Winter 2007

Modeling and Efficient Estimation of Intra-Family Correlations

Roy Sabo
Old Dominion University

MODELING AND EFFICIENT ESTIMATION OF INTRA-FAMILY CORRELATIONS

by

Roy Sabo
B.A. May 2000, Hamilton College

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

MATHEMATICS AND STATISTICS

OLD DOMINION UNIVERSITY
December 2007

Approved by:

N. Rao Chaganty (Director)

Dayanand Naik (Member)

Larry Lee (Member)

Larry Filer (Member)

ABSTRACT

MODELING AND EFFICIENT ESTIMATION OF INTRA-FAMILY CORRELATIONS

Roy Sabo
Old Dominion University, 2007
Director: Dr. N. Rao Chaganty

Familial data occur when observations are taken on multiple members of the same family. Due to relationships between these members, both genetic and by cohabitation, their response variables will likely exhibit some form of dependence. Most of the existing literature models this dependence with an equicorrelated structure. This structure is appropriate when the dependencies between family members are similar, such as in genetic studies, but not in cases where we expect the dependencies to differ, such as behavioral comparisons across different age groups. In this dissertation we first discuss an alternative structure based upon first-order autoregressive correlation. Specifically we create and compare various estimators based on existing and emerging methods of estimation. Asymptotic and small-sample properties are discussed, as is hypothesis testing.

The second part of this dissertation involves a slightly more complicated version of autoregressive familial correlation, where we now model heterogeneous intra-class variances. Again we create and compare various estimators and discuss both their asymptotic and small-sample properties.

In the final part of this dissertation we discuss the nuclear family model, basing the familial dependence on an equicorrelated structure. Note that while this correlation structure has been extensively studied in the case of heterogeneous variance, we model homogenous variance and use a new method for estimating the parameters. Noteworthy here is that we apply a linear transformation to simplify both the correlation matrix and the correlation parameter estimators. As before, we generate estimators and compare their asymptotic performance.

ACKNOWLEDGMENTS

I wish to express the deepest thanks and gratitude to my advisor Professor N. Rao Chaganty for his numerous suggestions and patient guidance throughout the development of this dissertation. His help and encouragement have proven essential both to the completion of this work and to my academic development. For these reasons I am eternally grateful.

I would like to thank Professors Dayanand Naik, Larry Lee and Larry Filer for serving on my dissertation committee. Their helpful comments, suggestions and expertise have improved the quality of this work substantially. Special thanks are also due to Dr. Naik for his immense encouragement.

I would also like to thank my fellow students, friends, professors and staff members in the Department of Mathematics and Statistics at Old Dominion University, especially Professor Dorrepaal, for providing a pleasant and stimulating environment. I extend thanks to the Department of Mathematics and Statistics for the financial support for my graduate studies, as well as to Dr. Dauer and Mike Lane of the Benthic Ecology lab for their support, guidance and friendship. Finally, I extend gratitude to Barbara Jeffrey and Gayle Tarkelsen, both of whom made my experience quite enjoyable.

I dedicate this thesis to MMM who is glad that I started,
and to my wife who is glad that I finished.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

I.1 Literature Review

Familial data arise in situations where a researcher is interested in the relationships among and between the measured responses of parents and children in the same family. As responses within these families (or groups, more generally) are most likely dependent, the estimation of correlations between parents and children (par-sib), between siblings (sib-sib) and to a lesser extent between parents (spouse-spouse) are of interest. The par-sib type correlation is known as inter-class correlation, while the sib-sib and spouse-spouse types are known as intra-class correlation.

One of the earliest treatments of intra-class correlation is found in the work of R. A. Fisher (1918 and 1925), who modeled intra-class correlation as the ratio of variance within a class to the total variance (the sum of variances within and between classes), which are estimated using conventional analysis of variance sums of squares. The idea is that large within-class variation indicates that observations in the same family are heterogeneous, and thus intra-class correlation is small. Testing in this ANOVA setting is equivalent to testing for the significance of within-class correlation. This method requires a balanced design, meaning that families have to be of the same size, and it was work by Fieller and Smith (1951) that expanded this method to account for unequal family sizes, or sibships.

Most of the early inter-class correlation estimators were moment-based, with some of the notable estimators being the pairwise, sib-mean, random-sib and ensemble estimators, as nicely summarized by Rosner, Donner and Hennekens (1977). Each are essentially extensions of the product-moment correlation coefficient, differing in approach as to which parent-child pairings to include. The pairwise estimator included all parent-child pairings, but also assumed that the child response variables were independent (note this assumption was only used to derive the inter-class correlation estimator). The sib-mean estimator sought to avoid this assumption by pairing the parental residual in each family with a residual incorporating the mean of the child

This dissertation follows the style of *Journal of the American Statistical Association.*

responses for that family, and the random-sib estimator paired the parental residual with a randomly chosen child residual. Since both of these estimators omitted information, the ensemble estimator was developed as an attempt to maintain the benefits of both estimators while diminishing their shortcomings. For each method intra-class correlation is estimated using Fisher's approach

The next class of estimators were maximum likelihood estimators. Elston (1975) showed that the pairwise estimator of inter-class correlation (essentially the product-moment correlation coefficient) was equivalent to the maximum likelihood estimator in the case where all sibship sizes are equal. However it was a study by Rosner (1979) that determined the MLE in the case of unequal sibship sizes, and Donner and Koval (1980) extended the MLE method to intra-class correlation. Note that these authors did not obtain closed form estimators for either correlation parameter, and as such used the Newton-Raphson method to find simultaneous solutions. Mak and Ng (1981) improved upon this methodology, but it was Srivastava (1984) who greatly elaborated and improved upon the methodology for both the inter- and intra-class cases by using a transformation to simplify estimation and obtain closed-form estimates.

Covariates can also be measured on each individual with a goal of model building. Though the literature presented above pioneered the estimation of familial correlation parameters, it does not include models with covariates. Some early works incorporating covariates into the MLE approach were by Stanish and Taylor (1983), who found intra-class estimators for the analysis of covariance (ANCOVA) model, and Munoz, Rosner and Carey (1986), who developed a regression model for the case of heterogeneous intra-class correlations between families. The study by Paul (1990) broadened this approach into a generalized model complete with covariates and maximum likelihood estimators for family specific means, variances and intra-class correlation parameters. Paul also showed that most previous models were simply special cases of this general model. By including covariates, a natural consequence would then be to utilize generalized linear models (GLM) for parameter estimation.

More recent work in the field of familial correlation has been predominantly concerned with the genetic relationships between family members. Many of these works use ANOVA modeling to analyze genetic behavior, such as Guo and Wang (2002) and McArdle and Prescott (2005). Another example of the ANOVA approach is done in Rabe-Hasketh *et al.* (2007), who used a mixed model approach to estimate variance components in the case of twins. Other recent studies are, for example, Magnus *et al.* (2001), who studied the genetic relationships between parental and child birth weights in the nuclear family case, and Pawitan *et al.* (2004), who studied both genetic and environmental determinants of binary traits using mixed models and likelihood-based inference for extended families.

