with missing data in which subjects are considered to be whole plots, the F statistic for testing certain hypotheses involving only subplot parameters is exact if $\boldsymbol{\Sigma}$ has the Huynh-Feldt linear structure. An important condition for this result to hold is that the model must be fully parameterized in the whole plot with a degree of freedom for each of the N whole-plot parameters. If $\boldsymbol{\Sigma}$ has Huynh-Feldt structure but the other conditions are not fulfilled then the following theorem provides an exact result under certain conditions.

Let **H** be a constant specified block diagonal matrix in which each block is a $p_i x(p_i-1)$ Helmert matrix satisfying (2.3.4), i.e.,

$$\mathbf{H} = \begin{bmatrix} H_1 & & & \\ & H_2 & & \\ & & \ddots & \\ & & & H_N \end{bmatrix} \qquad (2.4.4)$$

Thus **H** is a p.x(p.-N) matrix which also satisfies (2.3.4).

Theorem 2.4.2

Assume GLMM($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\boldsymbol{\Sigma}$) with normality. Assume that $\boldsymbol{\gamma} = \mathbf{L}\text{Vec}(\boldsymbol{\beta}')$ is testable, where **L** is a cxqp know constant matrix with rank(**L**) = c ≤ qp. A test statistic for testing H_0: $\boldsymbol{\gamma} = \mathbf{0}$ is given by

$$F = \frac{\hat{\boldsymbol{\gamma}}'[\mathbf{L}((\mathbf{X}\otimes\mathbf{I}_p)'\mathbf{M}'\mathbf{H}\mathbf{H}'\mathbf{M}(\mathbf{X}\otimes\mathbf{I}_p))^{-}\mathbf{L}')^{-1}\hat{\boldsymbol{\gamma}}/c}{S_H/[(p.-N)-r(p-1)]} \qquad (2.4.5)$$

where $S_H = (\mathbf{y} - (\mathbf{X}\otimes\mathbf{I}_p)\text{Vec}(\hat{\boldsymbol{\beta}}'))'\mathbf{M}'\mathbf{H}\mathbf{H}'\mathbf{M}(\mathbf{y} - (\mathbf{X}\otimes\mathbf{I}_p)\text{Vec}(\hat{\boldsymbol{\beta}}'))$ (2.4.6)

If $\mathbf{H}'\mathbf{M}(\mathbf{I}_N\otimes\boldsymbol{\Sigma})\mathbf{M}'\mathbf{H} = \lambda\mathbf{I}_{p.-N}$ then this test statistic has an exact F distribution with c and (p.-N) - r(p-1) degrees of freedom and noncentrality parameter cF_a.

Proof Given GLMM($\mathbf{Y};\mathbf{X}\boldsymbol{\beta},\Sigma$) let $\mathbf{y}^{+} = \mathbf{H}'\mathbf{M}\text{Vec}(\mathbf{Y}')$. Then $\mathbf{y}^{+} \sim N(\mathbf{H}'\mathbf{M}(\mathbf{X}\otimes\mathbf{I}_p)\text{Vec}(\boldsymbol{\beta}'),\mathbf{H}'\mathbf{M}(\mathbf{I}_N\otimes\Sigma)\mathbf{M}'\mathbf{H})$. If $\mathbf{H}'\mathbf{M}(\mathbf{I}_N\otimes\Sigma)\mathbf{M}'\mathbf{H} = \lambda\mathbf{I}_{p.N}$ we have the univariate case as described in section 2.1 and a univariate test statistic, equal to F (2.4.5) may be constructed. □

Corollary 2.4.1

If Σ has the Huynh-Feldt structure and **H** and **M** are constant matrices as described above, (2.4.4) and (2.4.1), then $\mathbf{H}'\mathbf{M}_i\,(\mathbf{I}_N\otimes\Sigma)\mathbf{M}'\mathbf{H} = \lambda\mathbf{I}_{p.\text{-}N}$.

Proof Assume Σ has the Huynh-Feldt structure as given in (2.3.11) i.e., $\Sigma = (\boldsymbol{\alpha}\otimes \mathbf{1}') + (\boldsymbol{\alpha}\otimes\mathbf{1}')' + \lambda\mathbf{I}_p$.

$$\text{Then } \mathbf{M}(\mathbf{I}_N\otimes\Sigma)\mathbf{M}' = \begin{bmatrix} \mathbf{M}_1\Sigma\mathbf{M}_1' & & & \\ & \mathbf{M}_2\Sigma\mathbf{M}_2' & & \\ & & \cdots & \\ & & & \mathbf{M}_N\Sigma\mathbf{M}_N' \end{bmatrix}$$

For all i we have

$$\mathbf{M}_i\Sigma\mathbf{M}_i' = \mathbf{M}_i(\boldsymbol{\alpha}\otimes\mathbf{1}')\mathbf{M}_i' + \mathbf{M}_i(\boldsymbol{\alpha}\otimes\mathbf{1}')'\mathbf{M}_i' + \lambda\mathbf{M}_i\mathbf{M}_i'.$$

If the $\mathbf{M}_i$ are selector matrices as described in section (2.2), they will act to remove rows and columns corresponding to missing data.

Thus $\mathbf{M}_i(\boldsymbol{\alpha} \otimes \mathbf{1}')\mathbf{M}_i'$ and $\mathbf{M}_i(\boldsymbol{\alpha} \otimes \mathbf{1}')'\mathbf{M}_i'$ may be smaller than $(\boldsymbol{\alpha} \otimes \mathbf{1}')$ but it will retain the correct structure. Also $\mathbf{M}_i\mathbf{M}_i' = I_{p_i}$ so we get

$$\mathbf{M}_i\boldsymbol{\Sigma}\mathbf{M}_i' = \mathbf{M}_i(\boldsymbol{\alpha} \otimes \mathbf{1}')\mathbf{M}_i' + \mathbf{M}_i(\boldsymbol{\alpha} \otimes \mathbf{1}')'\mathbf{M}_i' + \lambda I_{p_i}.$$

Thus $\mathbf{M}_i\boldsymbol{\Sigma}\mathbf{M}_i'$ has the Huynh-Feldt structure.

Then we can use the results of corollary 2.3.1 which states that if a covariance matrix has the Huynh-Feldt structure, the corresponding Helmert matrix will diagonalize the covariance matrix. Thus we have that $H_i'\mathbf{M}_i\boldsymbol{\Sigma}\mathbf{M}_i'H_i = \lambda I_{p_i-1}$, for all i. Thus $\mathbf{H}'\mathbf{M}(\mathbf{I}_N \otimes \boldsymbol{\Sigma})\mathbf{M}'\mathbf{H} = \lambda I_{p.-N}$.

2.5 Extension of Results to the Mixed Model

In this section we will extend the previous results for the GLMM to the more general case where it is assumed that each subject's data is distributed according to a normal distribution with unique mean and covariance matrices. The only restrictions on the covariance matrices is that they are required to have the Huynh-Feldt structure with constant λ. We will then extend these results to the mixed model.

Let $\mathbf{Y}_i$ represent a n_ix1 vector of data from the ith subject and assume that $\mathbf{Y}_i \sim \text{NID}(\mathbf{A}_i\mathbf{\Phi},\mathbf{\Sigma}_i)$, where $\mathbf{\Sigma}_i$ has the Huynh-Feldt structure with constant λ, $\mathbf{A}_i$ is a known constant design matrix and $\mathbf{\Phi}$ is a qx1 vector of parameters.

Let $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \ldots, \mathbf{Y}_k')'$ be the Nx1 vector of data from k subjects, where $N = \sum_{i=1}^{k} n_i$. Then $\mathbf{Y} \sim N(\mathbf{A\Phi},\mathbf{\Sigma})$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \cdots \\ \mathbf{A}_k \end{bmatrix} \qquad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & & & \\ & \mathbf{\Sigma}_2 & & \\ & & \cdots & \\ & & & \mathbf{\Sigma}_k \end{bmatrix}$$

and each $\mathbf{\Sigma}_i$ has Huynh-Feldt structure with constant λ.

Theorem 2.5.1

Assume $\mathbf{Y} \sim N(\mathbf{A\Phi},\mathbf{\Sigma})$ as described above. Let $\boldsymbol{\gamma} = \mathbf{L\Phi}$ be a testable secondary parameter(i.e., the inverse in (2.5.1) exists), where $\mathbf{L}$ is a cxq matrix of constants with rank(L) = c ≤ q. A test statistic for testing H_0: $\boldsymbol{\gamma} = 0$ is given by

$$F = \frac{\hat{\boldsymbol{\gamma}}'[L(A'HH'A)^{-}L']^{-1}\hat{\boldsymbol{\gamma}}/c}{S/(N-k)-(q-1)} \tag{2.5.1}$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & & & \\ & \mathbf{H}_2 & & \\ & & \cdots & \\ & & & \mathbf{H}_k \end{bmatrix} \tag{2.5.2}$$

and each H_i is a $(n_i-1)\mathrm{x}n_i$ Helmert matrix satisfying (2.3.4),

$$\mathbf{S} = (\mathbf{H'Y}-\mathbf{H'A}\hat{\mathbf{\Phi}})'(\mathbf{H'Y}-\mathbf{H'A}\hat{\mathbf{\Phi}}),$$

$$= (\mathbf{Y}-\mathbf{A}\hat{\mathbf{\Phi}})'\mathbf{HH'}(\mathbf{Y}-\mathbf{A}\hat{\mathbf{\Phi}}), \tag{2.5.3}$$

$$\hat{\mathbf{\Phi}} = [(\mathbf{H'A})'(\mathbf{H'A})]^{-}(\mathbf{H'A})'\mathbf{H'Y}, \tag{2.5.4}$$

and $\hat{\boldsymbol{\gamma}} = \mathbf{L}\hat{\mathbf{\Phi}}$. (2.5.5)

This test statistic has an exact F distribution with c and (N-k)-(q-1) degrees of freedom and noncentrality parameter cF_a.

Proof We will transform the data vector $\mathbf{Y}$ by $\mathbf{H}$ (2.5.2). This is equivalent to transforming each subject's data by $\mathbf{H}_i$. Thus we have that $\mathbf{H}_i'\mathbf{Y}_i \sim NID(\mathbf{H}_i'\mathbf{A}_i\mathbf{\Phi},\mathbf{H}_i'\mathbf{\Sigma}_i\mathbf{H}_i)$. Since each $\mathbf{\Sigma}_i$ has the Huynh-Feldt structure we get that $\mathbf{H}_i'\mathbf{\Sigma}_i\mathbf{H}_i = \lambda\mathbf{I}_{n_i}$. Thus we have that $\mathbf{H'Y} \sim N(\mathbf{H'A\Phi},\lambda\mathbf{I})$ which is the general linear univariate model. We can then use univariate results to get the test statistic given above.□

Given the results above if we no longer make the assumption that $\mathbf{\Sigma}_i$ has Huynh-Feldt structure but instead assume that there exists some matrix $\mathbf{H}_i$ such that $\mathbf{H}_i'\mathbf{\Sigma}_i\mathbf{H}_i = \lambda\mathbf{I}_{p_i}$ for all i, then the F statistic given in Theorem 2.5.1 will still be exact.

Let $\mathbf{Y}_i$ represent a n_ix1 vector of data from the ith subject. The mixed model with linear covariance structure is given by the following:

$$\mathbf{Y}_i = \mathbf{A}_i\mathbf{\Phi} + \mathbf{B}_i\mathbf{d}_i + \mathbf{e}_i \qquad (2.5.6)$$

where $\mathbf{A}_i$ is an $n_i$xp design matrix for the fixed effects,

- $\mathbf{\Phi}$ is a px1 vector of unknown, constant, population parameters,
- $\mathbf{B}_i$ is an $n_i$xq known constant within-subject design matrix corresponding to the random effects $\mathbf{d}_i$,
- $\mathbf{d}_i$ is a qx1 vector of unknown random subject effects,
- $\mathbf{e}_i$ is a n_ix1 vector of random error terms.

We assume that $\mathbf{d}_i \sim NID(\mathbf{0},\mathbf{D})$ independent of $\mathbf{e}_i \sim NID(\mathbf{0},\sigma^2\mathbf{I}_{n_i})$. Thus $E(\mathbf{Y}_i) = \mathbf{A}_i\mathbf{\Phi}$ and $Var(\mathbf{Y}_i) = \mathbf{\Sigma}_i = \mathbf{B}_i\mathbf{D}\mathbf{B}_i' + \lambda\mathbf{I}_{n_i}$.

Let $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \ldots, \mathbf{Y}_k')'$ where k is the number of subjects.
Then $E(\mathbf{Y}) = \mathbf{A\Phi}$, where $\mathbf{A} = (\mathbf{A}_1', \mathbf{A}_2', \ldots, \mathbf{A}_k')'$
and $\mathrm{Var}(\mathbf{Y}) = \mathbf{B}(\mathbf{I}_k \otimes \mathbf{D})\mathbf{B}' + \lambda \mathbf{I}_N$, where

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & & & \\ & \mathbf{B}_2 & & \\ & & \ddots & \\ & & & \mathbf{B}_k \end{bmatrix} .$$

With the normality assumptions above, the results of Theorem 2.5.1 are applicable to the mixed model if each Σ_i has Huynh-Feldt structure with constant λ. The problem will be that most often when Σ_i has the structure $\mathbf{B}_i\mathbf{D}\mathbf{B}_i' + \lambda \mathbf{I}_{n_i}$ it will not have the Huynh-Feldt structure and so different results will be needed for this case.

Let $\mathbf{H}_i$ be a matrix such that $\mathbf{H}_i'\mathbf{B}_i = \mathbf{0}$. An example of such a matrix is one that has orthonormal columns such that the space spanned by the columns of $\mathbf{H}_i$ is identical to the space spanned by $(\mathbf{I} - \mathbf{B}_i(\mathbf{B}_i'\mathbf{B}_i)^{-1}\mathbf{B}_i')$.

Theorem 2.5.2

Let $\mathbf{Y}$ have the mixed model structure described in the previous section. Let $\mathbf{H}_i$ be a matrix such that $\mathbf{H}_i{}'\mathbf{B}_i=\mathbf{0}$ and let $\mathbf{H}$ be a block diagonal matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & & & \\ & \mathbf{H}_2 & & \\ & & \cdots & \\ & & & \mathbf{H}_k \end{bmatrix}$$

Assume that the secondary parameter $\boldsymbol{\gamma} = \mathbf{L}\boldsymbol{\Phi}$ is testable, where $\mathbf{L}$ is a cxp constant matrix with rank = c ≤ p. Then a test statistic for H_o: $\boldsymbol{\gamma} = \mathbf{0}$ is given by

$$F = \frac{\hat{\boldsymbol{\gamma}}'[\mathbf{L}((\mathbf{H}'\mathbf{A})'(\mathbf{H}'\mathbf{A}))^{-}\mathbf{L}']^{-1}\hat{\boldsymbol{\gamma}}/c}{S/(N\text{-rank}(\mathbf{H}'\mathbf{A}))} \tag{2.5.7}$$

where
$$\begin{aligned} \mathbf{S} &= (\mathbf{H}'\mathbf{Y}-\mathbf{H}'\mathbf{A}\hat{\boldsymbol{\Phi}})'(\mathbf{H}'\mathbf{Y}-\mathbf{H}'\mathbf{A}\hat{\boldsymbol{\Phi}}) \\ &= (\mathbf{Y}-\mathbf{A}\hat{\boldsymbol{\Phi}})'\mathbf{H}\mathbf{H}'(\mathbf{Y}-\mathbf{A}\hat{\boldsymbol{\Phi}}), \end{aligned} \tag{2.5.8}$$

and $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\gamma}}$ are given by (2.5.4-5). This test statistic has an exact F distribution with c and N-rank($\mathbf{H}'\mathbf{A}$) degrees of freedom and noncentrality parameter cF_a.

Proof Transform $\mathbf{Y}$ by $\mathbf{H}'$. This is equivalent to transforming each $\mathbf{Y}_i$ by $\mathbf{H}_i'$. We can write $\mathbf{H}'\mathbf{Y}_i$ in the following way

$$\begin{aligned}\mathbf{H}_i'\mathbf{Y}_i &= \mathbf{H}_i'\mathbf{A}_i\boldsymbol{\Phi} + \mathbf{H}_i'\mathbf{B}_i\mathbf{d}_i + \mathbf{H}_i'\mathbf{e}_i \\ &= \mathbf{H}_i'\mathbf{A}_i\boldsymbol{\Phi} + \mathbf{H}_i'\mathbf{e}_i.\end{aligned}$$

Then $E(\mathbf{H}_i'\mathbf{Y}_i) = \mathbf{H}_i'\mathbf{A}_i\boldsymbol{\Phi}$

and
$$\begin{aligned}\mathrm{Var}(\mathbf{H}_i'\mathbf{Y}_i) &= \mathbf{H}_i'(\mathbf{B}_i\mathbf{D}\mathbf{B}_i' + \lambda\mathbf{I})\mathbf{H}_i \\ &= \lambda\mathbf{H}_i'\mathbf{H}_i \\ &= \lambda\mathbf{I}_{n_i}.\end{aligned}$$

We thus have that $E(\mathbf{H}'\mathbf{Y}) = \mathbf{H}'\mathbf{A}\boldsymbol{\Phi}$ and $\mathrm{Var}(\mathbf{H}'\mathbf{Y}) = \lambda\mathbf{I}_N$ and since $\mathbf{Y}$ follows a multivariate normal distribution, $\mathbf{H}'\mathbf{Y}$ also follows a multivariate normal distribution with the above mean and covariance matrices. Since $\mathrm{Var}(\mathbf{H}'\mathbf{Y}) = \lambda\mathbf{I}_N$ we have a univariate model as described in section 2.1 and a univariate test statistic is constructed based on this theory. □

Next we will develop two approximate F tests for the mixed model. An approximate statistic for testing hypotheses about the fixed effects in the mixed model will be developed using the weighted least squares approach and ordinary (not REML) ML estimators for $\boldsymbol{\tau}$ and σ^2.

Assume that we have the mixed model described previously where $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \ldots, \mathbf{Y}_k')'$ and $\mathbf{A} = (\mathbf{A}_1', \mathbf{A}_2', \ldots, \mathbf{A}_k')'$. In this model $\mathbf{Y} \sim N(\mathbf{A\Phi}, \mathbf{\Sigma})$ where

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & & & \\ & \mathbf{\Sigma}_2 & & \\ & & \cdots & \mathbf{\Sigma}_k \end{bmatrix}$$

and $\mathbf{\Sigma}_i = \mathbf{B}_i \mathbf{D} \mathbf{B}_i' + \sigma^2 \mathbf{I}_{n_i}$.

Let $\mathbf{\Sigma}_i = \sigma^2 \mathbf{\Sigma}_i^\star$ where $\mathbf{\Sigma}_i^\star = \mathbf{B}_i \mathbf{D}^\star \mathbf{B}_i' + \mathbf{I}_{n_i}$ and $\mathbf{D}^\star = (1/\sigma^2)\mathbf{D}$.

We can then make the following transformation, using the fact that $\mathbf{\Sigma} = \sigma^2 \mathbf{\Sigma}^\star$,

$$\mathbf{Z}^\star = \mathbf{\Sigma}^{\star^{-\frac{1}{2}}} \mathbf{Y} \sim N(\mathbf{\Sigma}^{\star^{-\frac{1}{2}}} \mathbf{A\Phi}, \sigma^2 \mathbf{I}_N). \tag{2.5.9}$$

We can construct an F test based on univariate w.l.s. theory using the transformed data:

$$F^\star = \frac{(\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\mathbf{\Sigma}}^{\star^{-1}}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}}) \;/\; \sigma^2 c}{\hat{\sigma}^2_{wls} \;/\; \sigma^2}.$$

Substituting $\hat{\Sigma}^{\star -1} = \sigma^2 \hat{\Sigma}^{-1}$, we get

$$F^{\star} = \frac{(\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}})\ \hat{\sigma}^2/\sigma^2 c}{\hat{\sigma}^2_{wls}/\sigma^2}$$

where $\hat{\sigma}^2_{wls} = \mathbf{Y}'[\hat{\Sigma}^{\star -1} - \hat{\Sigma}^{\star -1}\mathbf{A}(\mathbf{A}'\hat{\Sigma}^{\star -1}\mathbf{A})^{-1}\mathbf{A}'\hat{\Sigma}^{\star -1}]\mathbf{Y} / df$

$$= \mathbf{Y}'[\hat{\sigma}^2\hat{\Sigma}^{-1} - \hat{\sigma}^2\hat{\Sigma}^{-1}\mathbf{A}(\mathbf{A}'\hat{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{A}'\hat{\Sigma}^{-1}\hat{\sigma}^2]\mathbf{Y} / df$$

$$= \hat{\sigma}^2\mathbf{Y}'[\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1}\mathbf{A}(\mathbf{A}'\hat{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{A}'\hat{\Sigma}^{-1}]\mathbf{Y}/df.$$

Then

$$F^{\star} = \frac{(\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}})\ / \ c}{\mathbf{Y}'[\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1}\mathbf{A}(\mathbf{A}'\hat{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{A}'\hat{\Sigma}^{-1}]\mathbf{Y}/df} . \qquad (2.5.10)$$

To the extent that $\hat{\Sigma} \dashrightarrow \Sigma$, this statistic follows an approximate asymptotic F distribution with noncentrality parameter $a = (\mathbf{L\Phi})'[\mathbf{L}(\mathbf{A}'\Sigma^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L\Phi})$ and c numerator degrees of freedom. The appropriate choice for denominator degrees of freedom is not clear. Simulation studies were performed with two choices of denominator degrees of freedom: N-rank(**A**) and N-rank[**A**||**B**]. The results will be discussed in Chapter 4.

The next test will be derived from the canonical form of the mixed model. This is based on unpublished work of Martin (1987). Consider the transformation $\mathbf{Z} = \mathbf{T}'\mathbf{Y}$ where $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \ldots, \mathbf{Y}_k')'$ is the Nx1 vector of data and $\mathbf{T}$ is an NxN orthogonal matrix. We can write this equation in partitioned form as follows:

$$\begin{bmatrix} Z_A \\ Z_B \\ Z_E \end{bmatrix} = \begin{bmatrix} T_A' \\ T_B' \\ T_E' \end{bmatrix} \mathbf{Y} = \begin{bmatrix} T_A'\mathbf{Y} \\ T_B'\mathbf{Y} \\ T_E'\mathbf{Y} \end{bmatrix} \qquad (2.5.11)$$

where $\mathbf{T}_A = \{ \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_p \}$ forms an orthonormal basis for V_p the space spanned by the columns of **A**,

$T_B = \{ \mathbf{t}_{p+1}, \ldots, \mathbf{t}_{p+b} \}$ forms an orthonormal basis for $V_{B|A}$, the space spanned by $\mathbf{B}|\mathbf{A}$ (Arnold 1981 pg. 34),

$T_E = \{ \mathbf{t}_{p+b+1}, \ldots, \mathbf{t}_N \}$ forms an orthonormal basis for V_E, the error space.

If $\mathbf{Y}$ has the structure described in (2.5.6)

$$\text{then } E(Z) = \mathbf{T}'\mathbf{A}\boldsymbol{\Phi} = \begin{bmatrix} T_A'\mathbf{A} \\ T_B'\mathbf{A} \\ T_E'\mathbf{A} \end{bmatrix} \boldsymbol{\Phi} = \begin{bmatrix} U_A \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \boldsymbol{\Phi} \qquad (2.5.12)$$

where $\mathbf{U}_A = \mathbf{T}_A'\mathbf{A}$ and because of orthogonality $\mathbf{T}_B'\mathbf{A} = \mathbf{0}$ and $\mathbf{T}_E'\mathbf{A} = \mathbf{0}$. Also we have that

$$\Sigma_z = \text{Var}(\mathbf{Z}) = \begin{bmatrix} \mathbf{U}_{AB}(\mathbf{I}_k \otimes \mathbf{D})\mathbf{U}_{AB}' + \sigma^2\mathbf{I}_p & \mathbf{U}_{AB}(\mathbf{I}_k \otimes \mathbf{D})\mathbf{U}_B' & \mathbf{0} \\ \mathbf{U}_B(\mathbf{I}_k \otimes \mathbf{D})\mathbf{U}_{AB}' & \mathbf{U}_B(\mathbf{I}_k \otimes \mathbf{D})\mathbf{U}_B' + \sigma^2\mathbf{I}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma^2\mathbf{I}_e \end{bmatrix} \tag{2.5.13}$$

where $\mathbf{U}_{AB} = \mathbf{T}_A'\mathbf{B}$ and $\mathbf{U}_B = \mathbf{T}_B'\mathbf{B}$.

We can write Σ_z in linear form by using $\mathbf{D} = \sum_{g=1}^{m-1} \tau_g \mathbf{G}_g$. Then

$$\Sigma_z = \sum_{g=1}^{m-1} \tau_g \mathbf{G}_g^{\star} + \sigma^2\mathbf{I} \tag{2.5.14}$$

where $\mathbf{G}_g^{\star} = \begin{bmatrix} \mathbf{U}_{AB}(\mathbf{I}_k \otimes \mathbf{G}_g)\mathbf{U}_{AB}' & \mathbf{U}_{AB}(\mathbf{I}_k \otimes \mathbf{G}_g)\mathbf{U}_B' & \mathbf{0} \\ \mathbf{U}_B(\mathbf{I}_k \otimes \mathbf{G}_g)\mathbf{U}_{AB}' & \mathbf{U}_B(\mathbf{I}_k \otimes \mathbf{G}_g)\mathbf{U}_B' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$.

We will calculate an estimate of σ^2 from the pure error space. Note that $E(\mathbf{Z}_E) = \mathbf{0}$ and $\text{Var}(\mathbf{Z}_E) = \sigma^2\mathbf{I}_e$, where $e = N\text{-rank}[\mathbf{A}||\mathbf{B}]$. We can thus use $\mathbf{Z}_E$ to calculate the estimate

$$\hat{\sigma}^2_{REML} = (\mathbf{Z}_E'\mathbf{Z}_E)/e. \tag{2.5.15}$$

M.L. estimates of $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_{m-1})$ can then be calculated from $\mathbf{Z}_A$ and $\mathbf{Z}_B$ using theory discussed previously:

$$\hat{\boldsymbol{\tau}} = [\langle tr(\hat{\Sigma}_{AB}^{-1} G_g^{\star} \hat{\Sigma}_{AB}^{-1} G_h^{\star}) \rangle_{gh}]^{-1} [\langle tr(\hat{\Sigma}_{AB}^{-1} G_g^{\star} \hat{\Sigma}_{AB}^{-1} \hat{C} \rangle_g]$$

$$\text{where } \hat{\Sigma}_{AB} = \hat{Var}\begin{bmatrix} Z_A \\ Z_B \end{bmatrix} \tag{2.5.16}$$

$$\text{and} \quad \hat{C} = \left\{ \begin{bmatrix} Z_A \\ Z_B \end{bmatrix} - \begin{bmatrix} U_A \hat{\Phi} \\ 0 \end{bmatrix} \right\} \left\{ \begin{bmatrix} Z_A \\ Z_B \end{bmatrix} - \begin{bmatrix} U_A \hat{\Phi} \\ 0 \end{bmatrix} \right\}' .$$

The m.l. estimate of $\boldsymbol{\Phi}$ is

$$\hat{\boldsymbol{\Phi}} = (U' \hat{\Sigma}_z^{-1} U)^{-1} U' \hat{\Sigma}_z^{-1} \mathbf{Z} \tag{2.5.17}$$

where

$$U = \begin{bmatrix} U_A \\ 0 \\ 0 \end{bmatrix} .$$

This expression can be reduced by writing

$$\hat{\Sigma}_z^{-1} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} & 0 \\ \mathbf{W}_{21} & \mathbf{W}_{22} & 0 \\ 0 & 0 & (1/\sigma^2) I_e \end{bmatrix} \tag{2.5.18}$$

where $\mathbf{W}_{11} = [\ \mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}]^{-1}$

$\mathbf{W}_{12} = -\mathbf{V}_{11}^{-1}\mathbf{V}_{12}[\ \mathbf{V}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}]^{-1}$

$\mathbf{W}_{22} = [\ \mathbf{V}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}]^{-1}$

and the $\mathbf{V}_{ij}$ represent the blocks of $\hat{\boldsymbol{\Sigma}}_Z$ in the partition. Using this partition we can write

$$\hat{\boldsymbol{\Phi}} = (\mathbf{U}_A'\mathbf{W}_{11}\mathbf{U}_A)^{-1}(\mathbf{U}_A'\mathbf{W}_{11}\mathbf{Z}_A + \mathbf{U}_A'\mathbf{W}_{12}\mathbf{Z}_B) \qquad (2.5.19)$$

where the estimates of $\mathbf{W}$ are calculated by using the m.l. estimates from (2.5.16) in equation (2.5.13). As before, $\hat{\boldsymbol{\Phi}}$ in (2.5.19) depends upon $\hat{\boldsymbol{\tau}}$ (through $\hat{\boldsymbol{\Sigma}}_Z$) and $\hat{\boldsymbol{\tau}}$ (2.5.16) depends upon $\hat{\boldsymbol{\Phi}}$ (through $\hat{\mathbf{C}}$). As usual, an iterative algorithm is required to compute the joint solutions of these equations.

In the following paragraphs we give the motivation for the derivation of an approximate F statistic for testing H_o: $\mathbf{L\Phi} = \mathbf{0}$ using the canonical version of the mixed model. First consider the case where we have GLUM($\mathbf{Y}$:$\mathbf{X}\boldsymbol{\beta}$,$\sigma^2\mathbf{V}$) where $\mathbf{V}$ is known, with secondary parameter $\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\beta}$. Assume that the hypothesis H_o: $\boldsymbol{\theta} = \mathbf{0}$ is testable. The likelihood ratio statistic for testing this hypothesis is given by

$$F = \frac{(\mathbf{C}\hat{\boldsymbol{\beta}})'[\mathbf{C}(\mathbf{X}'(\mathbf{V}\sigma^2)^{-1}\mathbf{X})^{-1}\mathbf{C}']^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}})\ /\ \text{rank}(\mathbf{C})}{\hat{\sigma}^2\ /\ \sigma^2} \qquad (2.5.20)$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$ and $\hat{\sigma}^2 = \mathbf{Y}'[\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{Y}$. This statistic has an exact F distribution with c = rank($\mathbf{C}$) and e degrees of freedom and noncentrality parameter $a = (\mathbf{C}\boldsymbol{\beta})'[\mathbf{C}(\mathbf{X}'(\mathbf{V}\sigma^2)^{-1}\mathbf{X})^{-1}\mathbf{C}']^{-1}(\mathbf{C}\boldsymbol{\beta})$.

We then make the following approximations and substitutions:

$$\hat{\boldsymbol{\Sigma}}_Z \simeq \sigma^2\mathbf{V}\ ,\ \hat{\boldsymbol{\Phi}} \simeq \hat{\boldsymbol{\beta}}\ ,\ \hat{\sigma}^2_{reml} \simeq \hat{\sigma}^2\ ,\ \mathbf{L} = \mathbf{C}\ ,\ \mathbf{A} = \mathbf{X},$$

where $\hat{\boldsymbol{\Phi}}$ is given in (2.5.19), $\hat{\sigma}^2_{reml}$ is given in (2.5.15), and $\hat{\boldsymbol{\Sigma}}_Z$ is the m.l. estimator of $\boldsymbol{\Sigma}_Z$ using $\hat{\sigma}^2_{reml}$ and (2.5.16). Substituting these into (2.5.20) we get the following formula:

$$F_1 = \frac{(\mathbf{L}\hat{\boldsymbol{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\boldsymbol{\Sigma}}_Z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\boldsymbol{\Phi}})\ /\ c}{\hat{\sigma}^2_{reml}/\sigma^2} \qquad (2.5.21)$$

where c = rank($\mathbf{L}$). Note that this formula contains an unknown parameter σ^2 and is therefore not very useful unless we know σ^2, an unlikely occurrence.

Based on a result given in Andrade and Helms (1984), and with appropriate regularity conditions, the numerator of F_1 follows an asymptotic chi-square distribution with c degrees of freedom and noncentrality parameter $a = (\mathbf{L}\boldsymbol{\Phi})'[\mathbf{L}(\mathbf{A}'\boldsymbol{\Sigma}_Z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\boldsymbol{\Phi})$.

The denominator of F_1 has an exact chi-square distribution with e degrees of freedom and is independent of the numerator because the numerator is computed from $\mathbf{Z}_A$ and $\mathbf{Z}_B$ while σ^2_{reml} is computed from $\mathbf{Z}_E$. Therefore this "statistic" follows an asymptotic F distribution with c and e degrees of freedom and noncentrality parameter $a = (\mathbf{L}\boldsymbol{\Phi})'[\mathbf{L}(\mathbf{A}'\boldsymbol{\Sigma}_Z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\boldsymbol{\Phi})$.

Because we don't know σ^2 we will consider a second statistic

$$F_2 = \frac{(\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{A}'(\hat{\Sigma}_z\sigma^2/\hat{\sigma}^2_{AB})^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}})/c}{\hat{\sigma}^2_{reml}/\sigma^2} \qquad (2.5.22)$$

where $\hat{\mathbf{\Phi}}$ and $\hat{\Sigma}_z$ are as given previously and $\hat{\sigma}^2_{AB}$ is the maximum likelihood estimate of σ^2 calculated from $\mathbf{Z}_A$ and $\mathbf{Z}_B$.

Recall that $\hat{\sigma}^2_{reml}$ is calculated from $\mathbf{Z}_E$ which is independent of $\mathbf{Z}_A$ and $\mathbf{Z}_B$. Thus the numerator and denominator of F_2 are independent. Since $\hat{\sigma}^2_{AB} \dashrightarrow \sigma^2$ in the limit, one could argue that the numerator is asymptotically equivalent to the numerator of (2.5.21). The denominator has a small sample chi-square distribution and thus one could argue that F_2 follows an asymptotic F distribution with c and e degrees of freedom and noncentrality parameter $a = (\mathbf{L}\mathbf{\Phi})'[\mathbf{L}(\mathbf{A}'\Sigma_z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\mathbf{\Phi})$. The σ^2 in the numerator and the denominator of F_2 will cancel leaving us with the following:

$$F_2 = \frac{\hat{\sigma}^2_{AB}\,(\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\Sigma}_z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}})/c}{\hat{\sigma}^2_{reml}} \quad . \qquad (2.5.23)$$

We thus have a statistic without the unknown σ^2. However $\hat{\sigma}^2_{AB}$ is typically based on far fewer degrees of freedom then $\hat{\sigma}^2_{reml}$ and it is not independent of $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\tau}}$. Thus a more reasonable estimator is $\hat{\sigma}^2_{reml}$ which we have used in the denominator. If we substitute $\hat{\sigma}^2_{reml}$ for $\hat{\sigma}^2_{AB}$ in (2.5.23) the numerator and denominator will no longer be independent but it seems reasonable that the gain in precision will be worth this loss. The loss of independence would make it difficult to prove exact results about the distribution of the test statistic and so we will work with the approximate asymptotic distribution. Making this substitution we get the following:

$$F_3 = (\mathbf{L}\hat{\boldsymbol{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\boldsymbol{\Sigma}}_z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\boldsymbol{\Phi}})/c. \qquad (2.5.24)$$

This statistic follows an approximate asymptotic F distribution with c and e degrees of freedom and noncentrality parameter $a = (\mathbf{L}\boldsymbol{\Phi})'[\mathbf{L}(\mathbf{A}'\boldsymbol{\Sigma}_z^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\boldsymbol{\Phi})$.

This statistic can be further simplified for computational purposes by using the partition described in (2.5.18). We then get

$$F_{reml} = (\mathbf{L}\hat{\boldsymbol{\Phi}})'[\mathbf{L}(\mathbf{U}_A'\mathbf{W}_{11}\mathbf{U}_A)^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\boldsymbol{\Phi}})/c \qquad (2.5.25)$$

and the noncentrality parameter becomes $a = (\mathbf{L}\boldsymbol{\Phi})'[\mathbf{L}(\mathbf{U}_A'\mathbf{W}_{11}\mathbf{U}_A)^{-1}\mathbf{L}']^{-1}(\mathbf{L}\boldsymbol{\Phi})$. It is important to note here that although $\hat{\sigma}^2_{reml}$ is no longer visible in (2.5.25) it is still used in the calculation of $\mathbf{W}_{11}$.

Chapter 3

Background and Methodology for Simulation Studies

3.1 Review of Test Statistics

In the preceeding chapters we have described several approximate F tests for the mixed model. Simulation studies were conducted to evaluate and compare the performance of these proposed tests and several alternative testing methods. In this section we will describe these studies and their results.

The test statistics that were evaluated are described next.

REML F test

The REML F test was described in Chapter 2. To review briefly: the following test statistic is proposed for testing H_o: $\mathbf{L\Phi} = \mathbf{0}$,

$$F_{REML} = (\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{U}_A{}'\mathbf{W}_{11}\mathbf{U}_A)^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}}) \;/\; \mathrm{rank}(\mathbf{L}) \qquad (3.1.1)$$

where $\hat{\mathbf{\Phi}} = (\mathbf{U}_A{}'\mathbf{W}_{11}\mathbf{U}_A)^{-1}(\mathbf{U}_A{}'\mathbf{W}_{11}\mathbf{Z}_A + \mathbf{U}_A{}'\mathbf{W}_{12}\mathbf{Z}_B).$ (3.1.2)

We hypothesize that this test statistic approximately follows an F distribution with c(=rank(L)) and N-rank[**A**||**B**] degrees of freedom and noncentrality parameter

$$a = (\mathbf{L\Phi})' [\mathbf{L}(\mathbf{U}_A'\mathbf{W}_{11}\mathbf{U}_A)^{-1}\mathbf{L}']^{-1}(\mathbf{L\Phi}). \quad (3.1.3)$$

WLS F test

The weighted least squares F test was described in Chapter 2. To review briefly: the following test statistic is proposed for testing H_0: $\mathbf{L\Phi} = \mathbf{0}$,

$$F_{WLS} = \frac{(\mathbf{L}\hat{\mathbf{\Phi}})'[\mathbf{L}(\mathbf{A}'\hat{\mathbf{\Sigma}}^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L}\hat{\mathbf{\Phi}}) \;/\; \mathrm{rank}(\mathbf{L})}{\mathbf{Y}'[\hat{\mathbf{\Sigma}} - \hat{\mathbf{\Sigma}}^{-1}\mathbf{A}(\mathbf{A}'\hat{\mathbf{\Sigma}}^{-1}\mathbf{A})^{-1}\mathbf{A}'\hat{\mathbf{\Sigma}}^{-1}]\mathbf{Y} \;/(N-p)} \,. \quad (3.1.4)$$

We hypothesize that F_{WLS} approximately follows an F distribution with c(=rank(L)) and N-p degrees of freedom, where p = rank(**A**), and noncentrality parameter

$$a = (\mathbf{L\Phi})'[\mathbf{L}(\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L\Phi}). \quad (3.1.5)$$

WLS F test with adjusted degrees of freedom

The degrees of freedom for the WLS F test given above are correct if $\mathbf{\Sigma}$ is known. Since we don't know $\mathbf{\Sigma}$ it must be estimated in most cases. We propose to adjust the denominator degrees of freedom by using the denominator degrees of freedom from the REML F test. This test statistic will be referred to as F_{WLS2}.

Likelihood Ratio test (Andrade and Helms 1984)

The likelihood ratio test of Andrade and Helms was described in the literature review. This test will be included in the simulation study for comparative purposes. To review briefly: For testing H_o: $\mathbf{L\Phi} = \mathbf{0}$ we may use the test statistic

$$\lambda = \prod_{i=1}^{k} \left[\frac{|\hat{\Sigma}_i|}{|\hat{\Sigma}_{io}|} \right] \tag{3.1.6}$$

where $\hat{\mathbf{\Sigma}}_i$ is the unconstrained MLE of $\mathbf{\Sigma}_i$ and $\hat{\mathbf{\Sigma}}_{oi}$ is the MLE of $\mathbf{\Sigma}_i$ under the null hypothesis. Under H_o, $-2\log\lambda$ asymptotically follows a chi-square distribution with c(=rank(**L**)) degrees of freedom.

Under the alternative hypothesis, H_a: $\mathbf{L\Phi} \neq \mathbf{0}$, $-2\log\lambda$ asymptotically follows a noncentral chi-square distribution with c degrees of freedom and noncentrality parameter

$$a = (\mathbf{L\Phi})'[\mathbf{L}(\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{L}']^{-1}(\mathbf{L\Phi}). \tag{3.1.7}$$

Box-adjusted Ordinary Least Squares F test

An alternative method of analysis for the mixed model is to ignore the correlation within subjects and use ordinary least squares. In effect we would then be assuming the following model:

$$\mathbf{Y} = \mathbf{A\Phi} + \mathbf{e} \tag{3.1.8}$$

where $\mathbf{Y}$ is an Nx1 vector of observations,

$\mathbf{A}$ is an Nxp design matrix,

$\mathbf{\Phi}$ is a px1 vector of parameters to be estimated

and $\mathbf{Y} \sim N(\mathbf{A\Phi}, \sigma^2 \mathbf{I}_N)$.

An F test for testing H_0: $\mathbf{L\Phi} = \mathbf{0}$ is derived by using the results for the general linear univariate model as described in Section 2.1. The test statistic is:

$$F_{OLS} = \frac{(\mathbf{L}\hat{\mathbf{\Phi}})'(\mathbf{L}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\mathbf{\Phi}})/c}{S/(N\text{-}p)} \tag{3.1.9}$$

where $\hat{\mathbf{\Phi}} = (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{Y}$, c = rank($\mathbf{L}$), p = rank($\mathbf{A}$) and $S = (\mathbf{Y}\text{-}\mathbf{A}\hat{\mathbf{\Phi}})'(\mathbf{Y}\text{-}\mathbf{A}\hat{\mathbf{\Phi}})$. Under the model (3.1.8) and the above assumptions this statistic follows an exact F distribution with c and (N-p) degrees of freedom and noncentrality parameter

$$a = (\mathbf{L\Phi})'(\mathbf{L}(\mathbf{A}'\mathbf{A})^{-}\mathbf{L}')^{-1}(\mathbf{L\Phi})/\sigma^2. \tag{3.1.10}$$

This method ignores an important aspect of the data, that is the correlation within subjects. In effect, we are assuming that $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$ which is not true for the mixed model. However, using Theorem 6.1 from Box (1954a), under H_0 the "adjusted F statistic"

$$F_{Box} = \frac{c}{(N-p)b} F_{OLS} \tag{3.1.11}$$

approximately follows a central F distribution with h_1 and h_2 degrees of freedom, where

$$h_1 = \frac{[tr(\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}L'[L(\mathbf{A}'\mathbf{A})^{-1}L']^{-1}L(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Sigma})]^2}{tr[(\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}L'[L(\mathbf{A}'\mathbf{A})^{-1}L']^{-1}L(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Sigma})^2]} \tag{3.1.12}$$

$$h_2 = \frac{[tr[(\mathbf{I}-\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\boldsymbol{\Sigma}]]^2}{tr[(\mathbf{I}-\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\boldsymbol{\Sigma}]^2} \tag{3.1.13}$$

and

$$b = \frac{tr(\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}L'[L(\mathbf{A}'\mathbf{A})^{-1}L']^{-1}L(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Sigma})}{tr[(\mathbf{I}-\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\boldsymbol{\Sigma}]} \,. \tag{3.1.14}$$

To calculate h_1, h_2 and b we will use the estimate of $\boldsymbol{\Sigma}$ from the mixed model with linear covariance structure. We hypothesize that under the alternative hypothesis F_{Box} approximately follows a noncentral F distribution with h_1 and h_2 degrees of freedom and noncentrality parameter $a = (\mathbf{L\Phi})'(L(\mathbf{A}'\mathbf{A})^{-}L')^{-1}(\mathbf{L\Phi})/\sigma^2$.

3.2 Objectives of the Simulation Study

The objectives of the simulation study were, for the designs and models studied:

1) to establish the extent to which the test statistics follow approximate F distributions in both the central and noncentral cases,
2) to calculate and compare Type I error rates and power,
3) to compare the results from the F approximations to the results from the likelihood ratio test.

The simulation study consisted of evaluating the test statistics under four experimental designs: a longitudinal design, a longitudinal design with missing data, a linked cross-sectional design, and a linked cross-sectional design with missing data. These designs are described in more detail in the next section.

Within each experimental design, the test statistics were evaluated under the null hypothesis and under 3 alternative hypotheses. The test statistics were calculated using data generated from the model described below.

The simulation studies were done within the context of a practical study design. We considered a class of hypothetical studies based on the results of a longitudinal study of pulmonary function in children that was analyzed by Fairclough and Helms (1984,1985), and is reviewed below. This class of hypothetical studies was used to select a model and the experimental designs that were used in the simulation study.

3.3 Review of Longitudinal Study of Pulmonary Function in Children

The objectives of the study analyzed by Fairclough and Helms were 1) to determine if overall level and rate of growth of pulmonary function in later childhood is related to occurrence of early age lower respiratory illness (LRI) and 2) to determine if that relationship changed with physical growth, as measured by height. Also it was of interest to study the variation within and among individuals in order to determine if this was related to a history of early LRI's or to physical growth. The objectives of the study indicated that a longitudinal design was appropriate.