I.2 Correlation Structures

Common to most of these treatments is the assumption that all inter-class and intra-class correlations are equicorrelated. A simplified example incorporating homogeneous intra-class variance, is to assume that each family consists of one parent and $t-1$ children, so that we design the $t \times t$ variance-covariance matrix for this family as follows

$$\Sigma_e(\phi, \lambda) = \phi R_e(\lambda) = \phi \begin{pmatrix} 1 & \rho & \rho & & \rho \\ \rho & 1 & \alpha & \cdots & \alpha \\ \rho & \alpha & 1 & & \alpha \\ & \vdots & & \ddots & \vdots \\ \rho & \alpha & \alpha & \cdots & 1 \end{pmatrix} \quad (1.2.1)$$

where ϕ is a scale or variance parameter and $\lambda = (\rho, \alpha)$ is the vector of correlation parameters, where ρ is the par-sib (inter-class) correlation and α is the sib-sib (intra-class) correlation. This correlation structure assumes that the correlation is constant for all parent-child and child-child combinations. According to Hand and Crowder (1996), the equi-correlated structure is appropriate when there is no reason to believe that some pairs of observations should have stronger correlations than other pairs. Based on this observation, we expect the equi-correlated structure to be suited for response variables of traits that are largely genetic, for familial data with age-independent response variables, or where the ages of *all* children are somewhat homogeneous. For example, height measurements on parents and their adult children are bound to exhibit correlation as they all have similar genetic profiles, and that correlation should be constant across pairings since adult children have reached their mature height. Other examples exist for groups of genetically unrelated people, such as coworkers or classmates.

However, one can easily imagine a scenario where this assumption is invalid. For instance, a family participating in a pediatric study would most likely include young children of differing ages, where small age differences could result in large physical differences. Taking again stature as an example, we would expect (on average) siblings to exhibit greater correlation in height if their ages are closer rather than farther apart, and we would expect the same relationship to exist between parents and their children. However, the correlation should decrease as the age difference between family members increases (both within and between classes), especially if the siblings are not yet adults. Thus, for response variables that are age-dependent, the equicorrelated pattern (1.2.1) seems inappropriate.

A non-family example could be taking water sedimentation levels at a series of locations where a freshwater stream empties into a saltwater body. In this case, a measurement at the mouth of the freshwater stream is the source (parent), and each successive measurement further away from that source is a series of destinations (children). For destinations close to the mouth we would expect a high degree of correlation in sedimentation levels as the sediment from the freshwater stream would dominate the existing sediment environment of that destination. However, for destinations far away from the mouth of the freshwater stream, we would expect the local sediment environment to dominate. Here we would expect the dependence relationship between the source and destinations to decrease as you move further into the saltwater body. This is also an instance where an equi-correlated structure seems inappropriate.

A model exhibiting an exponentially decaying correlation pattern would be more appropriate here, where $\rho^{|a_1-a_i|}$ is the correlation between the parent and the ith child (with ages a_1 and a_i, respectively), and $\alpha^{|a_i-a_j|}$ is the correlation between the ith and jth children (with ages a_i and a_j, respectively). A more general model incorporating age-differences is the Markov or generalized Markov structures. Though these candidate models allow a certain degree of flexibility, they are very complicated and difficult to apply to the present situation. Thus we use a simplified model that incorporates an exponentially decaying pattern, namely the first order autoregressive structure. The variance-covariance matrix for this pattern has the following appearance

$$\Sigma(\phi, \lambda) = \phi R(\lambda) = \phi \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & & \rho^{t-1} \\ \rho & 1 & \alpha & \alpha^2 & \cdots & \alpha^{t-2} \\ \rho^2 & \alpha & 1 & \alpha & & \alpha^{t-3} \\ & \vdots & & & \ddots & \vdots \\ \rho^{t-1} & \alpha^{t-2} & \alpha^{t-3} & \alpha^{t-4} & \cdots & 1 \end{pmatrix}. \quad (1.2.2)$$

Note here that ρ is the basis for inter-class correlation, which decreases with each subsequent parent-child pairing, and a similar pattern holds for the child pairings, where α is the basis for intra-class correlation. The structure described in (1.2.2) models a simple form of what we may expect to find within a given family if the response variable is age-dependent. By first examining this AR(1) structure, we will then be able to extend the work to more complicated structures. More on alternative age-dependent structures is discussed in Chapter V.

Much of the literature on familial correlation specifies heterogeneous variances between classes. This essentially means that the variance in a particular class is not necessarily equal to the variance in another class, and so the two variances are treated as separate parameters. Noteworthy examples of this are found in Elston (1975), Rosner, Donner and Hennekens (1977), Rosner (1979) and Srivastava (1984). We again use the simple assumption that a family consists of one parent and $t-1$ children and we design the $t \times t$ variance-covariance matrix for the equicorrelated structure as follows

$$\begin{aligned} \Sigma_e(\Phi, \lambda) &= D(\Phi) R_e(\lambda) D(\Phi) \\ &= \begin{pmatrix} \phi_p & \sqrt{\phi_p\phi_s}\rho & \sqrt{\phi_p\phi_s}\rho & \cdots & \sqrt{\phi_p\phi_s}\rho \\ \sqrt{\phi_p\phi_s}\rho & \phi_s & \phi_s\alpha & \cdots & \phi_s\alpha \\ \sqrt{\phi_p\phi_s}\rho & \phi_s\alpha & \phi_s & \cdots & \phi_s\alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{\phi_p\phi_s}\rho & \phi_s\alpha & \phi_s\alpha & \cdots & \phi_s \end{pmatrix} \end{aligned} \quad (1.2.3)$$

where $D(\Phi) = diag(\phi_p^{1/2}, \phi_s^{1/2}, \cdots, \phi_s^{1/2})$ is a $t \times t$ diagonal matrix of scale parameters, ϕ_p is the parent variance, ϕ_s is the sibling variance, and $\lambda = (\rho, \alpha)$ is defined as before. Note again the equi-correlated structure indicates a somewhat homogeneous class of siblings. We can also incorporate heterogeneous variance into the autoregressive correlation structure, as the following matrix shows.

$$\Sigma(\lambda,\Phi) = D(\Phi)R(\lambda)D(\Phi)$$
$$= \begin{pmatrix} \phi_p & \sqrt{\phi_p\phi_s}\rho & \sqrt{\phi_p\phi_s}\rho^2 & \cdots & \sqrt{\phi_p\phi_s}\rho^{t-1} \\ \sqrt{\phi_p\phi_s}\rho & \phi_s & \phi_s\alpha & \cdots & \phi_s\alpha^{t-2} \\ \sqrt{\phi_p\phi_s}\rho^2 & \phi_s\alpha & \phi_s & \cdots & \phi_s\alpha^{t-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{\phi_p\phi_s}\rho^{t-1} & \phi_s\alpha^{t-2} & \phi_s\alpha^{t-3} & \cdots & \phi_s \end{pmatrix} \quad (1.2.4)$$

More generally, we can model the variance-covariance matrix for a separate classes as done by Elston (1975), who used the following model

$$\Sigma(\Phi,\Lambda) = \begin{pmatrix} \Sigma_1 & \Sigma_{12} & \cdots & \Sigma_{1a} \\ \Sigma_{21} & \Sigma_2 & \cdots & \Sigma_{2a} \\ & \vdots & \ddots & \vdots \\ \Sigma_{a1} & \Sigma_{a2} & \cdots & \Sigma_a \end{pmatrix} \quad (1.2.5)$$

where Φ is a vector of variance parameters and Λ is a vector of correlation parameters. Let Σ_i be an $m_i \times m_i$ matrix whose diagonal elements are ϕ_i and whose off-diagonal elements are $\phi_i\rho_i, i = 1,\cdots,a$. We also let Σ_{ij} be an $m_i \times m_j$ matrix (correspond to m_i members in class i and m_j members in class j) whose elements are all $\phi_i^{1/2}\phi_j^{1/2}\rho_{ij}, i,j = 1,\cdots,a, i \neq j$. Note that this structure can accommodate any number of classes of any size sibship. Also implicit in this model is an equicorrelated structure within and between each class. The most common forms of the familial variance-covariance matrix have only two classes, such as (1.2.1) and (1.2.3). In these cases there is only one parent in the first class and any number of children in the second. However, if two parents are involved, we need three classes as we cannot assume (1) that the parents are uncorrelated, and (2) that the correlations between each parent and the children are equal. If we assume, for our purposes, that the three class variances are equal, then we model the variance-covariance matrix as follows.