LRI history from birth to age 3 was collected for each subject. This included the occurrence of tracheobronchitis, bronchiolitis and pneumonia. Children were then followed from age 3 to approximately age 12, with pulmonary function measurements scheduled for every 3 months. As might be expected, a large portion of the data was missing for a variety of reasons. It was assumed that the data were missing at random. The data set analyzed by Fairclough and Helms consisted of 534 measurements from 50 black children.

The response of interest was $V_{max50\%}$, a measure of small airway function. This was modeled using a mixed model with linear covariance structure as described in the literature review. The methodology used to fit the model has also been described in the literature review.

The results of the analysis indicated that 1) there was no association between early tracheobronchitis and $V_{max50\%}$, 2) bronchiolitis was associated with a constant decrement in mean $V_{max50\%}$ for the heights studied (90-130 cm), 3) pneumonia was associated with an increase in $V_{max50\%}$ growth rate over the heights studied. Within-subject variability was constant as subjects grew from 95 to 125 cm and between-subject variability increased by a factor of 2. Correlation of measurements made at height differences of 5 cm also increased with increasing height.

The final model used to describe the data included fixed effects for intercept and slope, an increment to intercept for those children who had had bronchiolitis and an increment to slope for those children who had had pneumonia. The random effects included in the model were increments to the intercept and slope for each child. The variability of the increment to the intercept was greater for those children who had fewer LRI's. Estimates and standard errors for these parameters are given in Table 3.3.1.

3.4 Designs for a Study of Pulmonary Function in Children

A future study based on these results may be of a confirmatory nature or may be designed to study effects not taken into consideration by Fairclough and Helms. For example, it may be of interest to study the effect of race or it may be of interest to examine environmental factors that may affect pulmonary function, such as air pollution or exposure to cigarette smoke.

The class of hypothetical studies that we considered was assumed to have the same general objectives as the study reviewed above. The response of interest was $V_{max50\%}$, which was assumed to be linearly related to height, and it was assumed that a mixed model with linear covariance structure was appropriate.

The study design was limited to females. We assumed that 30% of the subjects were black and the remaining 70% of the subjects were white. Based on the previous data, we assumed that 46% of the subjects had had bronchiolitis and 25% had had pneumonia. The models contained seperate fixed effects parameters for blacks and whites. The random effects and covariance component parameters, however, were the same for both blacks and whites. The null hypothesis was that the fixed effect parameters for blacks and whites were the same.

The data used in the simulation study was generated from the model described below. The parameter estimates from Fairclough and Helms, given in Table 3.3.1, were used as the "true" values.

The complete longitudinal and linked cross-sectional designs used in the simulation study are illustrated in Figures 3.4.1 and 3.4.2. Each of the proposed designs has 25 subjects and was designed to collect data over the age range 3-10 years. Figure 3.4.1 illustrates the pure longitudinal design in which one cohort is followed for seven years, resulting in a total of 325 data points. Figure 3.4.2 illustrates the LCS design with five cohorts of 5 subjects each. The study would take 3 years to complete and a total of 125 data points would be collected. For the designs with missing data we assumed that approximately 20% of the data were missing completely at random.

Here a "cohort" was defined as a group of subjects born in the same year and having the same measurement schedule as specified by the design. For the power calculations cohort main effects and interactions were assumed to be negligible.

3.5 Methods

3.5.1 Construction of Design Matrices

We first constructed the design matrices for the complete longitudinal design. The other design matrices were then constructed from these.

To construct a design matrix for each subject we generated a sequence of 15 heights which represented growth from age 3 to age 10. To do this we randomly chose an initial height (corresponding to age 3 years) from a normal distribution with mean = 93 cm (the mean height for females at age 3 years) and st.dev. = 3.5 cm. This distribution was obtained from standard growth charts (Tanner, et.al., 1966).

We then randomly chose increments to represent growth for six-month periods. These increments were chosen from a normal distribution with mean = 3 cm and st. dev. = 1.134 cm. This distribution was also derived from the standard growth charts. If the chosen increment was less than 0 a new number was chosen. The increments were then cumulatively added to the starting values.

The design matrices were constructed so that the first four columns corresponded to fixed effects for black children and the remaining four columns corresponded to fixed effects for white children. Typical design matrices for the longitudinal design are given for a black child and for a white child in Table 3.5.1. The heights have been centered by subtracting the mean height (=114.8 cm). Table 3.5.2 contains the randomly chosen starting heights of the 25 subjects.

We used a uniform random number generator (SAS function RANUNI) to randomly assign race and "disease history" to subjects. The composition of the data analyzed by Fairclough and Helms was used as a model for this. Each subject had a 70% chance of being white, a 46% chance of having had bronchiolitis and a 25% chance of having had pneumonia. The actual distribution of subjects is given in Table 3.5.3. Of the 25 subjects, 17(68%) were white, 5 subjects (20%) had had bronchiolitis and 8(32%) had had pneumonia.

Each subject in the complete longitudinal design had 15 measurements giving a total of 375 measurements for this design. The number of measurements for each study is given in Table 3.5.4. The measurements for the longitudinal design are cross-tabulated by race and disease status in Table 3.5.5.

The data "to be missing" was chosen at random using the SAS uniform random number generator RANUNI. Each row of the design matrix, i.e., each measurement, had a 20% chance of being "missing". For the longitudinal design we actually had 21.6% missing data (81 of 375 measurements). As shown in Table 3.5.4, 24 of the 25 subjects had some missing data. The average number of measurements was 11.8 with a minimum of 7 and a maximum of 15. There were a total of 294 measurements for this study. The measurements are cross-tabulated by race and disease status in Table 3.5.6. As can be seen, the incomplete data have approximately the same percentage distributions with respect to race and disease status as the complete data.

The design matrices for the linked cross-sectional design were constructed by selecting certain specified rows from the complete longitudinal design matrices. This was done by first dividing the subjects into cohorts and then, for each subject, selecting rows from the complete longitudinal design matrix according to which cohort the subject was in.

For the linked cross-sectional design shown in Figure 3.4.2, where there are 5 cohorts and 7 measurements per subject the design matrices were constructed as follows:

1) The subjects were divided into equal sized cohorts of 5 subjects each. The first 5 subjects were be assigned to cohort A, the second 5 to cohort B, etc.;
2) Rows from the full longitudinal design matrix were selected according to the subject's cohort. For those subjects in cohort A the first 7 rows were used, for those in cohort B rows 3 through 9 were used, for those in cohort C rows 5 through 11 were used, for those in cohort D rows 7 through 13 were used, and for those in cohort E rows 9 through 15 were used.

Thus each subject had 7 measurements giving a total of 175 measurements for this study. The measurements are cross-tabulated by race and disease status in Table 3.5.7.

For the incomplete linked cross-sectional design we again had planned to randomly select 20% of the data to be missing using the method described previously. In actuality there was 21.1% missing data (37 of 175 measurements). As shown in Table 3.5.4, 20 of the 25 subjects had some missing data. The average number of measurements was 5.5 with a minimum of 3 measurements and a maximum of 7. There were a total of 138 measurements in this study. These are cross-tabulated by race and disease status in Table 3.5.8. As can be seen, the incomplete data have approximately the same percentage distributions with respect to race and disese status as the complete data.

3.5.2 Generation of Data

The values of the dependent variable, $V_{max50\%}$, were generated using a mixed model with linear covariance structure as described in Chapter 2 and the design matrices described in Section 3.5.1.

To review briefly, the mixed model is given by

$$\mathbf{Y}_i = \mathbf{A}_i\mathbf{\Phi} + \mathbf{B}_i\mathbf{d}_i + \mathbf{e}_i \qquad (3.5.1)$$

where

$\mathbf{e}_i$ is an n_ix1 vector of random error terms,and

$\mathbf{d}_i$ is a 2x1 vector of random subject effects.

The design matrix for fixed effects for the ith subject is the n_ix8 matrix, $\mathbf{A}_i$, described above. The design matrix for random effects for the ith subject, $\mathbf{B}_i$, is an n_ix2 matrix, where the first column is a column of 1's and the second column contains the centered heights. The vector of fixed effects, $\mathbf{\Phi} = (\mathbf{\Phi}_b', \mathbf{\Phi}_w')'$, consisted of the following:

$$\mathbf{\Phi}_b = \begin{cases} \phi_{int(b)} & = \text{intercept for blacks} = 1.97, \\ \phi_{slope(b)} & = \text{slope for blacks} = 0.0198, \\ \phi_{BrInt(b)} & = \text{increment to intercept due to the} = -0.33 \\ & \quad \text{occurrence of bronchiolitis in blacks,} \\ \phi_{PnSlope(b)} & = \text{increment to the slope due to the} = 0.0118 \\ & \quad \text{occurrence of pneumonia in blacks,} \end{cases}$$

$$\mathbf{\Phi}_w = \begin{cases} \phi_{int(w)} & = \text{intercept for whites} = 1.97, \\ \phi_{slope(w)} & = \text{slope for whites .} = 0.0198. \\ \phi_{BrInt(w)} & = \text{increment to intercept due to} = -0.33 \\ & \quad \text{occurrence of bronchiolitis in whites,} \\ \phi_{PnSlope(w)} & = \text{increment to slope due to} = 0.0118 \\ & \quad \text{occurrence of pneumonia in whites.} \end{cases}$$

The parameter values given are those under the null hypothesis.

The vector of random effects $\mathbf{d}_i$ consisted of the following:

$$\mathbf{d}_i = \begin{cases} d_{1i} = \text{increment to intercept for ith subject,} \\ d_{2i} = \text{increment to slope for ith subject.} \end{cases}$$

The vector of variance components $\boldsymbol{\tau}$ consisted of the following:

$$\boldsymbol{\tau} = \begin{cases} \tau_1 = \mathrm{Var}(d_1) = 0.146, \\ \tau_2 = \mathrm{Var}(d_2) = 0.000184, \\ \tau_3 = \mathrm{Cov}(d_1,d_2) = 0.00265, \\ \tau_4 = \text{within subject error } (\sigma^2) = 0.118. \end{cases}$$

τ_3 corresponds to a correlation of 0.0094 between the random increment to intercept and the random increment to slope.

We assumed that $\mathbf{d}_i \sim \mathrm{NID}(\mathbf{0},\mathbf{D})$ independent of $\mathbf{e}_i \sim \mathrm{NID}(\mathbf{0},\sigma^2\mathbf{I}_{n_i})$ where

$$\mathbf{D} = \begin{bmatrix} \tau_1 & \tau_3 \\ \tau_3 & \tau_2 \end{bmatrix} = \begin{bmatrix} 0.146 & 0.00265 \\ 0.00265 & 0.000184 \end{bmatrix} \qquad (3.5.2)$$

and $\sigma^2 = \tau_4 = 0.118$. We generated the vectors $\mathbf{e}_i$ and $\mathbf{d}_i$ for each subject using the SAS function RANNOR with fixed seeds chosen from a table of uniform random digits. Equation 3.5.1 was then used to calculate $\mathbf{Y}_i$.

To generate data from the canonical model described in Chapter 2 we used the following equations:

$$Z_A = U_A\Phi + U_{AB}\mathbf{d} + \mathbf{e}_A \tag{3.5.3}$$

$$Z_B = U_B\mathbf{d} + \mathbf{e}_B \tag{3.5.4}$$

where $\mathbf{d}' = (\mathbf{d}_1', \mathbf{d}_2', \ldots, \mathbf{d}_k')$ and the $\mathbf{d}_i$'s are the same as given in the previous section,

$\mathbf{e}_A$ is a px1 vector of random error terms,

$\mathbf{e}_B$ is a kqx1 vector of random error terms.

For these models $\mathbf{e}_A \sim N(\mathbf{0}, \sigma^2\mathbf{I}_p)$ and $\mathbf{e}_B \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{kq})$, and $\mathbf{U}_A, \mathbf{U}_{AB}$, $\mathbf{U}_B$ and $\mathbf{\Phi}$ are as given in Chapter 2.

$\mathbf{U}_A, \mathbf{U}_{AB}$, and $\mathbf{U}_B$ were calculated in the following way:

1) Form $\mathbf{X} = \mathbf{A}||\mathbf{B}$;

2) Form the sum of squares and cross-products matrix $\mathbf{X}'\mathbf{X}$.

3) Take the Cholesky square root of $\mathbf{X}'\mathbf{X}$ to get $\mathbf{U}$:

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_A & \mathbf{U}_{AB} \\ \mathbf{0} & \mathbf{U}_B \end{bmatrix}$$

Then $\mathbf{X}'\mathbf{X} = \mathbf{U}'\mathbf{U}$. We formed $\mathbf{X}'\mathbf{X}$ in Proc Matrix and then used the function HALF to get the Cholesky square root.

The random vectors $\mathbf{e}_A$ and $\mathbf{e}_B$ were generated using the SAS function RANNOR with fixed seeds chosen from a table of uniform random digits. The $\mathbf{d}_i$'s used in the computation of $\mathbf{Y}$ were also used to compute $\mathbf{Z}_A$ and $\mathbf{Z}_B$.

Data was generated under the null hypothesis and under three alternative hypotheses. These are described below. The null hypothesis that was that the intercept and slope for blacks were the same as the intercept and slope for whites. This can be written as H_o: $\mathbf{L\Phi} = \mathbf{0}$, where

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix} .$$

The values of $\mathbf{\Phi}$ under the null hypothesis are given in Table 3.5.9.

The alternative hypothesis can be written as H_a: $\mathbf{L\Phi} \neq \mathbf{0}$. Various specifications of the alternative hypothesis were possible. For the power calculations we chose to consider alternatives in which the main effects of intercept and slope differ for blacks and whites but the other fixed effects remain the same. That is,

$$\begin{aligned} H_a: \quad & \phi_{int(b)} - \phi_{int(w)} \neq 0, \\ & \phi_{slope(b)} - \phi_{slope(w)} \neq 0, \\ & \phi_{BrInt(b)} - \phi_{BrInt(w)} = 0, \\ & \phi_{PnSlope(b)} - \phi_{PnSlope(w)} = 0. \end{aligned}$$

We varied the values of $\phi_{int(w)}$ and $\phi_{slope(w)}$ to form the three alternative hypotheses. These values are given in Table 3.5.9 where the alternative hypotheses are labeled as H_{A1}, H_{A2} and H_{A3}. These values were chosen as points of a 2x2 factorial design. The criteria for choosing the points was that the hypothesized power for the longitudinal design was greater than 0.90 and the number of iterations required to estimate the power at that point was not prohibitive. The method for determining the number of iterations needed is discussed in the next section.

3.5.3 Estimation of Type I Error Rate and Power

The number of iterations needed to estimate Type I error rate was chosen so that the probability of detecting a 0.05 difference between the hypothesized value (0.05) and the true value was 0.95. That is, we determined the number of iterations needed to detect a true Type I error rate of 0.10 or more with 95% power when the hypothesized Type I error rate was 0.05. Sample size (number of iterations) was based upon a one tail alternative and the normal approximation to the binomial distribution (Colton, 1974, Chapter 5).

The number of iterations needed to estimate power was chosen so that a 95% confidence interval for power would be 0.10 wide. That is, if we were interested in 90% power we wanted enough iterations for the 95% confidence interval for the power to be (0.85,0.95). Sample size (number of iterations was based on the normal approximation to the binomial distribution and a confidence interval $\pi \pm 1.96$ S.E., where π is the hypothesized power. The number of iterations needed for estimating Type I error rate and power under the alternative hypotheses for the four designs is given in Table 3.5.10.

To calculate Type I error rate we performed 340 iterations. Each iteration consisted of the following steps:

1) Data was generated under the null hypothesis, H_o: $\mathbf{L\Phi} = \mathbf{0}$, using the parameter values given in Table 3.5.9;
2) The parameter estimates were calculated using an iterative procedure and the test statistics were calculated;
3) For each test statistic it was determined if the null hypothesis was accepted or rejected at $\alpha = 0.05$ by comparing the test statistic to the appropriate value from the hypothesized distribution given in Section 3.1.

After the iterations were completed we calculated the Type I error rate for each test by dividing the number of rejections by the total number of iterations, i.e.,

$$\hat{\pi}_o = \frac{\text{Number of rejections}}{\text{Number of iterations}} . \qquad (3.5.5)$$

The standard error of this estimate is

$$\text{S.E.} = \left[\frac{\hat{\pi}_o(1-\hat{\pi}_o)}{\text{Number of iterations}} \right]^{\frac{1}{2}} . \qquad (3.5.6)$$

We estimated power under the 3 alternative hypotheses given in Table 3.5.9. The number of iterations needed for each alternative hypothesis under each design is given in Table 3.5.10. Each iteration consisted of the following steps:

1) Data was generated under the alternative hypothesis H_o: $\mathbf{L\Phi} \neq \mathbf{0}$;
2) Parameter estimates were calculated using an iterative procedure and the test statistics were calculated;
3) For each test statistic it was determined if the null hypothesis was accepted or rejected at $\alpha = 0.05$ by comparing the test statistic to the appropriate value from the hypothesized distribution given in Section 3.1.

After the iterations were completed we calculated the power for each test by dividing the number of rejections by the total number of iterations, i.e.,

$$\hat{\pi}_A = \frac{\text{Number of rejections}}{\text{Number of iterations}} . \qquad (3.5.7)$$

The standard error of the estimated power is

$$\text{S.E.} = \left[\frac{\hat{\pi}_A(1-\hat{\pi}_A)}{\text{Number of iterations}} \right]^{\frac{1}{2}} . \qquad (3.5.8)$$

Chapter 4.

Results of Simulation Studies

4.1 Introduction

This chapter is organized into four sections, one for each of the experimental designs that were used. Corresponding to each section there are three tables and 10 figures. The first two tables in each section give the means and standard errors of the parameter estimates under the four hypotheses for the original mixed model and for the canonical form of the mixed model, respectively. These tables are included as a check on the iterative procedures used to fit the models.

The third table in each section gives the observed Type I error rate and power and their standard errors for each of the five test statistics. Also given are the Type I error rate and power from the hypothesized distributions of the test statistics which are given in Section 3.1.

To check the distribution of the test statistics we constructed an "F-plot" for the approximate F statistics and a "Chi-square plot" for the LRT statistic. These plots are direct analogues of a "normal plot" (Draper and Smith, 1982, p. 88). If the test statistic follows the hypothesized F distribution the points in the F-plot should approximately follow a straight line with slope equal to one. A slope that is greater than one indicates that the observed values of the test statistic tend to be greater than the values from the hypothesized distribution. A slope that is less than one would indicate the opposite.

The F-plots for each of the four hypotheses are presented separately. In these plots we have drawn two sets of horizontal and vertical reference lines. The first set of lines (those closest to the origin) correspond to the critical values of the F distribution. We can use these two lines to indicate the observed and hypothesized power. The percentage of points above the lower horizontal reference line represents the observed power while the percentage of points to the right of the leftmost vertical reference line represents the hypothesized power. If the points lie approximately on the line " $y = x$ " the observed power and the hypothesized power should be close in value. A large number of points in the upper left "quadrant" or the lower right "quadrant" of the plot indicates a substantial difference between the observed power and the hypothesized power. The second set of reference lines is provided as an aid in constructing the " $y = x$ " line for comparison. Note that in this type of plot the letter "A" represents one point, "B" represents two points, etc.

The critical value for the F_{Box} test statistic depends on the data and was computed separately in each iteration from Box's correction factor for the degrees of freedom. However, for the F-plot the degrees of freedom (h_1, h_2) and correction factor (b) were calculated using the formulas given in Section 3.1 and substituting the values of τ and σ^2 given in Table 3.5.9. Only the second set of reference lines is given on these plots since the critical value changed for each iteration.

For the sake of completeness we have also included the average degrees of freedom and correction factor from the simulation studies. These are given in Tables 4.1.2 - 4.1.5. The average values were close to the values used to construct the F-plots although the average value of h_2 was consistently higher than the value of h_2 used in the F-plot.

4.2 Results for Complete Longitudinal Design

The results of the simulation study for the complete longitudinal design are given in Tables 4.2.1 - 4.2.3 and Figures 4.2.1 - 4.2.19. The method of scoring was used to estimate the parameters for both forms of the mixed model. As can be seen from Tables 4.2.1 and 4.2.2 and in comparison with the "true" values given in Table 3.5.9, the method of scoring provided accurate estimates of the parameters for both forms of the model and under all of the hypotheses. All $\hat{\mathbf{D}}$ matrices were positive definite and the method converged in all cases.

Table 4.2.3 gives the results for the five proposed test statistics. For this design F_{REML} gave the most accurate Type I error rate and power values. The F_{Box} statistic gave an accurate Type I error rate but the power was very low. The other test statistics gave observed Type I error rates that were substantially larger than 0.05.

The F_{REML} test statistic with 4 and 333 degrees of freedom gave an observed Type I error rate of 0.076 which was within two standard errors of the hypothesized value of 0.05. Under each of the alternative hypotheses the observed power was within one standard error of the hypothesized power.

The F-plots for checking the distribution of F_{REML} are given in Figures 4.2.1 - 4.2.4. In each of the four plots the points approximately follow a straight line. Under the null hypothesis the line seems to have a slope slightly greater than one and there are some points in the upper left quadrant. This suggests that there is probably some inflation of the Type I error rate even though the observed value was not significantly different from the hypothesized value.

Under the first alternative hypothesis the slope is very close to one while under the second and third alternative hypotheses the slopes are greater than one. As mentioned previously, when the slope is greater than one it is expected that the observed power would be greater than the hypothesized power. This indicates that the test is conservative although in these cases the test appears to be only slightly conservative.

The F_{WLS} statistic with 4 and 367 degrees of freedom gave an observed Type I error rate of 0.19 and the F_{WLS2} statistic with 4 and 333 degrees of freedom gave an observed Type I error rate of 0.16. Both of these values were more than two standard errors away from the hypothesized value of 0.05. The observed power values were all within two standard errors of the hypothesized values with the exception of

F_{WLS} under the first alternative hypothesis. In this case the observed power was substantially larger than the hypothesized power.

The F-plots for F_{WLS} and F_{WLS2} are given in Figures 4.2.5-4.2.12. The patterns for the two statistics are similar. Under each of the four hypotheses the points deviate from the " y = x " line with slope equal to one. Under the null hypothesis the points approximately follow a straight line with slope greater than one and there are a number of points in the upper left quadrant. Under the alternative hypotheses the points begin to curve upwards as the values increase. Thus for both statistics we conclude that there is a significant departure from the hypothesized distribution.

The LRT statistic gave an observed Type I error rate of 0.14. This was more than two standard errors greater than the hypothesized value of 0.05. The observed power values were all within two standard errors of the hypothesized values.

The χ^2-plots for the LRT statistic are given in Figures 4.2.13 - 4.2.16. Under each hypothesis the points approximately follow straight lines. Under the null hypothesis this line has a slope greater than one, which indicates that the Type I error rate is inflated. Under the second and third alternative hypotheses the slope is slightly less than one. Under the first alternative hypothesis the points followed quite closely the straight line " y = x ". Although the last three plots indicate that the hypothesized distribution fits well under the alternative hypotheses, the test statistic is not of much practical use because of the inflated Type I error rate.

The F_{Box} statistic gave an observed Type I error rate of 0.062. This was not significantly different than the hypothesized value of 0.05. The observed power values under the first and second alternative hypotheses were significantly lower than the hypothesized power values. This indicates that if the hypothesized distribution were used for power computations the actual power would be less than the specified power. The observed power values were also much lower than the observed values given by the other statistics.

The F-plots for the F_{Box} statistic are given in Figures 4.2.17 - 4.2.19. A plot is not given for the third alternative hypothesis because the noncentrality parameter became so large that technically the plot could not be constructed. Under the null hypothesis the points approximately follow a straight line with slope equal to one. Under the first and second alternative hypotheses the points follow straight lines with slopes much less than one. This may indicate that the actual distribution is proportional to the hypothesized distribution. It may also indicate that there is a problem with the noncentrality parameter. These issues will be discussed in more detail in a later section.

In summary, for the longitudinal design the F_{REML} statistic gave the most accurate estimates of Type I error rate and power. The F_{Box} statistic accurately estimated the Type I error rate but had very low power. The other statistics all had severely inflated Type I error rates. The F_{REML} statistic followed its hypothesized distribution quite closely while F_{WLS} and F_{WLS2} did not.

4.3 Results for Longitudinal Design with Missing Data

The results of the simulation study for the longitudinal design with 20% randomly missing data are given in Tables 4.3.1 - 4.3.3 and Figures 4.3.1 -4.3.19. The method of scoring was used to estimate the parameters for both forms of the mixed model. As can be seen from Tables 4.3.1 and 4.3.2 and in comparison with the true values given in Table 3.5.9, the method of scoring provided accurate estimates of the parameters for both forms of the model and under all of the hypotheses. All $\hat{\mathbf{D}}$ matrices were positive definite and the method converged in all cases.

Table 4.3.3 gives the results for the five proposed test statistics. F_{REML} gave the most accurate Type I error rate and power for this design. The LRT and F_{Box} statistics also gave accurate Type I error rates but power was inflated for the LRT statistic and extremely low for the F_{Box} statistic. F_{WLS} and F_{WLS2} both gave observed Type I error rates that were substantially larger than 0.05.

The F_{REML} statistic with 4 and 252 degrees of freedom gave an observed Type I error rate of 0.071 which is within two standard errors of the hypothesized value of 0.05. Under each of the alternative hypotheses, the observed power was within two standard errors of the hypothesized power.

The F-plots for checking the distribution of F_{REML} are given in Figures 4.3.1 - 4.3.4. The patterns seen here are similar to those seen in the complete longitudinal design. In each of the four plots the points approximately follow a straight line. Under the null hypothesis the line seems to have a slope slightly greater than one. As in the complete longitudinal design this indicates that there is probably some inflation of the Type I error rate. Under the alternative hypotheses the points approximately follow straight lines with slopes equal to or slightly greater than one. In these cases the observed and hypothesized power were not significantly different.

The F_{WLS} statistic with 4 and 286 degrees of freedom gave an observed Type I error rate of 0.12 and the F_{WLS2} statistic with 4 and 252 degrees of freedom gave an observed Type I error rate of 0.091. Both of these values were more than two standard errors away from the hypothesized value of 0.05. For both statistics the observed power values under the first alternative hypothesis were substantially larger than the hypothesized values. The observed power values under the second and third alternative hypotheses were within two standard errors of the hypothesized values.

The F-plots for F_{WLS} and F_{WLS2} are given in Figures 4.3.5 - 4.3.12. The patterns for the two statistics are similar. Under the null hypothesis the points diverge from the straight line with slope equal to one. The divergence becomes more extreme under the alternative hypotheses. Again we conclude that there is a significant departure from the hypothesized distribution.

The LRT statistic gave an observed Type I error rate of 0.079. This was not significantly different from the hypothesized value of 0.05. The observed power for the first alternative hypothesis was substantially larger than the hypothesized power. The observed power values for the remaining hypotheses were within two standard errors of the hypothesized values.

The χ^2-plots for the LRT statistic are given in Figures 4.3.13 - 4.3.16. The patterns seen here are similar to those seen in the complete longitudinal design. Under each hypothesis the points approximately follow straight lines. Under the null hypothesis this line has a slope greater than one which indicates that the Type I error rate is inflated although not significantly in this case. Under the second and third alternative hypotheses the slope is slightly less than one. Under the first alternative hypothesis the points follow quite closely a straight line with slope equal to one although there are some points in the upper left quadrant which accounts for the higher observed power. The plots indicate that the hypothesized distribution fits well although there is some inflation of the Type I error rate.

The F_{Box} statistic gave an observed Type I error rate of 0.068. This was not significantly different from the hypothesized value of 0.05. The observed power values under the first and second alternative hypotheses were significantly lower than the hypothesized power values. The observed values were also much lower than the observed values given by the other statistics.

The F-plots for the F_{Box} statistic are given in Figures 4.3.17 - 4.3.19. A plot is not given for the third alternative hypothesis because the noncentrality parameter became so large that technically the plot could not be constructed. The patterns are similar to those seen in the complete longitudinal design. Under the null hypothesis the points approximately follow a straight line with slope equal to one. Under the first and second alternative hypotheses the points follow a straight line with slope equal to one. Under the first and second alternatives the points follow straight lines with slopes much less than one.

In summary the results for the longitudinal design with missing data are similar to the results for the complete longitudinal design. The F_{REML} gave the most accurate estimates of Type I error rate and power and followed its hypothesized distribution quite closely. F_{WLS} and F_{WLS2} had inflated Type I error rates and the F-plots indicated that there were significant departures from the hypothesized distributions. The LRT statistic gave a more accurate estimate of Type I error rate for this design. The F_{Box} statistic again gave an accurate Type I error rate but had very low power.

4.4 Results for Linked Cross-Sectional Design

The results of the simulation study for the linked cross-sectional design are given in Tables 4.4.1 - 4.4.3 and Figures 4.4.1 - 4.4.20. In this design we encountered a problem with the method of scoring when used to fit the original form of the mixed model. The process frequently failed to converge and frequently produced unreasonable parameter estimates. This seemed to be associated with the reduction in the total number of measurements. Because of this we used the EM algorithm for the linked cross-sectional designs. This algorithm produced reasonable parameter estimates and converged according to our criteria. The method of scoring was used for the canonical form of the mixed model and produced satisfactory results.

As can be seen from Tables 4.4.1 and 4.4.2 and in comparison with the "true" parameter values given in Table 3.5.9, the two iterative methods performed reasonably well.

Table 4.4.3 gives the results for the five proposed test statistics. For this design both F_{REML} and F_{WLS2} gave accurate Type I error rates, with F_{REML} being slightly more accurate. F_{WLS} gave an observed Type I error rate that was substantially larger than 0.05, as did the LRT statistic and the F_{Box} statistic.

The F_{REML} test statistic with 4 and 133 degrees of freedom gave an observed Type I error rate of 0.050 which was exactly equal to the hypothesized value of 0.05. Under each of the alternative hypotheses the observed power was within two standard errors of the hypothesized power values.

The F-plots for checking the distribution of F_{REML} are given in Figures 4.4.1 - 4.4.4. Under the null hypothesis the points approximately follow a straight line with slope equal to one. Under the alternative hypothesis the points approximately follow straight lines with slopes slightly greater than one.

The F_{WLS} statistic with 4 and 167 degrees of freedom gave an observed Type I error rate of 0.14 which was more than two standard errors away from the hypothesized value of 0.05. The observed power values under the first and second alternative hypotheses were substantially larger than the hypothesized power values.

The F-plots for F_{WLS} are given in Figures 4.4.5 - 4.4.8. Under each of the hypotheses the points diverge upward from the straight line with slope equal to one. We conclude that there is significant departure from the hypothesized distribution.

The F_{WLS2} statistic with 4 and 133 degrees of freedom gave an observed Type I error rate of 0.068 which was within two standard errors of the hypothesized value of 0.05. The observed power values were all within two standard errors of the hypothesized power values.

The F-plots for F_{WLS2} are given in Figures 4.4.9 - 4.4.12. Under each of the hypotheses the points approximately follow a straight line. Under the null and first alternative hypotheses the slope of this line is close to one while under the remaining two hypotheses the slope is slightly greater than one. From this we conclude that F_{WLS2} closely followed its hypothesized distribution.

The LRT statistic gave an observed Type I error rate of 0.094. This was more than two standard errors greater than the hypothesized value of 0.05. The observed power values were all more than two standard errors greater than the hypothesized values.

The χ^2-plots for the LRT statistic are given in Figures 4.4.13 - 4.4.16. Under each hypothesis the points approximately follow straight lines. Under the null hypothesis the slope of this line is greater than one and the points in the upper left quadrant indicate that Type I error rate is inflated. Under the alternative hypotheses the slopes are greater than one and there are points in the upper left quadrant. This is consistent with the observed power being greater than the hypothesized power.

The F_{Box} statistic gave an observed Type I error rate of 0.088 which was more than two standard errors greater than the hypothesized value of 0.05. The observed power values were all substantially lower than the hypothesized power values.

The F-plots for the F_{Box} statistic are given in Figures 4.4.17 - 4.4.20. The patterns are similar to those seen in the previous designs. Under the null hypothesis the points approximately follow a straight line with slope greater than one. Under the alternative hypotheses the points diverge drastically from the line with slope equal to one.

In summary, for the linked cross-sectional design the F_{REML} statistic and the F_{WLS2} statistic both gave accurate estimates of the Type I error rate and power. The F_{REML} statistic gave a slightly more accurate estimate. Both statistics followed their hypothesized distributions closely. The remaining statistics had inflated Type I error rates.

4.5 Results for the Linked Cross-Sectional Design with Missing Data

The results of the simulation study for the linked cross-sectional design with missing data are given in Tables 4.5.1 - 4.5.3 and Figures 4.5.1 - 4.5.20. The EM algorithm was used to estimate the parameters to fit the original form of the mixed model and the method of scoring was used to fit the canonical form of the mixed model. As can be seen in Tables 4.5.1 and 4.5.2 and in comparison with the "true" parameter values given in Table 3.5.9, both methods provided accurate estimates of the parameters.

Table 4.5.3 gives the results for the five proposed test statistics. For this design F_{REML} gave the most accurate Type I error rate. The other test statistics had observed Type I error rates that were substantially larger than 0.05.

The F_{REML} statistic with 4 and 96 degrees of freedom gave an observed Type I error rate of 0.059 which was within two standard errors of the hypothesized value of 0.05. Under the first alternative hypothesis the observed power was significantly lower than the hypothesized power. There was no significant difference between the observed power and the hypothesized power under the second and third alternative hypotheses.

The F-plots for F_{REML} are given in Figures 4.5.1 - 4.5.4. Under the null hypothesis the points approximately follow a straight line with slope close to one. Under the first alternative hypothesis many of the points are below the line and in the lower right quadrant which is consistent with the observed power being less than the hypothesized power. Under the second and third alternative hypotheses the points approximately follow straight lines with slopes slightly greater than one, although there is some divergence from this as the observed values become larger.

The F_{WLS} statistic with 4 and 130 degrees of freedom gave an observed Type I error rate of 0.16. This is more than two standard errors away from the hypothesized value of 0.05. The observed power under the first and second alternative hypotheses was more than two standard errors above the hypothesized power.

The F-plots for F_{WLS} are given in Figures 4.5.5 - 4.5.8. Under each of the hypotheses the points approximately follow straight lines with slopes greater than one. Under each of the hypotheses there are some points in the upper left quadrant. This is consistent with the observed values being greater than the hypothesized values. We conclude that there is significant departure from the hypothesized distribution.

The F_{WLS2} statistic with 4 and 96 degrees of freedom gave an observed Type I error rate of 0.088 which was significantly different from the hypothesized value of 0.05. The observed power under the first alternative hypothesis was significantly less than the hypothesized power.

The F-plots for F_{WLS2} are given in Figures 4.5.9 - 4.5.12. Under the null hypothesis the points approximately follow a straight line with a slope greater than one. There are some points in the upper left hand quadrant which is consistent with the observed power being greater than the hypothesized power. Under the first alternative hypothesis the points approximately follow a straight line but there are some points in the lower right quadrant. This is consistent with the observed power being less than the hypothesized power. Under the second and third alternative hypotheses the points diverge upward from the straight line with slope equal to one as the values become larger.

The LRT statistic gave an observed Type I error rate of 0.14. This was more than two standard errors greater than the hypothesized value of 0.05. The observed power values under the second and third alternative hypotheses were significantly greater than the hypothesized values.

The χ^2-plots for the LRT statistic are given in Figure 4.5.13 - 4.5.16. Under the null hypothesis the points approximately follow a straight line with slope greater than one. There is a substantial percentage of points in the upper left hand quadrant which is consistent with the observed power being greater than the hypothesized power. Under the second alternative hypothesis the points approximately follow a straight line with slope slightly greater than one. Under the second and third alternative hypotheses the points diverge from the straight line.

The F_{Box} statistic gave an observed Type I error rate of 0.103 which was more than two standard errors above the hypothesized value of 0.05. Under each of the alternative hypotheses the observed power was significantly less than the hypothesized power.

The F-plots for the F_{Box} statistic are given in Figures 4.5.17 - 4.5.20. The patterns are similar to those seen in the previous designs. Under the null hypothesis the points approximately follow a straight line with slope greater than one. Under the alternative hypotheses the points diverge drastically from the line with slope equal to one.

In summary, for the linked cross-sectional design with missing data the F_{REML} statistic gave the most accurate estimates of Type I error rate and power. The remaining test statistics had inflated Type I error rates.

4.6 Summary

The F_{REML} statistic consistently gave observed Type I error rates and power that were within two standard errors of the hypothesized values. In only one case, the linked cross-sectional design with missing data under the first alternative hypothesis, was the observed value significantly different than the hypothesized value. The F-plots for all designs indicated that the hypothesized distribution fit well.

The F_{WLS} statistic gave significantly inflated Type I error rates for all four designs. The statistic gave observed power values that were consistently higher than the hypothesized values. In half of the cases they were significantly higher. In all four designs, the F-plots indicated that there were significant departures from the hypothesized distribution.

The F_{WLS2} statistic also gave inflated Type I error rates. In 3 of the 4 designs this inflation caused the observed values to be significantly higher than 0.05. The statistic gave accurate estimates of power. In only 2 of 12 cases was the observed power significantly different from the hypothesized power. The F-plots indicated some departures from the hypothesized distribution.

The LRT statistic also gave inflated Type I error rates. In 3 of the 4 designs the observed Type I error rates were significantly higher than 0.05. The statistic gave observed power values that were higher than the hypothesized values for most cases. The difference was significant in 6 of the 12 cases.

The F_{Box} statistic gave Type I error rates that were only slightly inflated for the longitudinal designs. For the linked cross-sectional designs the observed Type I error rates were significantly larger than 0.05. The observed power values were significantly smaller than the hypothesized values in 10 of 12 cases. The F-plots for this statistic indicated that the hypothesized distribution did not fit well under the alternative hypotheses. This leads us to suspect that there may be a problem with the noncentrality parameter. The fact that the hypothesized power was greater than or equal to 0.95 in all cases while the observed power varied from 0.21 to 0.95 also seems to indicate a problem with the noncentrality parameter.

Chapter 5

Summary and Discussion

The purpose of this work was to extend results from the general linear univariate model and the general linear multivariate model to special cases of the mixed model with linear covariance structure. These extensions were then used to motivate approximate F statistics for the mixed model. Three approximate F statistics were proposed; one was based on the canonical form of the mixed model (F_{REML}) and two were based on weighted least squares (F_{WLS}, F_{WLS2}). In a simulation study the three statistics were compared to each other, to the likelihood ratio statistic (LRT) of Andrade and Helms (1984) and to an adjusted ordinary least squares test (F_{Box}).

In the simulation study we varied the designs and hypotheses under which the test statistics were evaluated. The designs that were used were a complete longitudinal design, a longitudinal design with missing data, a linked cross-sectional design and a linked cross-sectional design with missing data. Thus the results of our study may be used to compare the designs as well as the test statistics.

5.1 Comparison of Test Statistics

We found that the F_{REML} statistic produced the most accurate Type I error rate and had the closest conformance with the hypothesized distribution. The F_{WLS}, F_{WLS2} and LRT statistics produced inflated Type I error rates. The F_{Box} statistic gave accurate Type I error rates for some cases but the power was very low under some of the alternative hypotheses.

Given these results we would recommend the use of the F_{REML} statistic. However, one drawback to using this statistic is the amount of computer time required to calculate it. To calculate the F_{REML} statistic one must fit the canonical form of the mixed model, while to calculate the F_{WLS} and F_{WLS2} test statistics one must fit the original form of the mixed model. Both types of models require the use of an iterative procedure but the size of the matrices to be inverted will be different. For the original form of the mixed model these matrices are no bigger than $n_i x n_i$ where n_i is the number of measurements for the ith subject. In the longitudinal design n_i was 15.

For the canonical form of the mixed model one must take the inverse of an mxm matrix where m = q*k-p, where q is the number of random effects, k is the number of subjects and p is the rank of the design matrix (**A**). In the longitudinal design q was 2, k was 25 and p was 8 so we had to invert a 42x42 matrix.

The size of the matrix to be inverted for the canonical model is a function of the number of subjects and will increase as the number of subjects increases. For the original model the number of matrices to be inverted will increase but the size of them will not. The difference in computer time and memory space required can be substantial. In some test runs, using a dataset with 30 subjects and fitting a slightly different model, we found that fitting the canonical model took approximately 2 and a half times as long as fitting the original mixed model.

The F_{WLS2} statistic performed better than the F_{WLS} statistic. Since they required the same calculations we would recommend the use of F_{WLS2} over F_{WLS}. The advantage of using the F_{WLS2} statistic, over the F_{REML} statistic, is that as the number of subjects increases it is easier to calculate than the F_{REML} statistic. The disadvantage of using the F_{WLS2} statistic is that the Type I error rate is significantly inflated. The power, however, is accurate.

In the cases that we studied the LRT statistic did not perform as well as the other statistics. For this reason we would not recommend its use. Also to calculate this statistic one must fit two models, both times using an iterative process. This may require a substantial amount of computer time.

The F_{Box} statistic gave an accurate Type I error rate but very low power for some of the cases that we studied. The advantage of using this statistic is that it is relatively easy to calculate since an iterative process is not required. However, an estimate of Σ is needed to calculate the adjusted degrees of freedom and correction factor. The performance of the statistic will thus depend on the quality of this estimate. We used the estimate of Σ obtained from fitting the mixed model with linear covariance structure.

Another disadvantage of the F_{Box} statistic is that its exact distribution under the alternative hypotheses is unknown. The distribution that we hypothesized did not fit so power calculations done using this would not be very realistic.

5.2 Comparison of Designs

Since the F_{REML} statistic followed its hypothesized distribution closely we can use the noncentral F distribution to compute power for various designs. This will give us a way of comparing the four designs that we used under the conditions that we specified.

The design that gave the highest power, according to our simulation study, was the longitudinal design. This was expected since this design had the most observations (15 per subject). Missing data did not change the power drastically. Under the first alternative hypothesis the power dropped from 0.82 to 0.75, while under the second and third alternative hypotheses there was only a slight change: (0.95 to 0.96) and (0.93 to 0.94), respectively. (The powers presented here are the observed powers and thus are subject to the random variation of the data. This accounts for the fact that the power actually rose slightly when there was missing data.)

For the linked cross-sectional design (7 observations per subject) the power under the first alternative hypothesis was less than one-half that in the longitudinal design (15 observations per subject). For this hypothesis the power dropped from 0.82 to 0.30. Under the second alternative hypothesis the power dropped from 0.95 to 0.81 while under the third hypothesis the power increased slightly from 0.93 to 0.94.

Again missing data did not change the power drastically. Under the first alternative hypothesis the power dropped from 0.30 to 0.22, under the second alternative hypothesis the power changed from 0.81 to 0.84 and under the third alternative hypothesis the power went from 0.94 to 0.92.

Thus we find that the greatest difference in the power of the designs is under the first alternative hypothesis. Under the second and third alternative hypotheses there is not much difference in power.

5.3 Future Research

The following are areas where future research is needed:

1. Evaluate the distributions of the F_{REML} and F_{WLS2} test statistics under a variety of designs, models and parameter values.
2. Determine the exact asymptotic distributions of these test statistics under very general regularity conditions.
3. Find an approximate noncentral distribution for the F_{Box} statistic.

We anticipate that some of these tasks will pose difficult mathematical problems.

References

Afifi, A.A., Elashoff, R.M. (1966) "Missing Observations in Multivariate Statistics I. Review of the Literature", Journal of the American Statistical Association, 61, 595-604.

Afifi, A.A., Elashoff, R.M. (1967) "Missing Observations in Multivariate Statistics II. Point Estimation in Simple Linear Regression", Journal of the American Statistical Association, 62, 10-29.

Afifi, A.A., Elashoff, R.M. (1969a) "Missing Observations in Multivariate Statistics III. Large Sample Analysis of Simple Linear Regression", Journal of the American Statistical Association, 64, 337-358.

Afifi, A.A., Elashoff, R.M. (1969b) "Missing Observations in Multivariate Statistics IV. A Note on Simple Linear Regression", Journal of the American Statistical Association, 64, 359-365.

Anderson, T.W. (1957) "Maximum Likelihood Estimates for a Multivariate Normal Distribution when some Observations are Missing", Journal of the American Statistical Association, 52, 200-203.

Anderson, T.W. (1970) "Estimation of Covariance Matrices Which are Linear Combinations of Given Matrices". Essays in Probability and Statistics, 1-24. University of North Carolina Press, Chapel Hill, NC.

Anderson, T.W. (1973) "Asymptotically Efficient Estimation of Covariance Matrices with Linear Structure" The Annals of Statistics,1, 135-141.

Anderson, T.W. (1984) An Introduction to Multivariate Statistical Analysis, 2nd ed., John Wiley and Sons, Inc.,New York.

Andrade, D.F., Helms, R.W. (1984) "ML Estimation and LR tests for the Multivariate Normal Distribution with Patterned Mean and Covariance Matrix. Complete and Incomplete - Data Cases" . Institute of Statistics Mimeo Series, No. 1455. University of North Carolina at Chapel Hill.

Baltes, P.B. (1968) "Longitudinal and Cross-Sectional Sequences in the Study of Age and Generation Effects", Human Development, 11, 145-171.

Box, G.E.P. (1954a) "Some Theorems on Quadratic Forms Applied in the Study of Analysis Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification", Annals of Mathematical Statistics, 25, 290-302.

Box, G.E.P. (1954b) "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification", Annals of Mathematical Statistics, 25, 484-498.

Boyles, R.A. (1983) "On the Convergence of the EM Algorithm", Journal of the Royal Statistical Society B, 45, 47-50.

Brown, B.W. (1980) "The Crossover Experiment for Clinical Trials", Biometrics, 36, 69-79.

Cochran, W.G., Cox, G.M. (1957) Experimental Design, John Wiley and Son, Inc., New York.

Colton, T. (1974) Statistics in Medicine, Little, Brown and Co., Boston.

Cook, N.R., Ware, J.H. (1983) "Design and Analysis Methods for Longitudinal Research", Annual Review of Public Health, 4, 1-23.

Cox, D.R. (1958) Planning of Experiments, John Wiley and Sons Inc., New York.

Davidson, M. L. (1980) "The Multivariate Approach to Repeated Measures", Unpublished manuscript.

Dempster, A.P., Laird, N.M., Rubin, D.B. (1977) "Maximum Likelihood from Incomplete Data via the EM Algorithm", Journal of the Royal Statistical Society B, 39, 1-38.

Fairclough, D.L., Helms, R.W. (1984) "Mixed Effects Model Analyses of Incomplete Longitudinal Pulmonary Function Measurements in Children", Institute of Statistics Mimeo Series No. 1470, University of North Carolina at Chapel Hill.

Fairclough, D.L., Helms, R.W. (1986) "A Mixed Linear Model with Linear Covariance Structure: A Sensitivity Analysis of Maximum Likelihood Estimators", Journal of Statistical Computation and Simulation, 24, 205-236.

Fleiss, J.L. (1981) Statistical Methods for Rates and Proportions 2nd ed., John Wiley and Sons, Inc., New York.

Goldstein, H. (1979) The Design and Analysis of Longitudinal Studies, Academic Press Inc., New York.

Greenhouse,S.W., Geisser, S. (1959) "On Methods in the Analysis of Profile Data", Psychometrika, 24, 95-112.

Hartley, H.D., Hocking, R.R. (1971) "The Analysis of Incomplete Data", Biometrics, 27, 783-823.

Harville, D. (1976) "Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects", The Annals of Statistics, 4, 384-395.

Harville, D.A. (1977) " Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems", Journal of the American Statistical Association,72, 320-338.

Hocking, R.R., Smith, W.B. (1968) "Estimation of Parameters in the Multivariate Normal Distribution with Missing Observations", Journal of the American Statistical Association, 63, 159-173.

Huynh, H., (1978) "Some Approximate Tests for Repeated Measurement Designs", Psychometrika, 43, 161-175.

Huynh, H., Feldt, L.S. (1970) "Conditions Under Which Mean Square Ratios in Repeated Measurement Designs have Exact F Distributions", Journal of the American Statistical Association, 65, 1582-1589.

Huynh, H., Feldt, L.S. (1976) "Estimation of the Box Correction for Degrees of Freedom from Sample Data in Randomized Block and Split-Plot Designs", Journal of Educational Statistics, 1, 69-82.

Jennrich, R.I., Schluchter, M.D. (1985) "Unbalanced Repeated Measures Models with Structured Covariance Matrices", Biometrics, 42, 805-820.

Johnson, R.A., Wichern, D.W. (1982) Applied Multivariate Statistical Analysis, Prentice-Hall Inc., Englewood Cliffs, N.J.

Judge, G.G., Griffiths, W.E., Hill, R.C., Lee T.C. (1980) The Theory and Practice of Econometrics, John Wiley and Sons, Inc., New York.

Kleinbaum, D.G. (1970) "Estimation and Hypothesis Testing for Generalized Multivariate Linear Models", Ph.D. thesis, University of North Carolina, Chapel Hill, NC.

Kleinbaum, D.G. (1973a) "Testing Linear Hypotheses in Generalized Multivariate Linear Models", Communications in Statistics , 1(5), 433-457.

Kleinbaum, D.G. (1973b) "A Generalization of the Growth Curve Model Which Allows Missing Data", Journal of Multivariate Analysis, 3, 117-124.

Kleinbaum, D.G., Kupper, L.L., Morgenstern, H. (1982) Epidemiologic Research Principles and Quantitative Methods, Wadsworth, Inc., Belmont, Calif.

Laird, N.M., Ware, J.H. (1982) "Random-Effects Models for Longitudinal Data", Biometrics, 38, 963-974.

Machin, D. (1975) "On a Design Problem in Growth Studies", Biometrics, 31, 749-753.

Martin, M.J. (1987) "Computational Algorithms in the Analysis of Incomplete Longitudinal Data", Unpublished Doctoral Dissertation.

Mauchly, J.W. (1940) "Significance Test for Sphericity of a Normal n-Variate Distribution", The Annals of Mathematical Staristics, 11, 204-209.

Morrison, D.F. (1976) Multivariate Statistical Methods, McGraw-Hill, Inc., New York, NY.

Muller, K.E., Peterson, B.L. (1984) "Practical Methods for Computing Power in Testing the Multivariate General Linear Hypothesis", Computational Statistics and Data Analysis, 2, 143158.

Muller, K.E., Barton, C.N. (1987) " Approximate Power for Repeated Measures ANOVA Lacking Sphericity", In Review.

Orchard, T., Woodbury, M.A. (1972) " A Missing Information Principle: Theory and Applications", Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 1, 697-715.

Potthoff, R.F., Roy, S.N. (1964) "A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems", Biometrika, 51, 313-326.

Rao,M.N., Rao, C.R. (1966) "Linked Cross-Sectional Study for Determining Norms and Growth Rates - A Pilot Survey on Indian School-Going Boys", Sankyha, 28, 237-258.

Rubin, D.B. (1976) "Inference and Missing Data", Biometrika, 63, 581-592.

Schaie, K.W. (1965) "A General Model for the Study of Developmental Problems", Psychological Bulletion, 64, 92-107.

Scheffe, H. (1959) The Analysis of Variance, John Wiley and Sons, Inc., New York.

Schwertman, N.C., (1978) "A Note on the Geisser-Greenhouse Correction for Incomplete Data Split-Plot Analysis", Journal of the American Statistical Association, 73, 393-396.

Schwertman, N.C., Allen, D.M. (1979) "The Smoothing of an Indefinite Variance-Covariance Matrix", Journal of Statistical Computing and Simulation, 9, 183-194.

Searle, S.R. (1971) Linear Models, John Wiley and Sons, Inc., New York.

Searle, S.R. (1982) Matrix Algebra Useful for Statistics, John Wiley and Sons, Inc., New York,

Snedecor, G.W., Cochran. W.G. (1980) Statistical Methods 7th ed., The Iowa State University Press, Ames, Iowa.

Timm, N.H. (1970) "The Estimation of Variance-Covariance and Correlation Matrices from Incomplete Data", Psychometrika, 35, 417-437.

Timm, N.H. (1975) Multivariate Analysis with Applications in Education and Psychology, Wadsworth Publishing Company, Inc., Belmont CA.

van't Hof, M.A., Roede, M.J., Kowalski, C.J. (1977) "A Mixed Longitudinal Data Analysis Model", Human Biology, 49, 165-179.

Ware, J.H. (1985) "Linear Models for the Analysis of Longitudinal Studies", The American Statistician, 39, 95-101.

Wilkinson, (1958) "Estimation of Missing Values for the Analysis of Incomplete Data", Biometrics, 14, 257-286.

Woolson, R.F., Leeper, J.D., Clarke, W.R. (1978) "Analysis of Incomplete Data from Longitudinal and Mixed Longitudinal Studies", Journal of the Royal Statistical Society A, 141, 242-252.

Wu, C.F.J. (1983) "On the Convergence Properties of the EM Algorithm", The Annals Of Statistics, 11, 95-103.

Table 3.5.1

Typical Design Matrices for the Longitudinal Design

Black Subject with No History of Disease

$\phi_{int(b)}$	$\phi_{sl(b)}$	$\phi_{BrInt(b)}$	$\phi_{PnSl(b)}$	$\phi_{int(w)}$	$\phi_{sl(w)}$	$\phi_{BrInt(w)}$	$\phi_{PnSl(w)}$
1	-20	0	0	0	0	0	0
1	-16	0	0	0	0	0	0
1	-14	0	0	0	0	0	0
1	-11	0	0	0	0	0	0
1	-10	0	0	0	0	0	0
1	-5	0	0	0	0	0	0
1	-4	0	0	0	0	0	0
1	-1	0	0	0	0	0	0
1	2	0	0	0	0	0	0
1	7	0	0	0	0	0	0
1	12	0	0	0	0	0	0
1	14	0	0	0	0	0	0
1	18	0	0	0	0	0	0
1	21	0	0	0	0	0	0
1	24	0	0	0	0	0	0

White Subject with History of Pneumonia

$\phi_{int(b)}$	$\phi_{sl(b)}$	$\phi_{BrInt(b)}$	$\phi_{PnSl(b)}$	$\phi_{int(w)}$	$\phi_{sl(w)}$	$\phi_{BrInt(w)}$	$\phi_{PnSl(w)}$
0	0	0	0	1	-15	0	-15
0	0	0	0	1	-11	0	-11
0	0	0	0	1	-8	0	-8
0	0	0	0	1	-5	0	-5
0	0	0	0	1	-3	0	-3
0	0	0	0	1	-1	0	-1
0	0	0	0	1	2	0	2
0	0	0	0	1	6	0	6
0	0	0	0	1	11	0	11
0	0	0	0	1	14	0	14
0	0	0	0	1	18	0	18
0	0	0	0	1	22	0	22
0	0	0	0	1	25	0	25
0	0	0	0	1	27	0	27
0	0	0	0	1	31	0	31

Table 3.5.2

Initial Values of Height

Subject	Initial Height (cm.)
1	96
2	91
3	95
4	90
5	95
6	90
7	88
8	96
9	97
10	93
11	90
12	91
13	95
14	93
15	94
16	99
17	85
18	91
19	102
20	91
21	89
22	96
23	93
24	89
25	94

Table 3.5.3

Subjects Cross-Tabulated by Race and Disease Status

	Blacks		Whites		Total	
Disease Status	No.	%	No.	%	No.	%
Pneumonia	1	13	3	18	4	16
Bronchiolitis	1	13	0	0	1	4
Bronchiolitis and Pneumonia	0	0	4	24	4	16
None	6	75	10	59	16	64
Total	8 (32%)	100*	17 (68%)	100*	25 (100%)	100

* Percentages do not sum to 100 because of round off error.

Table 3.5.4

Average Number of Measurements Per Subject for Each Design

	Long.	Long. with missing data	LCS	LCS with missing data
Total no. Subjects	25	25	25	25
Total no. Subjects w/ Complete data	25	1	25	5
Planned no. Measurements per subject	15	15	7	7
Total no. Measurements	375	294	175	138
Average no. Measurements per subject	15	11.8	7	5.5
Minimum	15	7	7	3
Maximum	15	15	7	7

Table 3.5.5

Measurements Cross-Tabulated by Race and Disease Status

Longitudinal Design

	Blacks		Whites		Total	
Disease Status	No.	%	No.	%	No.	%
Pneumonia	15	13	45	18	60	16
Bronchiolitis	15	13	0	0	15	4
Bronchiolitis and Pneumonia	0	0	60	24	60	16
None	90	75	150	59	240	64
Total	120 (32%)	100*	255 (68%)	100*	375 (100%)	100

* Percentages do not sum to 100 because of round off error.

Table 3.5.6

Measurements Cross-Tabulated by Race and Disease Status

Longitudinal Design with Missing Data

	Blacks		Whites		Total	
Disease Status	No.	%	No.	%	No.	%
Pneumonia	14	14	35	18	49	17
Bronchiolitis	14	14	0	0	14	5
Bronchiolitis and Pneumonia	0	0	48	25	48	16
None	75	73	108	57	183	62
Total	103 (35%)	100*	191 (65%)	100*	294 (100%)	100

* Percentages do not sum to 100 because of round off error.

Table 3.5.7

Measurements Cross-Tabulated by Race and Disease Status

Linked Cross-Sectional Design

	Blacks		Whites		Total	
Disease Status	No.	%	No.	%	No.	%
Pneumonia	7	13	21	18	28	16
Bronchiolitis	7	13	0	0	7	4
Bronchiolitis and Pneumonia	0	0	28	24	28	16
None	42	75	70	59	112	64
Total	56 (32%)	100*	119 (68%)	100*	175 (100%)	100

* Percentages do not sum to 100 because of round off error.

Table 3.5.8

Measurements Cross-Tabulated by Race and Disease Status

Linked Cross-Sectional Design with Missing Data

	Blacks		Whites		Total	
Disease Status	No.	%	No.	%	No.	%
Pneumonia	5	11	16	17	21	15
Bronchiolitis	7	15	0	0	7	5
Bronchiolitis and Pneumonia	0	0	21	23	21	15
None	34	74	55	60	89	64
Total	46 (33%)	100	92 (67%)	100	138 (100%)	100*

* Percentages do not sum to 100 because of round off error.

Table 3.5.9

Parameter Values for Null and Alternative Hypotheses

Parameter	H_o	H_{A1}	H_{A2}	H_{A3}
$\phi_{int(b)}$	1.97	1.97	1.97	1.97
$\phi_{slope(b)}$	0.0198	0.0198	0.0198	0.0198
$\phi_{BrInt(b)}$	-0.33	-0.33	-0.33	-0.33
$\phi_{PnSlope(b)}$	0.0118	0.0118	0.0118	0.0118
$\phi_{int(w)}$	1.97	1.97	2.663	2.663
$\phi_{slope(w)}$	0.0198	0.0423	0.0198	0.0423
$\phi_{BrInt(w)}$	-0.33	-0.33	-0.33	-0.33
$\phi_{PnSlope(w)}$	0.0118	0.0118	0.0118	0.0118
τ_1	0.146	0.146	0.146	0.146
τ_2	0.00018	0.00018	0.00018	0.00018
τ_3	0.00265	0.00265	0.00265	0.00265
σ^2	0.118	0.118	0.118	0.118

Table 3.5.10

Number of Simulations Needed for Each Hypothesis and Design

Design	H_o	H_{A1}	H_{A2}	H_{A3}
Longitudinal	340	220	65	60
Longitudinal with missing data	340	260	75	65
Linked Cross-Sectional	340	340	210	110
Linked Cross-Sectional with missing data	340	320	240	130

Table 4.1.1

Values of h_1, h_2 and b used in the F-plot for F_{Box} Test Statistic

Design	Parameter Values
Longitudinal	$h_1 = 3.40$ $h_2 = 62.03$ $b = 0.080$
Longitudinal with missing data	$h_1 = 3.43$ $h_2 = 59.55$ $b = 0.092$
Linked Cross-Sectional	$h_1 = 3.68$ $h_2 = 46.36$ $b = 0.120$
Linked Cross-Sectional with missing data	$h_1 = 3.77$ $h_2 = 43.11$ $b = 0.135$

Table 4.1.2

Mean Degrees of Freedom (h_1,h_2) and Correction

Factor (b) for F_{Box} Statistic

Longitudinal Design

Hypothesis		Mean	N	St.Dev.	Min	Max
H_o	h_1	3.35	340	0.14	2.98	3.91
	h_2	74.35	340	22.61	37.90	205.71
	b	0.074	340	0.009	0.042	0.097
H_{A1}	h_1	3.35	220	0.13	2.99	3.77
	h_2	75.17	220	23.46	38.84	164.28
	b	0.074	220	0.010	0.045	0.098
H_{A2}	h_1	3.33	65	0.13	3.05	3.84
	h_2	72.69	65	22.07	40.60	137.70
	b	0.075	65	0.010	0.055	0.098
H_{A3}	h_1	3.32	60	0.11	3.14	3.62
	h_2	72.21	60	19.63	44.60	126.14
	b	0.075	60	0.009	0.054	0.094

Table 4.1.3

Mean Degrees of Freedom (h_1,h_2) and Correction

Factor (b) for F_{Box} Statistic

Longitudinal Design with Missing Data

Hypothesis		Mean	N	St.Dev.	Min	Max
H_o	h_1	3.37	340	0.12	2.99	3.78
	h_2	70.20	340	21.66	36.14	196.12
	b	0.085	340	0.012	0.044	0.114
H_{A1}	h_1	3.37	260	0.14	2.96	3.79
	h_2	69.93	260	19.49	40.81	142.84
	b	0.085	260	0.011	0.054	0.110
H_{A2}	h_1	3.33	75	0.11	3.08	3.60
	h_2	68.66	75	19.72	38.58	160.59
	b	0.085	75	0.012	0.050	0.112
H_{A3}	h_1	3.36	65	0.14	3.03	3.72
	h_2	71.61	65	23.18	37.55	140.08
	b	0.085	65	0.012	0.056	0.112

Table 4.1.4

Mean Degrees of Freedom (h_1,h_2) and Correction

Factor (b) for F_{Box} Statistic

Linked Cross-Sectional Design

Hypothesis		Mean	N	St.Dev.	Min	Max
H_o	h_1	3.60	340	0.32	2.71	3.99
	h_2	55.64	340	16.27	27.19	145.51
	b	0.110	340	0.016	0.049	0.158
H_{A1}	h_1	3.59	340	0.31	2.71	3.99
	h_2	57.19	340	18.60	28.28	135.78
	b	0.108	340	0.017	0.055	0.145
H_{A2}	h_1	3.59	210	0.31	2.72	3.99
	h_2	59.06	210	18.39	28.74	129.02
	b	0.106	210	0.017	0.061	0.144
H_{A3}	h_1	3.62	110	0.31	2.72	3.99
	h_2	59.81	110	18.80	29.74	111.91
	b	0.106	110	0.018	0.065	0.145

Table 4.1.5

Mean Degrees of Freedom (h_1,h_2) and Correction

Factor (b) for F_{Box} Statistic

Linked Cross-Sectional Design with Missing Data

Hypothesis		Mean	N	St.Dev.	Min	Max
H_o	h_1	3.61	340	0.31	2.76	3.99
	h_2	53.24	340	17.14	22.67	116.18
	b	0.121	340	0.021	0.060	0.170
H_{A1}	h_1	3.61	320	0.31	2.68	4.00
	h_2	51.96	320	16.82	24.69	124.81
	b	0.122	320	0.020	0.047	0.162
H_{A2}	h_1	3.64	240	0.31	2.77	3.99
	h_2	53.23	240	16.96	24.60	124.79
	b	0.121	240	0.020	0.046	0.163
H_{A3}	h_1	3.64	130	0.28	2.90	3.99
	h_2	52.89	130	15.20	28.89	94.50
	b	0.121	130	0.019	0.071	0.164

Table 4.2.1

Mean Parameter Estimates and Standard Errors For the Mixed Model

Using Method of Scoring

Longitudinal Design

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 220)	H_{A2} (N = 65)	H_{A3} (N = 60)
$\phi_{int(b)}$	1.97 (0.14)	1.98 (0.16)	1.95 (0.15)	1.98 (0.15)
$\phi_{slope(b)}$	0.020 (0.006)	0.020 (0.006)	0.019 (0.006)	0.020 (0.006)
$\phi_{BrInt(b)}$	-0.34 (0.40)	-0.34 (0.38)	-0.33 (0.39)	-0.31 (0.46)
$\phi_{PnSlope(b)}$	0.012 (0.015)	0.013 (0.016)	0.012 (0.014)	0.013 (0.017)
$\phi_{int(w)}$	1.97 (0.11)	1.96 (0.11)	2.65 (0.12)	2.66 (0.11)
$\phi_{slope(w)}$	0.020 (0.005)	0.042 (0.005)	0.020 (0.005)	0.042 (0.004)
$\phi_{BrInt(w)}$	-0.34 (0.22)	-0.32 (0.22)	-0.27 (0.23)	-0.35 (0.22)
$\phi_{PnSlope(w)}$	0.012 (0.007)	0.012 (0.008)	0.011 (0.007)	0.012 (0.007)
τ_1	0.12 (0.038)	0.12 (0.043)	0.13 (0.043)	0.12 (0.038)
τ_2	0.00015 (0.00006)	0.00015 (0.00006)	0.00014 (0.00005)	0.00014 (0.00007)
τ_3	0.0024 (0.0013)	0.0024 (0.0015)	0.0023 (0.0014)	0.0025 (0.0013)
σ^2	0.118 (0.009)	0.119 (0.009)	0.117 (0.010)	0.118 (0.008)

Table 4.2.2
Mean Parameter Estimates and Standard Errors For the Canonical form of the Mixed Model, Using the Method of Scoring
Longitudinal Design

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 220)	H_{A2} (N = 65)	H_{A3} (N = 60)
$\phi_{int(b)}$	1.98 (0.14)	1.98 (0.16)	1.95 (0.16)	1.98 (0.15)
$\phi_{slope(b)}$	0.020 (0.006)	0.020 (0.006)	0.019 (0.006)	0.020 (0.005)
$\phi_{BrInt(b)}$	-0.34 (0.39)	-0.34 (0.36)	-0.33 (0.40)	-0.31 (0.42)
$\phi_{PnSlope(b)}$	0.012 (0.015)	0.014 (0.016)	0.013 (0.017)	0.013 (0.016)
$\phi_{int(w)}$	1.97 (0.10)	1.97 (0.12)	2.65 (0.12)	2.66 (0.11)
$\phi_{slope(w)}$	0.020 (0.005)	0.042 (0.005)	0.019 (0.004)	0.042 (0.005)
$\phi_{BrInt(w)}$	-0.33 (0.22)	-0.33 (0.22)	-0.28 (0.23)	-0.34 (0.21)
$\phi_{PnSlope(w)}$	0.012 (0.007)	0.012 (0.008)	0.012 (0.007)	0.011 (0.007)
τ_1	0.13 (0.039)	0.13 (0.041)	0.14 (0.044)	0.13 (0.035)
τ_2	0.00019 (0.00006)	0.00019 (0.00006)	0.00019 (0.00005)	0.00019 (0.00006)
τ_3	0.0021 (0.0013)	0.0021 (0.0014)	0.0022 (0.0013)	0.0022 (0.0013)
σ^2	0.117 (0.009)	0.116 (0.009)	0.116 (0.010)	0.117 (0.009)

Table 4.2.3

Observed and Hypothesized Type I Error Rate and Power

Longitudinal Design

N = Number of iterations

SE = Standard error of observed power

Test Statistic	H_o (N = 340) Obs. (SE)	Hyp.	H_{A1} (N = 220) Obs. (SE)	Hyp.	H_{A2} (N = 65) Obs. (SE)	Hyp.	H_{A3} (N = 60) Obs. (SE)	Hyp.
F_{REML}	0.076 (0.014)	0.05	0.82 (0.026)	0.83	0.95 (0.026)	0.96	0.93 (0.032)	0.96
F_{WLS}	0.19* (0.021)	0.05	0.90* (0.020)	0.83	0.95 (0.026)	0.96	0.97 (0.023)	0.96
F_{WLS2}	0.16* (0.020)	0.05	0.87 (0.023)	0.83	0.95 (0.026)	0.96	0.97 (0.023)	0.96
LRT	0.14* (0.019)	0.05	0.86 (0.023)	0.84	0.95 (0.027)	0.96	0.95 (0.028)	0.96
F_{Box}	0.062 (0.013)	0.05	0.33* (0.032)	1.00	0.89* (0.038)	1.00	0.95 (0.028)	1.00

* The observed value was more than 2 standard errors from the hypothesized value.

Table 4.3.1

Mean Parameter Estimates and Standard Errors For the Mixed Model Using Method of Scoring Longitudinal Design with Missing Data

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 260)	H_{A2} (N = 75)	H_{A3} (N = 65)
$\phi_{int(b)}$	1.96 (0.15)	1.99 (0.15)	1.97 (0.14)	1.96 (0.14)
$\phi_{slope(b)}$	0.019 (0.006)	0.020 (0.006)	0.020 (0.006)	0.020 (0.006)
$\phi_{BrInt(b)}$	-0.33 (0.41)	-0.34 (0.41)	-0.30 (0.41)	-0.24 (0.42)
$\phi_{PnSlope(b)}$	0.012 (0.017)	0.011 (0.015)	0.010 (0.016)	0.008 (0.013)
$\phi_{int(w)}$	1.97 (0.12)	1.96 (0.11)	2.67 (0.11)	2.66 (0.10)
$\phi_{slope(w)}$	0.020 (0.005)	0.042 (0.006)	0.020 (0.005)	0.042 (0.006)
$\phi_{BrInt(w)}$	-0.35 (0.21)	-0.33 (0.25)	-0.34 (0.21)	-0.29 (0.23)
$\phi_{PnSlope(w)}$	0.012 (0.008)	0.012 (0.008)	0.011 (0.007)	0.012 (0.008)
τ_1	0.12 (0.043)	0.12 (0.038)	0.12 (0.044)	0.12 (0.047)
τ_2	0.00014 (0.00006)	0.00014 (0.00006)	0.00015 (0.00007)	0.00015 (0.00007)
τ_3	0.0025 (0.0014)	0.0024 (0.0013)	0.0025 (0.0015)	0.0026 (0.0014)
σ^2	0.118 (0.010)	0.117 (0.011)	0.116 (0.011)	0.119 (0.011)

Table 4.3.2

Mean Parameter Estimates and Standard Errors For the Mixed Model

Using Method of Scoring

Longitudinal Design with Missing Data

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 260)	H_{A2} (N = 75)	H_{A3} (N = 65)
$\phi_{int(b)}$	1.95 (0.15)	1.99 (0.15)	1.97 (0.14)	1.97 (0.14)
$\phi_{slope(b)}$	0.020 (0.006)	0.020 (0.006)	0.020 (0.006)	0.020 (0.006)
$\phi_{BrInt(b)}$	-0.32 (0.41)	-0.36 (0.40)	-0.30 (0.40)	-0.29 (0.40)
$\phi_{PnSlope(b)}$	0.013 (0.016)	0.011 (0.015)	0.010 (0.015)	0.007 (0.020)
$\phi_{int(w)}$	1.96 (0.12)	1.96 (0.11)	2.67 (0.10)	2.67 (0.11)
$\phi_{slope(w)}$	0.020 (0.005)	0.042 (0.005)	0.020 (0.004)	0.042 (0.005)
$\phi_{BrInt(w)}$	-0.34 (0.21)	-0.34 (0.23)	-0.34 (0.21)	-0.28 (0.22)
$\phi_{PnSlope(w)}$	0.012 (0.008)	0.012 (0.008)	0.011 (0.006)	0.011 (0.008)
τ_1	0.13 (0.042)	0.13 (0.040)	0.13 (0.045)	0.13 (0.046)
τ_2	0.00020 (0.00007)	0.00020 (0.00007)	0.00020 (0.00008)	0.00022 (0.00007)
τ_3	0.0021 (0.0014)	0.0021 (0.0012)	0.0022 (0.0016)	0.0022 (0.0013)
σ^2	0.117 (0.010)	0.117 (0.010)	0.117 (0.012)	0.117 (0.010)

Table 4.3.3

Observed and Hypothesized Type I Error Rate and Power

Longitudinal Design with Missing Data

N = Number of iterations

SE = Standard error of observed power

Test Statistic	H_o (N = 340) Obs. (SE)	Hyp.	H_{A1} (N = 260) Obs. (SE)	Hyp.	H_{A2} (N = 75) Obs. (SE)	Hyp.	H_{A3} (N = 65) Obs. (SE)	Hyp.
F_{REML}	0.071 (0.014)	0.05	0.75 (0.027)	0.79	0.96 (0.023)	0.95	0.94 (0.030)	0.96
F_{WLS}	0.12* (0.018)	0.05	0.90* (0.019)	0.79	0.97 (0.019)	0.95	0.97 (0.022)	0.96
F_{WLS2}	0.091* (0.016)	0.05	0.86* (0.021)	0.79	0.97 (0.019)	0.95	0.97 (0.022)	0.96
LRT	0.079 (0.015)	0.05	0.85* (0.022)	0.80	0.97 (0.020)	0.95	0.95 (0.027)	0.96
F_{Box}	0.068 (0.014)	0.05	0.23* (0.026)	1.00	0.87* (0.039)	1.00	1.00	1.00

* The observed value was more than 2 standard errors from the hypothesized value.

Table 4.4.1

Mean Parameter Estimates and Standard Errors For the Original Mixed Model, Using the EM Algorithm

Linked Cross-Sectional Design

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 340)	H_{A2} (N = 210)	H_{A3} (N = 110)
$\phi_{int(b)}$	1.98 (0.18)	1.95 (0.18)	1.99 (0.19)	1.98 (0.17)
$\phi_{slope(b)}$	0.020 (0.009)	0.019 (0.009)	0.020 (0.010)	0.020 (0.009)
$\phi_{BrInt(b)}$	-0.30 (0.41)	-0.32 (0.43)	-0.35 (0.41)	-0.29 (0.45)
$\phi_{PnSlope(b)}$	0.010 (0.026)	0.013 (0.027)	0.008 (0.027)	0.009 (0.024)
$\phi_{int(w)}$	1.97 (0.12)	1.97 (0.12)	2.65 (0.12)	2.66 (0.10)
$\phi_{slope(w)}$	0.019 (0.008)	0.042 (0.008)	0.020 (0.008)	0.043 (0.008)
$\phi_{BrInt(w)}$	-0.32 (0.24)	-0.34 (0.24)	-0.32 (0.25)	-0.31 (0.23)
$\phi_{PnSlope(w)}$	0.013 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)
τ_1	0.12 (0.048)	0.11 (0.046)	0.11 (0.045)	0.11 (0.044)
τ_2	0.00016 (0.00011)	0.00016 (0.00011)	0.00016 (0.00010)	0.00016 (0.00011)
τ_3	0.0022 (0.0018)	0.0021 (0.0019)	0.0020 (0.0018)	0.0018 (0.0019)
σ^2	0.116 (0.014)	0.116 (0.014)	0.117 (0.015)	0.116 (0.014)

Table 4.4.2

Mean Parameter Estimates and Standard Errors For the Canonical form of the Mixed Model, Using the Method of Scoring

Linked Cross-Sectional Design

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 340)	H_{A2} (N = 210)	H_{A3} (N = 110)
$\phi_{int(b)}$	1.97 (0.18)	1.96 (0.18)	1.97 (0.19)	1.98 (0.18)
$\phi_{slope(b)}$	0.019 (0.009)	0.019 (0.009)	0.020 (0.009)	0.020 (0.009)
$\phi_{BrInt(b)}$	-0.32 (0.41)	-0.32 (0.44)	-0.33 (0.43)	-0.27 (0.48)
$\phi_{PnSlope(b)}$	0.013 (0.027)	0.013 (0.026)	0.008 (0.028)	0.008 (0.026)
$\phi_{int(w)}$	1.97 (0.12)	1.97 (0.12)	2.66 (0.12)	2.67 (0.10)
$\phi_{slope(w)}$	0.019 (0.008)	0.043 (0.008)	0.019 (0.008)	0.043 (0.008)
$\phi_{BrInt(w)}$	-0.32 (0.24)	-0.34 (0.24)	-0.32 (0.25)	-0.32 (0.24)
$\phi_{PnSlope(w)}$	0.013 (0.013)	0.011 (0.012)	0.011 (0.011)	0.012 (0.012)
τ_1	0.14 (0.049)	0.13 (0.049)	0.13 (0.052)	0.13 (0.047)
τ_2	0.00042 (0.00016)	0.00039 (0.00015)	0.00038 (0.00017)	0.00041 (0.00016)
τ_3	0.0014 (0.0022)	0.0014 (0.0022)	0.0013 (0.0021)	0.0011 (0.0020)
σ^2	0.117 (0.015)	0.116 (0.014)	0.118 (0.015)	0.114 (0.014)

Table 4.4.3

Observed and Hypothesized Type I Error Rate and Power

Linked Cross-Sectional Design

N = Number of iterations

SE = Standard error of observed power

Test Statistic	H_o (N = 340) Obs. (SE)	Hyp.	H_{A1} (N = 340) Obs. (SE)	Hyp.	H_{A2} (N = 210) Obs. (SE)	Hyp.	H_{A3} (N = 110) Obs. (SE)	Hyp.
F_{REML}	0.050 (0.012)	0.05	0.30 (0.025)	0.33	0.81 (0.027)	0.84	0.94 (0.023)	0.93
F_{WLS}	0.14* (0.019)	0.05	0.47* (0.027)	0.33	0.89* (0.022)	0.84	0.95 (0.022)	0.93
F_{WLS2}	0.068 (0.014)	0.05	0.31 (0.025)	0.33	0.83 (0.026)	0.84	0.92 (0.026)	0.93
LRT	0.094* (0.016)	0.05	0.42* (0.027)	0.34	0.90* (0.021)	0.85	0.98* (0.013)	0.94
F_{Box}	0.088* (0.015)	0.05	0.22* (0.022)	0.98	0.78* (0.029)	1.00	0.95* (0.022)	1.00

* The observed value was more than 2 standard errors from the hypothesized value.

Table 4.5.1

Mean Parameter Estimates and Standard Errors For the Original form of the Mixed Model, Using the EM Algorithm Linked Cross-Sectional Design with Missing Data

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 320)	H_{A2} (N = 240)	H_{A3} (N = 130)
$\phi_{int(b)}$	1.96 (0.18)	1.98 (0.19)	1.97 (0.19)	1.95 (0.18)
$\phi_{slope(b)}$	0.020 (0.009)	0.020 (0.009)	0.021 (0.010)	0.020 (0.010)
$\phi_{BrInt(b)}$	-0.31 (0.44)	-0.30 (0.42)	-0.37 (0.44)	-0.29 (0.42)
$\phi_{PnSlope(b)}$	0.011 (0.029)	0.014 (0.027)	0.012 (0.026)	0.013 (0.025)
$\phi_{int(w)}$	1.96 (0.11)	1.96 (0.11)	2.66 (0.13)	2.66 (0.11)
$\phi_{slope(w)}$	0.019 (0.009)	0.043 (0.009)	0.020 (0.009)	0.042 (0.009)
$\phi_{BrInt(w)}$	-0.32 (0.25)	-0.32 (0.24)	-0.31 (0.25)	-0.33 (0.23)
$\phi_{PnSlope(w)}$	0.012 (0.014)	0.011 (0.013)	0.012 (0.014)	0.011 (0.012)
τ_1	0.10 (0.047)	0.11 (0.046)	0.11 (0.050)	0.11 (0.046)
τ_2	0.00017 (0.00013)	0.00017 (0.00011)	0.00016 (0.00012)	0.00016 (0.00010)
τ_3	0.0021 (0.0021)	0.0022 (0.0022)	0.0020 (0.0020)	0.0018 (0.0020)
σ^2	0.116 (0.017)	0.116 (0.017)	0.115 (0.016)	0.115 (0.017)

Table 4.5.2

Mean Parameter Estimates and Standard Errors For the Canonical form of the Mixed Model, Using the Method of Scoring Linked Cross-Sectional Design with Missing Data

N = Number of Iterations

Parameter	H_o (N = 340)	H_{A1} (N = 320)	H_{A2} (N = 240)	H_{A3} (N = 130)
$\phi_{int(b)}$	1.97 (0.20)	1.98 (0.19)	1.98 (0.20)	1.96 (0.19)
$\phi_{slope(b)}$	0.019 (0.010)	0.020 (0.010)	0.020 (0.009)	0.019 (0.010)
$\phi_{BrInt(b)}$	-0.33 (0.44)	-0.33 (0.44)	-0.39 (0.46)	-0.32 (0.41)
$\phi_{PnSlope(b)}$	0.011 (0.029)	0.015 (0.026)	0.013 (0.029)	0.011 (0.026)
$\phi_{int(w)}$	1.97 (0.12)	1.96 (0.11)	2.66 (0.12)	2.67 (0.12)
$\phi_{slope(w)}$	0.019 (0.008)	0.042 (0.010)	0.020 (0.009)	0.042 (0.009)
$\phi_{BrInt(w)}$	-0.32 (0.25)	-0.32 (0.25)	-0.29 (0.25)	-0.34 (0.24)
$\phi_{PnSlope(w)}$	0.011 (0.013)	0.013 (0.014)	0.012 (0.014)	0.012 (0.013)
τ_1	0.13 (0.049)	0.14 (0.051)	0.13 (0.051)	0.13 (0.049)
τ_2	0.00043 (0.00019)	0.00042 (0.00019)	0.00042 (0.00021)	0.00042 (0.00019)
τ_3	0.0014 (0.0024)	0.0015 (0.0024)	0.0012 (0.0026)	0.0013 (0.0025)
σ^2	0.118 (0.018)	0.118 (0.018)	0.119 (0.017)	0.121 (0.018)

Table 4.5.3

Observed and Hypothesized Type I Error Rate and Power

Linked Cross-Sectional Design with Missing Data

N = Number of iterations

SE = Standard error of observed power

Test Statistic	H_o (N = 340) Obs. (SE)	Hyp.	H_{A1} (N = 320) Obs. (SE)	Hyp.	H_{A2} (N = 240) Obs. (SE)	Hyp.	H_{A3} (N = 130) Obs. (SE)	Hyp.
F_{REML}	0.059 (0.013)	0.05	0.22 (0.023)	0.28	0.84 (0.024)	0.82	0.92 (0.025)	0.91
F_{WLS}	0.16* (0.020)	0.05	0.41* (0.028)	0.29	0.90* (0.019)	0.82	0.95 (0.020)	0.91
F_{WLS2}	0.088* (0.015)	0.05	0.23* (0.024)	0.28	0.80 (0.026)	0.82	0.89 (0.027)	0.91
LRT	0.14* (0.019)	0.05	0.35 (0.027)	0.30	0.89 (0.020)	0.84	0.96 (0.017)	0.92
F_{Box}	0.103* (0.016)	0.05	0.21* (0.023)	0.95	0.75* (0.028)	1.00	0.90* (0.026)	1.00

* The observed value was more than 2 standard errors from the hypothesized value.

Figure 1.3.1

Cross-Sectional Study of School Children in Grades 5 through 11

	School Year						
Cohort	5	6	7	8	9	10	11
A	*						
B		*					
C			*				
D				*			
E					*		
F						*	
G							*
							'78

Time of Measurement

* indicates measurement made at this point

Figure 1.3.2

Longitudinal Study of School Children in Grades 5 through 11

	School Year						
Cohort	5	6	7	8	9	10	11
A	★	★	★	★	★	★	★
	'78	'79	'80	'81	'82	'83	'84

Time of Measurement

★ indicates measurement made at this point.

Figure 1.3.3

Linked Cross-Sectional Study (with Overlap of Degree 2) of School Children in Grades 5 through 11.

	School Year						
Cohort	5	6	7	8	9	10	11
A	★	★	★				
B		★	★	★			
C			★	★	★		
D				★	★	★	
E					★	★	★
					'78	'79	'80

Time of Measurement

★ indicates measurements made at this point.

Figure 1.3.4

Linked Cross-Sectional Study (with Overlap of Degree 3) of
School Children in Grades 5 through 11.

School Year

Cohort	5	6	7	8	9	10	11
A	★-----	★-----	★-----	★			
B		★-----	★-----	★-----	★		
C			★-----	★-----	★-----	★	
D				★-----	★-----	★-----	★
				'78	'79	'80	'81

Time of Measurement

★ indicates a measurement made at this point.

Figure 1.3.5

Linked Cross-Sectional Design with Varying Degrees of Overlap

	School Year						
Cohort	5	6	7	8	9	10	11
A	*	*	*				
B		*	*	*			
C		*	*	*	*		
D				*	*	*	
E				*	*	*	*
Time of measurement				'78	'79	'80	'81

*indicates measurement made at this point

Figure 1.3.6

A Purposefully Incomplete Full-Period Design for a Six-Year Follow-up Study of School Children in Grades 5 through 11.

	School Year						
Cohort	5	6	7	8	9	10	11
A	★	★	★	★	★	★	★
B	★		★		★		★
C	★			★			★
D	★	★		★		★	★
	'78	'79	'80	'81	'82	'83	'84
	Time of Measurement						

★ indicates a measurement made at this point.