$$\Sigma(\phi,\lambda) = \phi\begin{pmatrix} \Sigma_1 & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_2 & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_3 \end{pmatrix} = \phi\begin{pmatrix} 1 & \gamma & \rho_1 & \rho_1 & & \rho_1 \\ \gamma & 1 & \rho_2 & \rho_2 & \cdots & \rho_2 \\ \rho_1 & \rho_2 & 1 & \alpha & & \alpha \\ \rho_1 & \rho_2 & \alpha & 1 & & \alpha \\ & \vdots & & & \ddots & \vdots \\ \rho_1 & \rho_2 & \alpha & \alpha & \cdots & 1 \end{pmatrix} \quad (1.2.6)$$

Here, $\Sigma_1 = \Sigma_2 = \phi$ and $\Sigma_3 = \phi((1-\alpha)I_{t-2} + \alpha J_{t-2})$, and $\Sigma_{12} = \phi\gamma$, $\Sigma_{13} = \phi\rho_1 e'_{t-2}$ and $\Sigma_{23} = \phi\rho_2 e'_{t-2}$, where e_{t-2} is a $(t-2) \times 1$ vector of ones. Note that (1.2.6) has the same correlation structure as used in Shoukri and Ward (1989), where the authors modeled heterogeneous variances, as opposed to the homogeneous intra-class variance modeled here. Implicit in this case is that we are assuming a family of size t, with two parents and $t-2$ children.

The variance-covariance structure (1.2.6) can be simplified using canonical reduction. Srivastava (1984) showed this for the one-parent case (1.2.3), and Khattree and Naik (1994) applied this procedure to the one-parent case where children exhibit a circular dependence structure. By using canonical reduction we can simplify the correlation matrix and ease the computational burden required to estimate the correlation parameters.

I.3 Estimation Procedures

As the dependence structures modeled in (1.2.1) and (1.2.3) have been well-studied, we will concentrate on (1.2.2) and (1.2.4), as well as (1.2.6), which to our knowledge has not been thoroughly analyzed. So with these familial correlation structures in mind, we want to study parameter estimation in a repeated measures setting. In the case of GLM, regression coefficients are usually of tantamount importance; however we will concentrate on estimating the correlation parameters. The maximum likelihood (MLE) method has already been proposed by numerous authors in the case of (1.2.3) (see above), is optimal if data are normally distributed, and serves as a natural starting point. A method independent of an assumed probability distribution is the method of moments (MoM), and as various moment estimators have already been developed, we will incorporate this procedure as well. Quasi-Least Squares (QLS) is an alternative distribution-free procedure that attempts to alleviate certain shortcomings in the moment estimating procedure. Thus, we would like to investigate the autoregressive familial correlation structure, as well as the equicorrelated structure in the nuclear family case, with an eye on gauging the performance of these three estimation procedures.

Let us assume that data is collected on n families, where, in the case of (1.2.2) or (1.2.4), $Y_i = (y_{i1}, y_{i2}, \cdots, y_{it})'$ is the $t_i \times 1$ vector of responses for family i, y_{i1} is the parental response and the remaining $y_{ij}, j = 2, \cdots, t_i$ belong to each of the $t_i - 1$ children. If we are discussing (1.2.6), then y_{i1} and y_{i2} are the parental responses and the remaining $y_{ij}, j = 3, \ldots, t_i$ belong to each of the $t_i - 2$ children. Further, we are assuming that the Y_i are continuous on $(-\infty, \infty)$. For our purposes, we assume the n sibships are of equal size, or $t_i = t$ for all $i = 1, \cdots, n$. Each individual has a $p \times 1$ vector of covariates $X_{ij} = (x_{ij1}, \cdots, x_{ijp})'$ such that

$$X_i = \begin{pmatrix} X_{i1} & X_{i2} & \cdots & X_{it} \end{pmatrix}'$$

is the $t \times p$ matrix of covariates for the ith family.

Based on standard GLM theory, we assume that $E(Y_i) = \mu_i$ and $\eta_i = g(\mu_i) = X_i\beta$, where β is a $p \times 1$ vector of regressor coefficients and $g(\cdot)$ is an invertible, monotone and differentiable link function such that $\mu_i = g^{-1}(X_i\beta)$. Also note that $V(Y_i) = A(\mu_i)^{\frac{1}{2}}\Sigma_i(\lambda, \Phi)A(\mu_i)^{\frac{1}{2}}$, where $A(\mu_i)$ is a $t \times t$ diagonal matrix of the form $diag(v(\mu_{i1}), \cdots, v(\mu_{it}))$, $v(\mu_{ij})$ being the variance function linking the variance of y_{ij} to its expected value μ_{ij}, $\Sigma(\lambda, \Phi)$ is of the form (1.2.2), (1.2.4) or (1.2.6), Φ is a vector of dispersion parameters, and λ is a vector of correlation parameters. Note that if μ_i (through g) is correctly specified, then the GLM estimates are consistent and asymptotically normal. Further, if the variance function $v(\cdot)$ is correctly specified then the GLM estimates have the smallest variance among all unbiased linear estimators. Though technically g can be any monotone function, we use the identity link function, which is allowable since our data are continuous on $(-\infty, \infty)$. Thus we model $E(Y_i) = \mu_i = X_i\beta$ and $v(\mu_{ij}) = 1$ for all $i = 1, \cdots, n$ and $j = 1, \cdots, t$ so that $A(\mu_i)$ is the identity and $V(Y_i) = \Sigma(\lambda, \Phi)$. We also let $\theta = (\beta, \lambda, \Phi)$ be the vector of all parameters.

For each estimating procedure, we use the same estimator for β,

$$\widehat{\beta} = \sum_{i=1}^{n} \left(X_i'\Sigma^{-1}(\widehat{\lambda}, \widehat{\Phi})X_i \right)^{-1} X_i'\Sigma^{-1}(\widehat{\lambda}, \widehat{\Phi})Y_i \tag{1.3.1}$$

where $\widehat{\Phi}$ and $\widehat{\lambda}$ are estimators of the variance and correlation parameters, respectively. Thus, each method differs only in how we estimate the variance and correlation parameters.

For the maximum likelihood estimation method (MLE) we assume that Y_i comes from a t-dimensional multivariate normal distribution with mean $X_i\beta$ and variance-covariance matrix $\Sigma(\lambda, \Phi)$. As we are assuming that the parameters are common to

n independent families, the likelihood function is then the product of n such pdf's.

$$L(Y_1,\cdots,Y_n|\theta) = \prod_{i=1}^{n} f(y_i|\theta) \tag{1.3.2}$$
$$= \prod_{i=1}^{n} (2\pi)^{-\frac{t}{2}} |\Sigma(\lambda,\Phi)|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(Y_i - X_i\beta)'\Sigma^{-1}(\lambda,\Phi)(Y_i - X_i\beta)\right]$$
$$= (2\pi)^{-\frac{nt}{2}} |\Sigma(\lambda,\Phi)|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}tr(\Sigma^{-1}(\lambda,\Phi)Z_n)\right]$$

where $Z_n = \sum_{i=1}^{n}(Y_i - X_i\beta)(Y_i - X_i\beta)' = \sum_{i=1}^{n} Z_i Z_i'$. The log-likelihood is found by taking the natural log of (1.3.2).

$$\begin{aligned} \ell &= \ln(L(Y_1,\cdots,Y_n|\theta)) \\ &= -\frac{nt}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma(\lambda,\Phi)| - \frac{1}{2}tr(\Sigma^{-1}(\lambda,\Phi)Z_n) \end{aligned} \tag{1.3.3}$$