Figure 3.4.1

A Longitudinal Design Covering Seven Years

	Age in Years							
Cohort	3	4	5	6	7	8	9	10

A ★--★--★--★--★--★--★--★--★--★--★--★--★--★--★

'87 '88 '89 '90 '91 '92 '93 '94

Time of Measurement

★ indicates measurement made at this point

Figure 3.4.2

Linked Cross-Sectional Design with Overlap of Degree 5 and Length 3 Years

Age in Years

```
Cohort   3    4    5    6    7    8    9    10

A        ★--★--★--★--★--★--★
B             ★--★--★--★--★--★--★
C                  ★--★--★--★--★--★--★
D                       ★--★--★--★--★--★--★
E                            ★--★--★--★--★--★--★

                            '87  '88  '89  '90
```

Time of Measurement

★ indicates measurement made at this point

FIGURE 4.2.1
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN

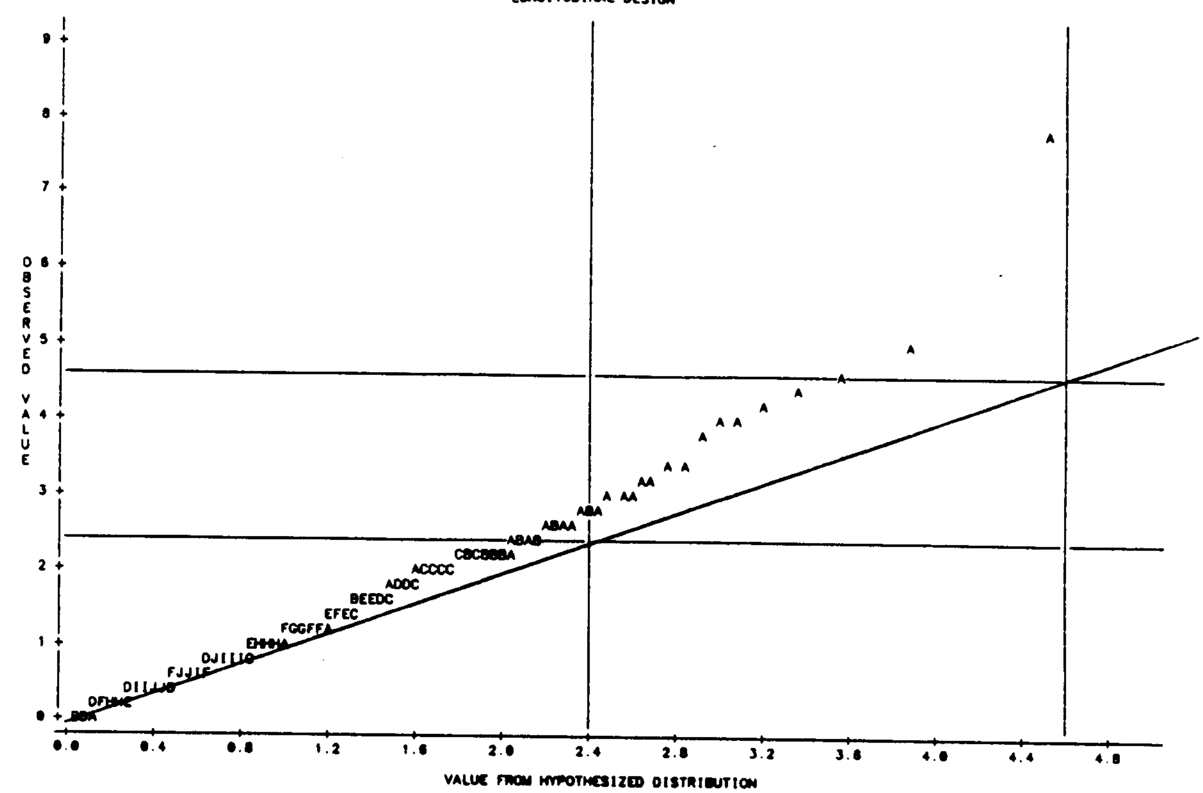

FIGURE 4.2.2
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

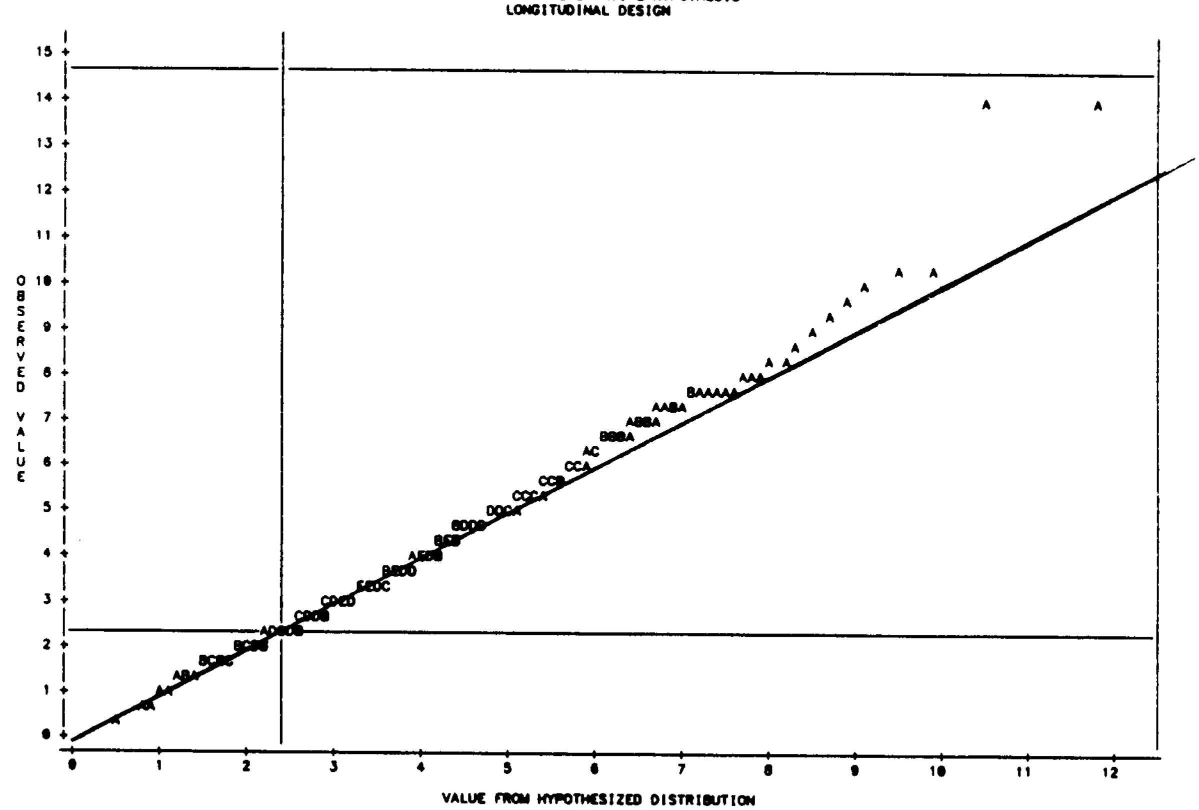

FIGURE 4.2.3
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

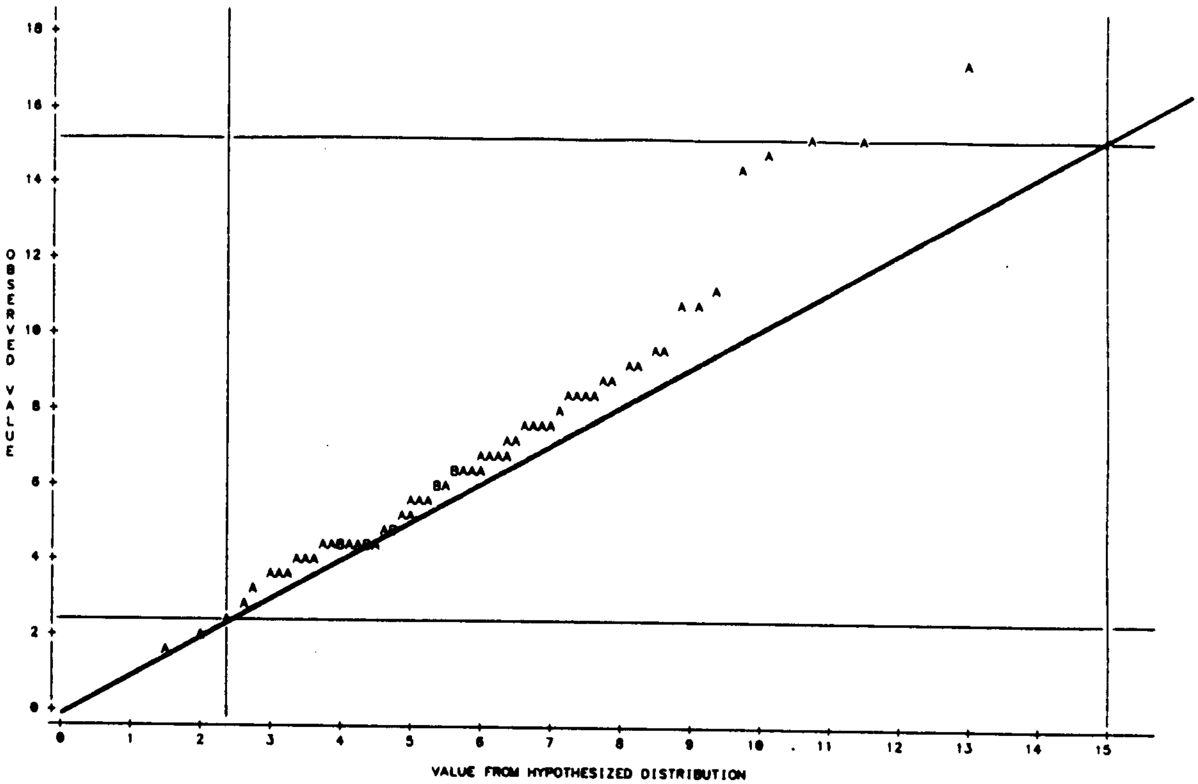

FIGURE 4.2.4
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

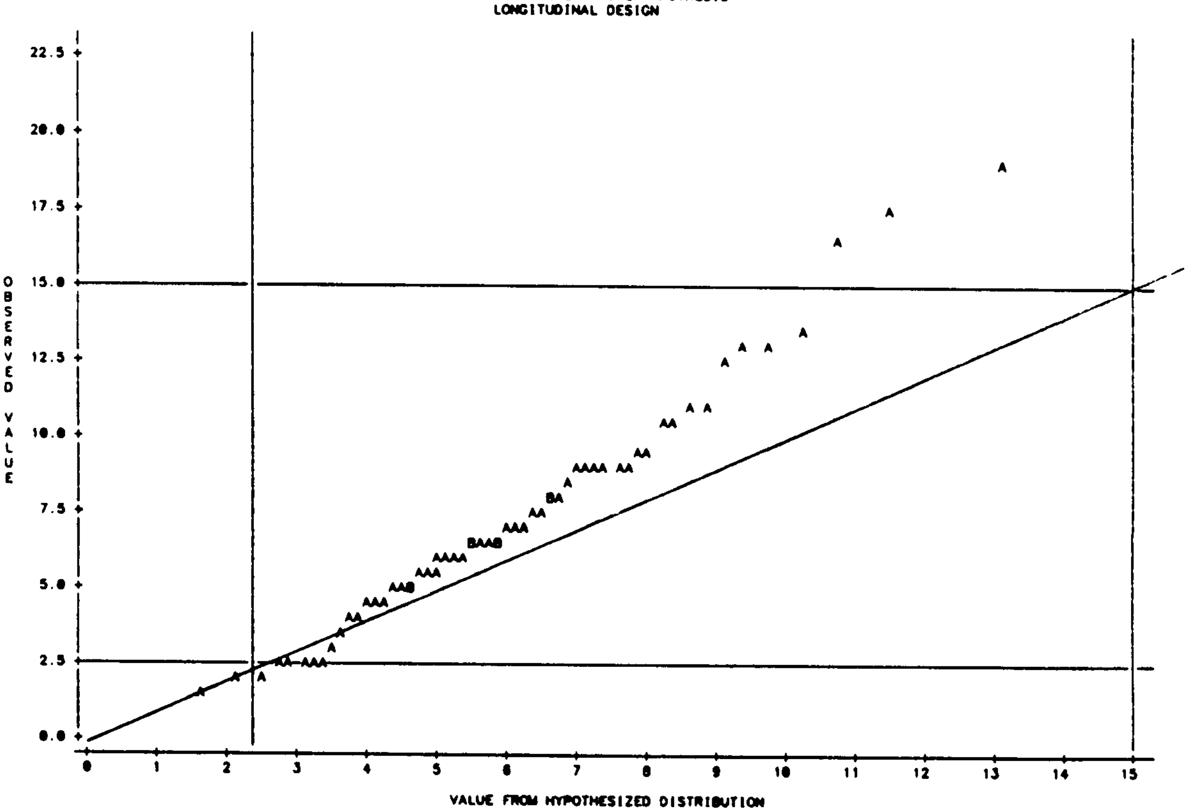

FIGURE 4.2.5
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN

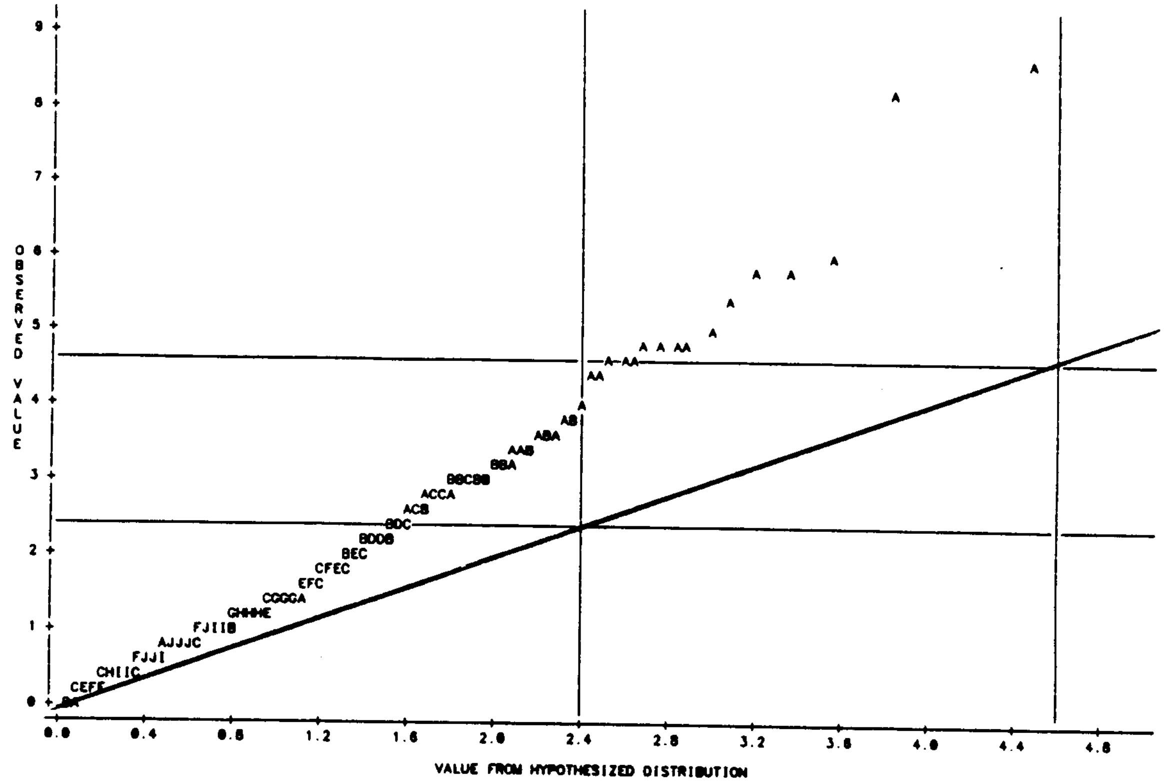

FIGURE 4.2.6
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

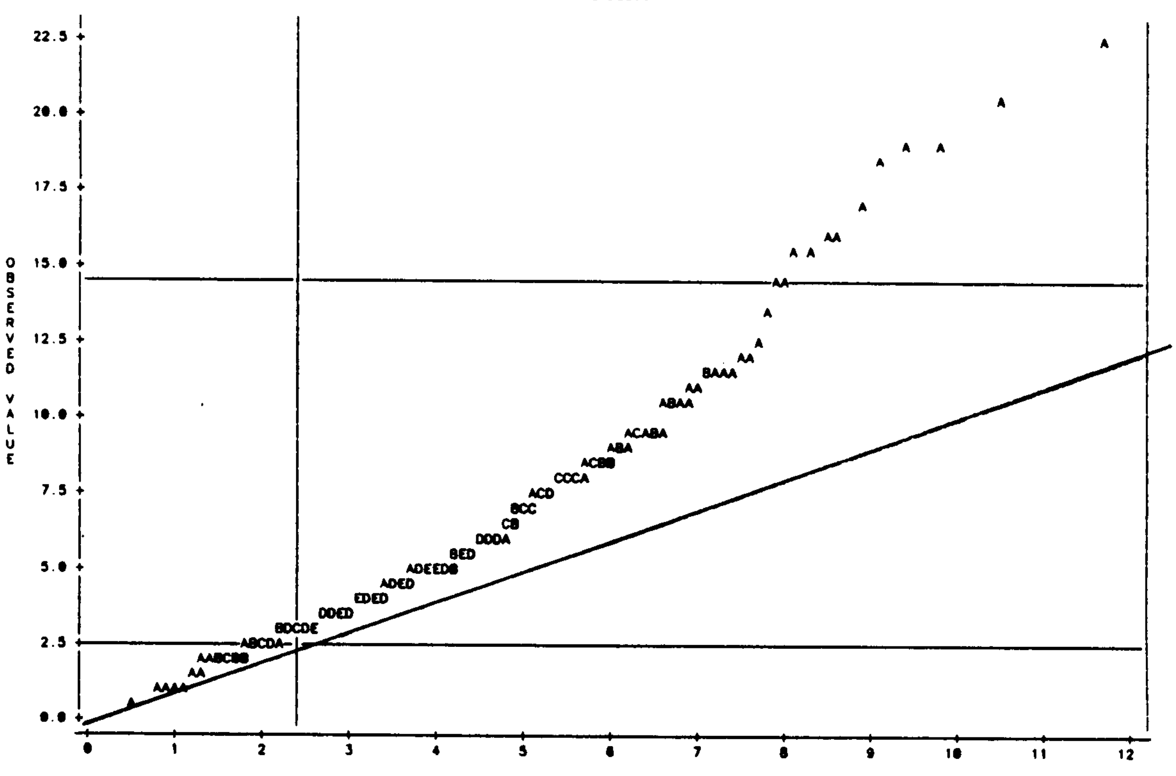

FIGURE 4.2.7
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

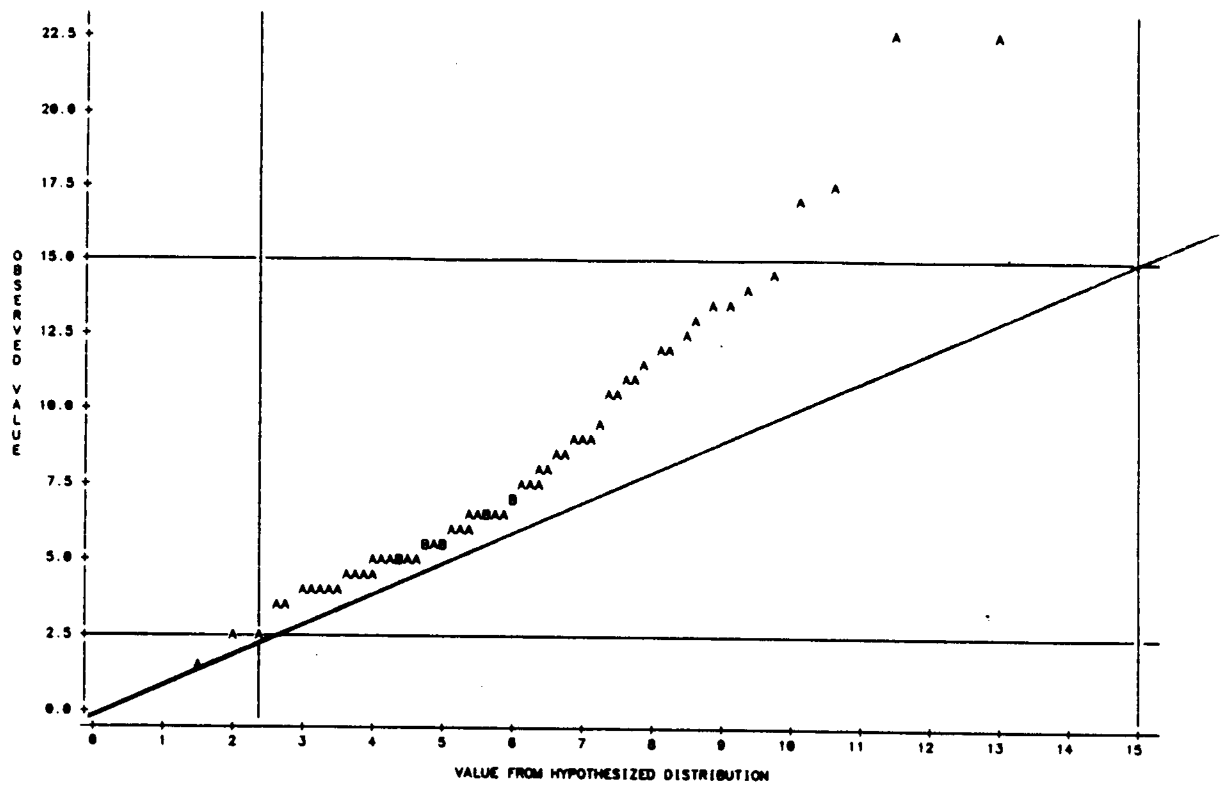

FIGURE 4.2.8
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

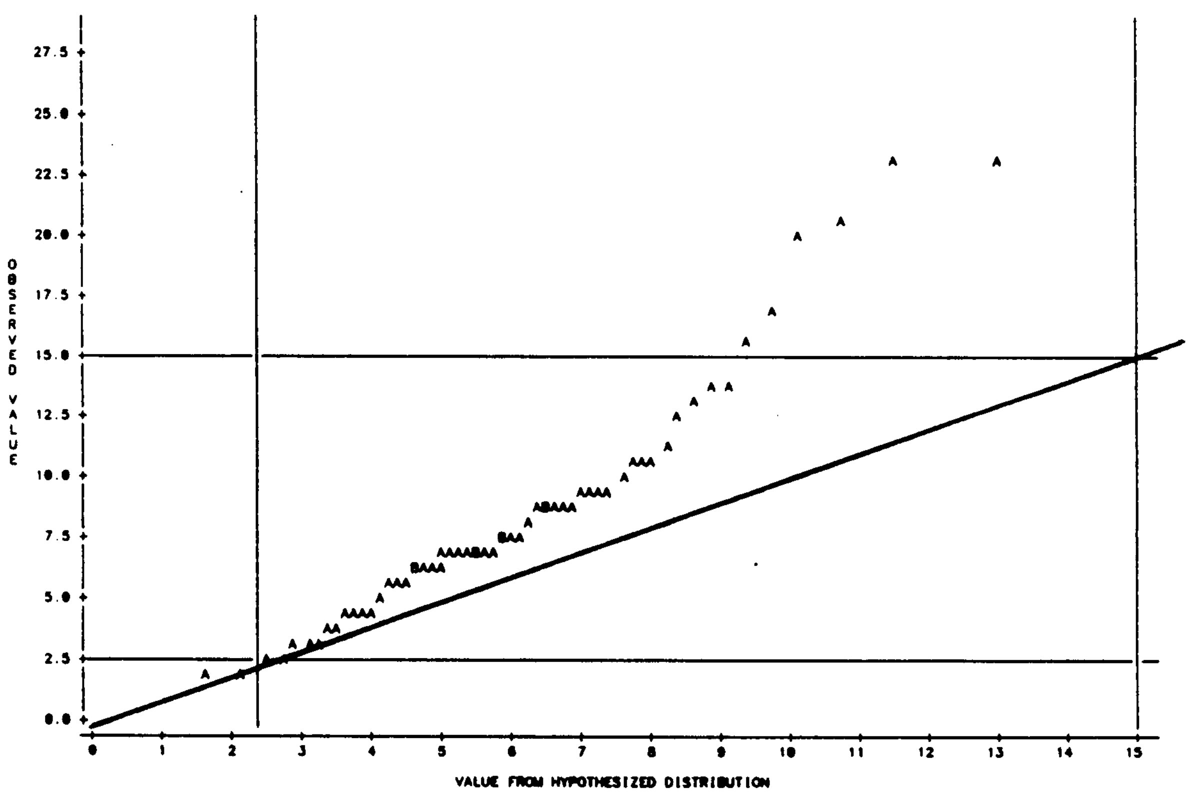

FIGURE 4.2.9
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN

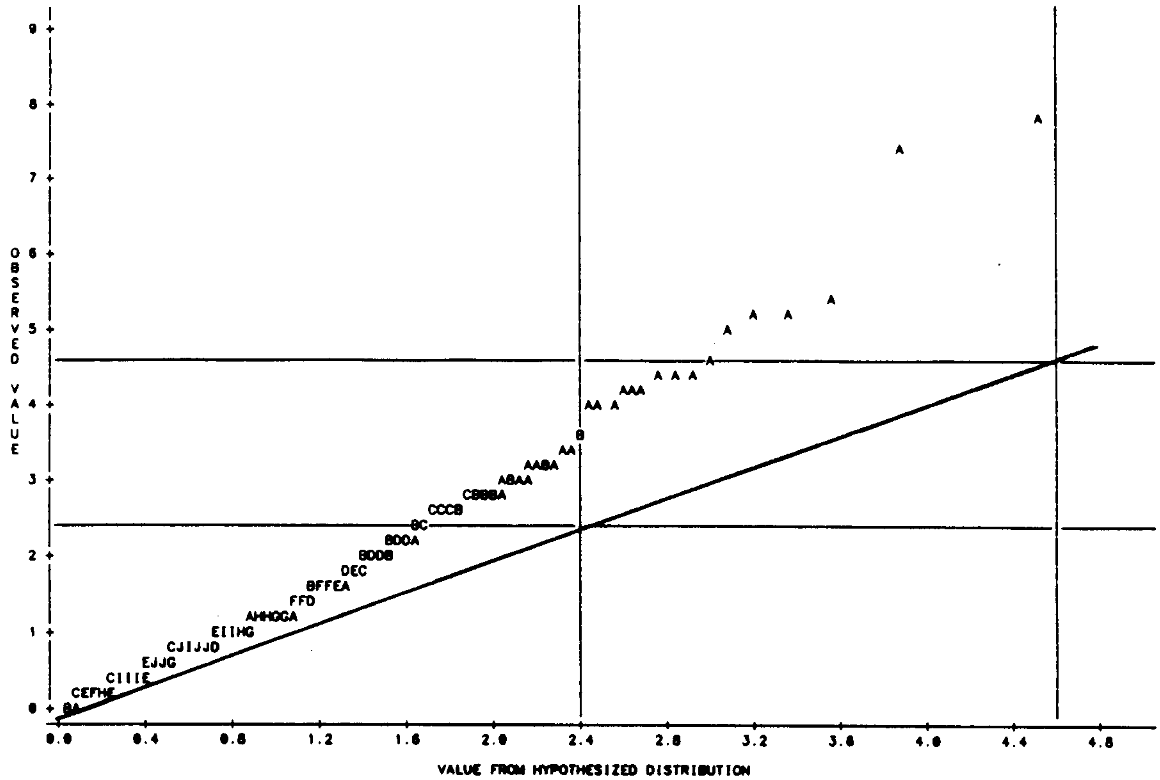

FIGURE 4.2.10
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

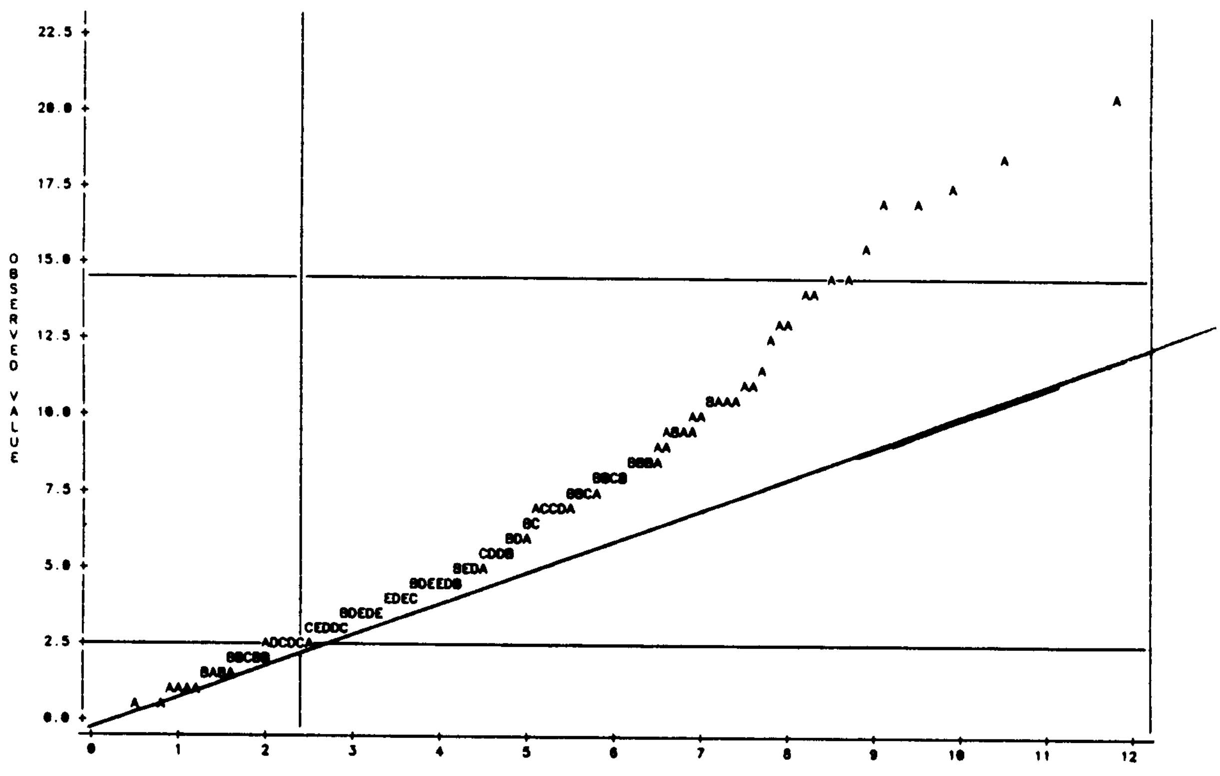

FIGURE 4.2.11
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

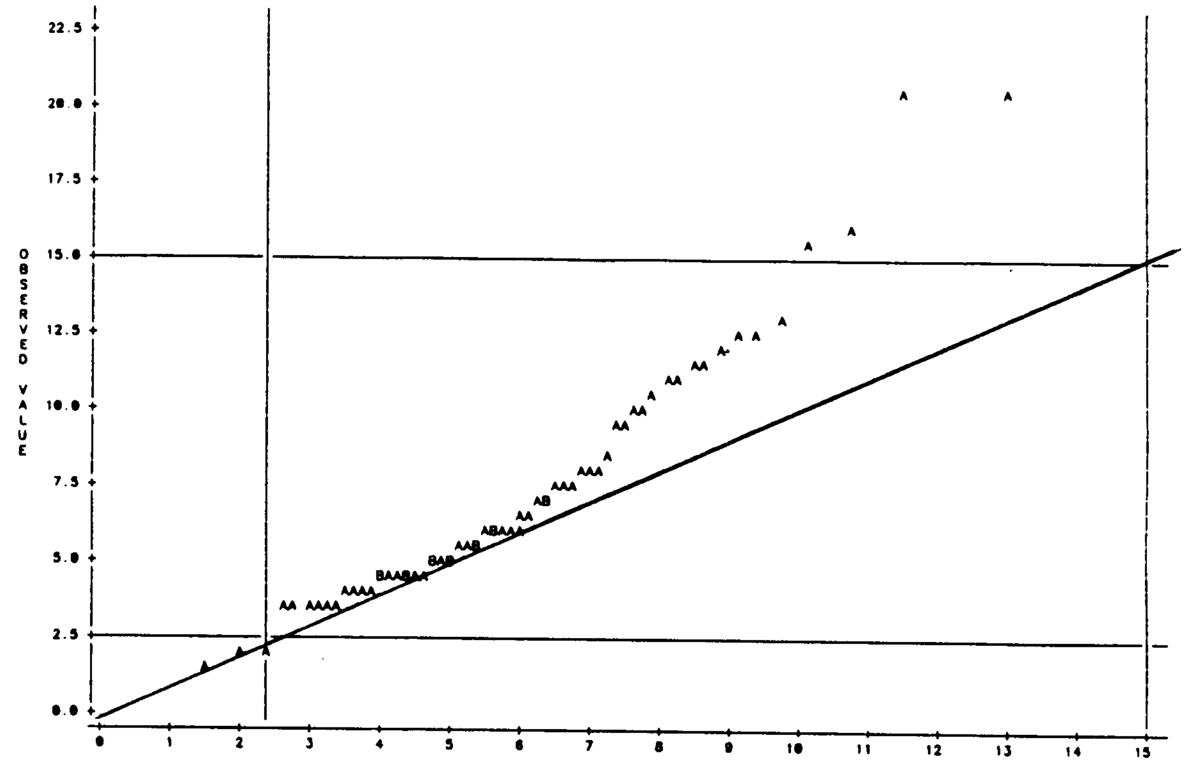

FIGURE 4.2.12
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

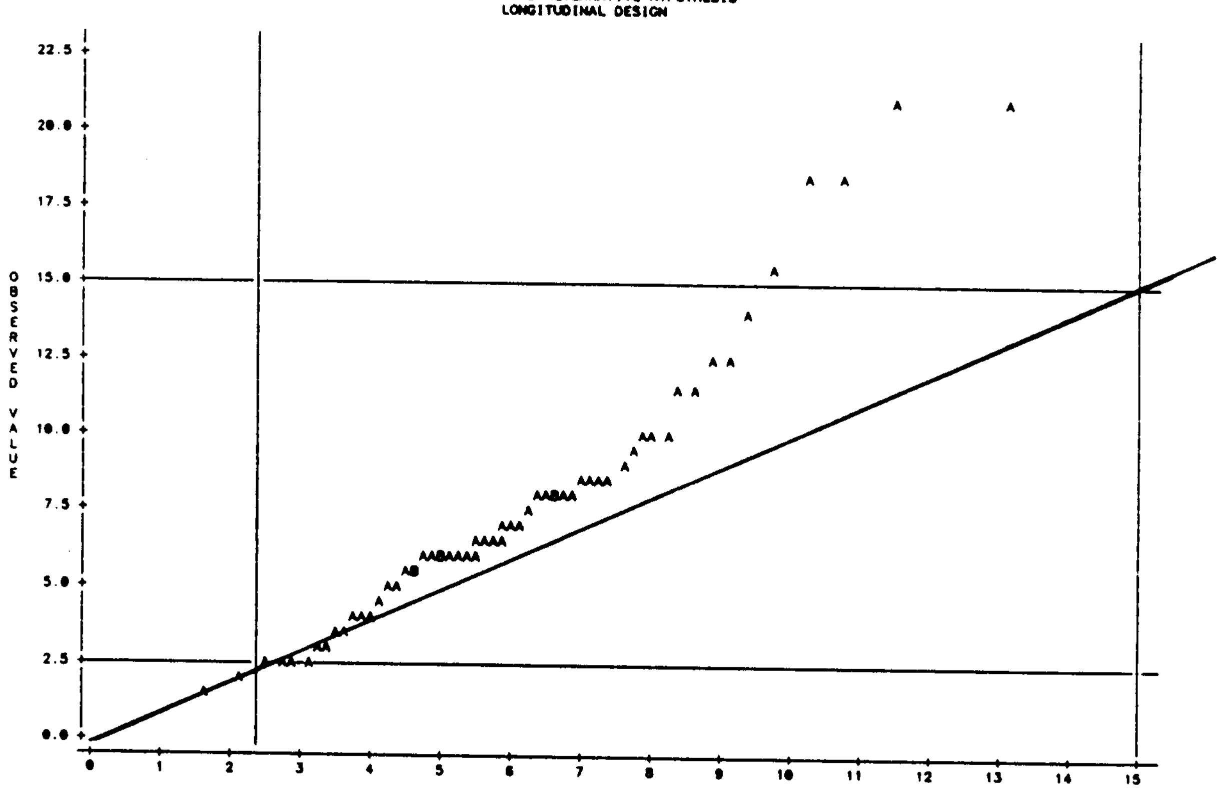

FIGURE 4.2.13
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN

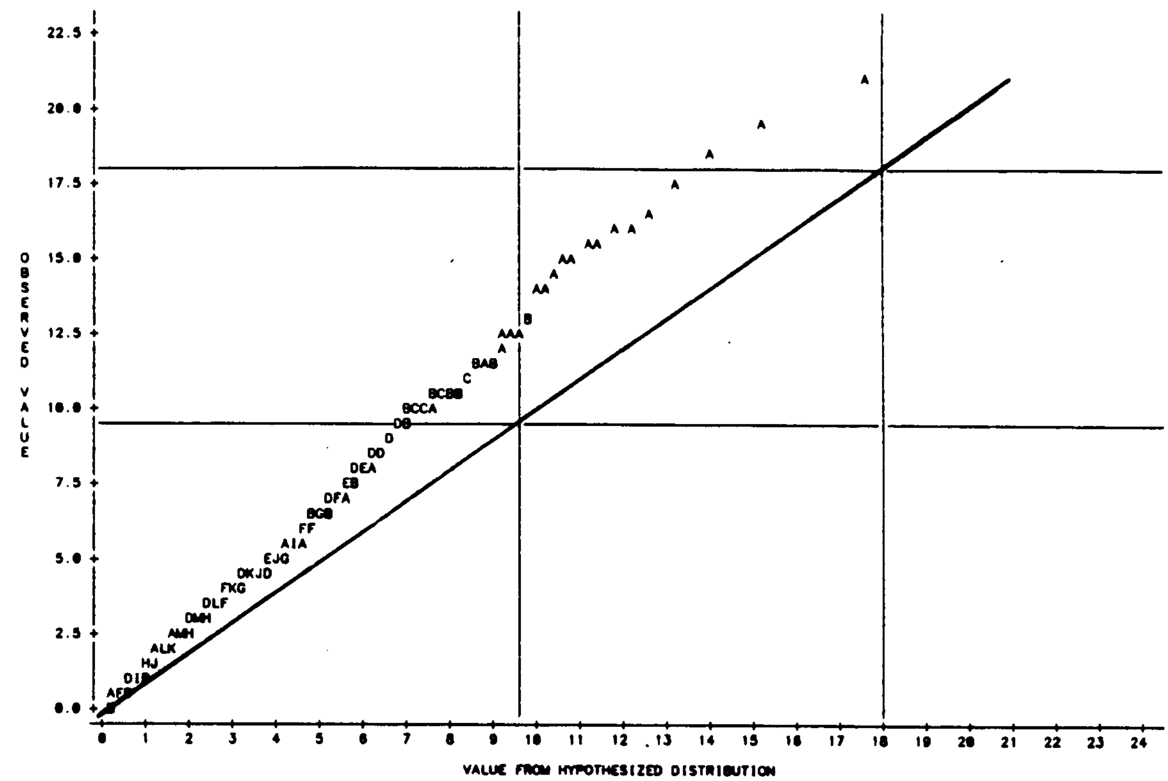

FIGURE 4.2.14
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

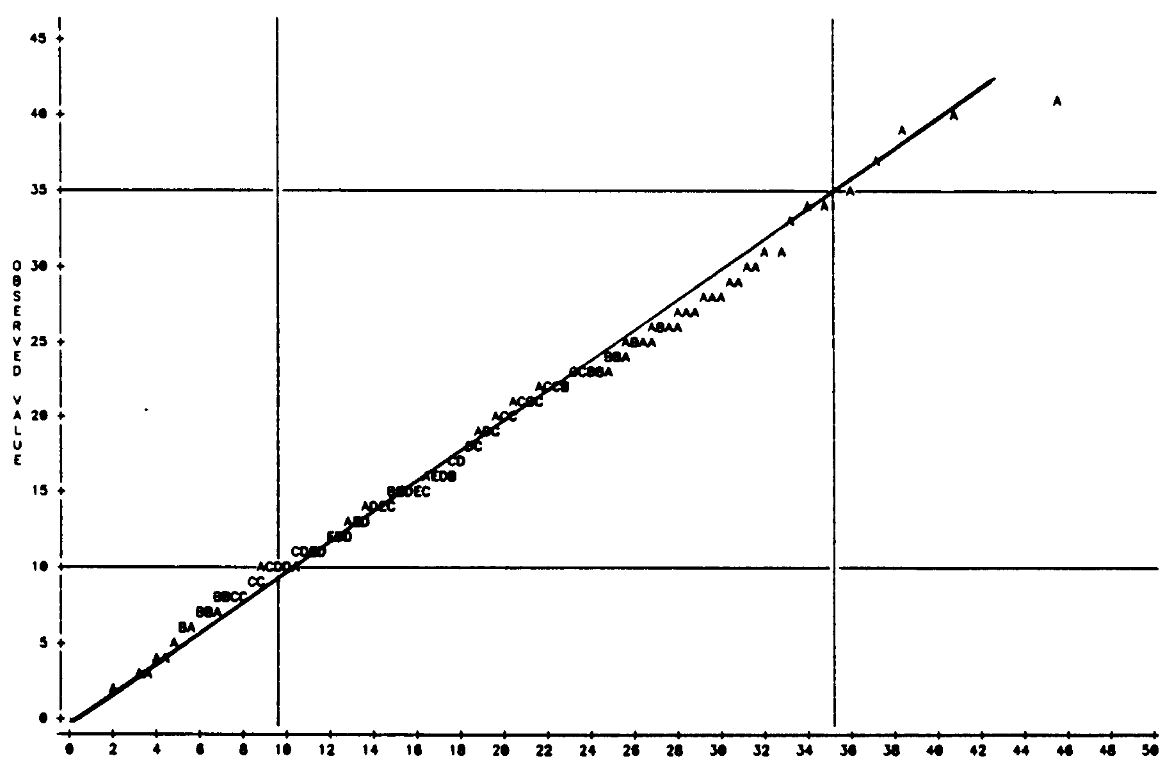

FIGURE 4.2.15
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

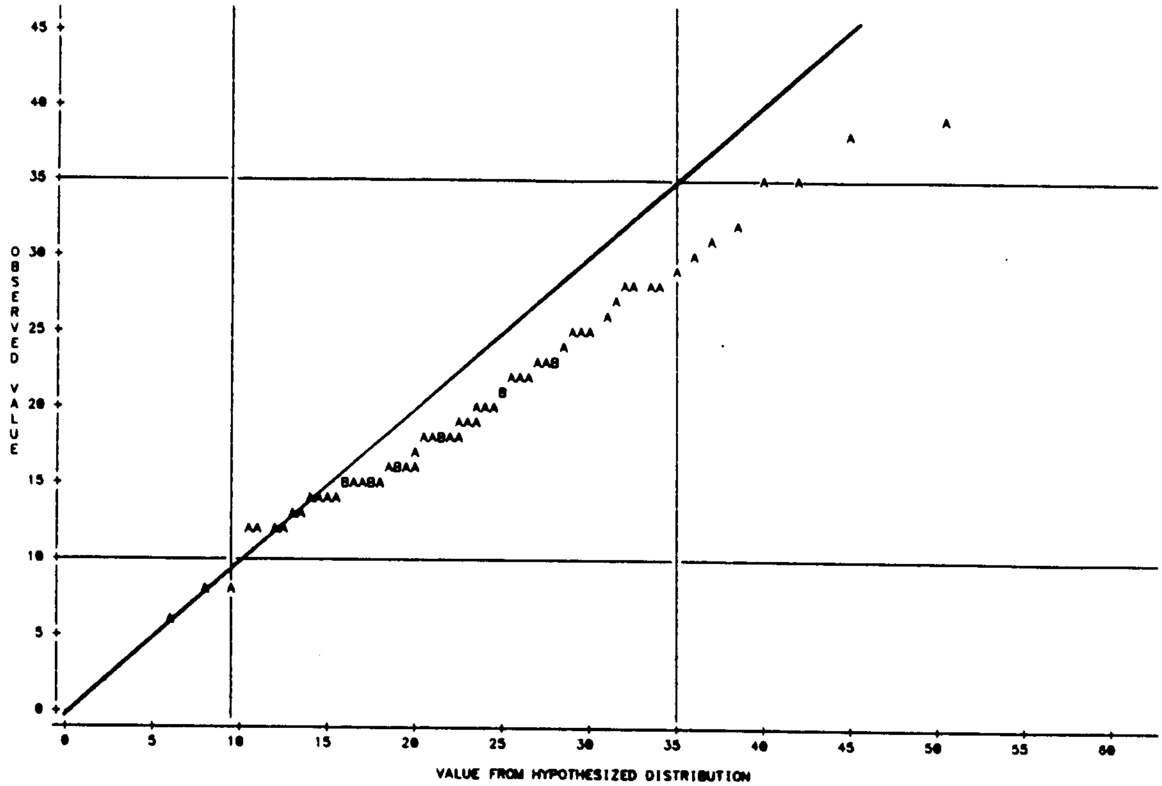