To find the MLE's of θ we need only take the derivative of (1.3.3) with respect to each parameter, set the resulting score equation equal to zero and solve for that parameter. The estimator $\widehat{\beta}$ has already been provided for the regression parameter, and for the variance parameters, recalling that $\frac{\partial \ln|A|}{\partial c} = tr(A^{-1}\frac{\partial A}{\partial c})$, we obtain the following estimating equation for Φ

$$\begin{aligned} &- \frac{n}{2}tr\left[\Sigma^{-1}(\lambda,\Phi)\frac{\partial\Sigma(\lambda,\Phi)}{\partial\Phi}\right] \\ &+\frac{1}{2}tr\left[\Sigma^{-1}(\lambda,\Phi)\frac{\partial\Sigma(\lambda,\Phi)}{\partial\Phi}\Sigma^{-1}(\lambda,\Phi)\widehat{Z}_n\right] = 0 \end{aligned} \tag{1.3.4}$$

where $\widehat{Z}_n$ is Z_n evaluated with $\widehat{\beta}$. In a similar fashion we obtain the following estimating equation for λ

$$\begin{aligned} &- \frac{n}{2}tr\left[\Sigma^{-1}(\lambda,\widehat{\Phi})\frac{\partial\Sigma(\lambda,\widehat{\Phi})}{\partial\lambda}\right] \\ &+\frac{1}{2}tr\left[\Sigma^{-1}(\lambda,\widehat{\Phi})\frac{\partial\Sigma(\lambda,\widehat{\Phi})}{\partial\lambda}\Sigma^{-1}(\lambda,\widehat{\Phi})\widehat{Z}_n\right] = 0. \end{aligned} \tag{1.3.5}$$

Typically, we iterate between $\widehat{\beta}$, $\widehat{\Phi}$ and $\widehat{\lambda}$ until convergence.

The method of moments (MoM) begins with a trial value β_0, which is typically found by solving (1.3.1) with an independent correlation structure (Hardin and Hilbe (2003)). This value is then used to compute residuals

$$Z_i = Y_i - X_i\beta_0, i = 1,\cdots,n.$$

Note that each Z_i is a $t \times 1$ vector of residuals $z_{ij}, j = 1, ..., t$. To obtain estimators for Φ and λ we find estimating equations for those parameters such that

$$\begin{aligned} Z_i'A(\Phi)Z_i - c &= 0, \qquad (1.3.6) \\ Z_i'A(\lambda)Z_i - d &= 0 \qquad (1.3.7) \end{aligned}$$

where c and d are constants. Solving equations (1.3.6) and (1.3.7) for Φ and λ, respectively, yield moment estimators for those parameters. These estimates are then used to solve (1.3.1) for $\widehat{\beta}$, which can in turn be used to recompute the residuals. This iterative process is continued until convergence of the parameters.

The Quasi-Least Squares method, as developed by Chaganty (1997), Shults and Chaganty (1998) and Chaganty and Shults (1999), is an extension of GLM that utilizes the quasi-score function (quasi-log-likelihood) to obtain consistent and efficient estimates not only of the regression parameters but for the correlation parameters as well. According to Wedderburn (1974), the quasi-log-likelihood function is proportional to a true likelihood function if the probability distribution of a random variable is known to belong to an exponential family, and otherwise retains key properties of a true likelihood function that gives QLS asymptotic properties similar to MLE. By specifying only the mean and variance for a random variable, parameter estimation is allowable even if use of the actual likelihood is prohibited. This eases computation in the case when the likelihood function is too complicated or is unknown.

For QLS we start with the quasi-log-likelihood function

$$S(\theta) = \sum_{i=1}^{n}(Y_i - X_i\beta)'\Sigma^{-1}(\lambda,\Phi)(Y_i - X_i\beta) = tr\left[\Sigma^{-1}(\lambda,\Phi)Z_n\right] \qquad (1.3.8)$$

Note that if the variance is homogeneous between classes, then we write

$$S(\theta) = \sum_{i=1}^{n}(Y_i - X_i\beta)'R^{-1}(\lambda)(Y_i - X_i\beta) = tr\left[R^{-1}(\lambda)Z_n\right].$$

Initially we minimize (1.3.8) with respect to β and λ. The *Step 1* regression parameter estimator $\widetilde{\beta}$ is (1.3.1) evaluated at $\widetilde{\Phi}$ and $\widetilde{\lambda}$, the *Step 1* estimates of Φ and λ, respectively. To obtain *Step 1* estimating equations for λ, we differentiate (1.3.8) and set equal to zero.

$$\frac{\partial S(\theta)}{\partial\lambda} = \frac{\partial}{\partial\lambda}\left(tr\left[\Sigma^{-1}(\lambda,\widetilde{\Phi})Z_n\right]\right) = tr\left[\frac{\partial\Sigma^{-1}(\lambda,\widetilde{\Phi})}{\partial\lambda}\widetilde{Z}_n\right] = 0 \qquad (1.3.9)$$

Here $\widetilde{Z}_n = \sum_{i=1}^{n} \widetilde{Z}_i \widetilde{Z}_i'$ is the quadratic form evaluated at our *Step 1* estimate for β. Iterating between (1.3.1) and (1.3.9) until convergence gives us our final *Step 1* estimates. It is well known (Chaganty and Shults (1999)) that the *Step 1* estimates of the correlation terms are biased, which becomes clear if we take the expectation of (1.3.9) evaluated with the *Step 1* estimates.

$$E\left(tr\left[\frac{\partial \Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial \lambda}\widetilde{Z}_n\right]\right) = tr\left[\frac{\partial \Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial \lambda}E(\widetilde{Z}_n)\right] = ntr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \lambda}R(\lambda)\right] \quad (1.3.10)$$

To eliminate this bias we equate (1.3.10) to zero and solve for λ for fixed $\widetilde{\lambda}$. The *Step 2* estimate $\widehat{\lambda}$ is asymptotically unbiased and efficient (Shults and Chaganty (1998)). Further, we obtain a *Step 2* estimate for β by substituting $\widehat{\lambda}$ into (1.3.1) to get $\widehat{\beta}$. If we assume homogeneous variance, then an estimate of $\widehat{\phi}$ is $\frac{1}{nt}S(\widehat{\theta})$; otherwise we use alternative estimators.

I.4 Overview of Thesis

This thesis is organized as follows. In Chapter II we focus on the autoregressive familial correlation structure with homogeneous variance described in (1.2.2). Specifically, we find basic properties of the correlation structure and estimators using the three estimating procedures discussed in Section I.3. We then examine the asymptotic and small-sample performance of those estimators, as well as highlight some basic hypothesis tests for the correlation parameters. In Chapter III we focus on the autoregressive correlation structure with heterogeneous variance described in (1.2.4). Here we also discuss basic properties and find estimators using moment estimators for the variance parameters, as well as examine the asymptotic and small-sample properties. In Chapter IV we concentrate on the nuclear equicorrelated familial structure described in (1.2.6). Here we examine canonical reduction of the correlation matrix, as well as find estimators and derive their asymptotic variance. Finally, we conclude in Chapter V, also illuminating topics for future research.

CHAPTER II

AR(1) STRUCTURE WITH HOMOGENEOUS VARIANCE

II.1 Introduction

In this Chapter we concentrate on the autoregressive familial correlation structure with homogeneous variance. The variance-covariance matrix, as described in Chapter I, is restated here.

$$\Sigma(\phi,\lambda) = \phi R(\lambda) = \phi \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & & \rho^{t-1} \\ \rho & 1 & \alpha & \alpha^2 & \cdots & \alpha^{t-2} \\ \rho^2 & \alpha & 1 & \alpha & & \alpha^{t-3} \\ & \vdots & & & \ddots & \vdots \\ \rho^{t-1} & \alpha^{t-2} & \alpha^{t-3} & \alpha^{t-4} & \cdots & 1 \end{pmatrix} \tag{2.1.1}$$

Here $\theta = (\beta, \lambda, \phi)$, where β is a $(k \times 1)$ vector of regression parameters, ϕ is the variance term, and λ is the vector of correlation parameters $\lambda = (\rho, \alpha)$, where ρ is the correlation between the parent and the first child and α is the correlation between all first-order child pairings (i.e. first and second, second and third, etc.). Recall that in (2.1.1) the correlation between the parent and children is first-order autoregressive based on ρ, and the correlation between the children is first-order autoregressive based on α.

The rest of this Chapter is outlined as follows. In Section II.2 we find the determinant and inverse of (2.1.1), as well as the positive-definite range. In Section II.3 we derive parameter estimators for each estimating procedure, and in Section II.4 we find asymptotic variances for those estimators and compare their asymptotic performance. In Section II.5 we compare the small-sample performance of the estimators in cases of both normally and non-normally distributed data. Lastly, we discuss hypothesis testing in Section II.6.

II.2 Properties of Correlation Matrix

Finding the inverse and determinant of (2.1.1) is simplified by partitioning the matrix as follows

$$R(\lambda) = \left(\begin{array}{c|cccc} 1 & \rho & \rho^2 & \cdots & \rho^{t-1} \\ \hline \rho & 1 & \alpha & \cdots & \alpha^{t-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{t-1} & \alpha^{t-2} & \alpha^{t-3} & \cdots & 1 \end{array}\right) = \begin{pmatrix} 1 & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad (2.2.1)$$

where $R_{12} = R_{21}'$ is a $1 \times (t-1)$ vector of inter-class correlations and R_{22} is the well known $(t-1) \times (t-1)$ first-order autoregressive matrix of intra-class correlations. Thus, $|R_{22}| = (1-\alpha^2)^{t-2}$ and $R_{22}^{-1} = \frac{1}{1-\alpha^2}\left[I_{t-1} + \alpha^2 C_2 - \alpha C_1\right]$, where C_1 is a $(t-1) \times (t-1)$ tri-diagonal matrix with 0's on the main diagonal and 1's on the off diagonals, and C_2 is a diagonal matrix with 1's on the main diagonal except for the first and last elements, which are both 0. We make use of these facts and the general forms for the inverse and determinant of a partitioned matrix to obtain the following results.

$$\begin{aligned} |R(\lambda)| &= |R_{22}||1 - R_{12}R_{22}^{-1}R_{21}| \\ &= \frac{(1-\alpha^2)^{t-3}}{1-\rho^2} \\ &\quad \times \left[(1-\alpha^2)(1-\rho^2) - (\rho^2 - \rho^{2t}) - \alpha^2(\rho^4 - \rho^{2t-2}) + 2\alpha(\rho^3 - \rho^{2t-1})\right] \end{aligned} \quad (2.2.2)$$

and

$$R^{-1}(\lambda) = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad (2.2.3)$$

where

$$\begin{aligned} B_{11} &= (1 - R_{12}R_{22}^{-1}R_{21})^{-1} \\ &= \frac{(1-\alpha^2)(1-\rho^2)}{(1-\alpha^2)(1-\rho^2) - (\rho^2 - \rho^{2t}) - \alpha^2(\rho^4 - \rho^{2t-2}) + 2\alpha(\rho^3 - \rho^{2t-1})} \\ B_{12} &= -B_{11}R_{12}R_{22}^{-1} \\ &= -\frac{B_{11}}{1-\alpha^2}\left[\begin{pmatrix} \rho \\ \rho^2 \\ \vdots \\ \rho^{t-1} \end{pmatrix}' + \alpha^2 \begin{pmatrix} 0 \\ \rho^2 \\ \vdots \\ \rho^{t-2} \\ 0 \end{pmatrix}' - \alpha \begin{pmatrix} \rho^2 \\ \rho + \rho^3 \\ \vdots \\ \rho^{t-3} + \rho^{t-1} \\ \rho^{t-2} \end{pmatrix}'\right] \end{aligned}$$

$$
\begin{aligned}
B_{21} &= B_{12}' \\
B_{22} &= R_{22}^{-1} + R_{22}^{-1} R_{21} B_{11} R_{12} R_{22}^{-1} \\
&= \frac{1}{1-\alpha^2}\left(I_{t-1} + \alpha^2 C_2 - \alpha C_1\right) \\
&\quad \times \left[I_{t-1} + \frac{B_{11}}{1-\alpha^2} \begin{pmatrix} \rho^2 & \rho^3 & \cdots & \rho^t \\ \rho^3 & \rho^4 & \cdots & \rho^{t+1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^t & \rho^{t+1} & \cdots & \rho^{2t-2} \end{pmatrix} \left(I_{t-1} + \alpha^2 C_2 - \alpha C_1\right) \right]
\end{aligned}
$$

In order for correlation estimates to be feasible they must be within a certain range that ensures matrix (2.1.1) is positive definite. Recall that a symmetric matrix (such as (2.1.1)) is positive definite if all its principal leading minors have positive determinants. Thus we can find positive definite ranges for ρ and α by creating inequalities where each principle minor is greater than zero and solving for the parameter values that satisfy the inequality. Of the first $t-1$ leading minors (of R_{22}), the determinant of the ith ($i < t$) is $(1-\alpha^2)^i$, meaning that $-1 < \alpha < 1$. Lastly, we set (2.2.2) greater than zero and solve for either α or ρ. Simplifying the resulting expression we get

$$(1-\alpha^2)(1-\rho^2) - (\rho^2 - \rho^{2t}) - \alpha^2(\rho^4 - \rho^{2t-2}) + 2\alpha(\rho^3 - \rho^{2t-1}) > 0. \qquad (2.2.4)$$

Solving for ρ, let $a = 1-\alpha^2$, $b = \alpha^2$, $c = 2\alpha$ and reorganize (2.2.4) to get

$$\rho^{2t} - c\rho^{2t-1} + b\rho^{2t-2} - b\rho^4 + c\rho^3 - (1+a)\rho^2 + a > 0. \qquad (2.2.5)$$

By selecting $t \geq 2$ and $-1 < \alpha < 1$ we find values of ρ such that the correlation matrix is positive definite by finding the real roots of (2.2.5) that lie between -1 and 1. Solving for α, on the other hand, let $a = -(1-\rho^2+\rho^4-\rho^{2t-2})$, $b = 2(\rho^3 - \rho^{2t-1})$ and $c = 1 - 2\rho^2 + \rho^{2t}$ and reorganize (2.2.4) to get

$$a\alpha^2 + b\alpha + c > 0. \qquad (2.2.6)$$

Note that since (2.2.6) is a quadratic equation, we find the roots with the quadratic formula

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

such that the upper admissible bound on α is

$$min\left[1, \frac{\rho^3 - \rho^{2t-1} + \sqrt{(\rho^3-\rho^{2t-1})^2 + (1-\rho^2+\rho^4-\rho^{2t-2})(1-2\rho^2+\rho^{2t})}}{1-\rho^2+\rho^4-\rho^{2t-2}}\right]$$

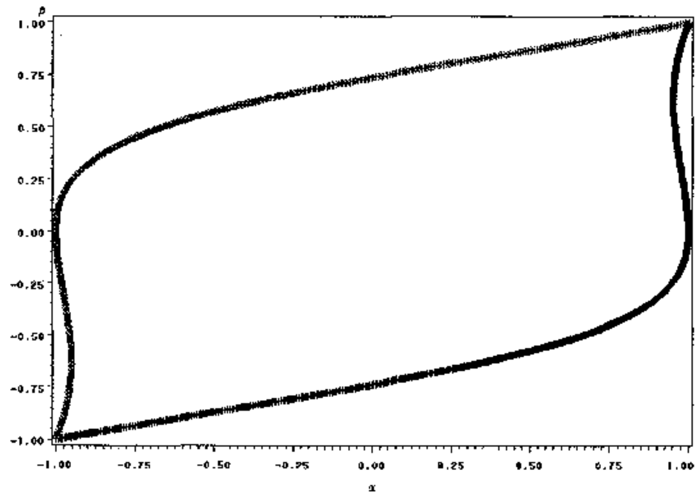

Figure 2.1: P.D. Range for ρ and α when $t = 4$

and the lower admissible bound on α is

$$max\left[-1, \frac{\rho^3 - \rho^{2t-1} - \sqrt{(\rho^3 - \rho^{2t-1})^2 + (1 - \rho^2 + \rho^4 - \rho^{2t-2})(1 - 2\rho^2 + \rho^{2t})}}{1 - \rho^2 + \rho^4 - \rho^{2t-2}}\right]$$

which are found by selecting $t \geq 2$ and $-1 < \rho < 1$. The admissible range is the same whether we solve for ρ or α, though solving for α is a much simpler task.