FIGURE 4.2.16
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

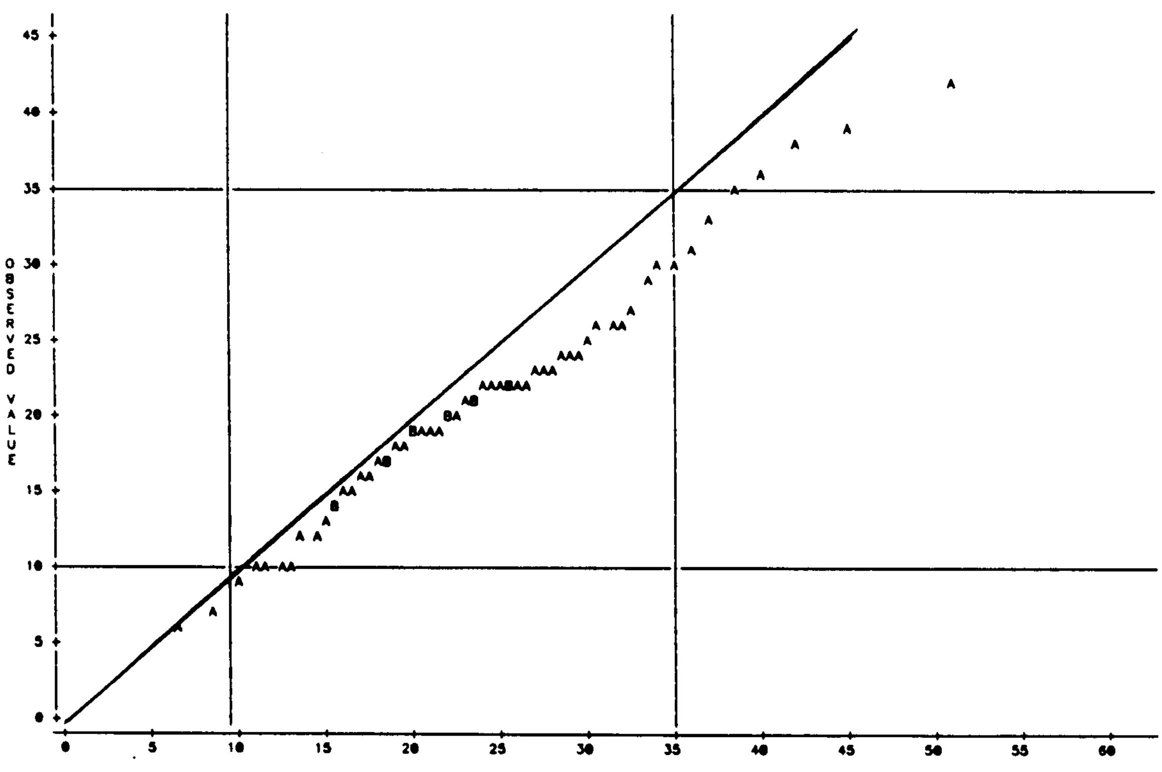

FIGURE 4.2.17
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN

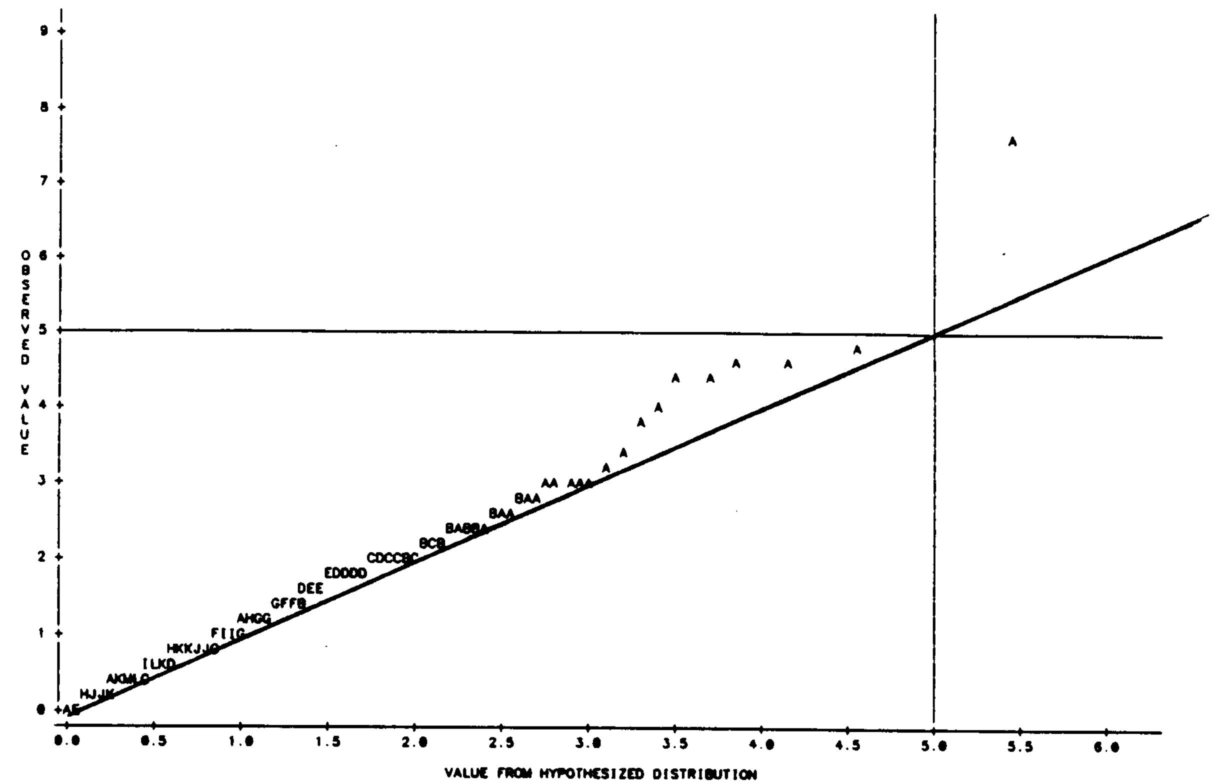

FIGURE 4.2.18
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

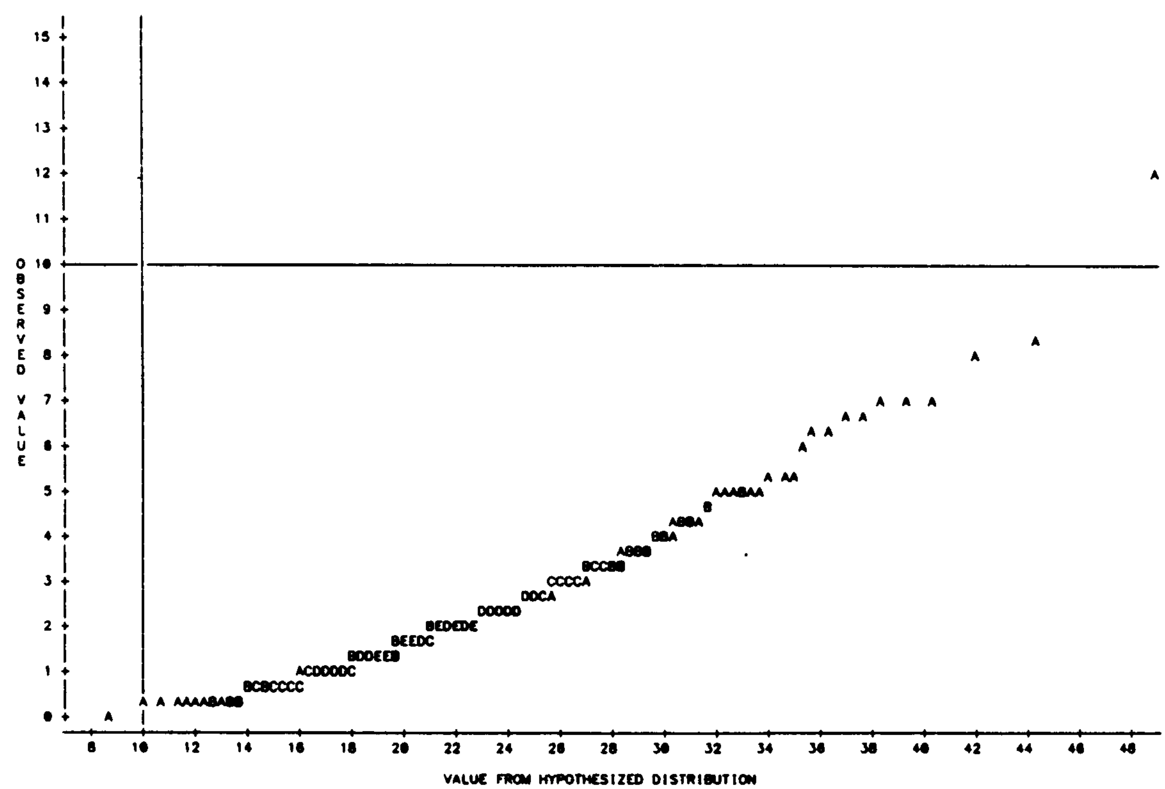

FIGURE 4.2.19
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN

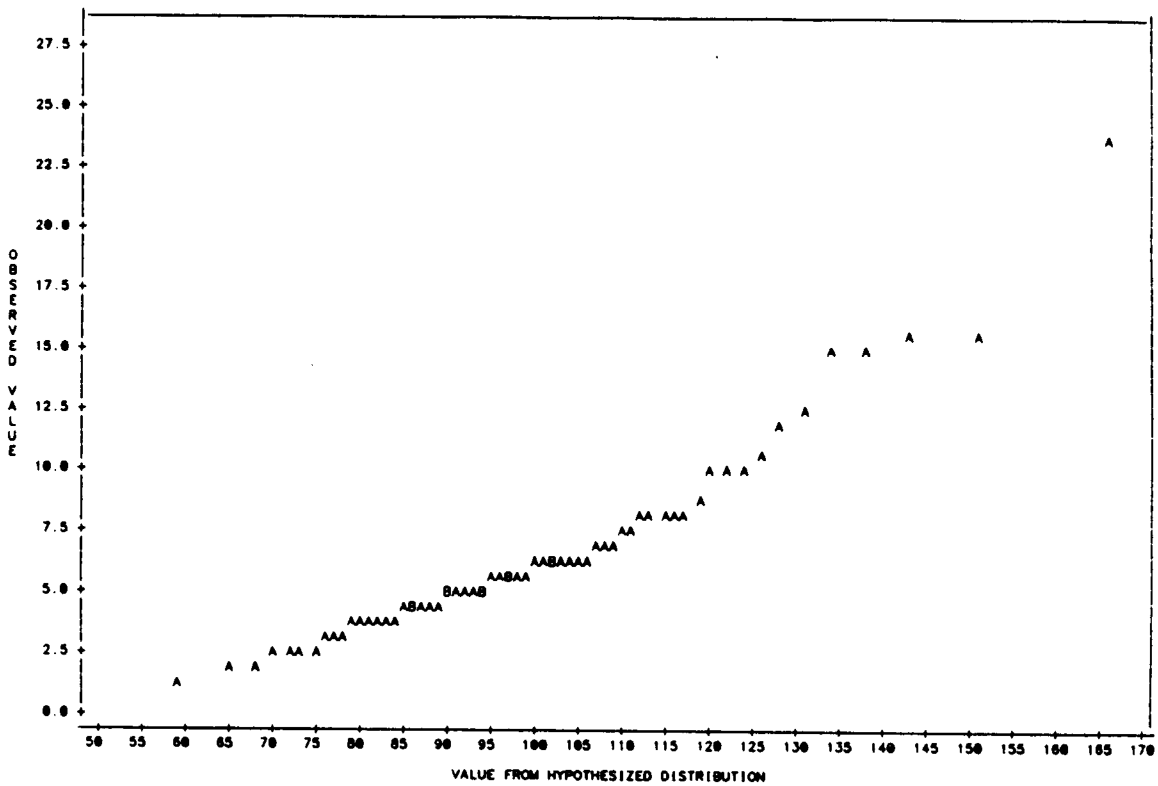

FIGURE 4.3.1
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

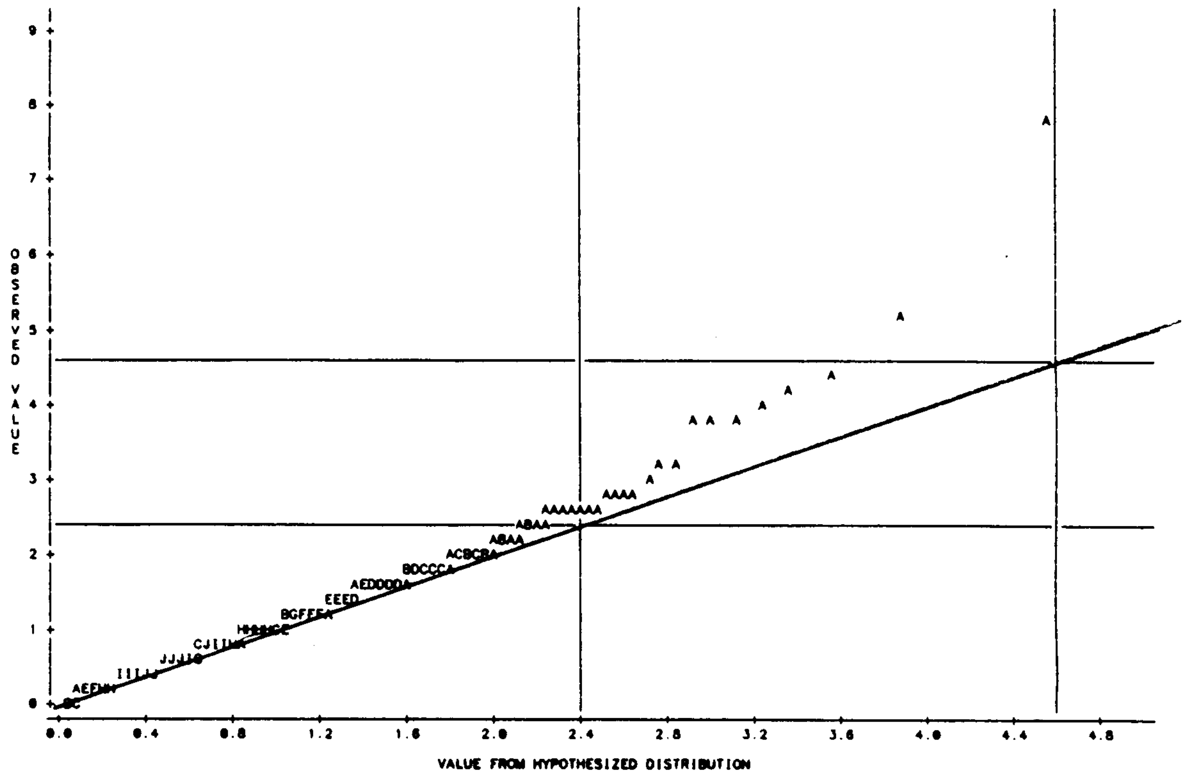

FIGURE 4.3.2
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

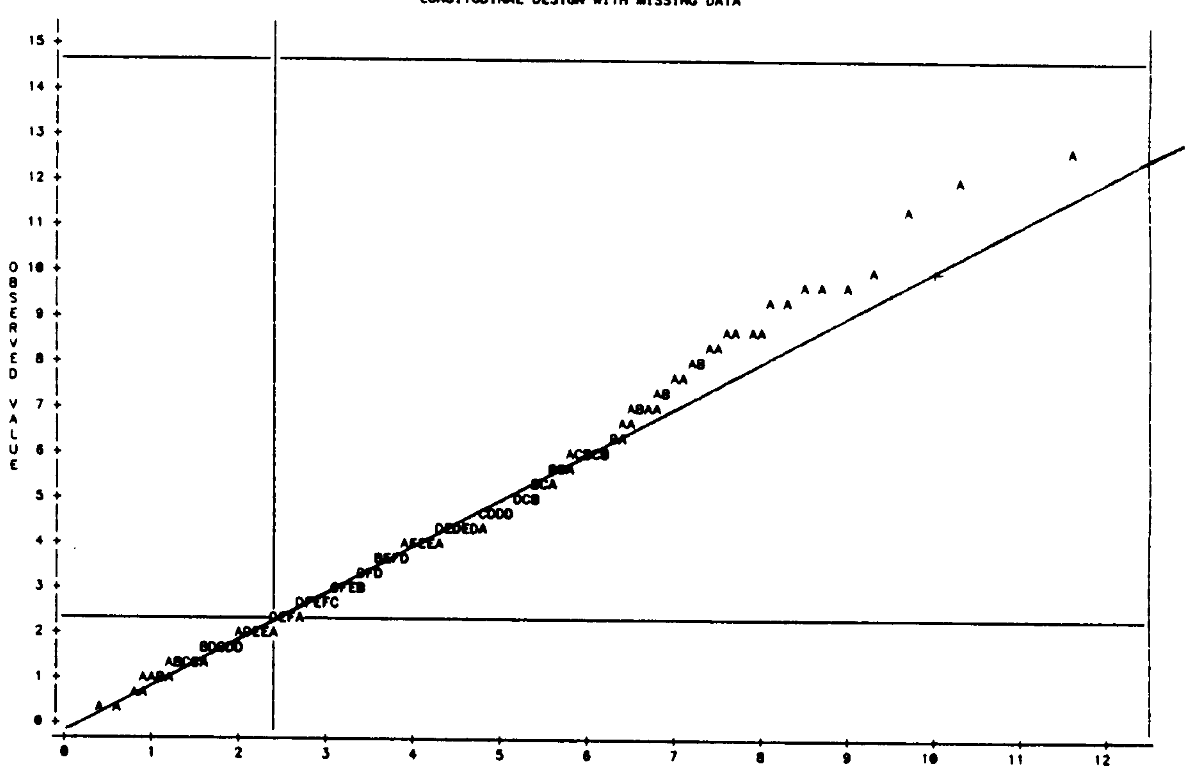

FIGURE 4.3.3
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

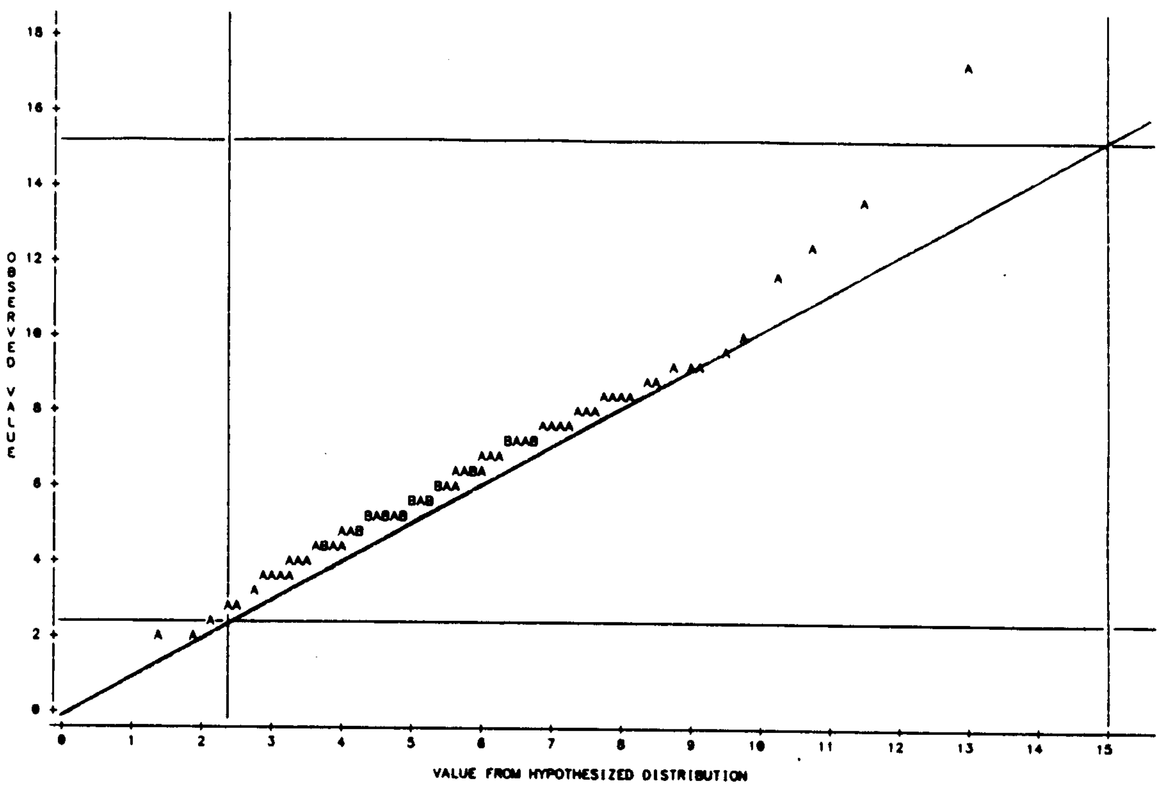

FIGURE 4.3.4
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

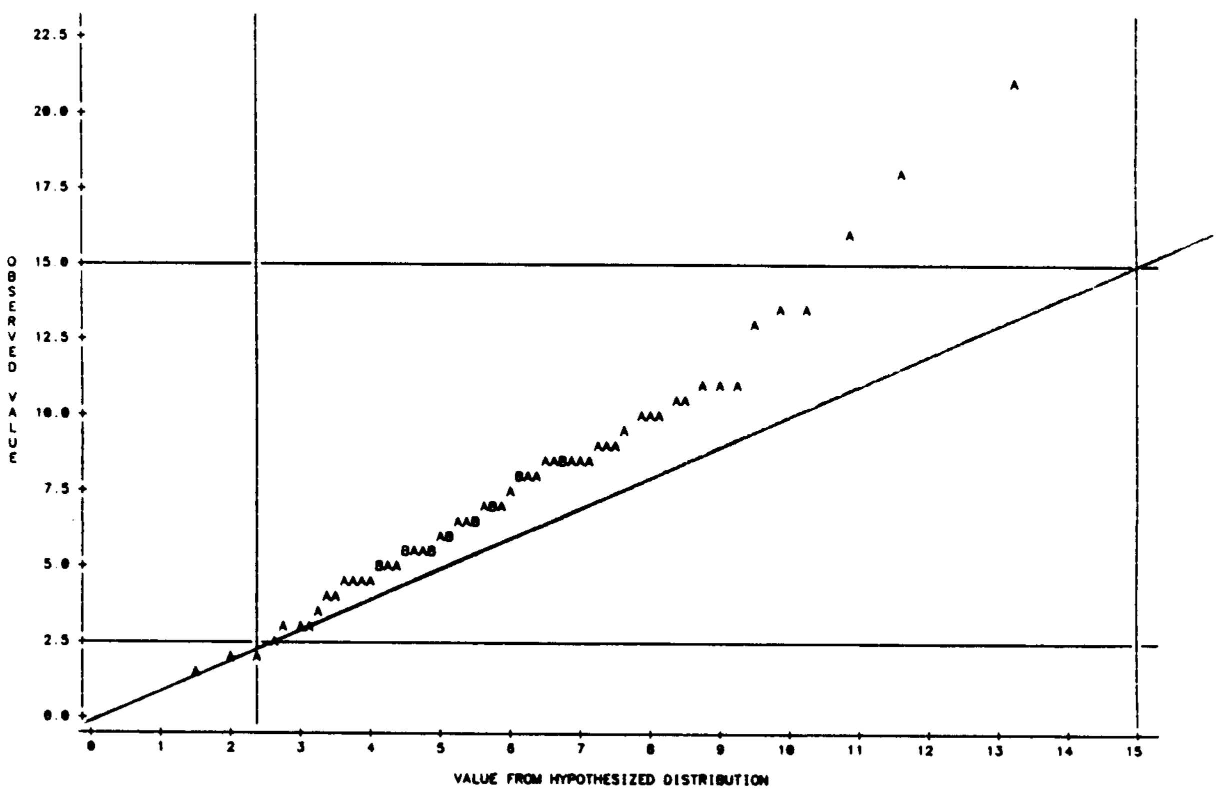

FIGURE 4.3.5
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

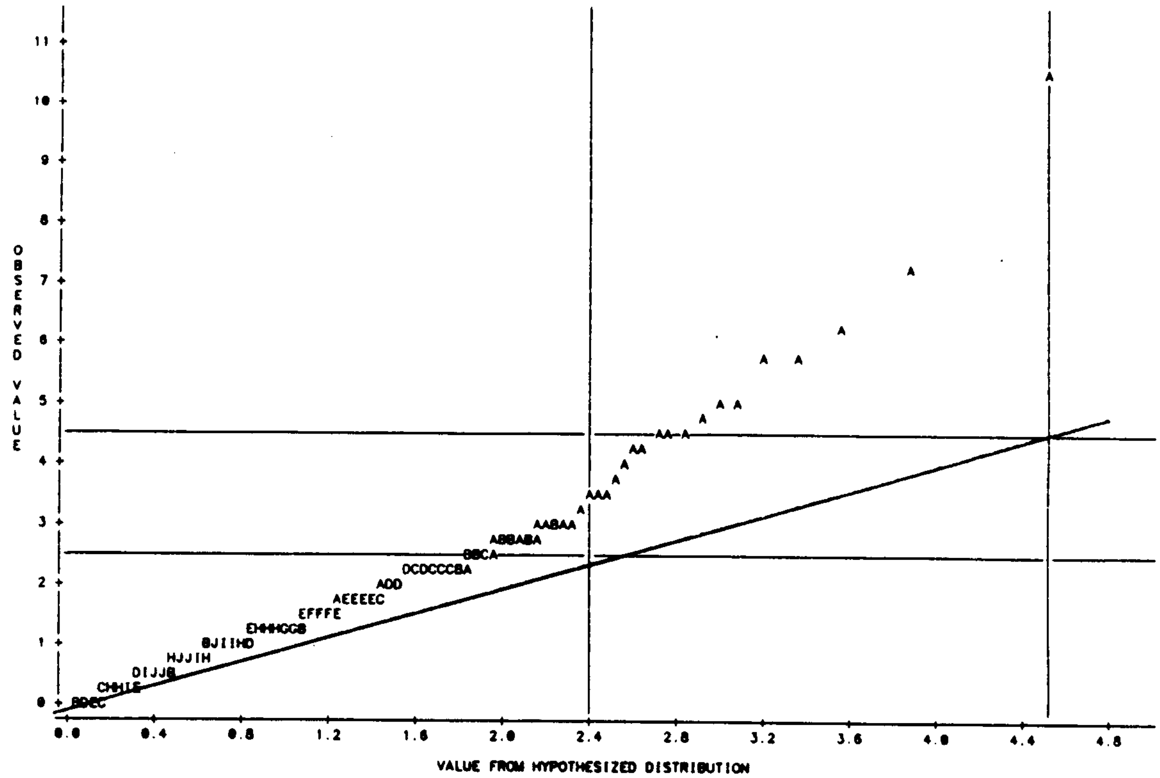

FIGURE 4.3.6
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

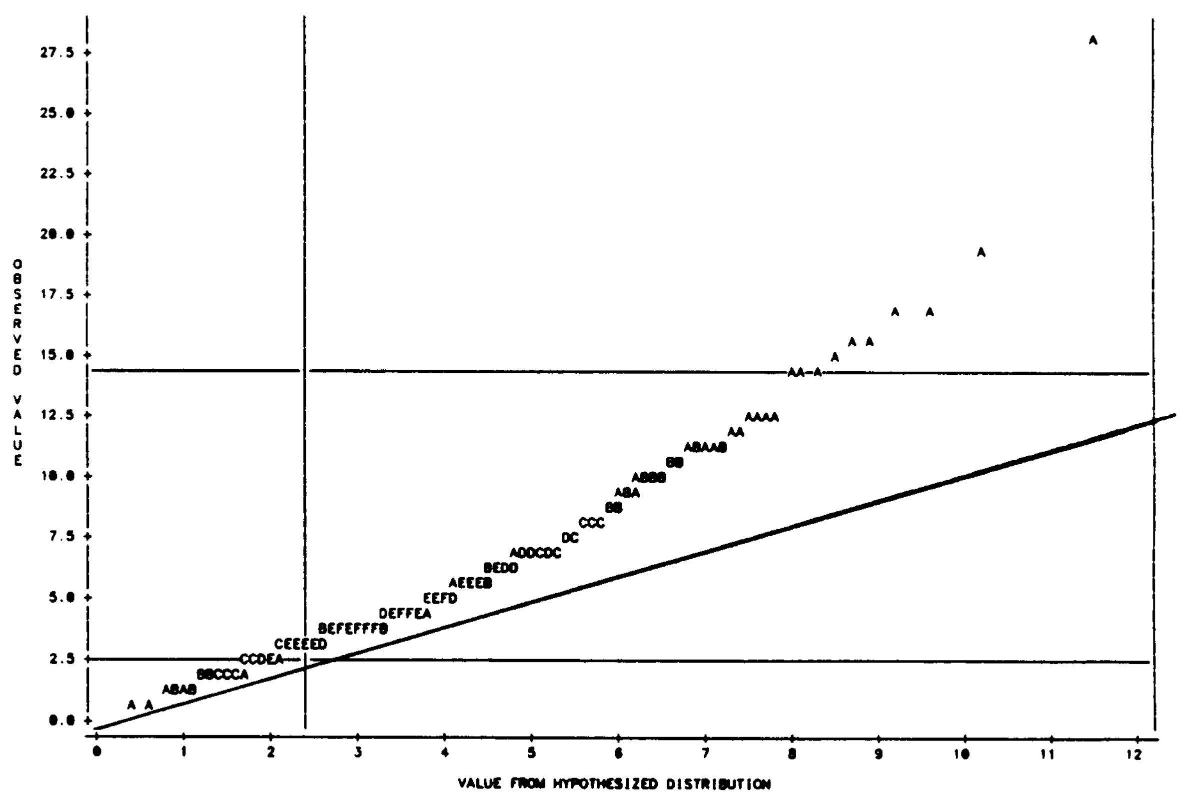

FIGURE 4.3.7
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

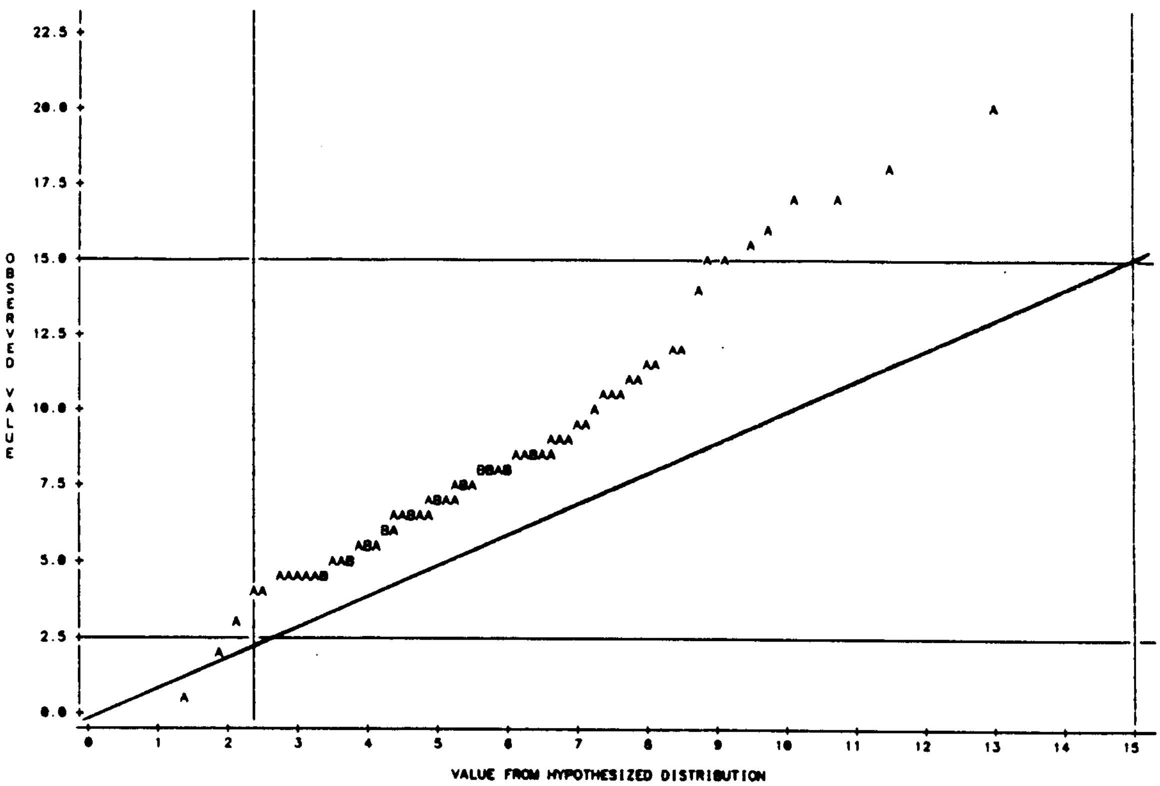

FIGURE 4.3.8
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

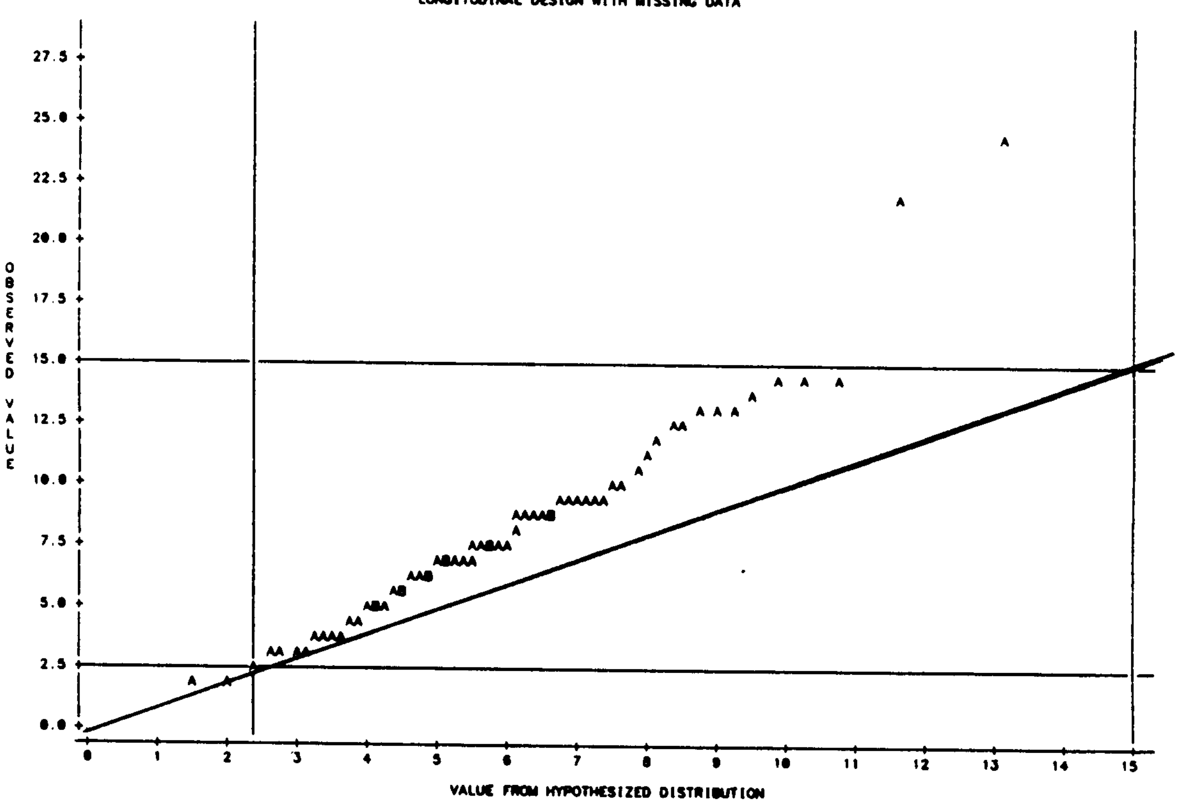

FIGURE 4.3.9
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

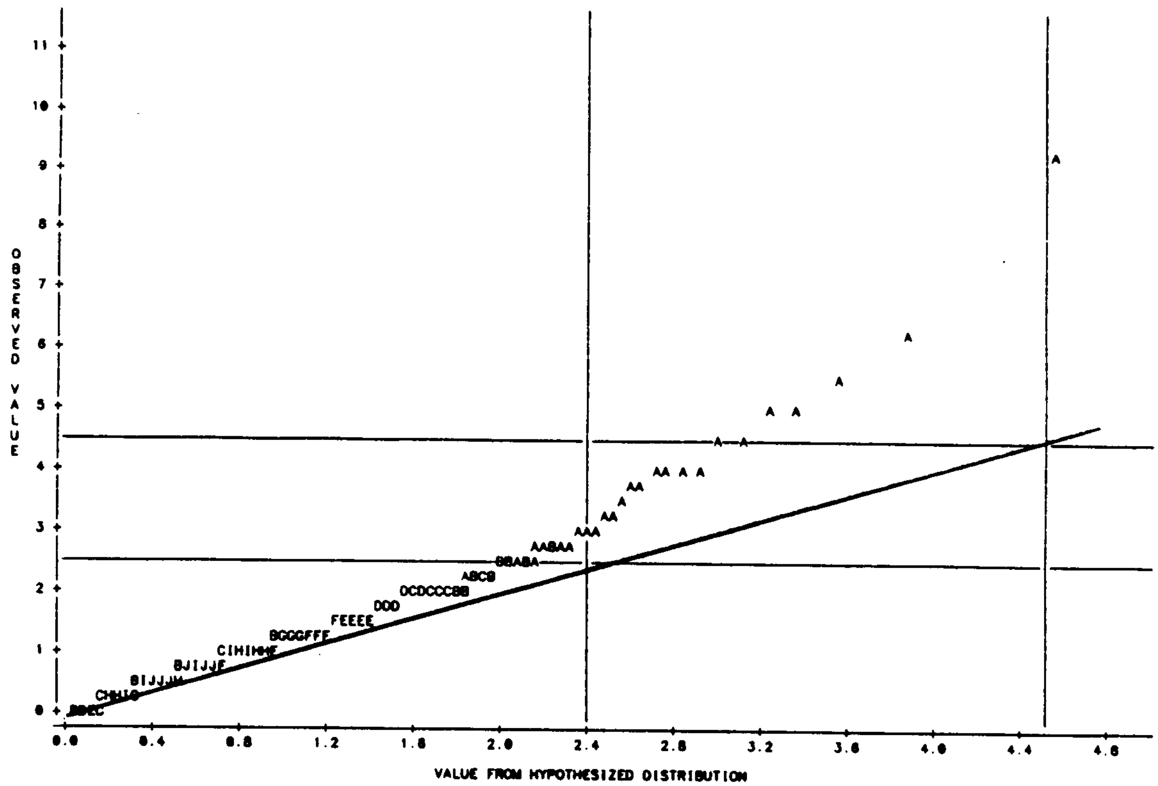

FIGURE 4.3.10
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

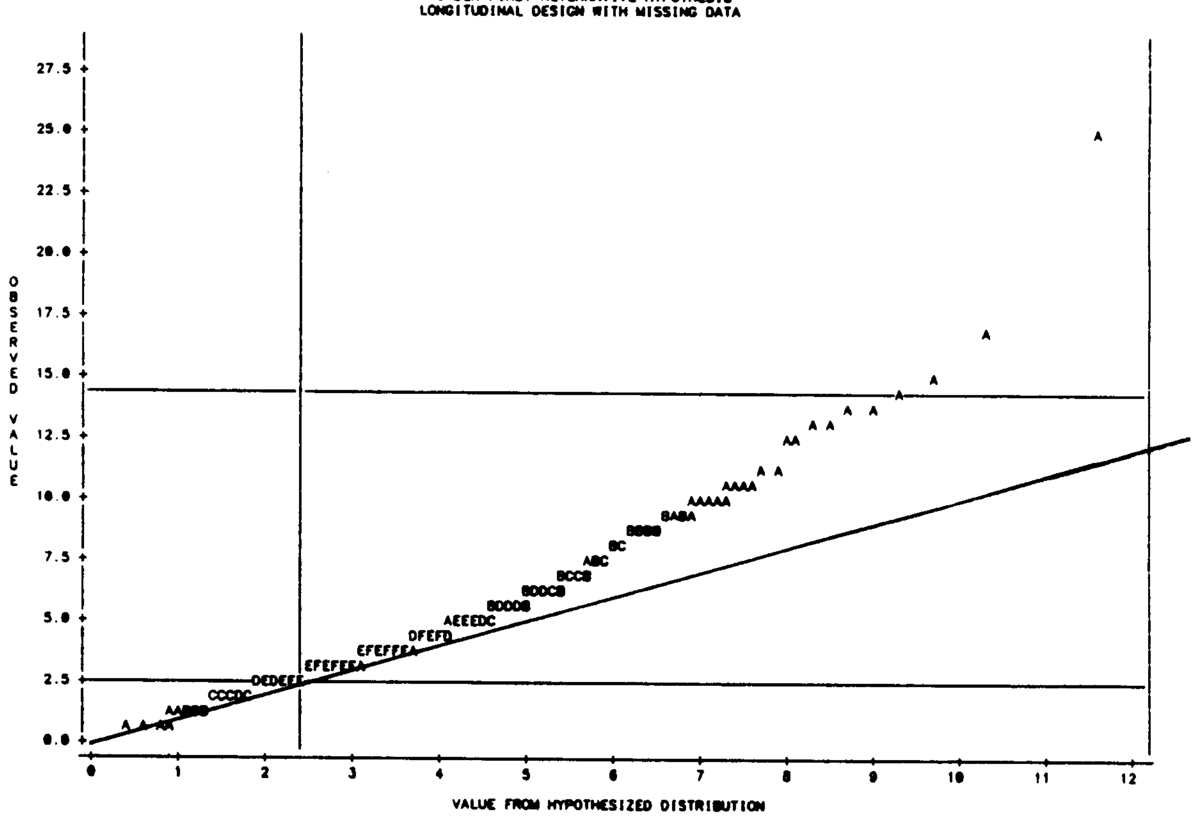

FIGURE 4.3.11
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

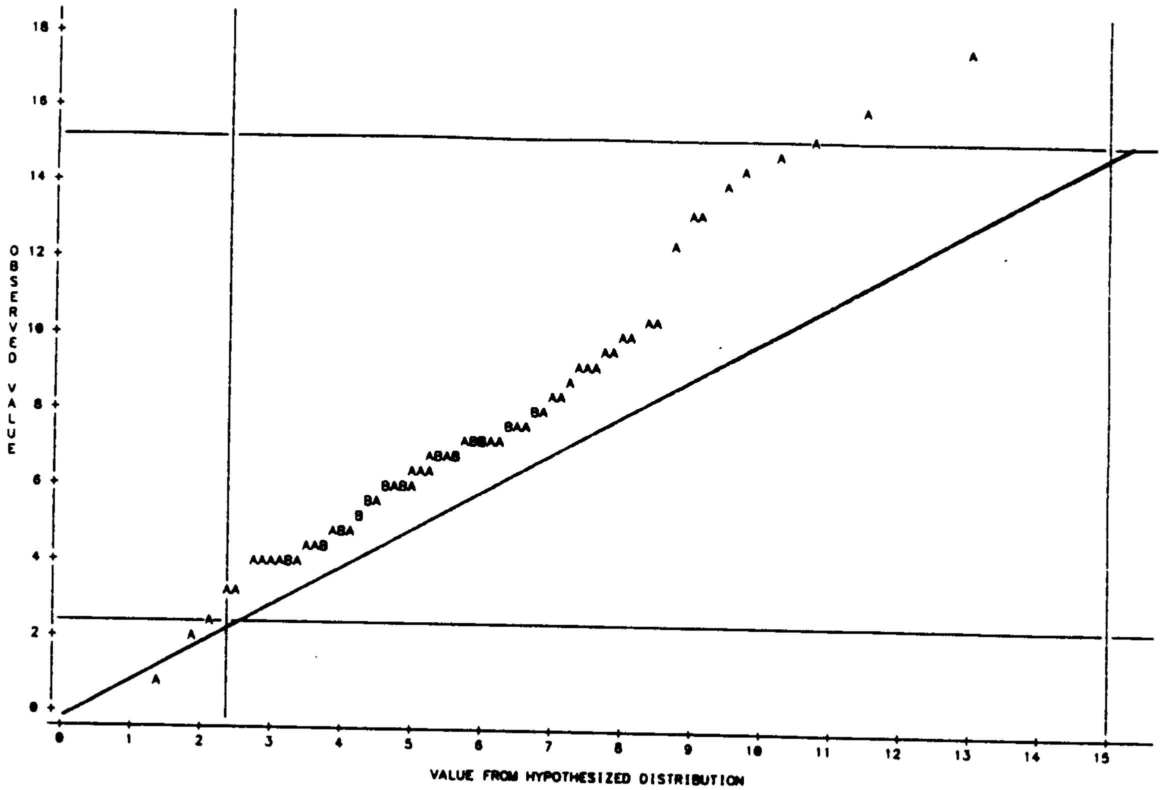

FIGURE 4.3.12
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

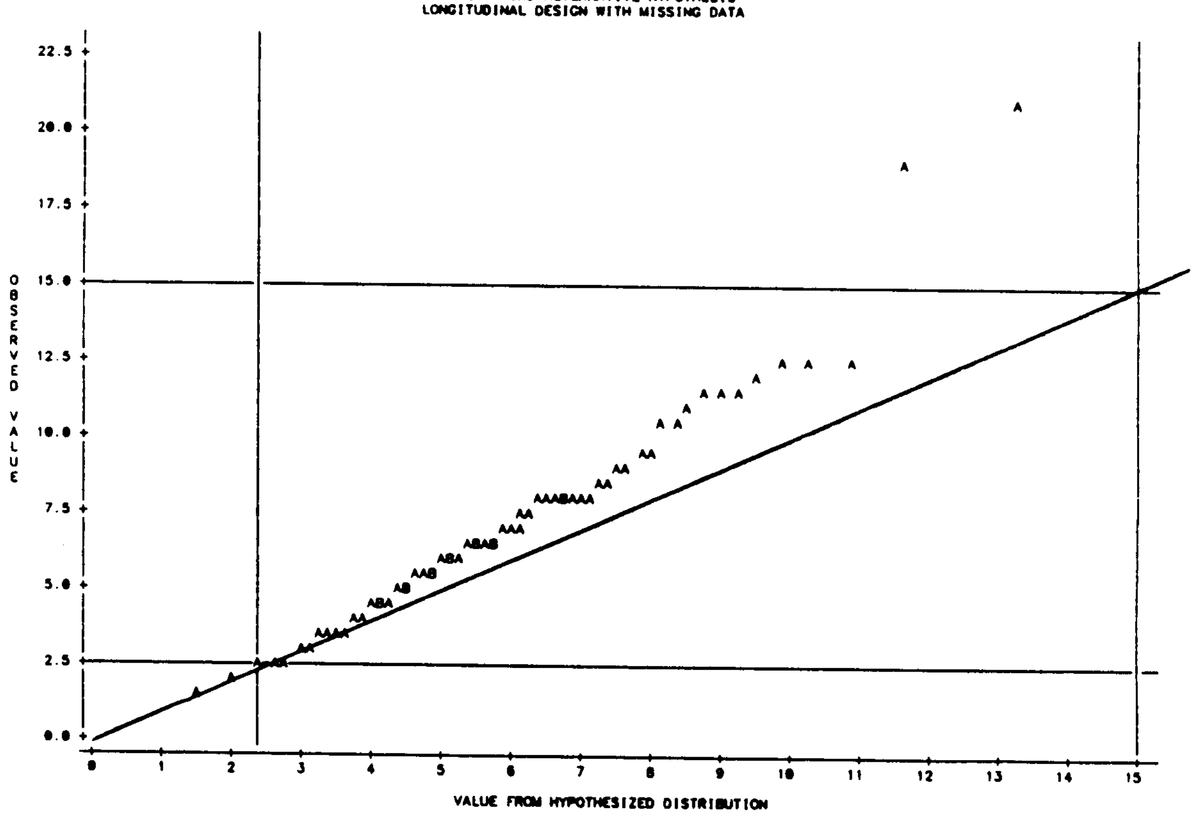

FIGURE 4.3.13
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

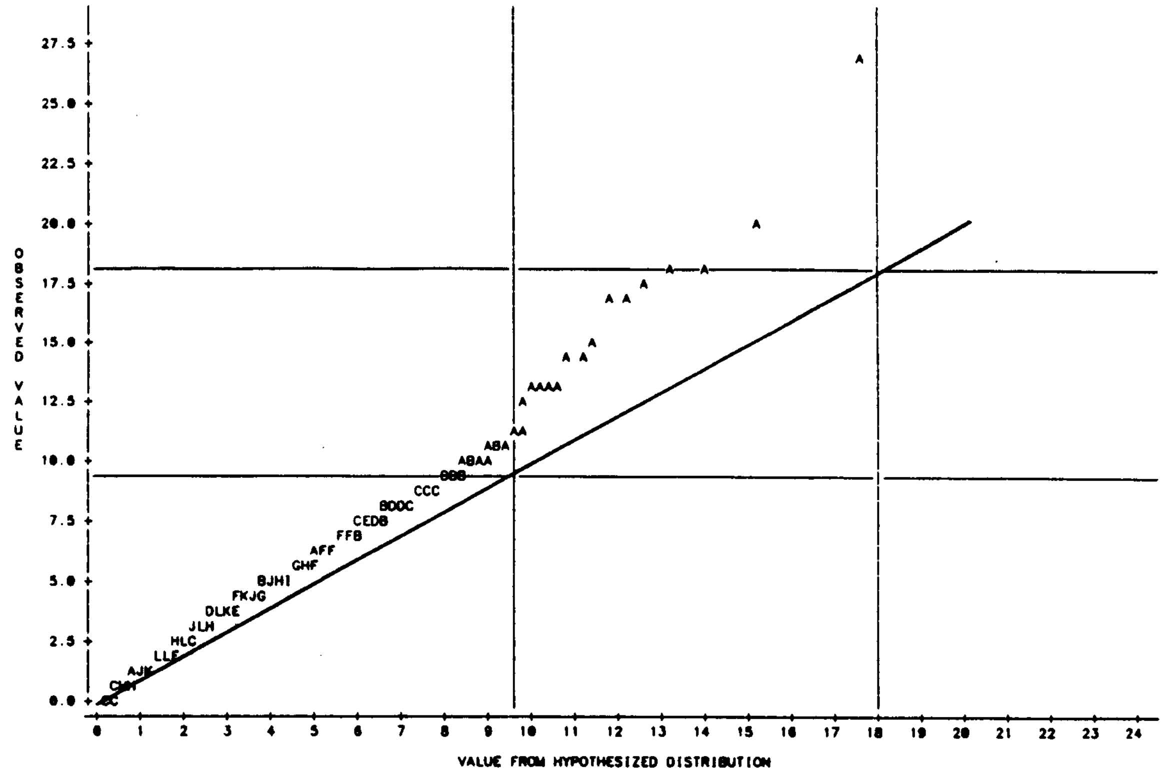

FIGURE 4.3.14
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

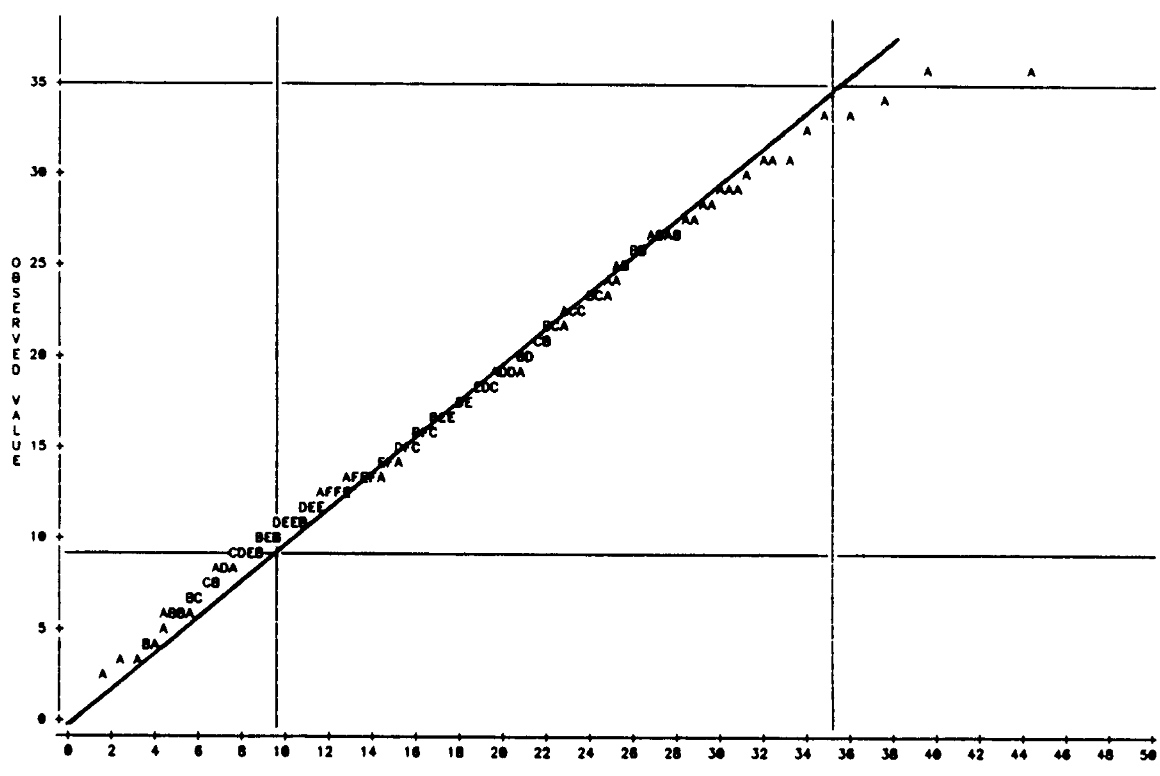

FIGURE 4.3.15
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

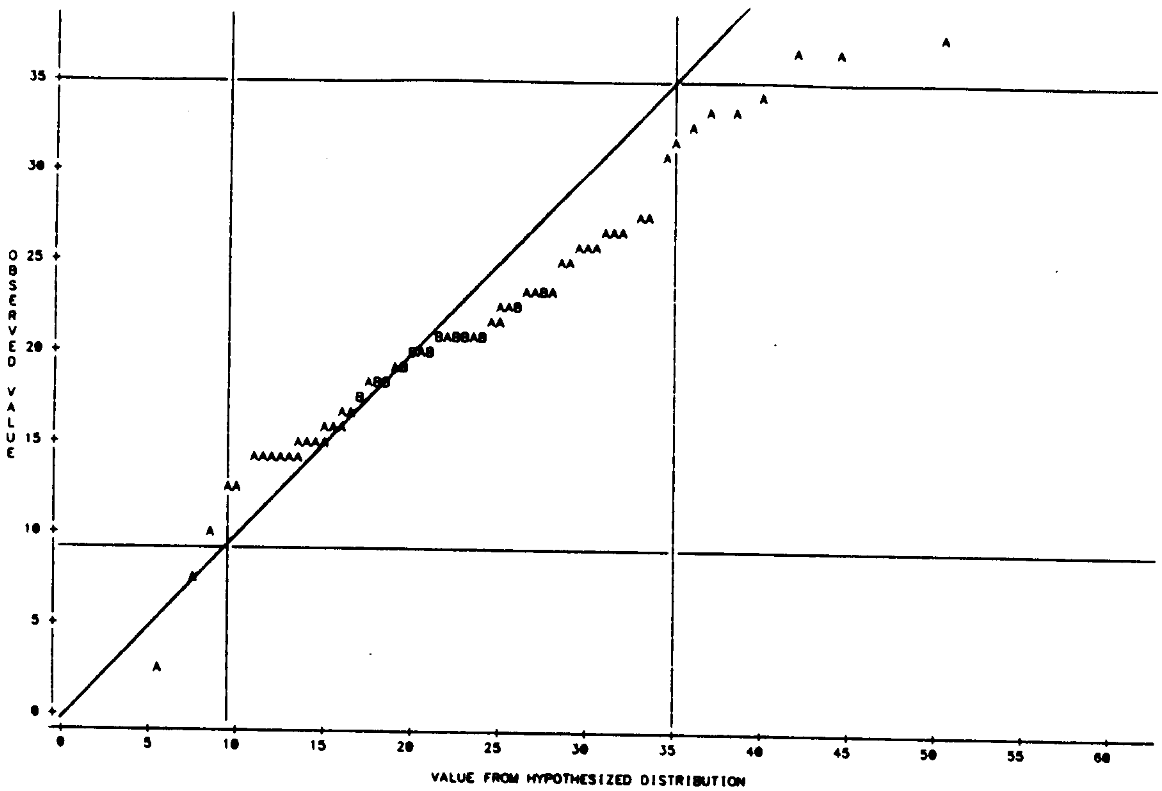

FIGURE 4.3.16
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

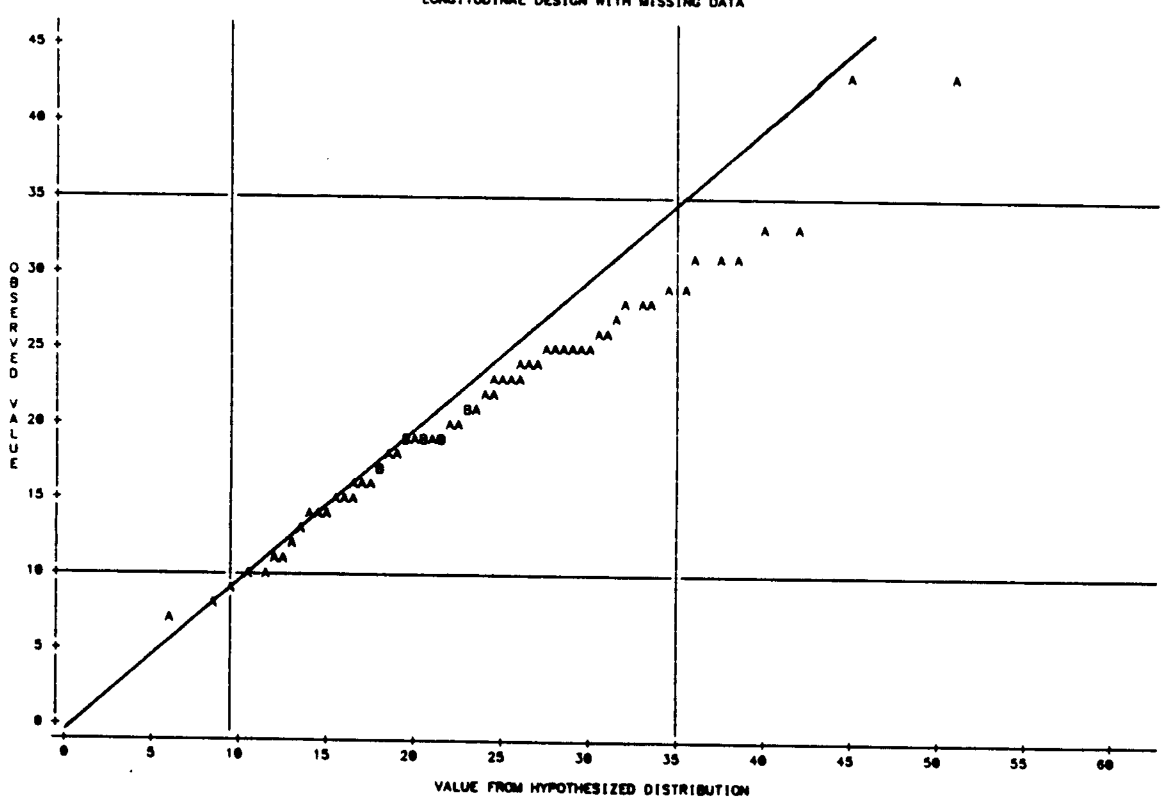

FIGURE 4.3.17
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER NULL HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