As an illustration, let $t = 4$, meaning that for each family we have one parent and three siblings. The plot of the positive definite range is shown in Figure 2.1. For reference, we can also let t approach ∞, at which point we get the positive definite range found in Figure 2.2. Notice that there is not much visual difference between the ranges shown in Figure 2.1 and Figure 2.2, though the two are not equal. Table 2.1 gives the upper and lower bounds for ρ over select values of α for both $t = 4$ and $t \to \infty$. As the table shows, the positive definite ranges are slightly wider for $t = 4$ than for $t \to \infty$, and it can be shown numerically that this is also the case for any $t < t + c$, where c is an arbitrary integer. Thus the positive definite range becomes slightly more restrictive as t increases.

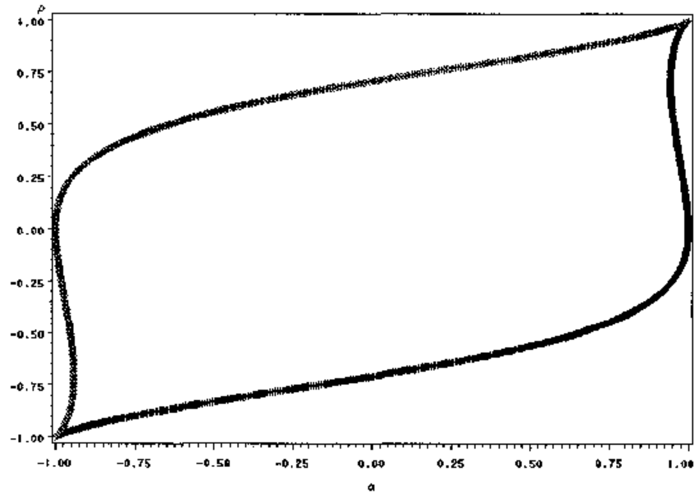

Figure 2.2: P.D. Range for ρ and α as $t \to \infty$

Table 2.1: P.D. Range for ρ when $t = 4$ and $t \to \infty$

α	$\rho^{\ell}_{t=4}$	$\rho^{\ell}_{t\to\infty}$	$\rho^{u}_{t=4}$	$\rho^{u}_{t\to\infty}$
-0.9	-0.97345	-0.94726	0.32163	0.32047
-0.5	-0.87010	-0.82827	0.57324	0.56032
-0.1	-0.76509	-0.73173	0.70855	0.68160
0.3	-0.64621	-0.62616	0.81839	0.77970
0.7	-0.47928	-0.47280	0.92144	0.88117

II.3 Parameter Estimation

In this Section we find estimators of θ for each estimating procedure. Note that closed form estimators are expressed where possible. For those estimators that do not simplify into a closed form, the expressions are left in trace or partitioned vector-matrix form, which are evaluated using (2.2.2), (2.2.3) and the derivatives listed in Appendix A.1.

For each estimating procedure we use the following estimators for β and ϕ

$$\widehat{\beta} = \left(X_i' R^{-1}(\widehat{\lambda}) X_i\right)^{-1} X_i' R^{-1}(\widehat{\lambda}) Y_i \tag{2.3.1}$$

$$\widehat{\phi} = \frac{1}{nt}\sum_{i=1}^{n}(Y_i - X_i\widehat{\beta}) R^{-1}(\widehat{\lambda})(Y_i - X_i\widehat{\beta}) = \frac{1}{nt} tr(R^{-1}(\widehat{\lambda})\widehat{Z}_n) \tag{2.3.2}$$

where $\widehat{\lambda} = (\widehat{\rho}, \widehat{\alpha})$ is the vector of correlation parameter estimators and $\widehat{Z}_n$ is evaluated at $\widehat{\beta}$. The matrix Z_n has the following partitioned form

$$Z_n = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$$

where

$$Z_{11} = \sum_{i=1}^{n} z_{i1}^2$$

$$Z_{12} = \begin{pmatrix} \sum_{i=1}^n z_{i1}z_{i2} & \sum_{i=1}^n z_{i1}z_{i3} \cdots & \sum_{i=1}^n z_{i1}z_{it} \end{pmatrix}$$

$$Z_{21} = Z_{12}'$$

$$Z_{22} = \begin{pmatrix} \sum_{i=1}^n z_{i2}^2 & \sum_{i=1}^n z_{i2}z_{i3} & & \sum_{i=1}^n z_{i2}z_{it} \\ \sum_{i=1}^n z_{i2}z_{i3} & \sum_{i=1}^n z_{i3}^2 & \cdots & \sum_{i=1}^n z_{i3}z_{it} \\ & \vdots & \ddots & \vdots \\ \sum_{i=1}^n z_{i2}z_{it} & \sum_{i=1}^n z_{i3}z_{it} & \cdots & \sum_{i=1}^n z_{it}^2 \end{pmatrix}.$$

II.3.1 Maximum Likelihood

For the maximum likelihood estimation method (MLE) we assume that Y_i comes from a t-dimensional multivariate normal distribution with mean $X_i\beta$ and variance-covariance matrix $\Sigma(\phi, \lambda) = \phi R(\lambda)$, as defined in (2.1.1). Using this variance-covariance matrix, the log-likelihood becomes

$$\ell = -\frac{nt}{2}\ln(2\pi) - \frac{nt}{2}\ln(\phi) - \frac{n}{2}\ln|R(\lambda)| - \frac{1}{2\phi} tr(R^{-1}(\lambda)Z_n). \tag{2.3.3}$$

To find MLE's of θ we need only take the derivative of (2.3.3) with respect to each parameter, set the resulting function equal to zero and solve for that parameter.

We obtain the following estimating equations for ρ and α, respectively

$$
\begin{aligned}
- & \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}\right] + \frac{1}{2\widehat{\phi}} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \quad (2.3.4)\\
\Leftrightarrow & -n\widehat{\phi}B_{12}\frac{\partial R_{21}}{\partial \rho} + B_{11}\frac{\partial R_{12}}{\partial \rho}B_{21}\widehat{Z}_{11} + B_{12}\frac{\partial R_{21}}{\partial \rho}B_{12}\widehat{Z}_{21}\\
& +B_{11}\frac{\partial R_{12}}{\partial \rho}B_{22}\widehat{Z}_{21} + B_{12}\widehat{Z}_{22}B_{22}\frac{\partial R_{21}}{\partial \rho} = 0\\
- & \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right] + \frac{1}{2\widehat{\phi}} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \quad (2.3.5)\\
\Leftrightarrow & -n\widehat{\phi}\, tr\left[B_{22}\frac{\partial R_{22}}{\partial \alpha}\right] + B_{12}\frac{\partial R_{22}}{\partial \alpha}B_{21}\widehat{Z}_{11} + 2B_{12}\frac{\partial R_{22}}{\partial \alpha}B_{22}\widehat{Z}_{21}\\
& +tr\left[B_{22}\frac{\partial R_{22}}{\partial \alpha}B_{22}\widehat{Z}_{22}\right] = 0
\end{aligned}
$$

where $\widehat{Z}_n$ is Z_n evaluated with $\widehat{\beta}$.

The MLE's are found by first choosing initial values $\lambda_0 = (\rho_0, \alpha_0)$ to estimate β using (2.3.1). We then use $\widehat{\beta}$ to update the residual matrix $\widehat{Z}_n$ and estimate ϕ using (2.3.2). These values are then used to simultaneously solve equations (2.3.4) and (2.3.5) using Newton-Raphson to obtain updated values of λ. This process is then repeated until convergence, those values being the MLE's: $\widehat{\theta}_\ell = (\widehat{\beta}_\ell, \widehat{\lambda}_\ell, \widehat{\phi}_\ell)$.

II.3.2 Method of Moments

For the method of moments (MoM) we obtain $\widehat{\lambda}$ by using variations of the product-moment estimators proposed by Hardin and Hilbe (2003) for the autoregressive case. For ρ, we use

$$\widehat{\rho}_m = \frac{\sum_{i=1}^{n}\widehat{z}_{i1}\widehat{z}_{i2}}{\frac{1}{t}\sum_{i=1}^{n}\sum_{j=1}^{t}\widehat{z}_{ij}^2} \quad (2.3.6)$$

where $\widehat{z}_{i1}$ and $\widehat{z}_{i2}$ are the residuals for the parent and first child, respectively, in the *ith* family. This residual pairing is included as it is the only pairing for which the expected value involves ρ raised to the first power. For α we use

$$\widehat{\alpha}_m = \frac{\frac{1}{(t-2)}\sum_{i=1}^{n}\sum_{j=2}^{t-1}\widehat{z}_{ij}\widehat{z}_{i,j+1}}{\frac{1}{t}\sum_{i=1}^{n}\sum_{j=1}^{t}\widehat{z}_{ij}^2}. \quad (2.3.7)$$

Here we only include child pairings of the first order (i.e. first and second children, but not the first and third, etc.) as only these pairings have an expected value that involve α raised to the first power. Note that (2.3.6) and (2.3.7) are obtained from the following two unbiased estimating equations

$$Z_i'A(\rho)Z_i = Z_i'\left[\frac{\rho}{t}I_t - \frac{1}{2}\begin{pmatrix} C_3 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}\right]Z_i = 0 \tag{2.3.8}$$

$$Z_i'A(\alpha)Z_i = Z_i'\left[\frac{\alpha}{t}I_t - \frac{1}{2(t-2)}\begin{pmatrix} 0 & \underline{0} \\ \underline{0} & C_1 \end{pmatrix}\right]Z_i = 0 \tag{2.3.9}$$

where $C_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and C_1 is defined in (2.2.2). Solving equations (2.3.8) and (2.3.9) for ρ and α, respectively, yield (2.3.6) and (2.3.7). These estimates are then used to solve (2.3.1) for $\widehat{\beta}_m$, which in turn is used to recompute the residuals. This iterative process is continued until convergence of the parameters. Upon convergence, we estimate ϕ with $\widehat{\phi}_m$ (2.3.2) using $\widehat{\beta}_m$ and $\widehat{\lambda}_m$. Thus, the MoM estimator is $\widehat{\theta}_m = (\widehat{\beta}_m, \widehat{\lambda}_m, \widehat{\phi}_m)$.

II.3.3 Quasi-Least Squares

For QLS we start with the quasi-log-likelihood function

$$S(\theta) = \sum_{i=1}^{n}(Y_i - X_i\beta)'R^{-1}(\lambda)(Y_i - X_i\beta) = tr\left[R^{-1}(\lambda)Z_n\right]. \tag{2.3.10}$$

Using the quasi-log-likelihood function (2.3.10) we obtain *Step 1* estimating equations for both ρ and α by differentiating (2.3.10) and setting equal to zero.

$$\begin{aligned}
\frac{\partial S(\theta)}{\partial \rho} &= \frac{\partial}{\partial \rho}\left(tr\left[R^{-1}(\lambda)\widetilde{Z}_n\right]\right) = tr\left[\frac{\partial R^{-1}(\lambda)}{\partial \rho}\widetilde{Z}_n\right] = 0 \qquad (2.3.11)\\
&\Leftrightarrow B_{11}\frac{\partial R_{12}}{\partial \rho}B_{21}\widetilde{Z}_{11} + B_{12}\frac{\partial R_{21}}{\partial \rho}B_{12}\widetilde{Z}_{21}\\
&\quad + B_{11}\frac{\partial R_{12}}{\partial \rho}B_{22}\widetilde{Z}_{21} + B_{12}\widetilde{Z}_{22}B_{22}\frac{\partial R_{21}}{\partial \rho} = 0\\
\frac{\partial S(\theta)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left(tr\left[R^{-1}(\lambda)\widetilde{Z}_n\right]\right) = tr\left[\frac{\partial R^{-1}(\lambda)}{\partial \alpha}\widetilde{Z}_n\right] = 0 \qquad (2.3.12)\\
&\Leftrightarrow B_{12}\frac{\partial R_{22}}{\partial \alpha}B_{21}\widetilde{Z}_{11} + 2B_{12}\frac{\partial R_{22}}{\partial \alpha}B_{22}\widetilde{Z}_{21}\\
&\quad + tr\left[B_{22}\frac{\partial R_{22}}{\partial \alpha}B_{22}\widetilde{Z}_{22}\right] = 0
\end{aligned}$$

Here $\widetilde{Z}_n = \sum_{i=1}^{n} \widetilde{Z}_i \widetilde{Z}_i'$ is the residual matrix evaluated with $\widetilde{\beta}$, which is found using (2.3.1). Solving (2.3.11) and (2.3.12) simultaneously for ρ and α using Newton-Raphson we get the initial *Step 1* estimate $\widetilde{\lambda} = (\widetilde{\rho}, \widetilde{\alpha})$. Iterating between (2.3.1) and both (2.3.11) and (2.3.12) until convergence gives the final *Step 1* estimates. It is well known (Chaganty and Shults (1999)) that the *Step 1* estimates of the correlation terms are biased, which becomes clear if we take the expectation of estimating equations (2.3.11) and (2.3.12) evaluated with the *Step 1* estimates.

$$\begin{aligned} E\left(tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\rho}} \widetilde{Z}_n\right]\right) &= tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\rho}} E(\widetilde{Z}_n)\right] \neq 0 && (2.3.13) \\ &\propto tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\rho}} R(\lambda)\right] \neq 0 \\ E\left(tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\alpha}} \widetilde{Z}_n\right]\right) &= tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\alpha}} E(\widetilde{Z}_n)\right] \neq 0 && (2.3.14) \\ &\propto tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\alpha}} R(\lambda)\right] \neq 0 \end{aligned}$$

To eliminate this bias, we equate these two expressions to zero and again simultaneously solve for ρ and α, as shown in equations (2.3.15) and (2.3.16).

$$\begin{aligned} & tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho} R(\lambda)\right] = 0 && (2.3.15) \\ \Leftrightarrow\quad & \widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\left[\widetilde{B}_{21} + \widetilde{B}_{22}R_{21}\right] + \widetilde{B}_{12}\left[R_{21}\widetilde{B}_{12} + R_{22}\widetilde{B}_{22}\right]\frac{\partial \widetilde{R}_{21}}{\partial \rho} = 0 \\ & tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha} R(\lambda)\right] = 0 && (2.3.16) \\ \Leftrightarrow\quad & \widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\left[\widetilde{B}_{21} + 2\widetilde{B}_{22}R_{21}\right] + tr\left[\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}\right] = 0 \end{aligned}$$

These *Step 2* estimates of the correlation parameters ($\widehat{\rho}_q$ and $\widehat{\alpha}_q$) are asymptotically unbiased and efficient (Shults and Chaganty (1998)). Further, we obtain a *Step 2* estimate for β by substituting $\widehat{\rho}_q$ and $\widehat{\alpha}_q$ into (2.3.1) to get $\widehat{\beta}_q$. We also estimate ϕ with $\widehat{\phi}_q$, which is (2.3.2) evaluated with $\widehat{\beta}_q$ and $\widehat{\lambda}_q$. The QLS estimates are $\widehat{\theta}_q = (\widehat{\beta}_q, \widehat{\lambda}_q, \widehat{\phi}_q)$.