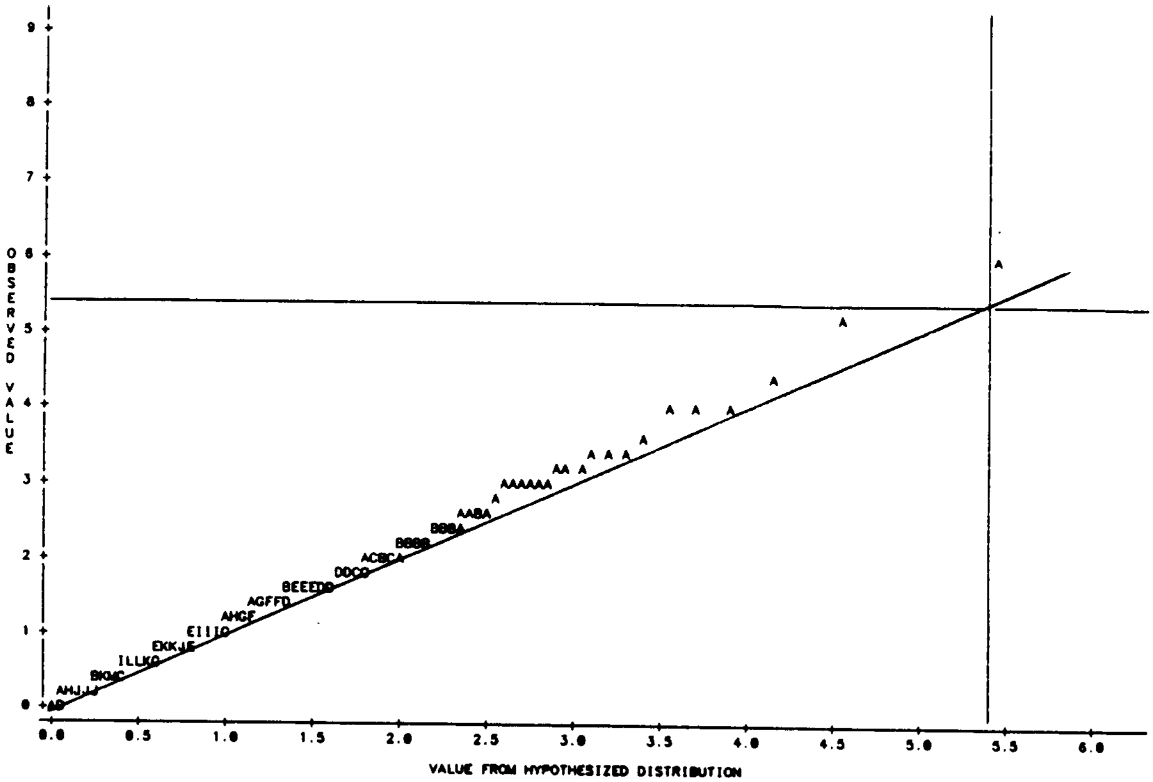

FIGURE 4.3.18
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

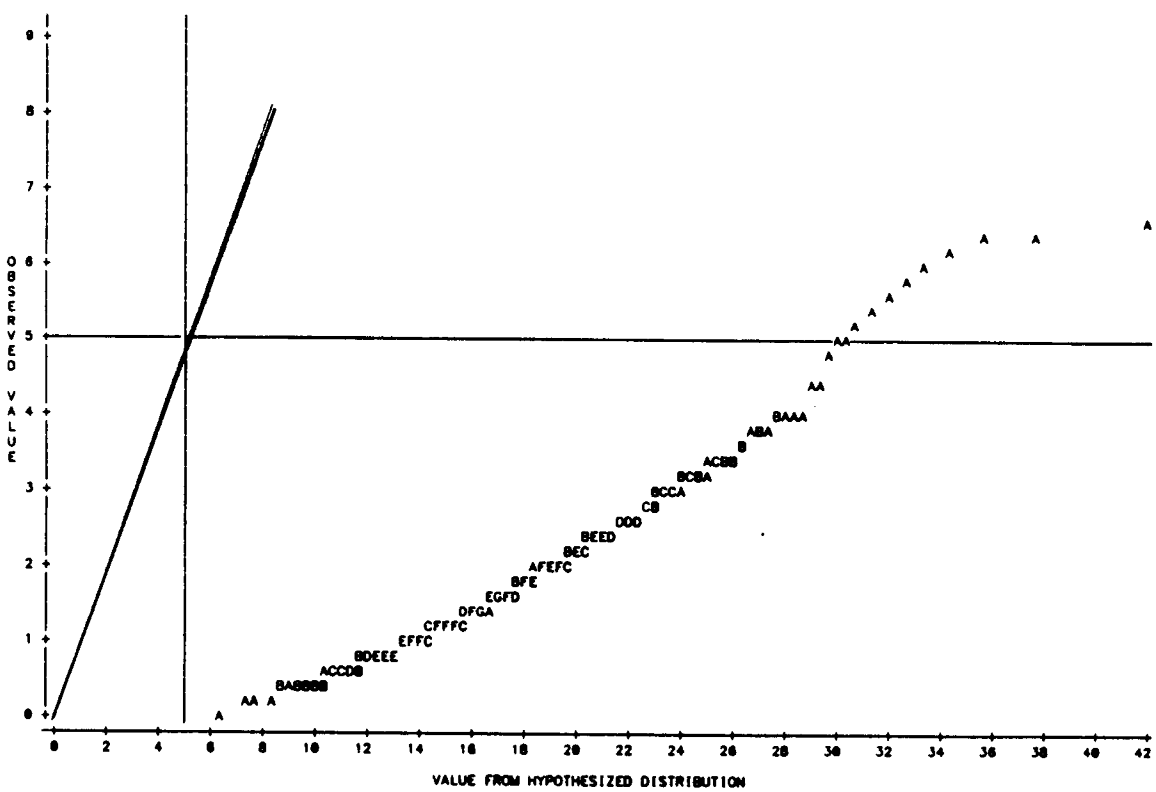

FIGURE 4.3.19
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LONGITUDINAL DESIGN WITH MISSING DATA

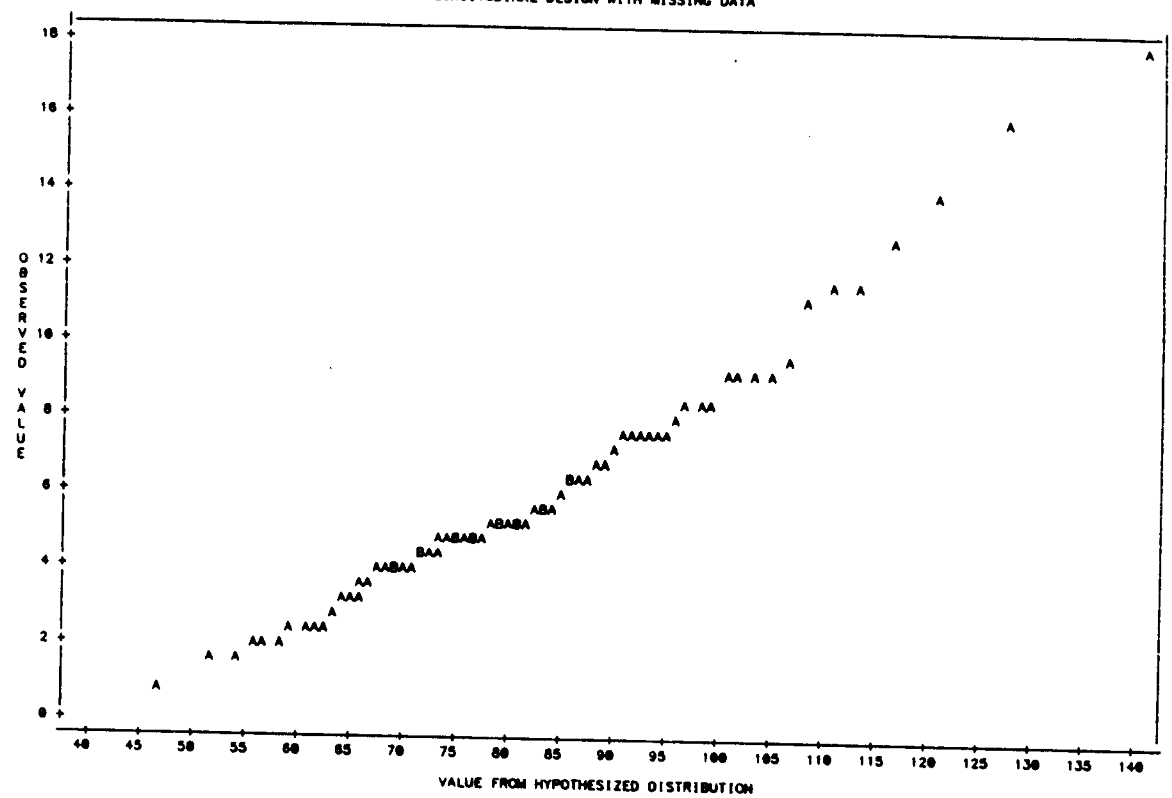

FIGURE 4.4.1
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

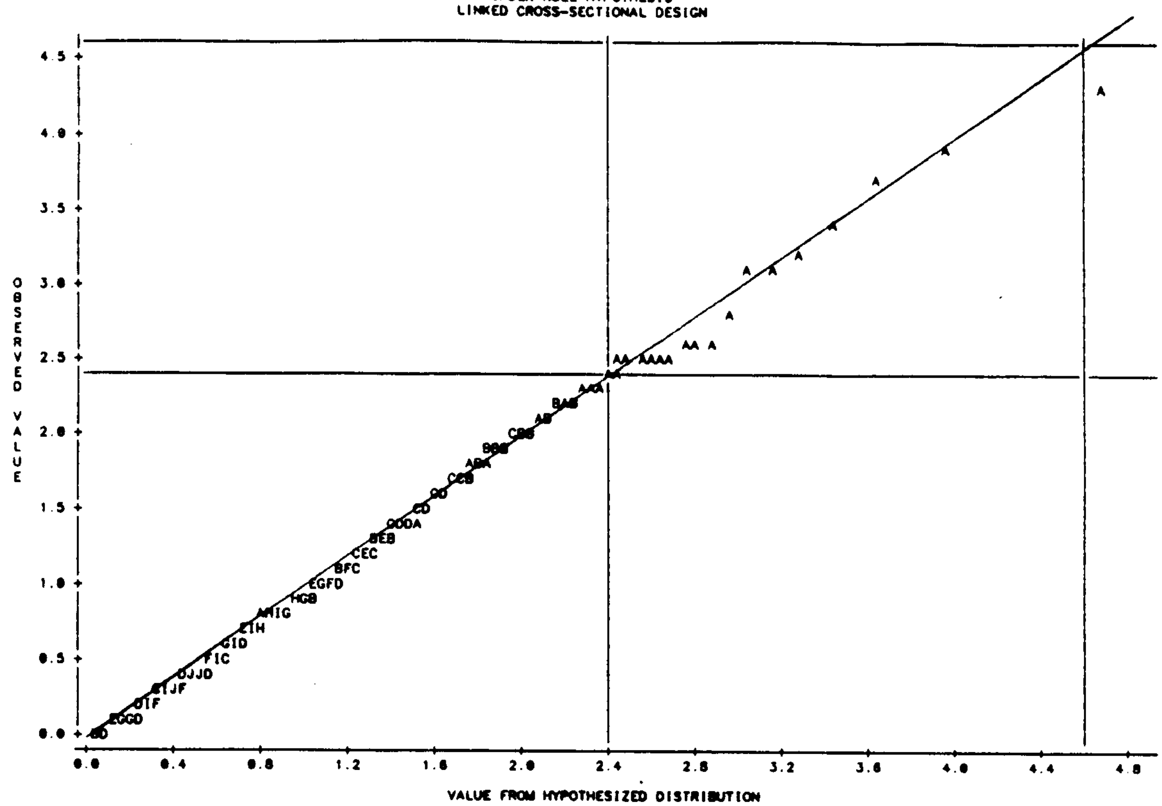

FIGURE 4.4.2
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

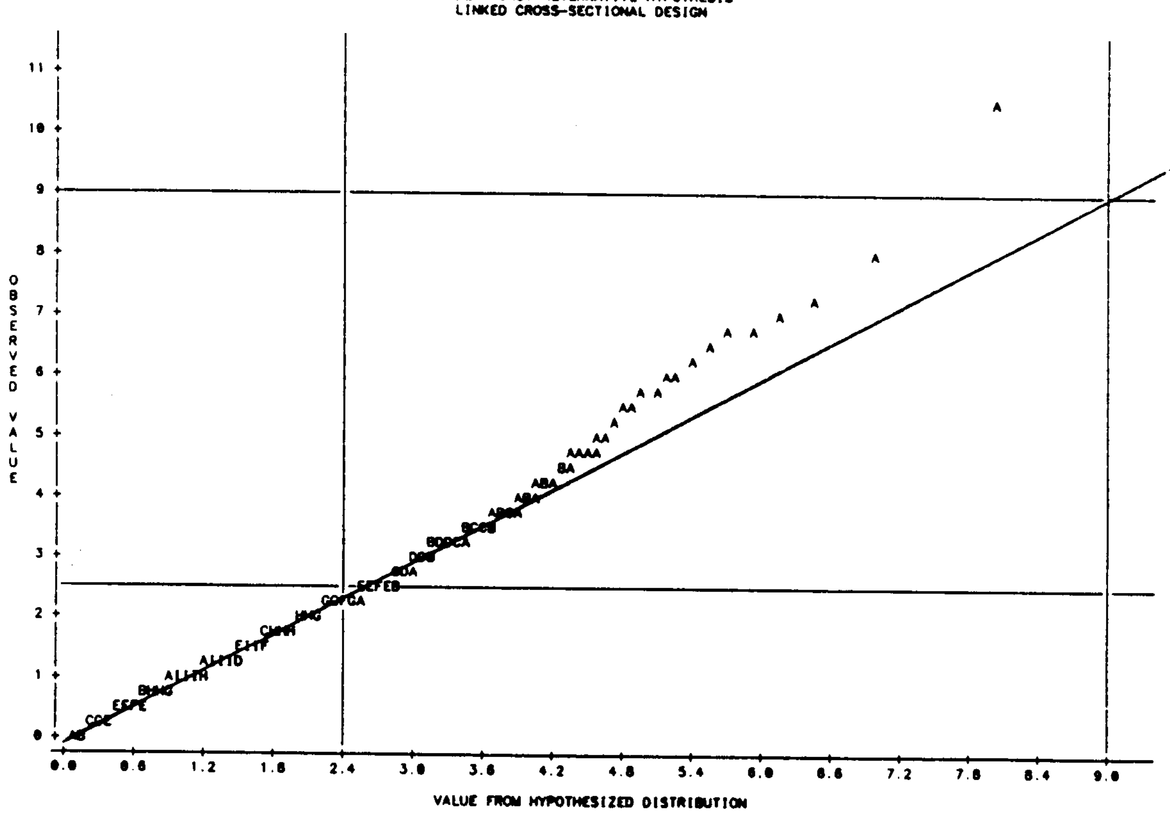

FIGURE 4.4.3
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

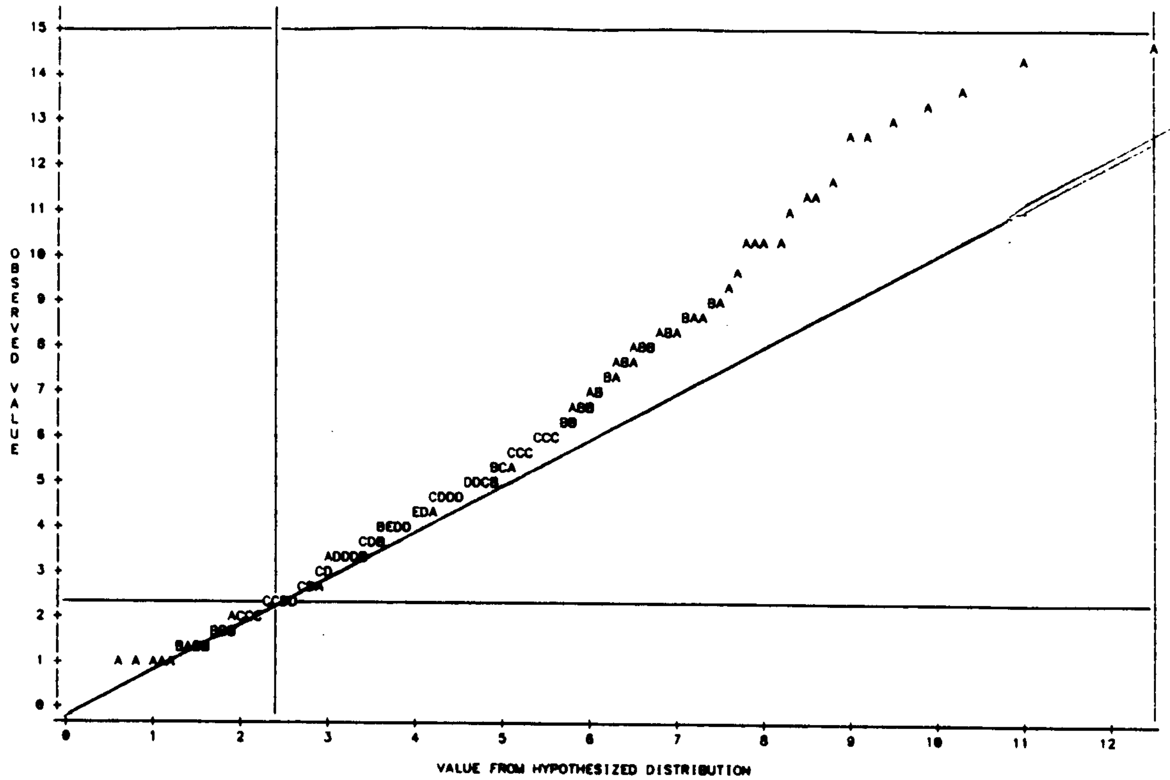

FIGURE 4.4.4
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

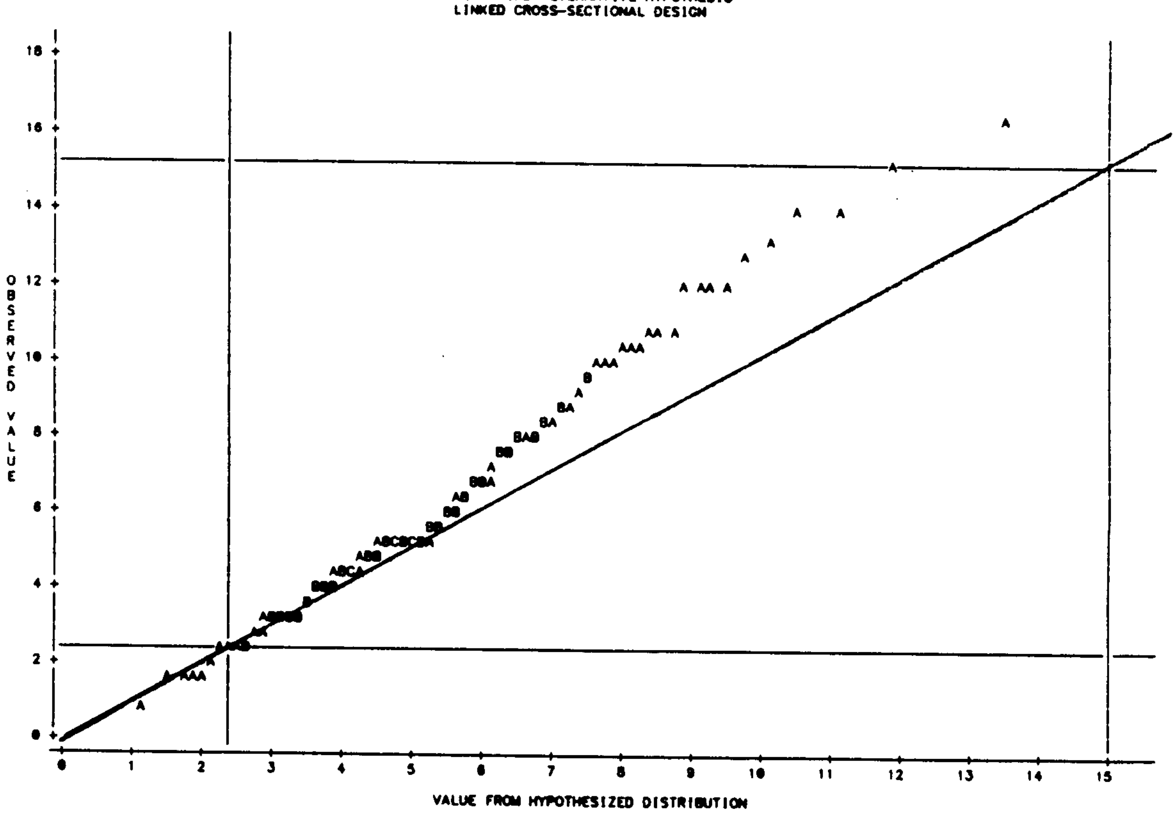

FIGURE 4.4.5
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

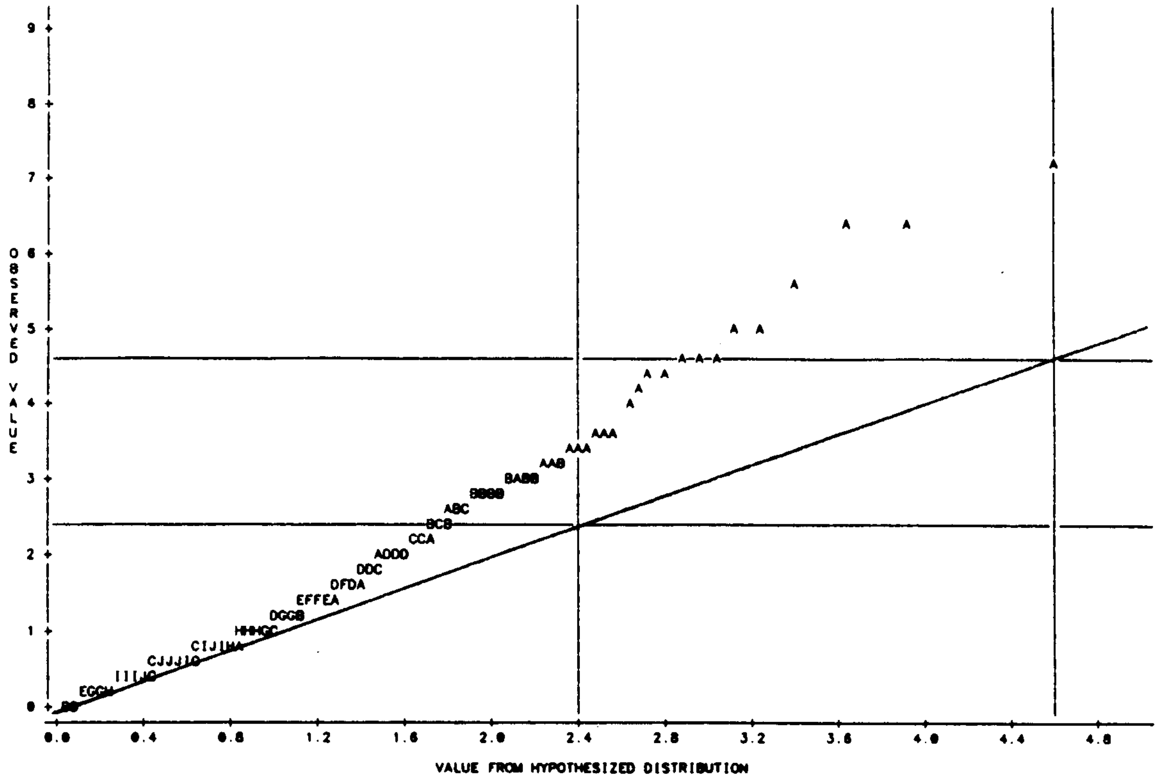

FIGURE 4.4.6
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

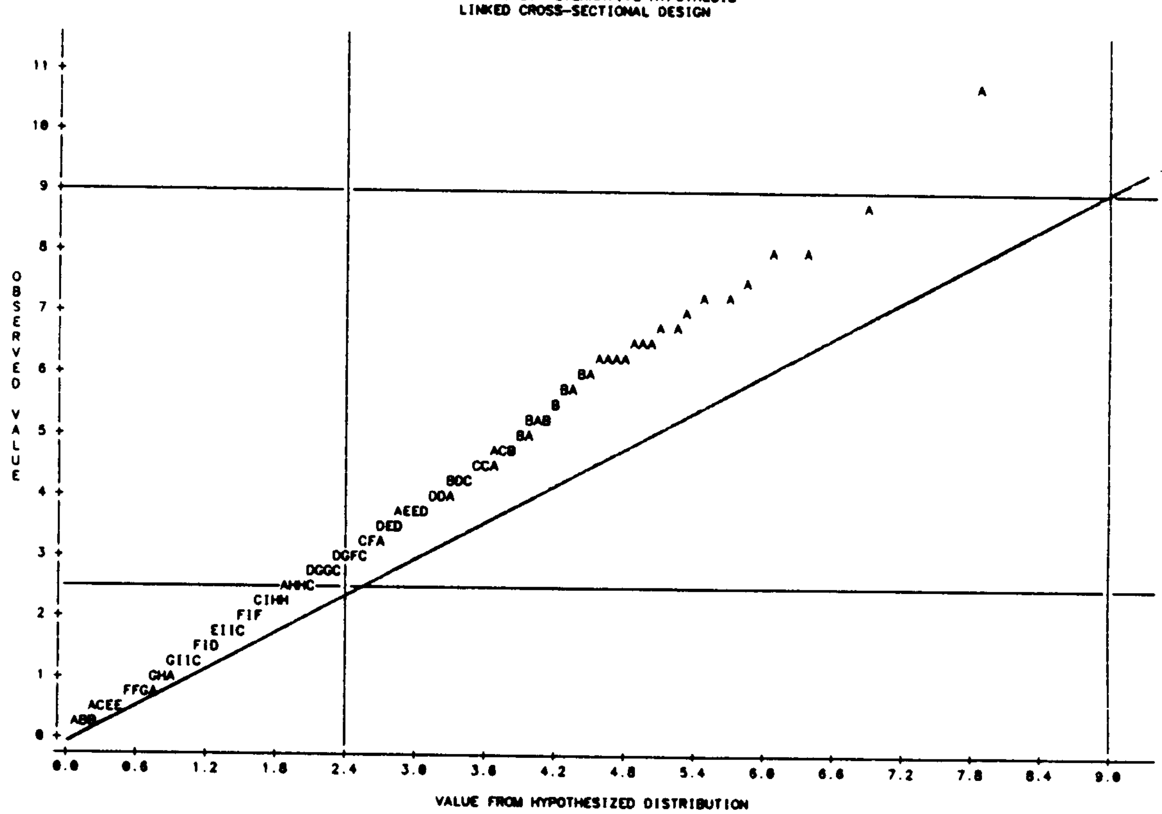

FIGURE 4.4.7
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

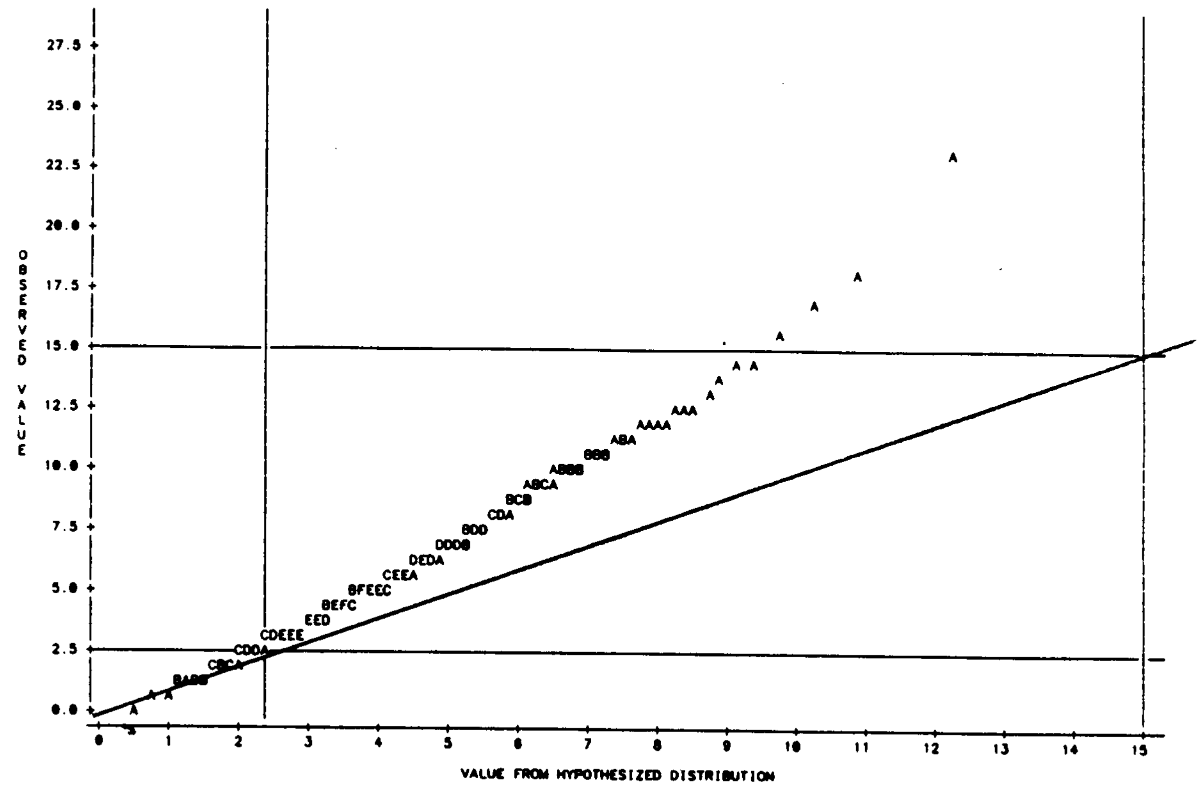

FIGURE 4.4.8
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

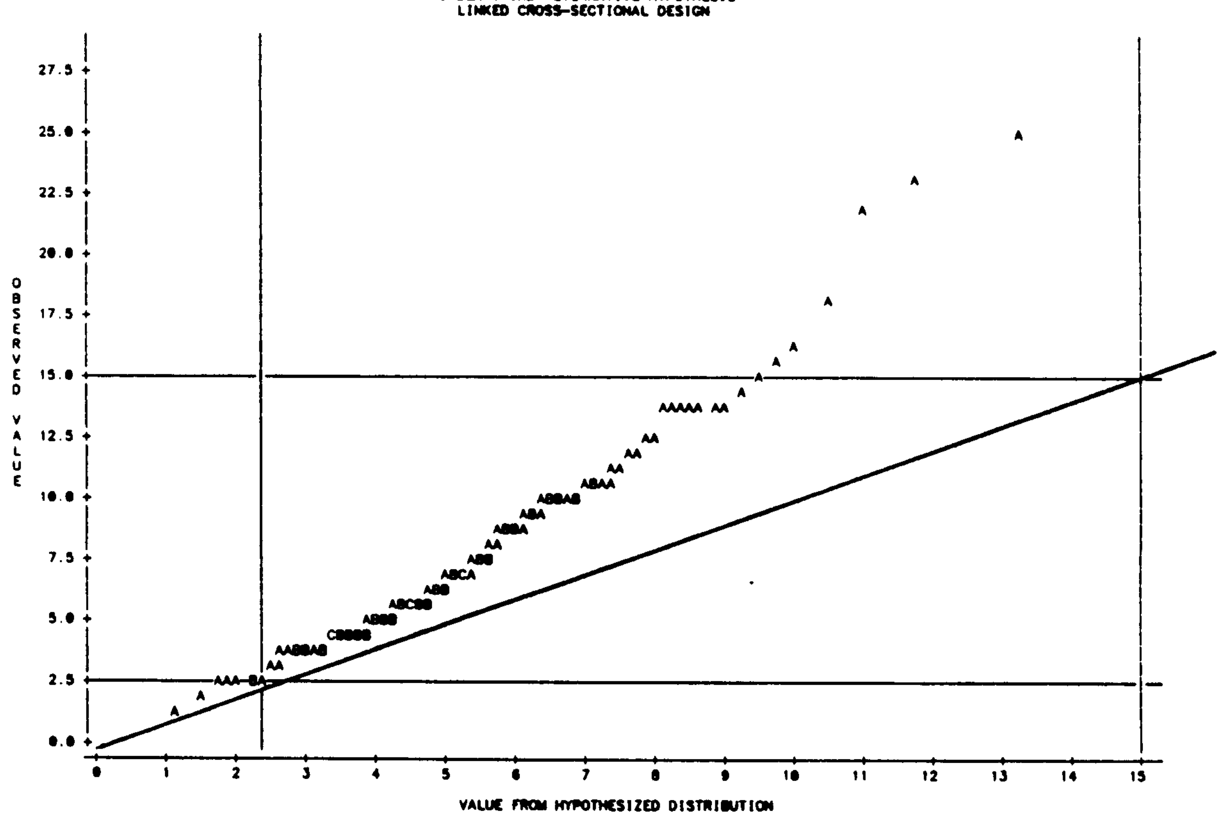

FIGURE 4.4.9
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

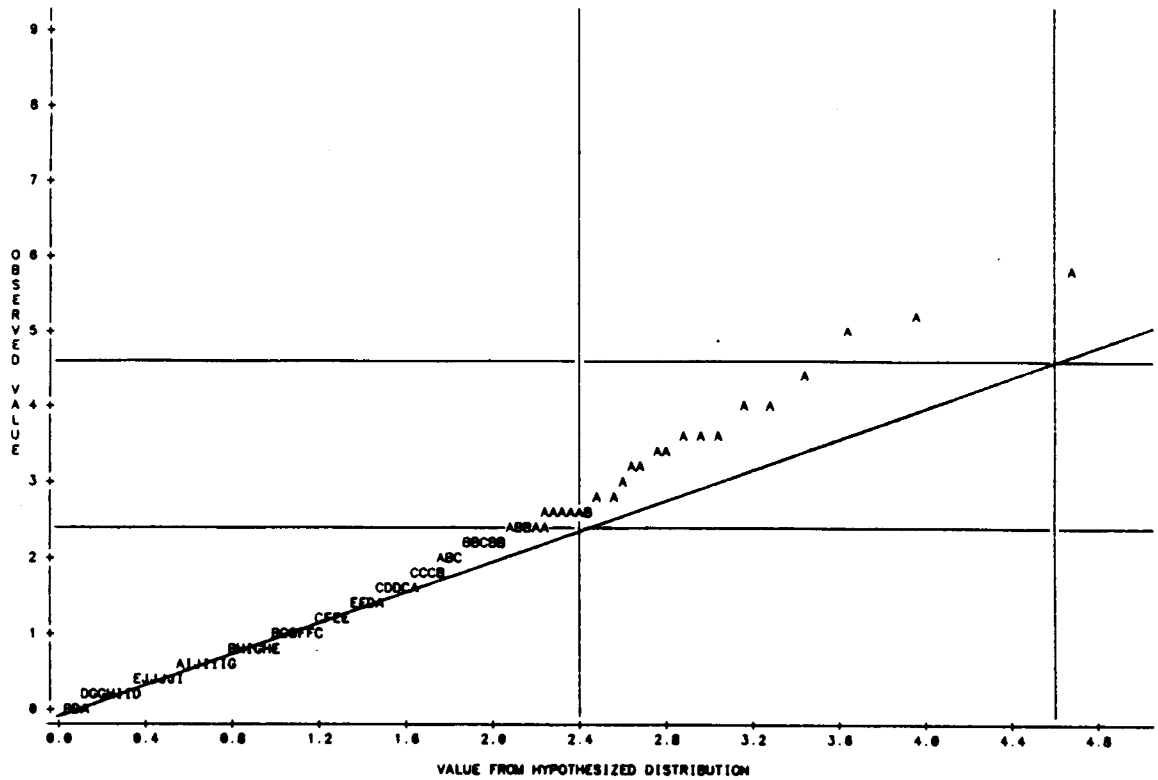

FIGURE 4.4.10
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

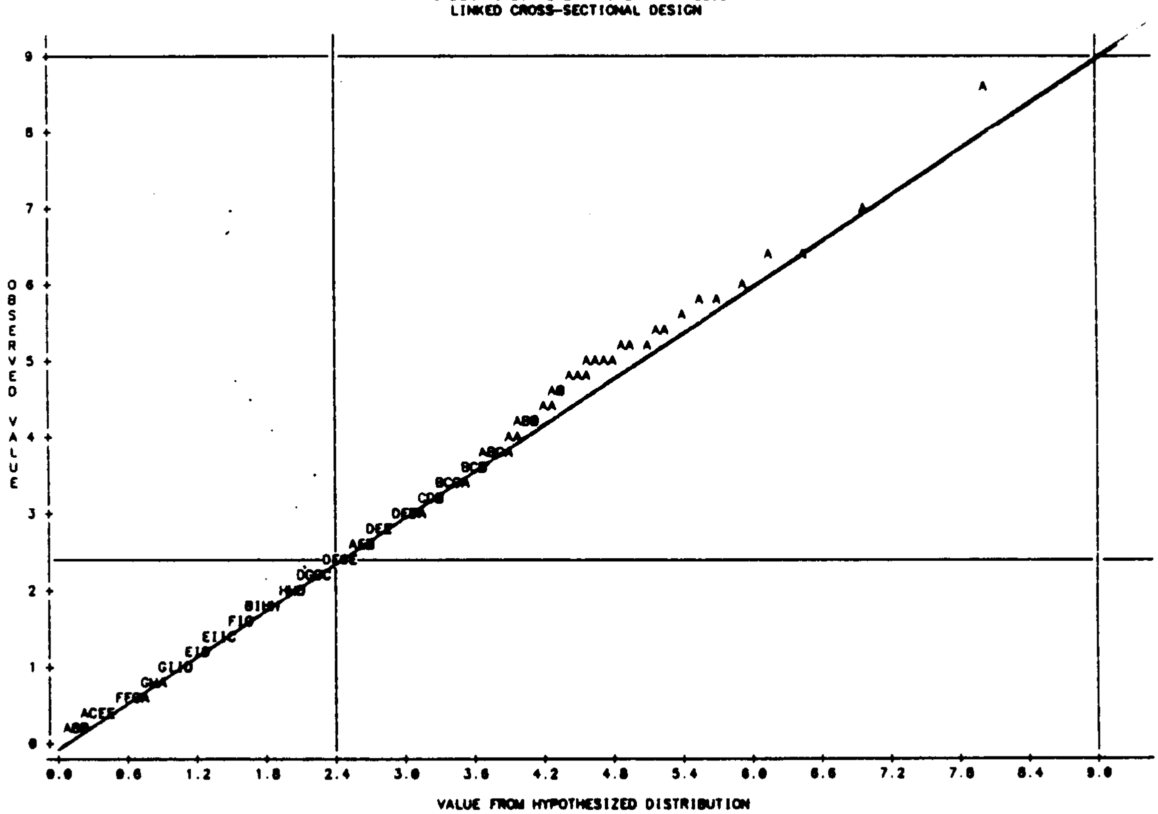

FIGURE 4.4.11
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

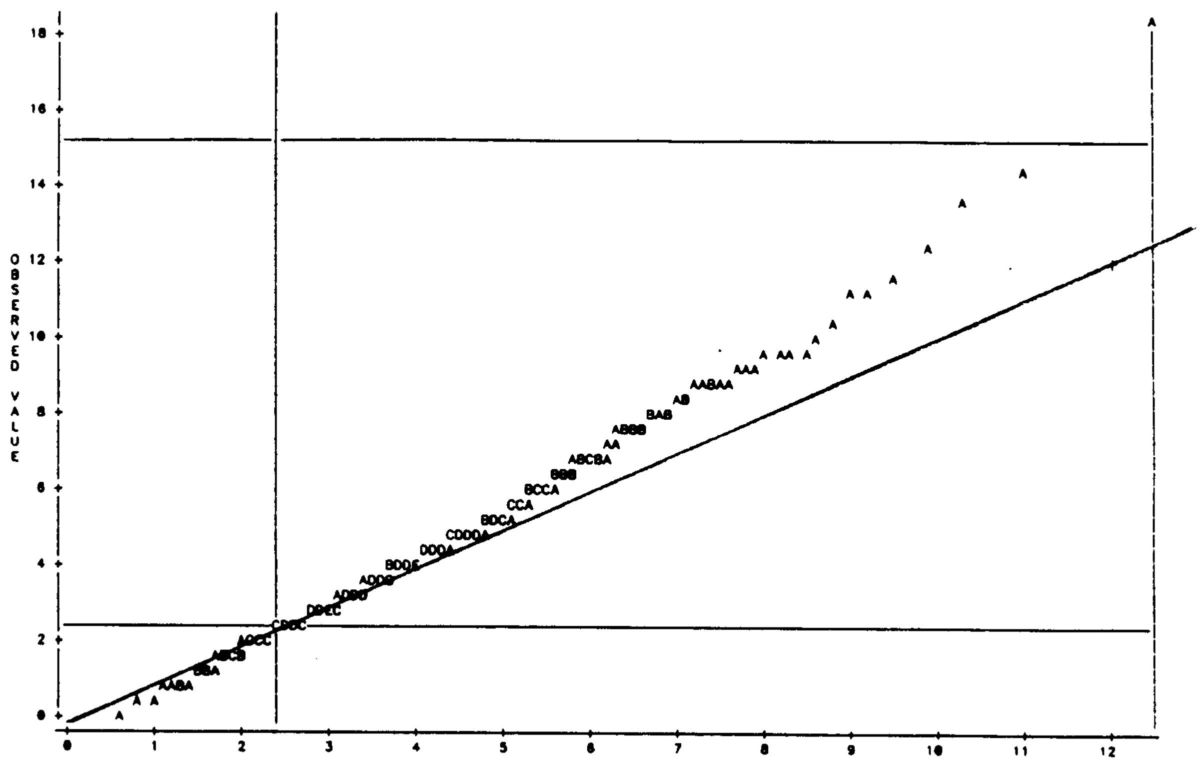

FIGURE 4.4.12
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

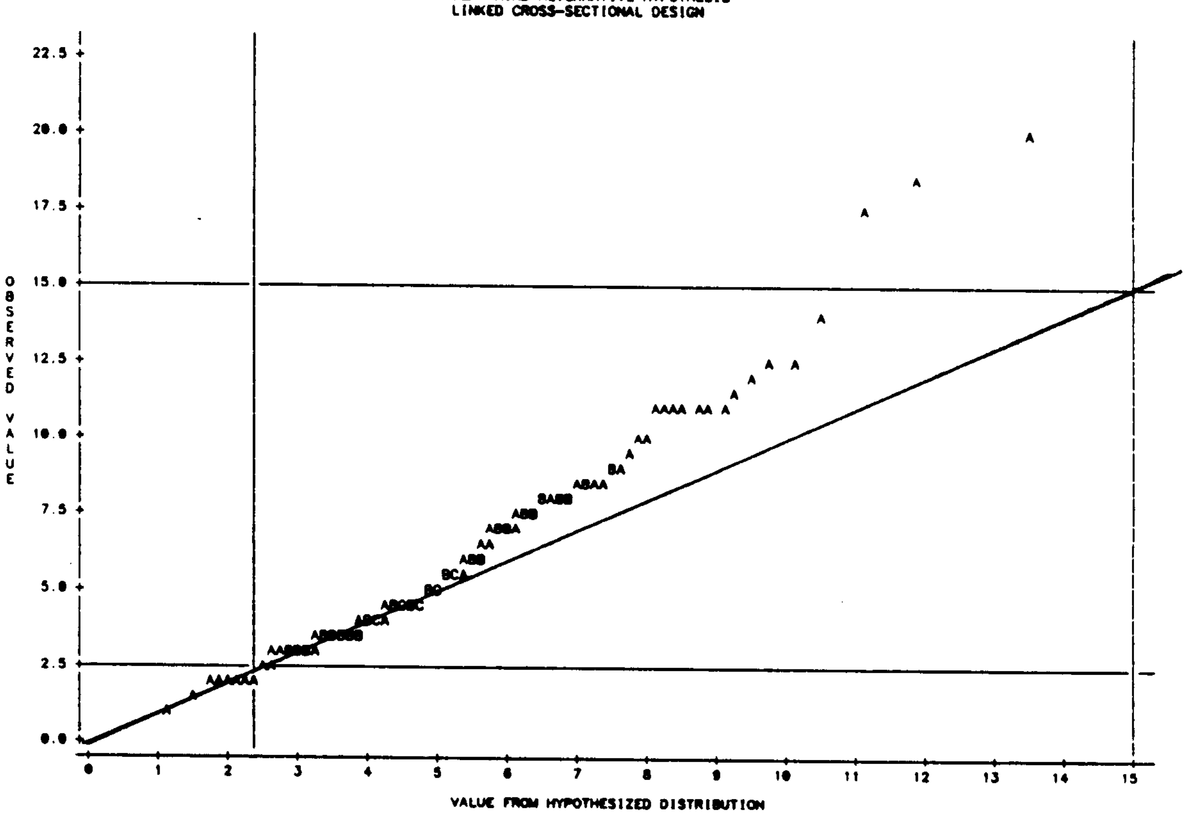

FIGURE 4.4.13
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

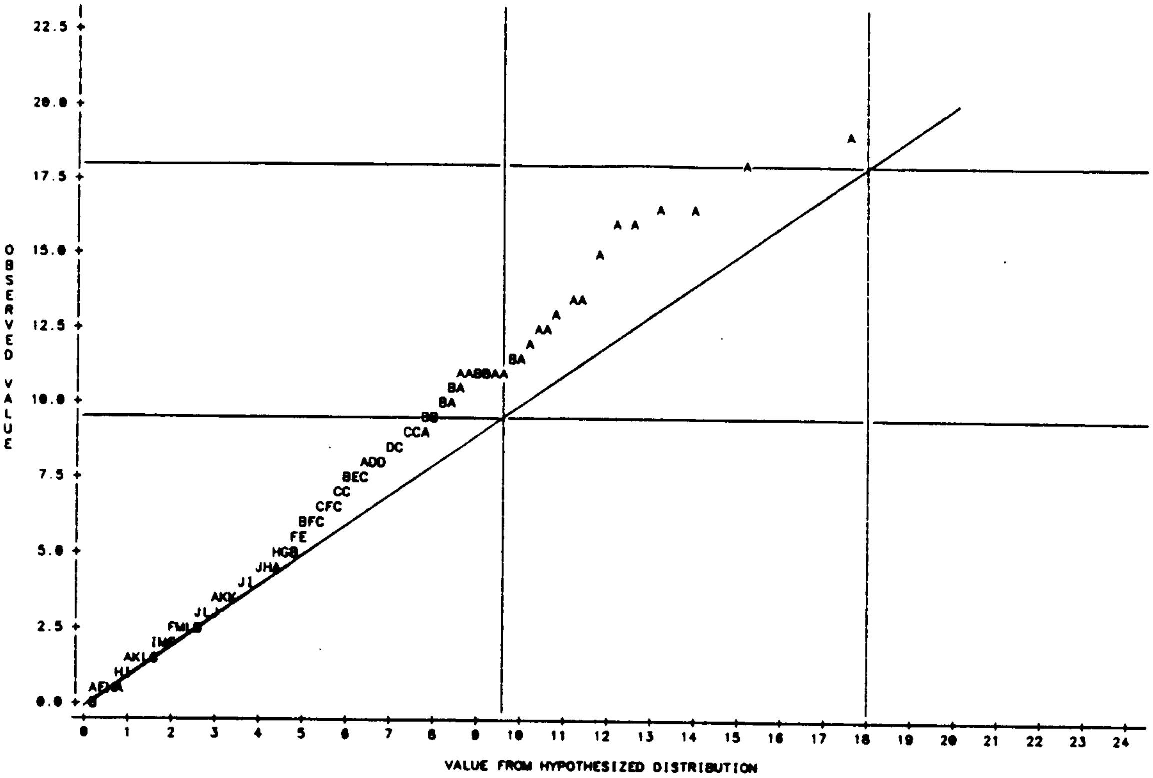

FIGURE 4.4.14
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

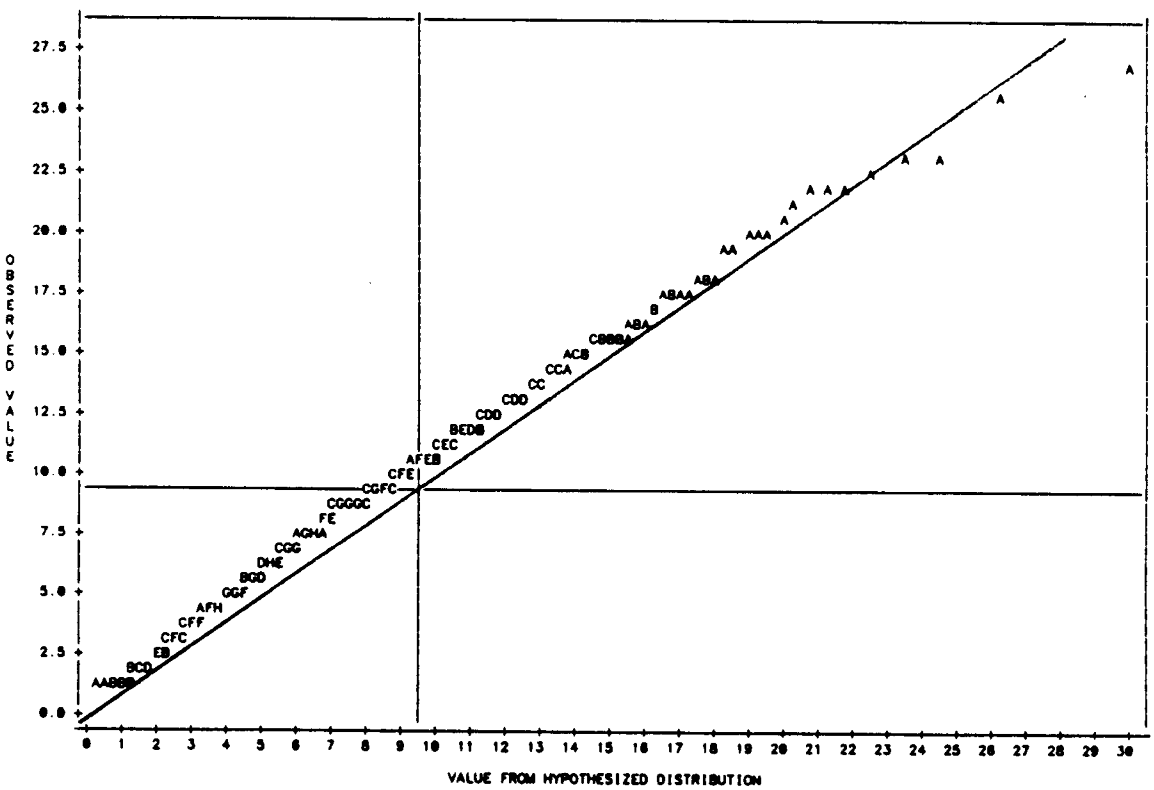

FIGURE 4.4.15
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

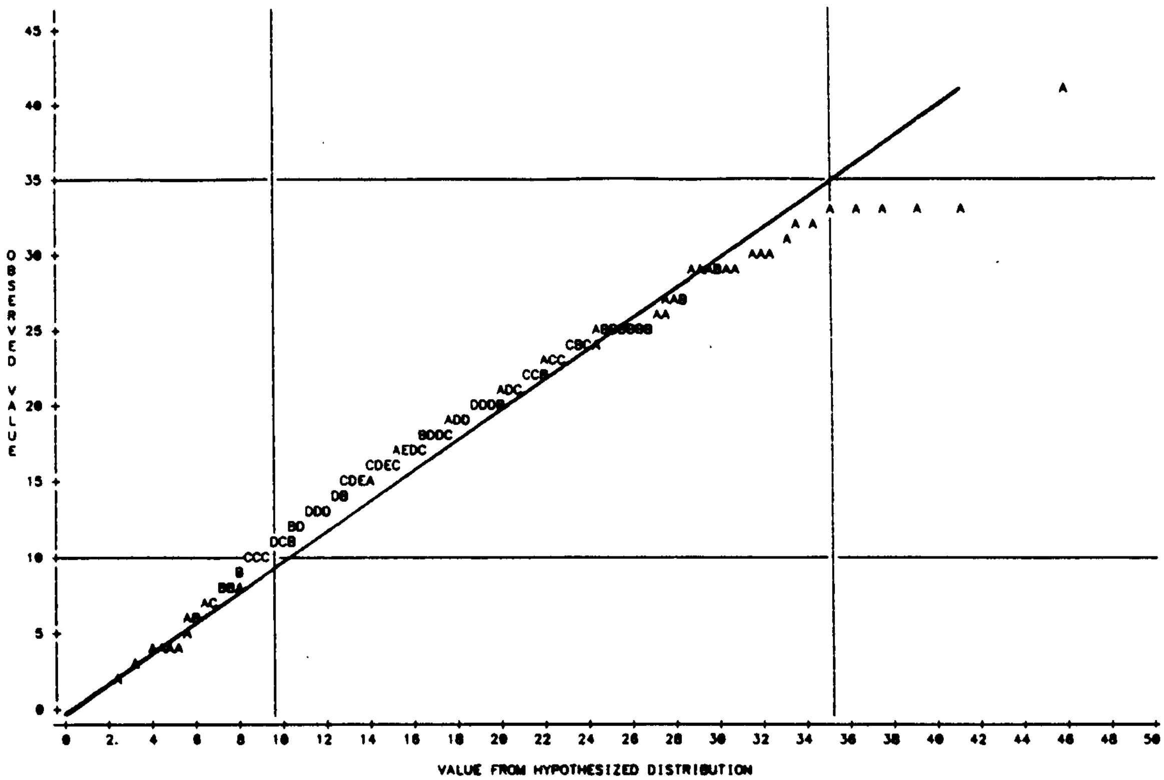

FIGURE 4.4.16
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINLED CROSS-SECTIONAL DESIGN

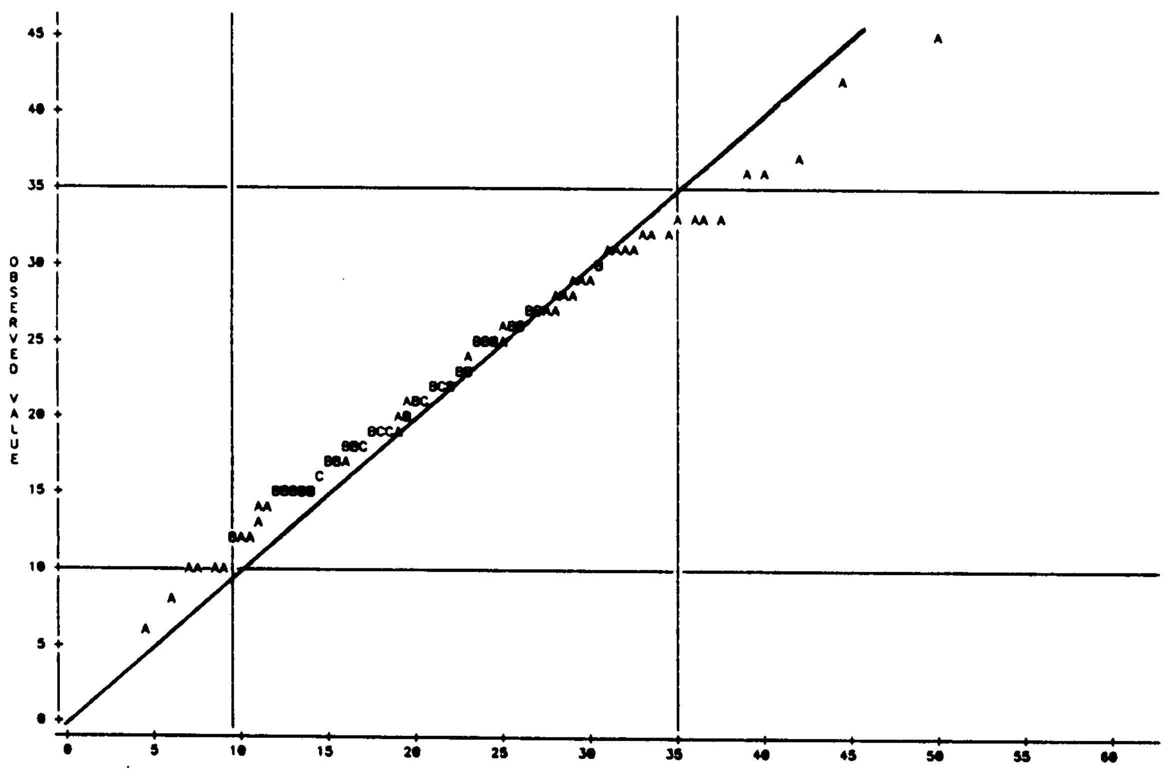

FIGURE 4.4.17
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

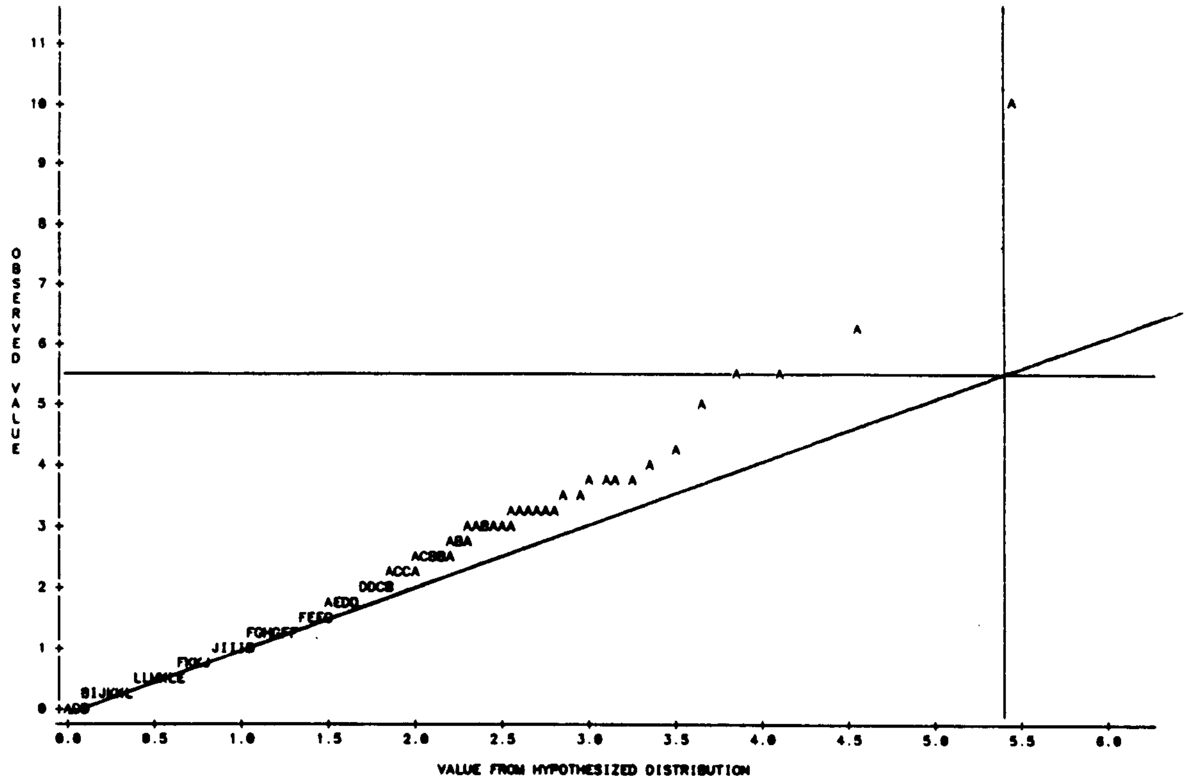

FIGURE 4.4.18
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

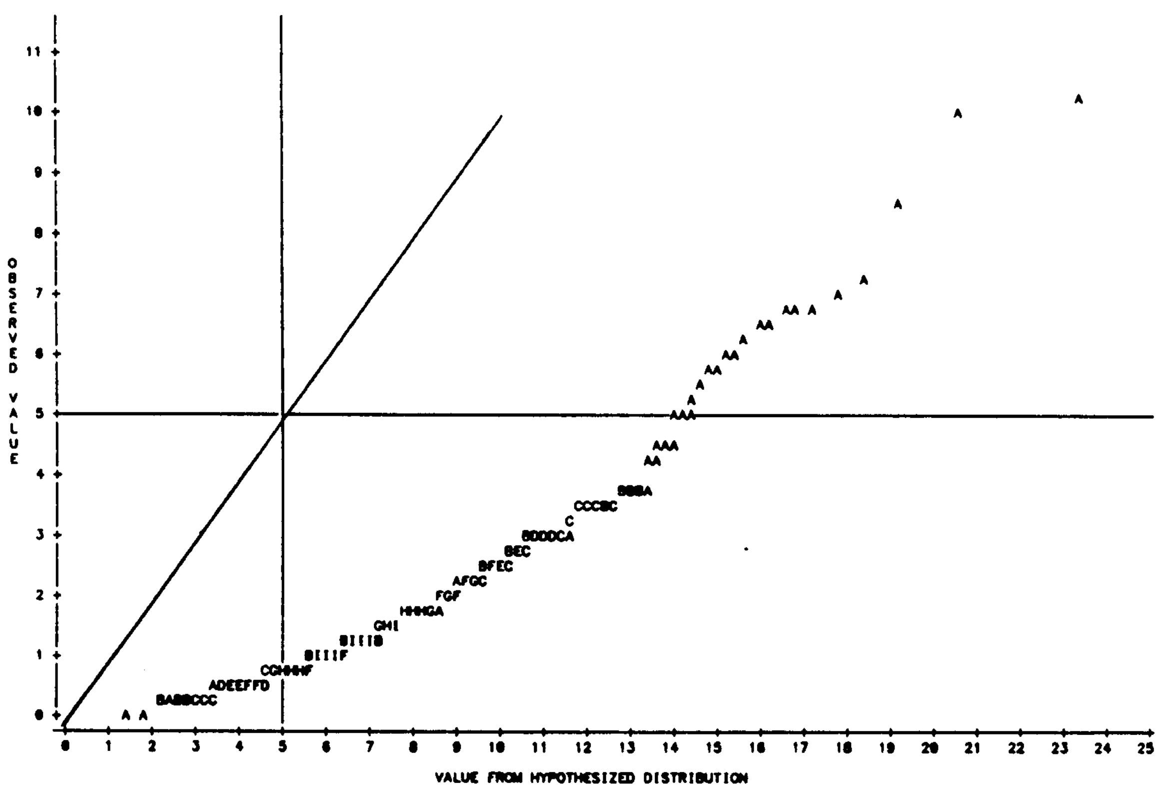

FIGURE 4.4.19
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

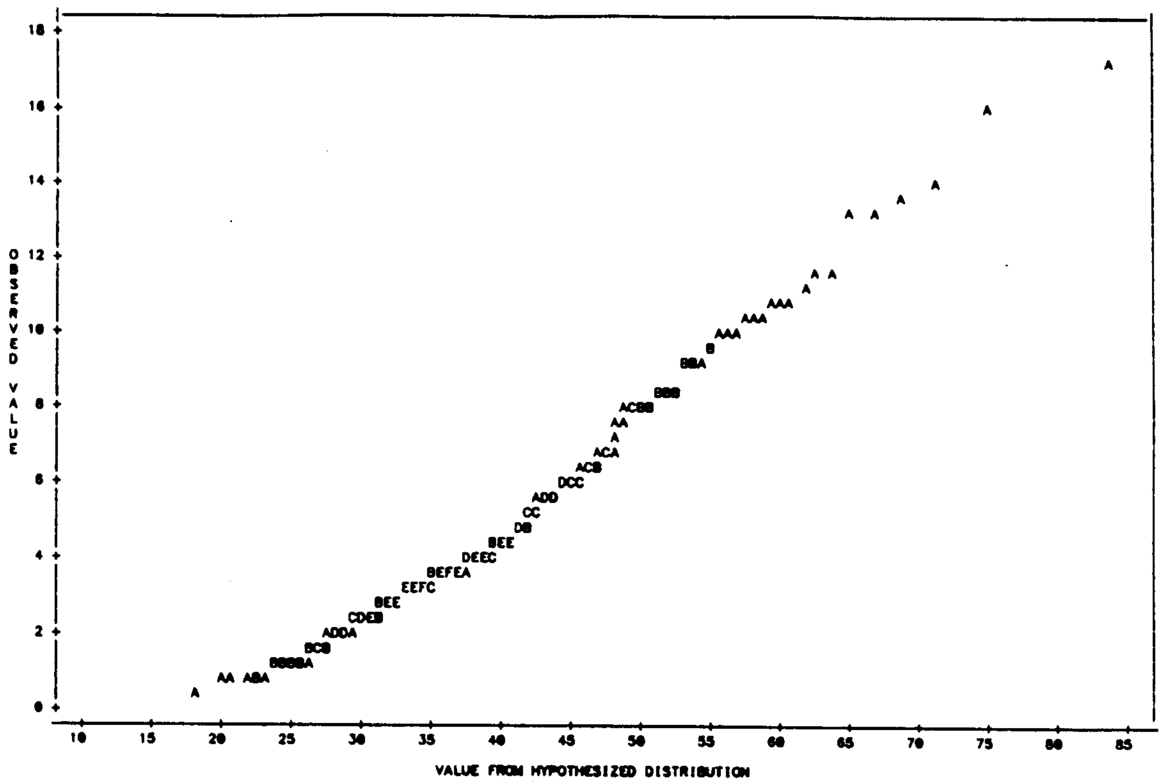

FIGURE 4.4.20
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN

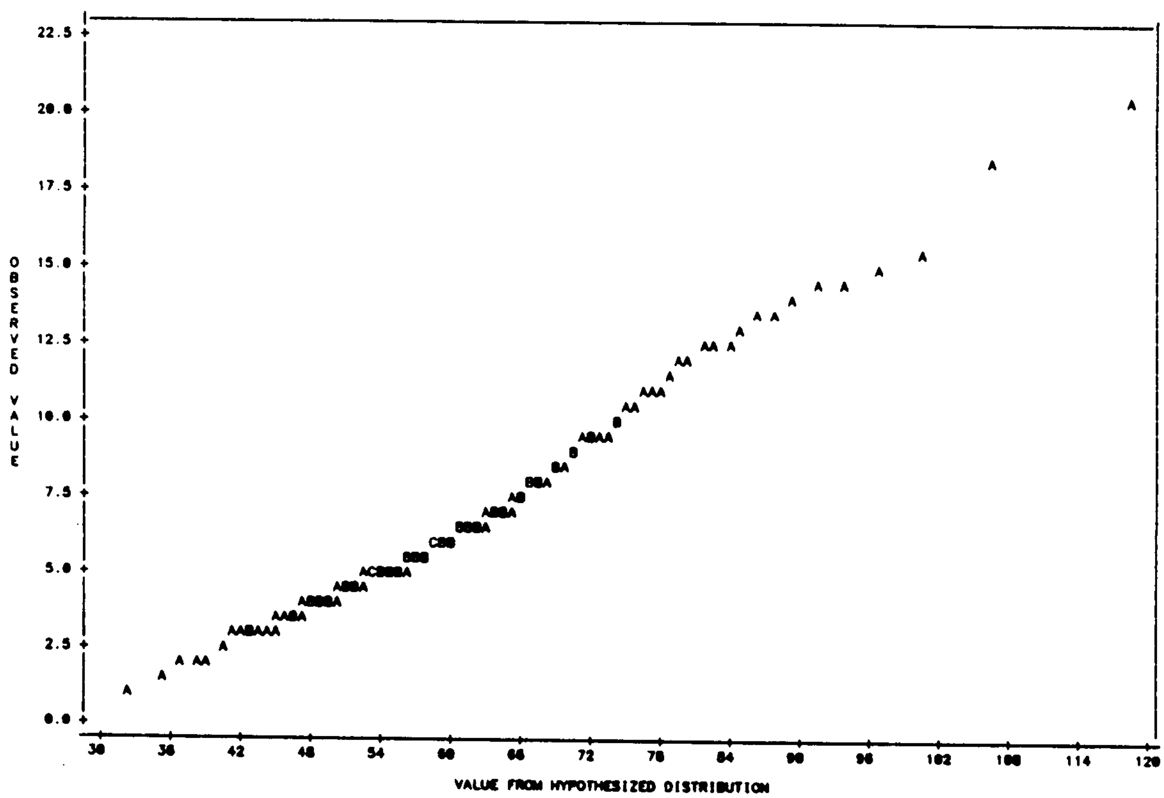

FIGURE 4.5.2
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

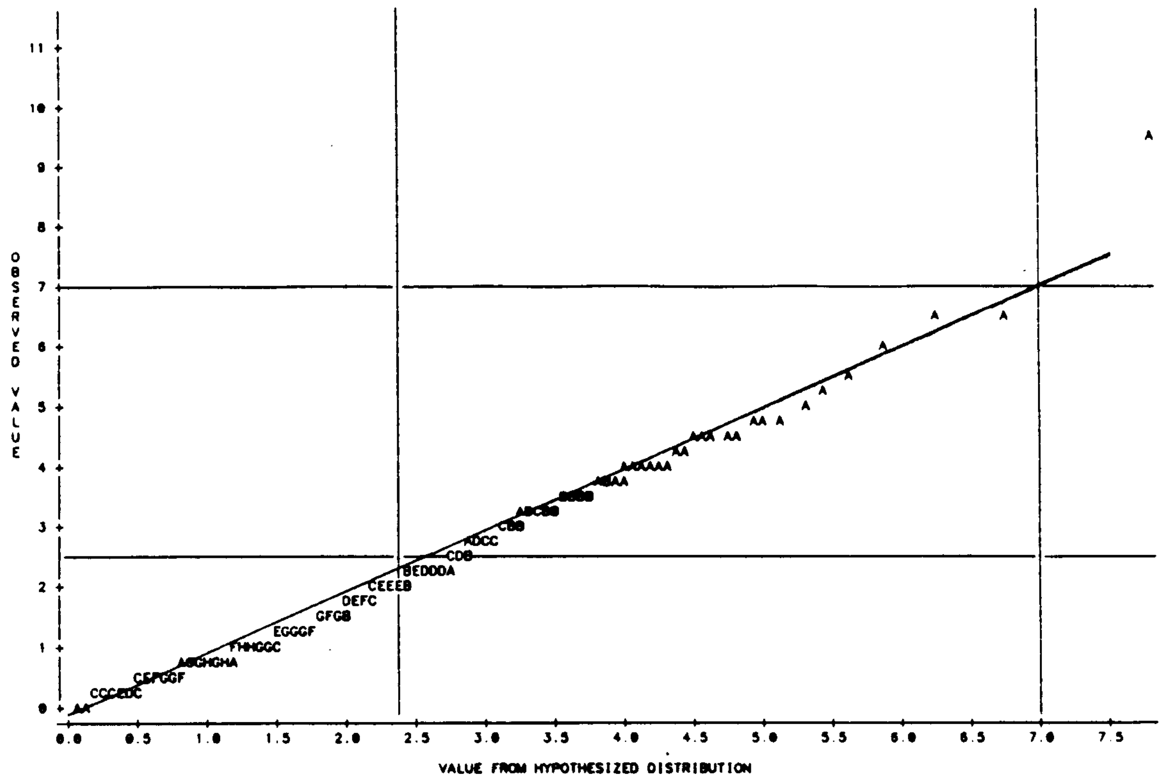

FIGURE 4.5.1
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

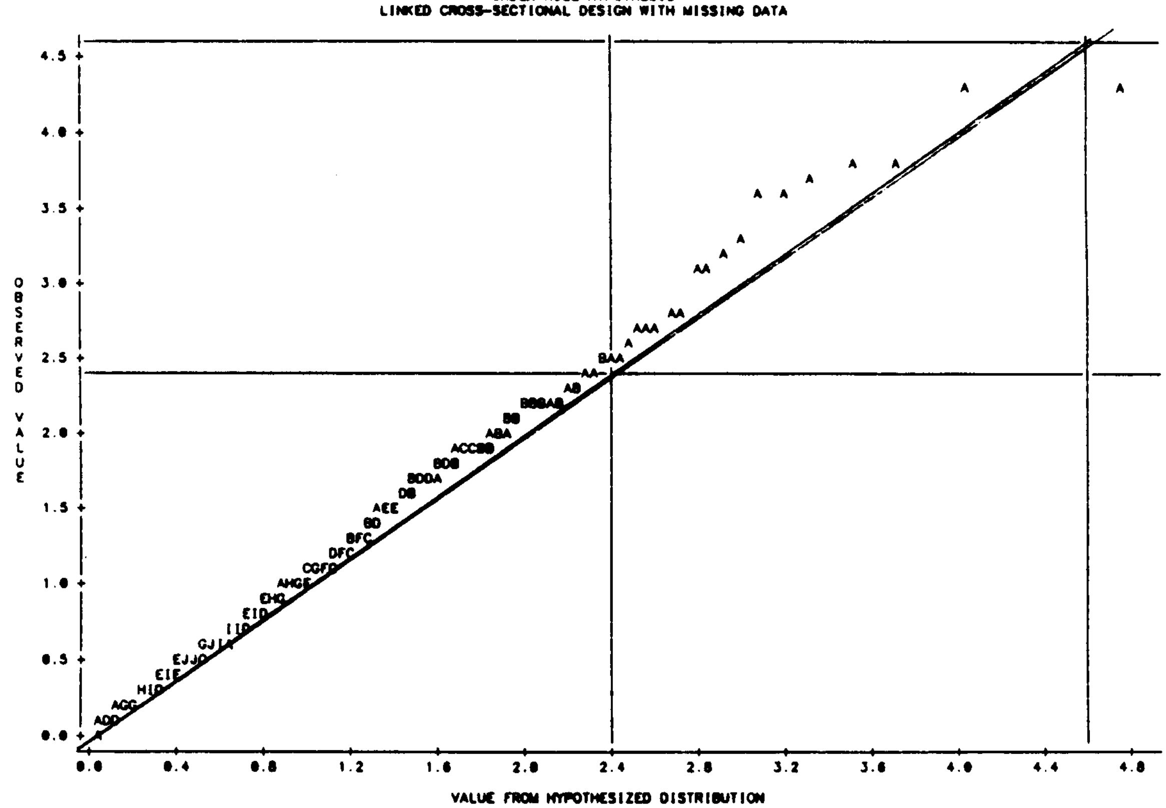

FIGURE 4.5.3
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

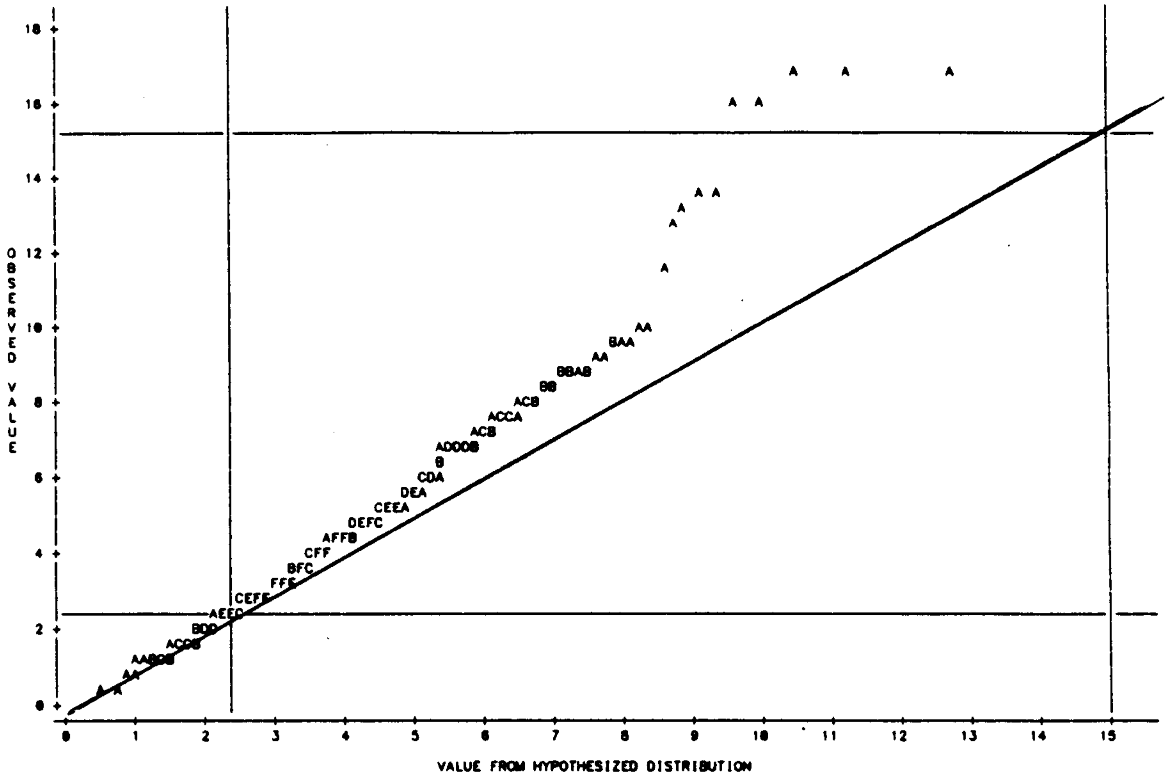

FIGURE 4.5.4
F-PLOT FOR THE DISTRIBUTION OF FREML STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

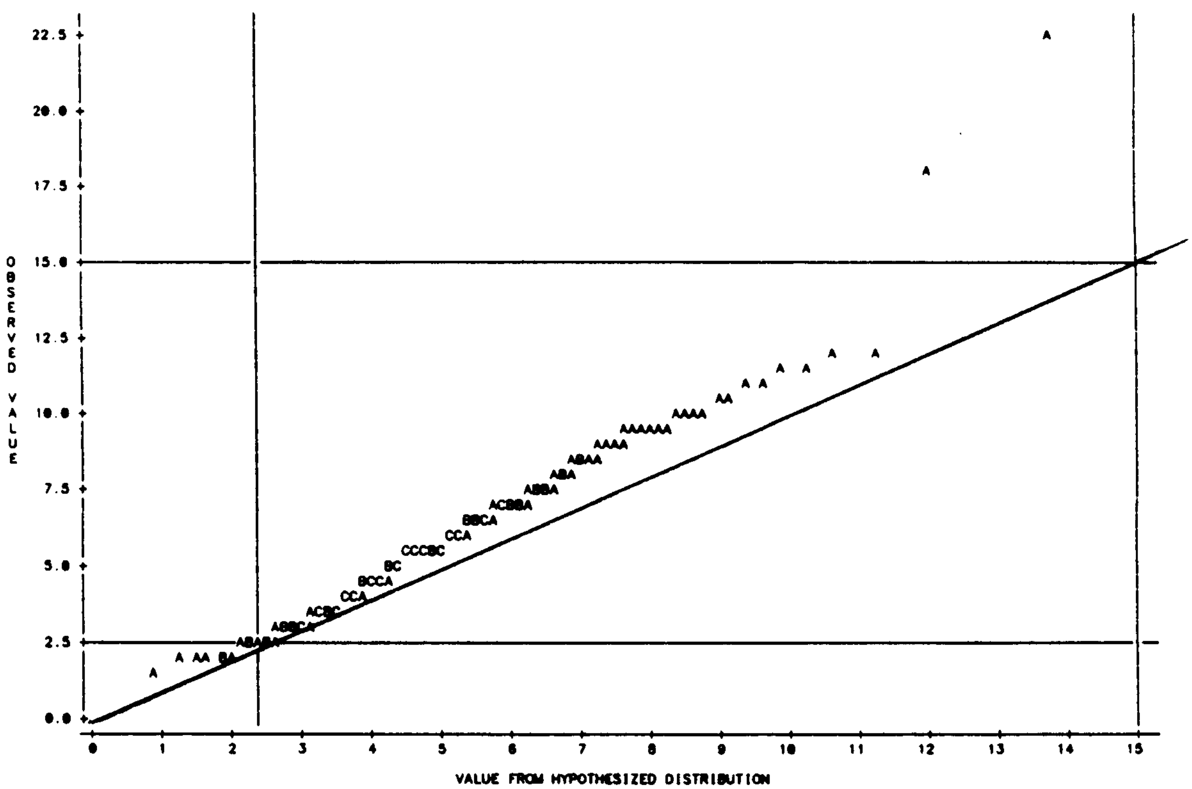

FIGURE 4.5.5
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

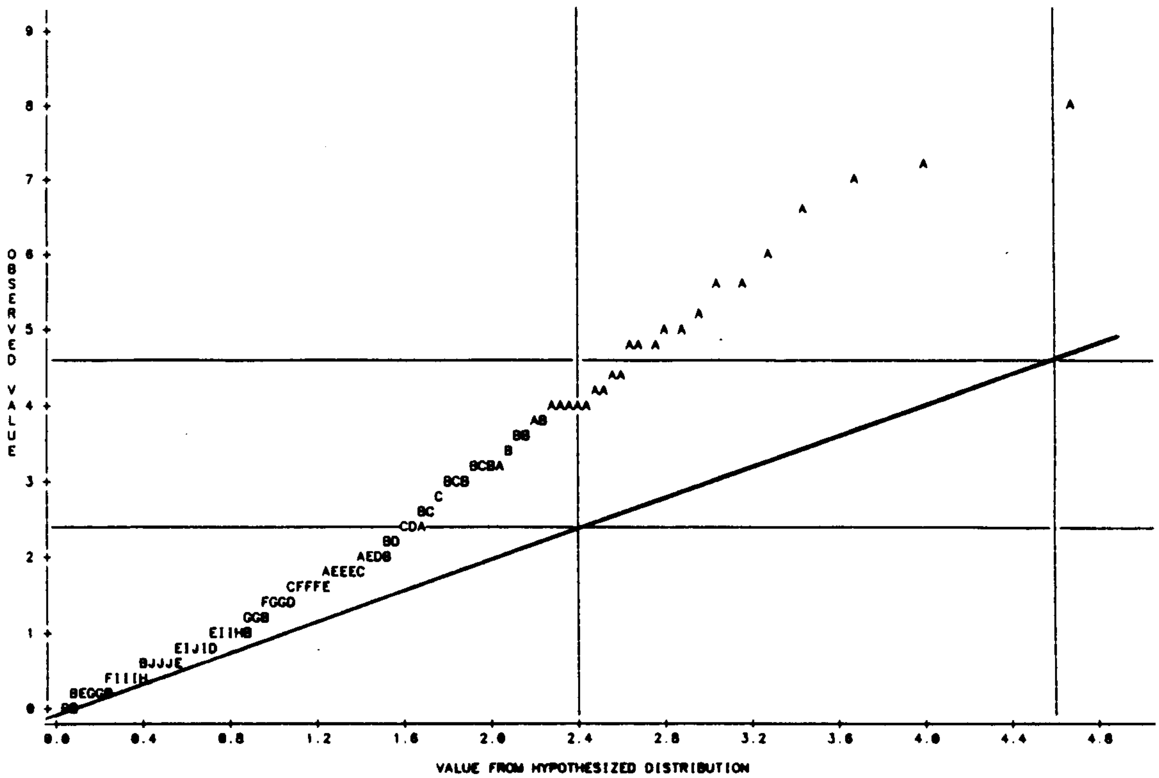

FIGURE 4.5.6
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

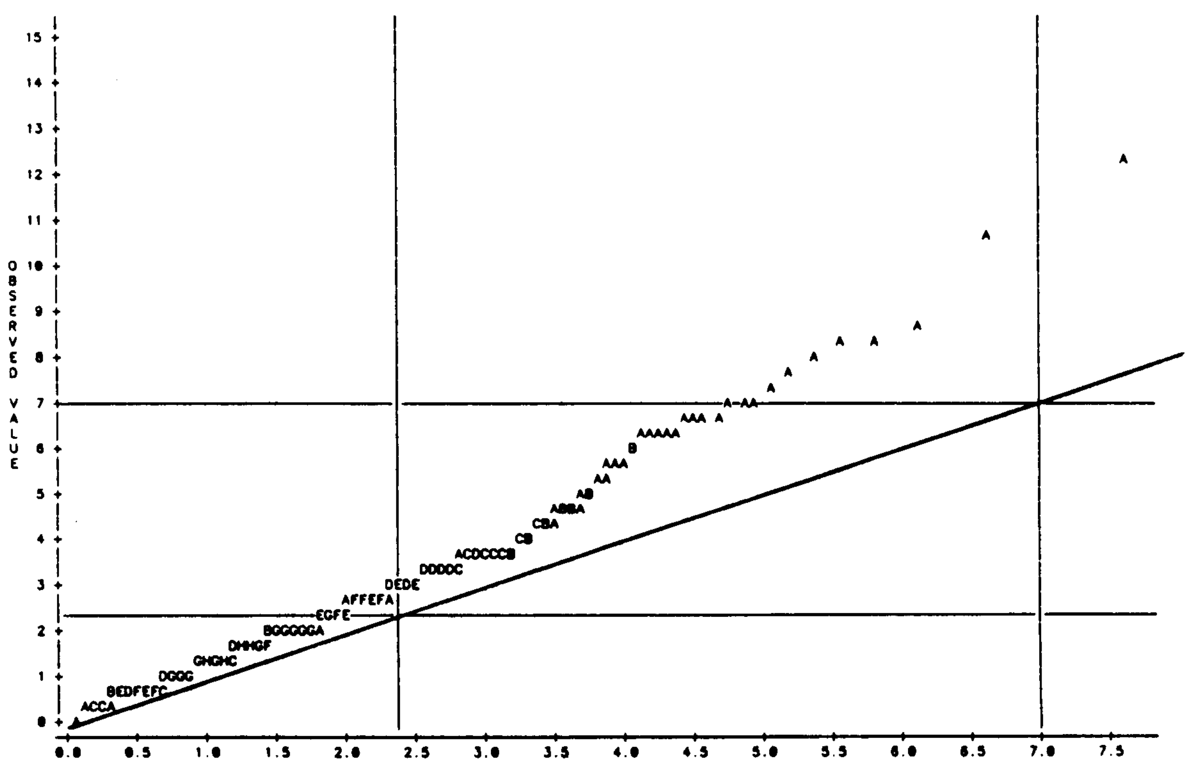

FIGURE 4.5.7
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

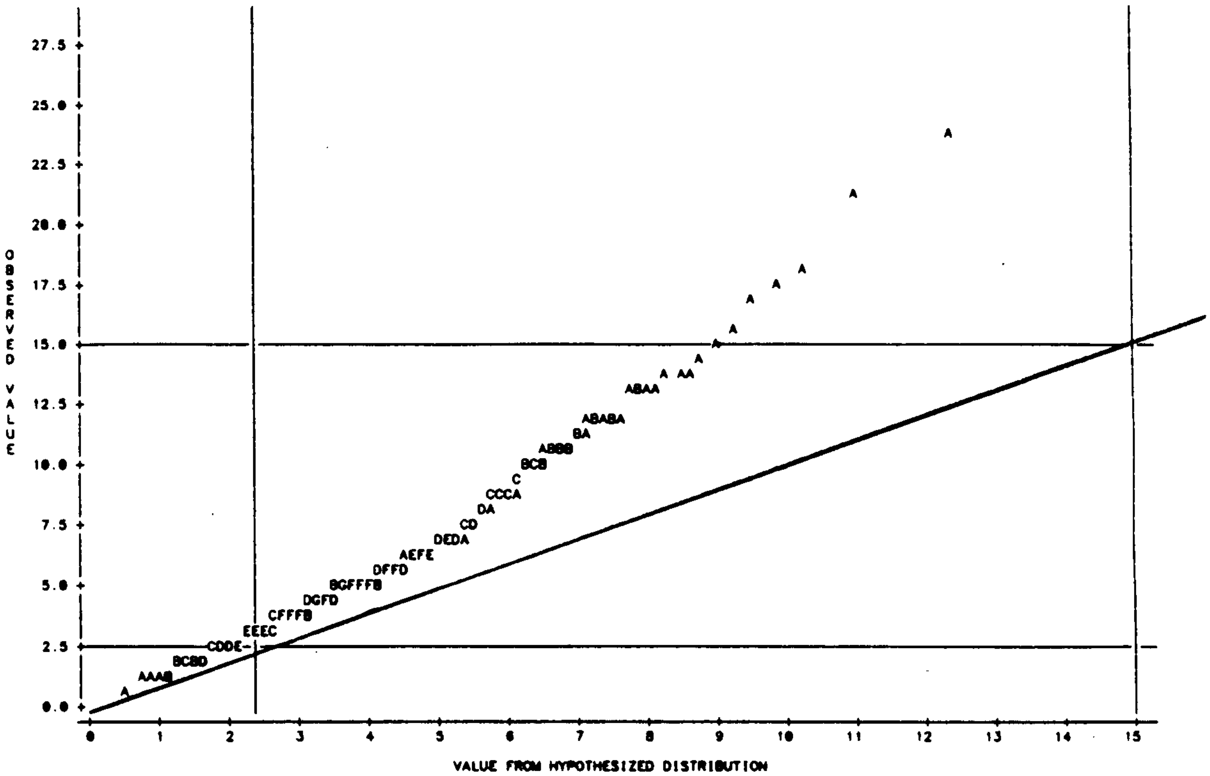

FIGURE 4.5.8
F-PLOT FOR THE DISTRIBUTION OF FWLS STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

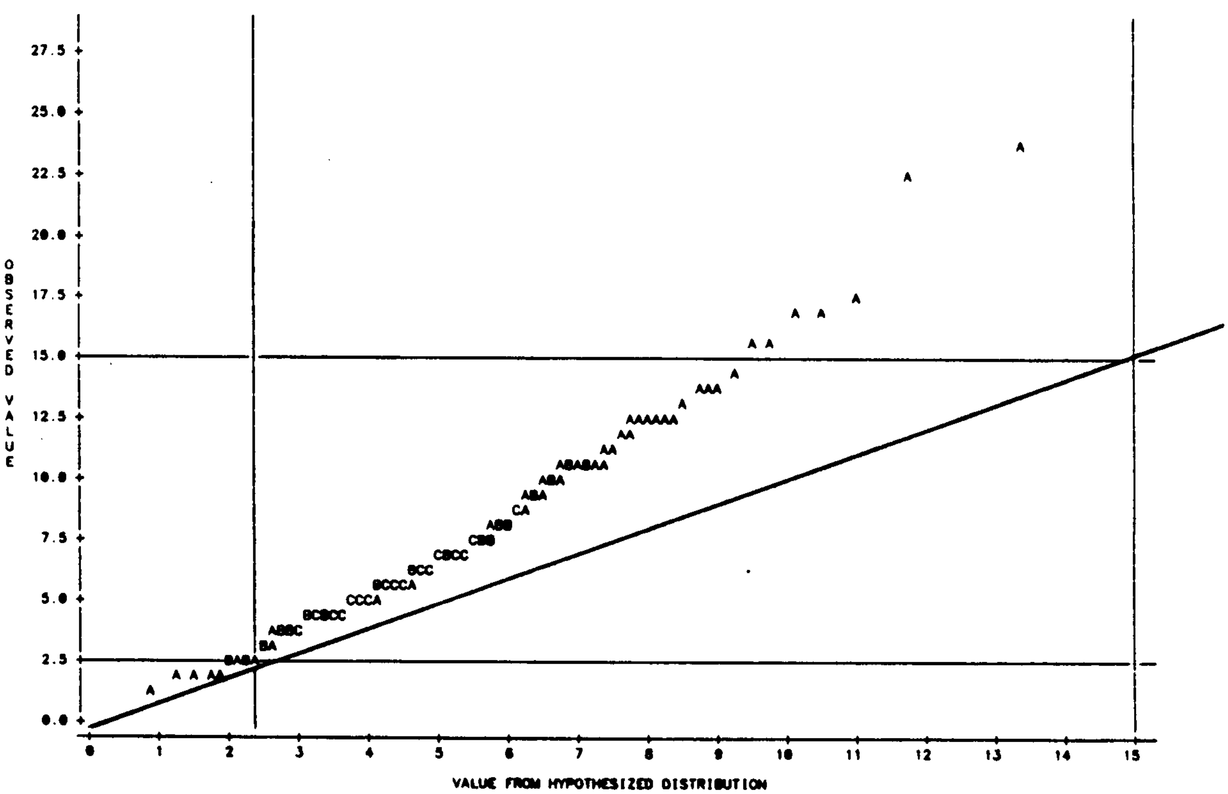

FIGURE 4.5.9
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

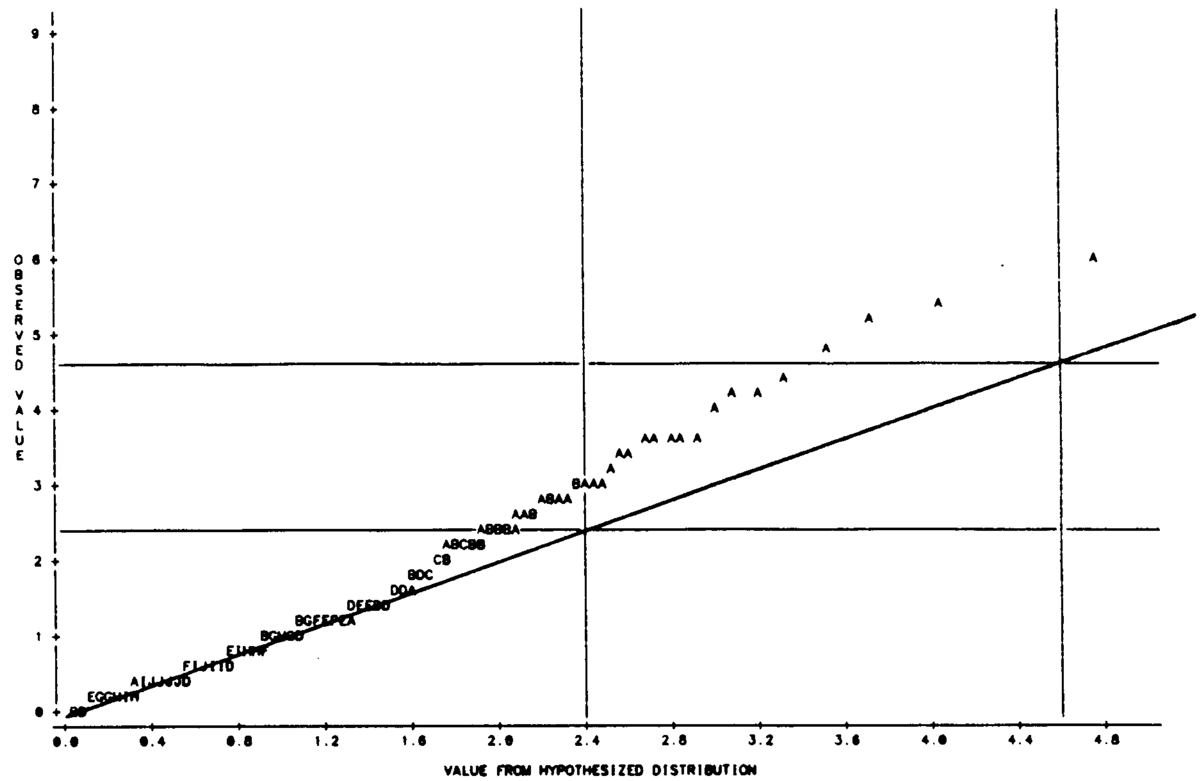

FIGURE 4.5.10
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

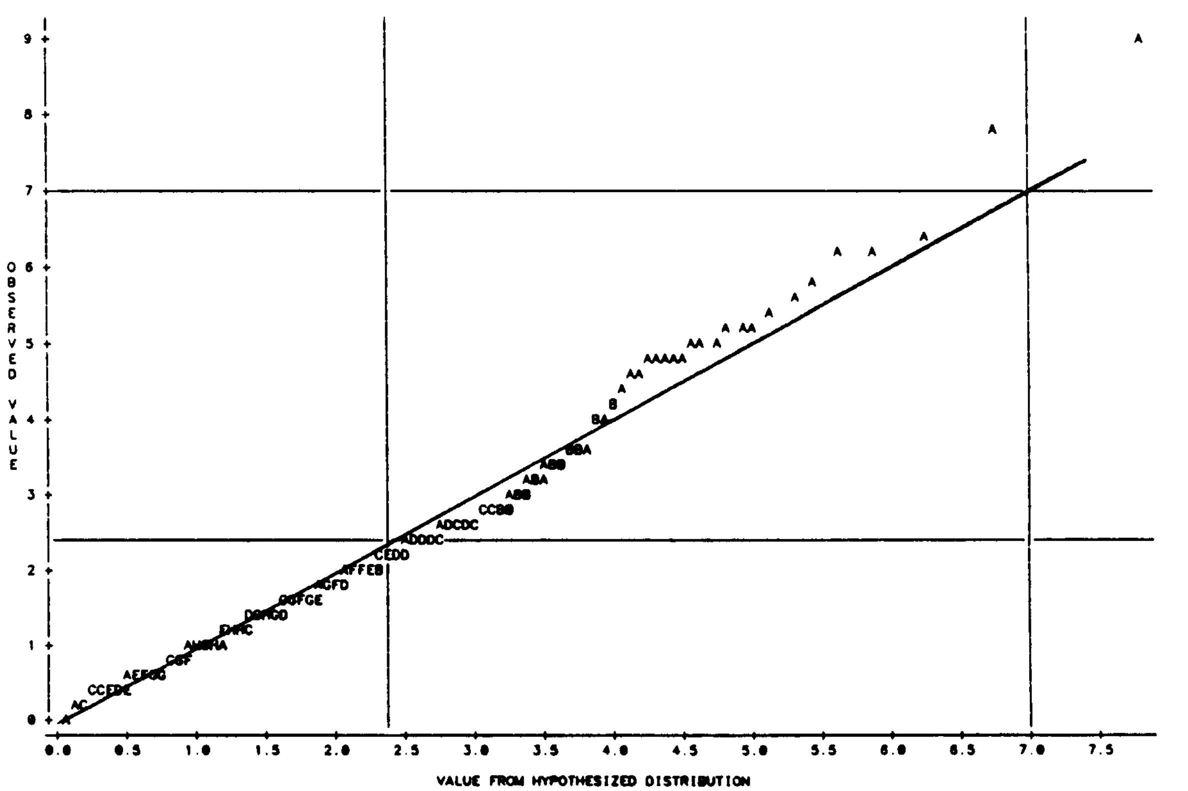

FIGURE 4.5.11
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

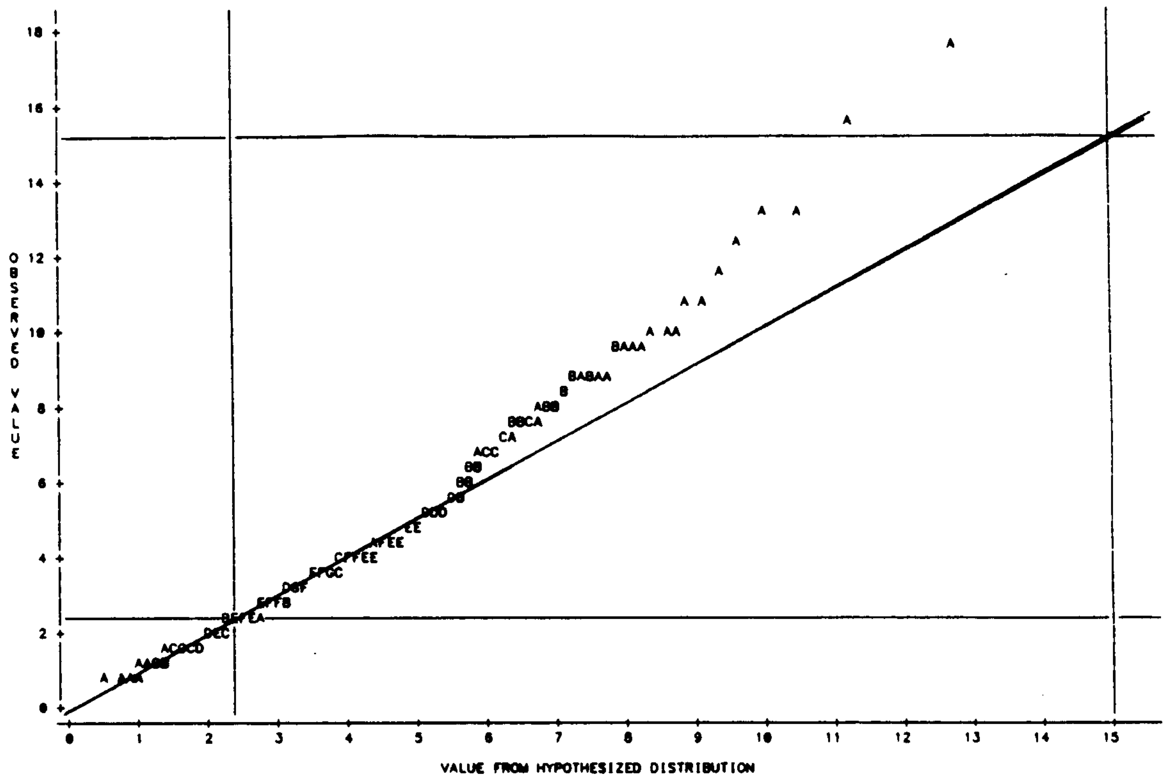

FIGURE 4.5.12
F-PLOT FOR THE DISTRIBUTION OF FWLS2 STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

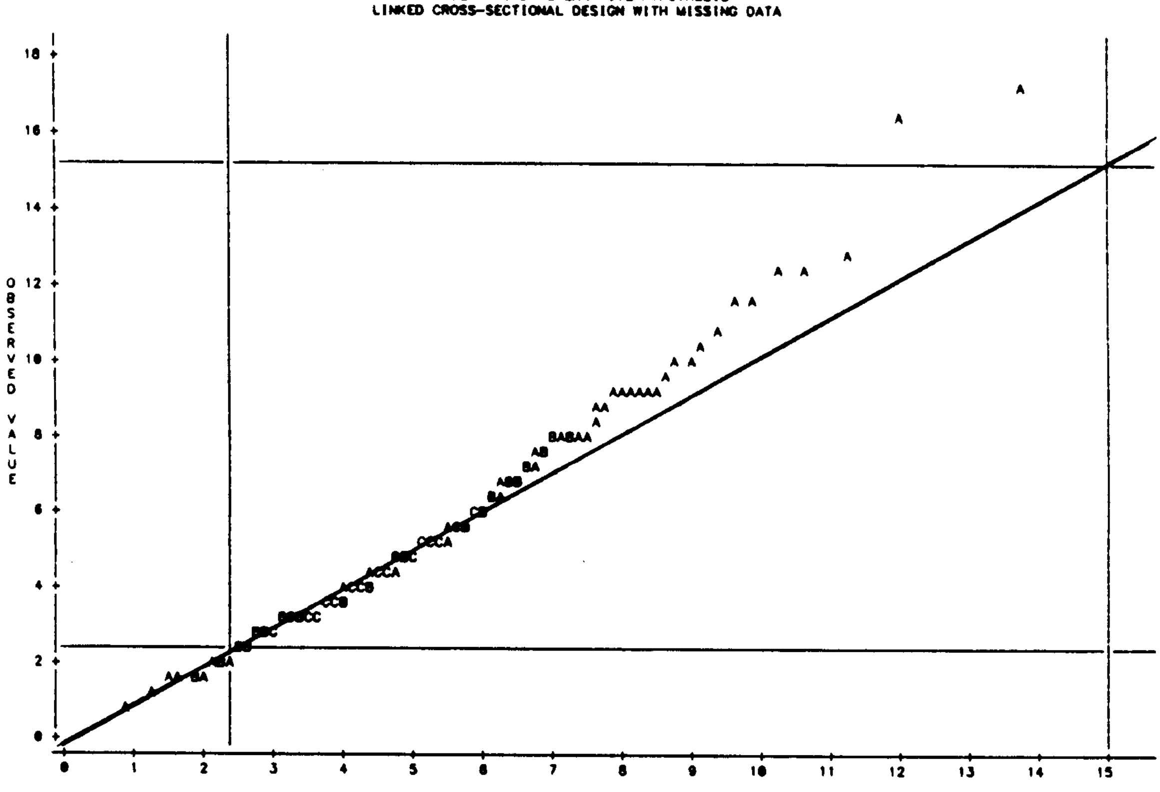

FIGURE 4.5.13
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

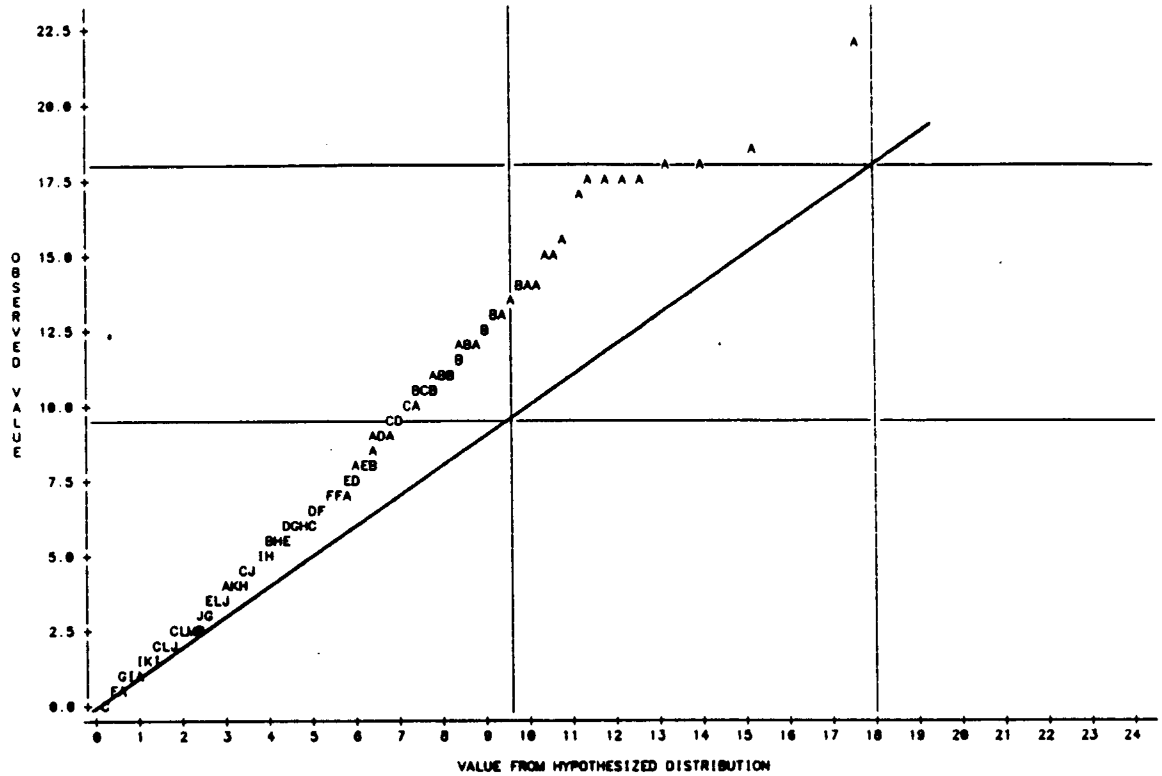

FIGURE 4.5.14
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

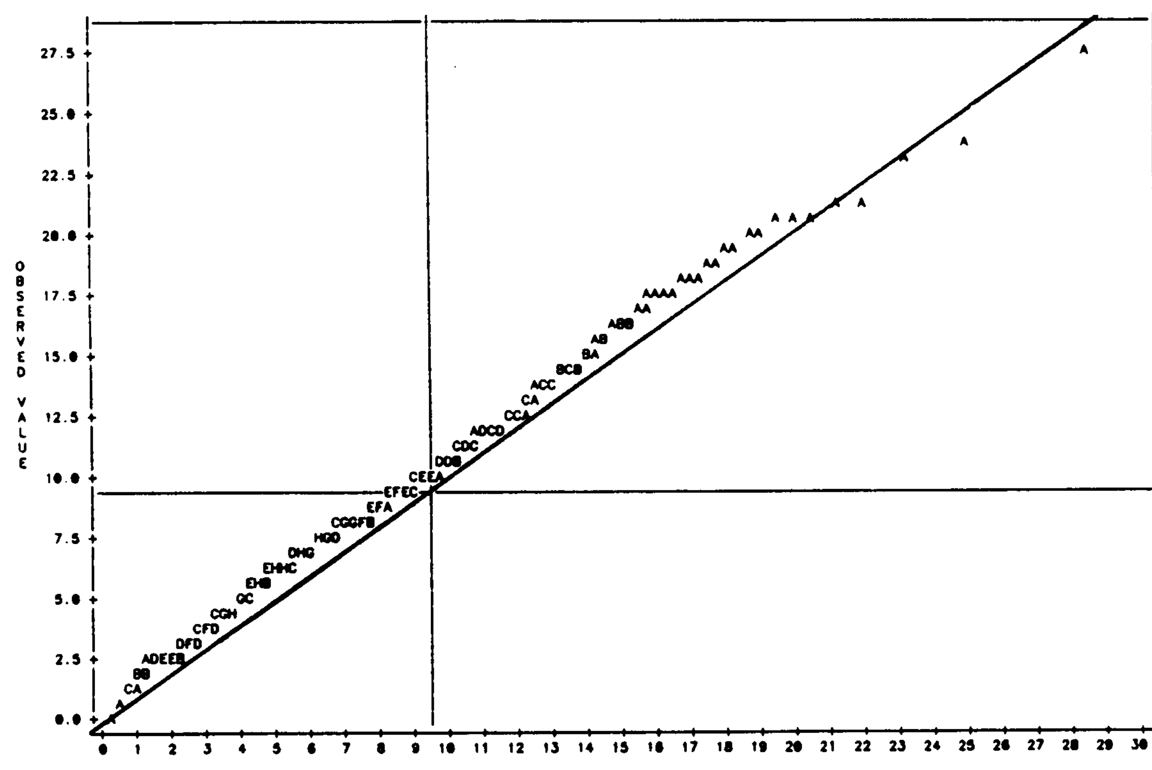

FIGURE 4.5.15
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

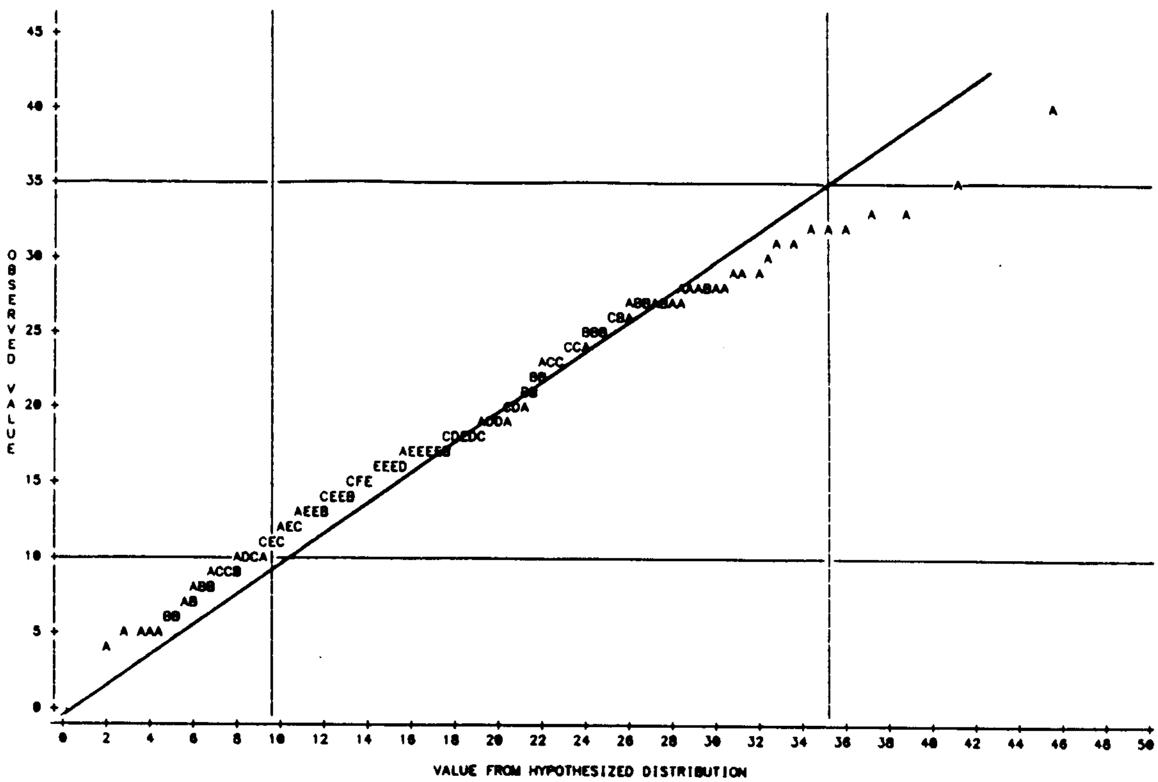

FIGURE 4.5.16
CHI-SQUARE PLOT FOR THE DISTRIBUTION OF LRT STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

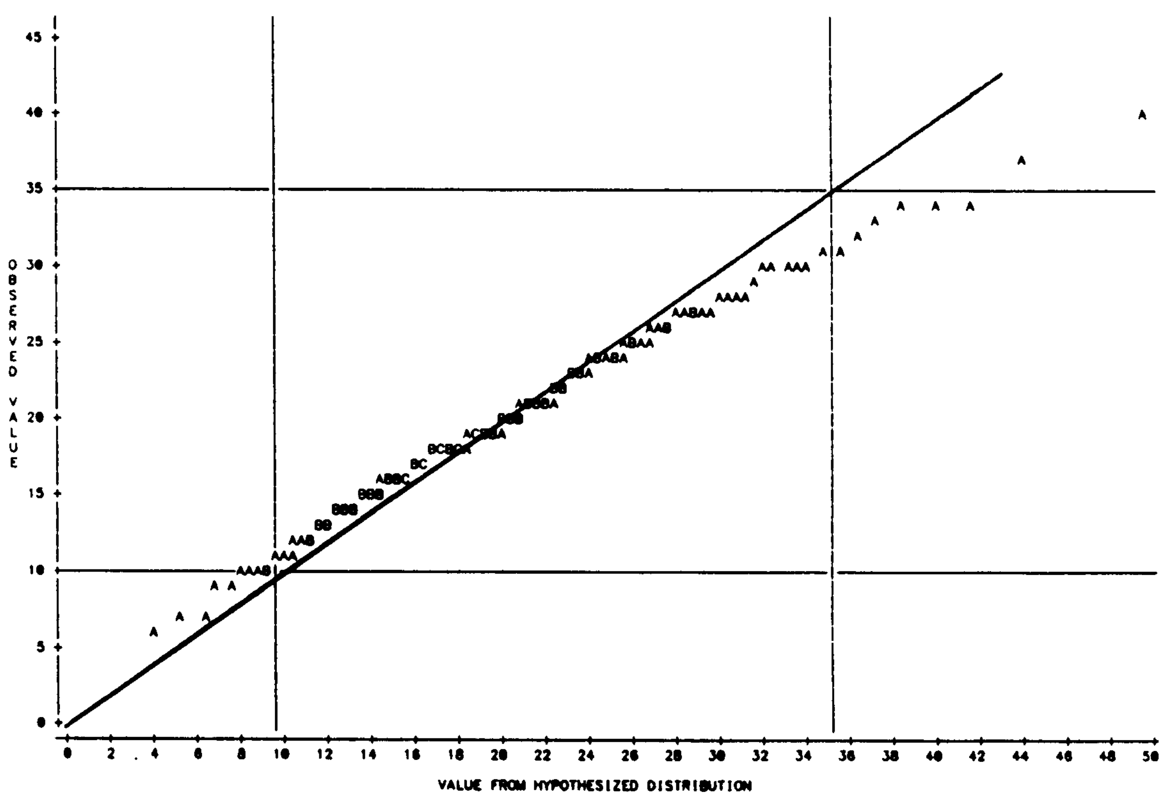

FIGURE 4.5.17
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER NULL HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

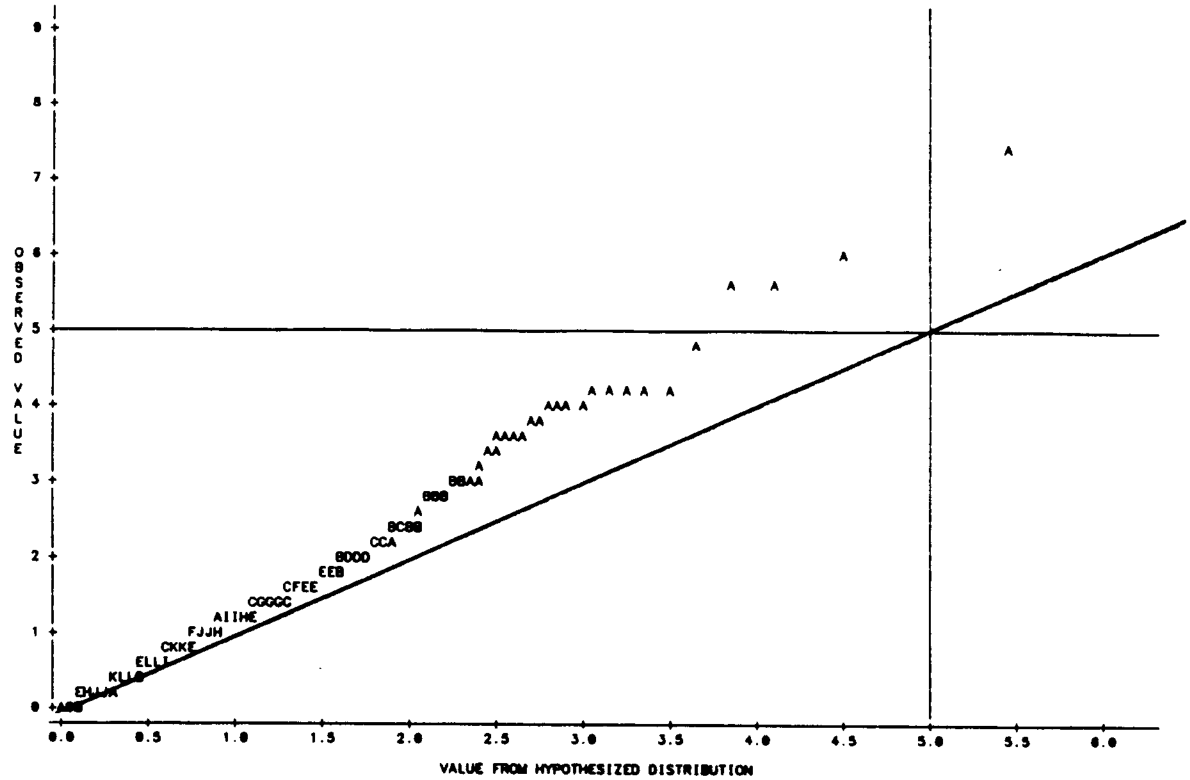

FIGURE 4.5.18
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER FIRST ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

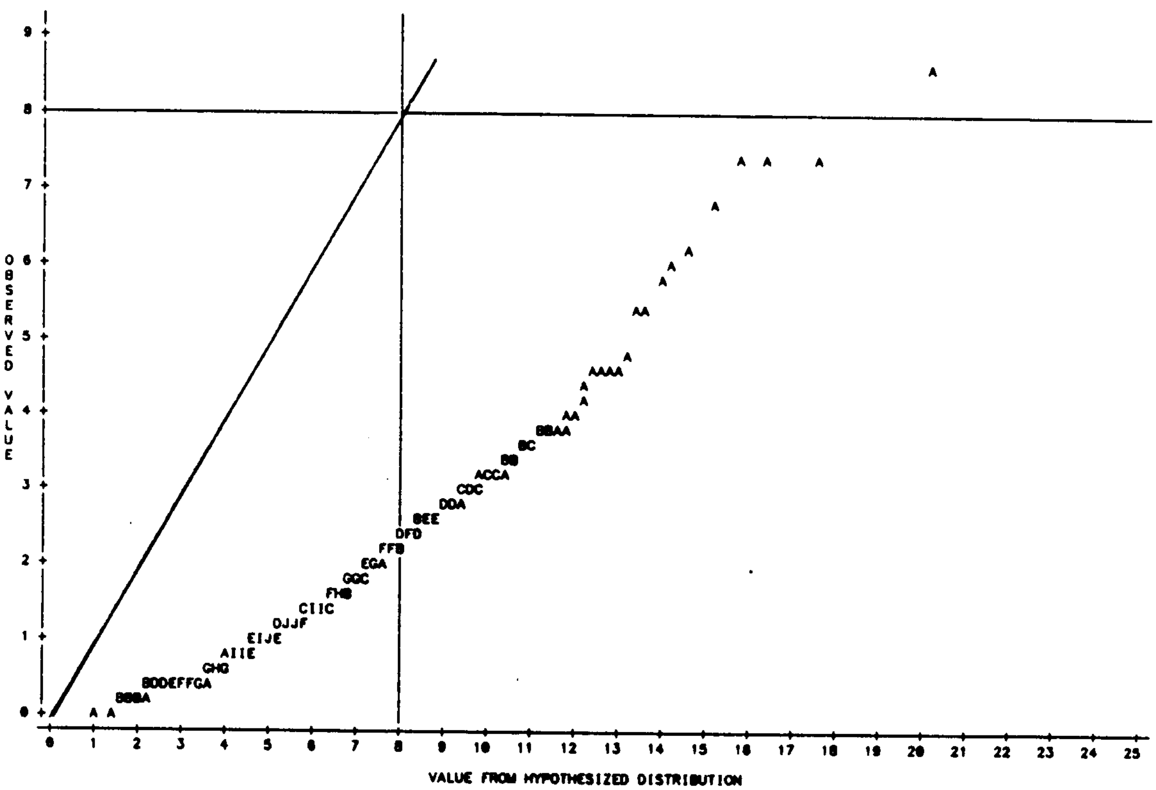

FIGURE 4.5.19
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER SECOND ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA

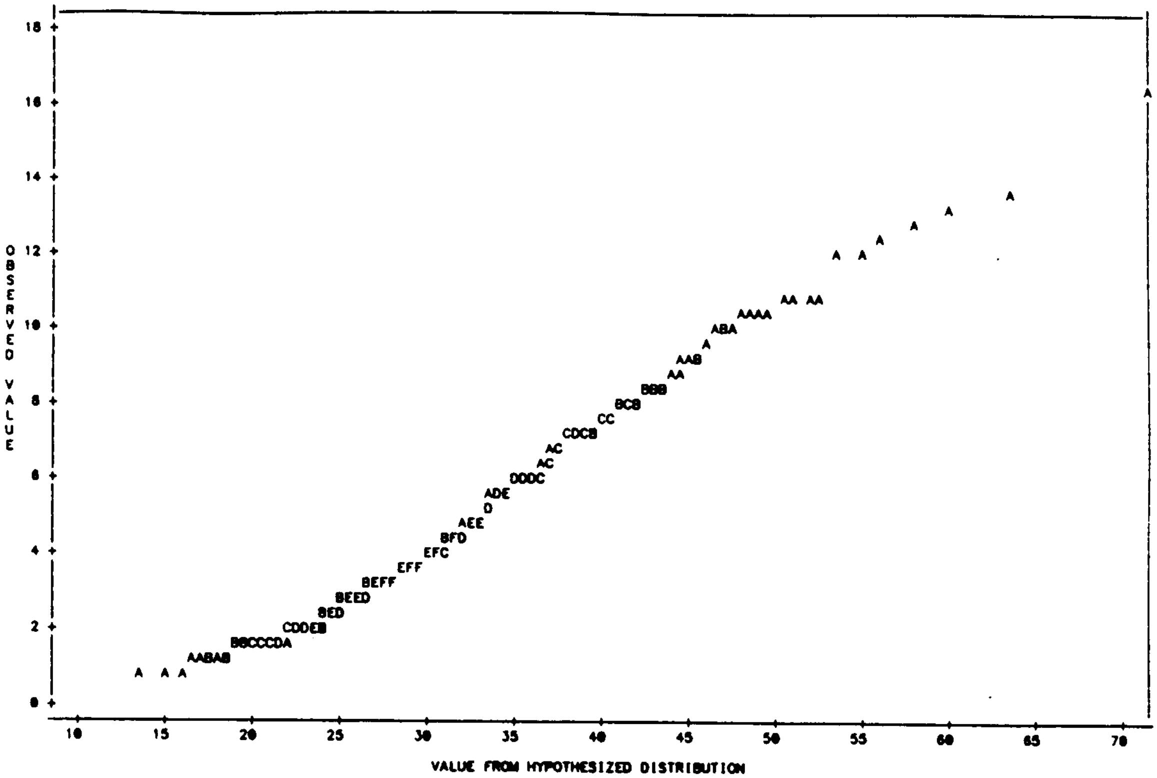

FIGURE 4.5.20
F-PLOT FOR THE DISTRIBUTION OF FBOX STATISTIC
UNDER THIRD ALTERNATIVE HYPOTHESIS
LINKED CROSS-SECTIONAL DESIGN WITH MISSING DATA