II.4 Asymptotic Variance and Performance

In this section we derive asymptotic variances for the estimation procedures described in Section II.3. The MLE procedure is straightforward, as we are assuming the residual vectors $\widehat{Z}_i = Y_i - X_i\widehat{\beta}$ are normally distributed. This allows us to make use of the log-likelihood function to find Fisher's information matrix $(I(\theta))$, where under the regularity conditions $-E(\frac{\partial^2 \ell}{\partial\theta\partial\theta'}) = Cov(\frac{\partial \ell}{\partial\theta}) = I(\theta)$. Since the asymptotic distributions for both the MoM and QLS methods depend on higher order moments, we must assume that the residuals are normally distributed for these methods as well. Note that this assumption is not needed for parameter estimation, but only for derivation of the asymptotic variances, and is justified by use of the Central Limit Theorem. For the MoM and QLS procedures, we make use of the following theorem by Joe (1997, p. 301), which states that under the regularity conditions

$$\sqrt{n}(\widehat{\theta} - \theta) \sim AMVN\left(\underline{0}, I_n^{-1}(\theta)M_n(\theta)(I_n^{-1}(\theta))'\right) \tag{2.4.1}$$

where $I_n(\theta) = -\frac{1}{n}\sum_{i=1}^n E\left[\frac{\partial h_i(\theta)}{\partial\theta}\right]$, $M_n(\theta) = \frac{1}{n}\sum_{i=1}^n Cov(h_i(\theta))$, and $h_i(\theta)$ is a vector of unbiased estimating equations for θ. Note that this theorem is a more general theorem for finding asymptotic variances than is Fisher's Information. If we apply this theorem to the MLE's, then $h_i(\theta) = \frac{\partial\ell}{\partial\theta}$, and $I_n(\theta) = -\frac{1}{n}\sum_{i=1}^n E\left(\frac{\partial^2\ell}{\partial\theta\partial\theta'}\right) = I(\theta)$, and $M_n(\theta) = \frac{1}{n}\sum_{i=1}^n Cov(\frac{\partial\ell}{\partial\theta}) = I(\theta) = I_n(\theta)$, so that $I_n^{-1}(\theta)M_n(\theta)(I_n^{-1}(\theta))' = I^{-1}(\theta)I(\theta)I^{-1}(\theta) = I^{-1}(\theta)$. Thus, using the multivariate normal log-likelihood function in (2.4.1) gives us the inverse of Fisher's Information matrix, which is what we obtained earlier.

II.4.1 Maximum Likelihood

Asymptotic variances and covariances for the maximum likelihood estimators are found by taking the negative expectation of the second derivative of the likelihood function with respect to θ. The resulting functions form the Fisher Information matrix. The diagonals of the inverse of this matrix are the asymptotic variances for the parameter estimators.

According to the Cramér's Theorem, we have

$$\sqrt{n}(\widehat{\theta}_\ell - \theta) \sim AMVN\left(\underline{0}, I_\ell^{-1}(\theta)\right). \tag{2.4.2}$$

It is straightforward to show that the information matrix $I_\ell(\theta)$ is of the following form

$$I_\ell(\theta) = \begin{pmatrix} I(\beta) & 0 & 0 & 0 \\ 0 & I(\rho) & I(\rho,\alpha) & I(\rho,\phi) \\ 0 & I(\rho,\alpha) & I(\alpha) & I(\alpha,\phi) \\ 0 & I(\rho,\phi) & I(\alpha,\phi) & I(\phi) \end{pmatrix} \tag{2.4.3}$$

where (recall that B_{11} is defined in (2.2.3))

$$\begin{aligned}
I(\beta) &= -E\left(\frac{\partial^2 \ell}{\partial \beta^2}\right) = \frac{1}{\phi}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
I(\rho) &= -E\left(\frac{\partial^2 \ell}{\partial \rho^2}\right) = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= \frac{2n^3 B_{11}}{(\phi(1-\alpha^2)(1-\rho^2)^2)^2}[(\rho - t\rho^{2t-1} + (t-1)\rho^{2t+1}) \\
&\quad +\alpha^2(2\rho^3 - (t-1)\rho^{2t-3} + (t-2)\rho^{2t-1}) \\
&\quad -\alpha(3\rho^2 - \rho^4 - (2t-1)\rho^{2t-2} + (2t-3)\rho^{2t})]^2 \\
&\quad +\frac{n}{2(1-\alpha^2)}\sum_{j=1}^{t-1}(j)^2\rho^{2j-2} + \frac{n\alpha^2}{2(1-\alpha^2)}\sum_{j=1}^{t-3}(j+1)^2\rho^{2j} \\
&\quad -\frac{n\alpha}{2(1-\alpha^2)}\sum_{j=1}^{t-2}(j)\rho^{2j-1} + \frac{nB_{11}}{2(1-\alpha^2)^2}\left(\sum_{j=1}^{t-1}(j)\rho^{2j-1}\right)^2 \\
&\quad +\frac{nB_{11}\alpha^2}{(1-\alpha^2)^2}\left(\frac{\rho - t\rho^{2t-1} + (t-1)\rho^{2t+1}}{(1-\rho^2)^2}\right) \\
&\quad \times\left(\frac{\rho - t\rho^{2t-1} + (t-1)\rho^{2t+1}}{(1-\rho^2)^2} - (\rho + (t-1)\rho^{2t-3})\right) \\
&\quad -\frac{nB_{11}\alpha}{(1-\alpha)^2}\left(\frac{\rho - t\rho^{2t-1} + (t-1)\rho^{2t+1}}{(1-\rho^2)^2}\right) \\
&\quad -\frac{nB_{11}\alpha}{(1-\alpha)^2}\left(\frac{3\rho^2 - \rho^4 - (2t-1)\rho^{2t-2} + (2t-3)\rho^{2t}}{(1-\rho^2)^2}\right) \\
&\quad -\frac{nB_{11}\alpha^3}{(1-\alpha^2)^2}\left(\frac{3\rho^2 - \rho^4 - (2t-1)\rho^{2t-2} + (2t-3)\rho^{2t}}{(1-\rho^2)^2}\right) \\
&\quad \times\left(\frac{2\rho^3 - \rho^5 - (t-1)\rho^{2t-3} + (t-2)\rho^{2t-1}}{(1-\rho^2)^2}\right) \\
&\quad +\frac{nB_{11}\alpha^2}{2(1-\alpha^2)^2}\left(\frac{3\rho^2 - \rho^4 - (2t-1)\rho^{2t-2} + (2t-3)\rho^{2t}}{(1-\rho^2)^2}\right)^2 \\
&\quad \frac{nB_{11}\alpha^4}{2(1-\alpha^2)^2}\left(\frac{2\rho^3 - \rho^5 - (t-1)\rho^{2t-3} + (t-2)\rho^{2t-1}}{(1-\rho^2)^2}\right)^2
\end{aligned} \tag{2.4.4}$$

$$
\begin{aligned}
I(\alpha) &= -E\left(\frac{\partial^2\ell}{\partial\alpha^2}\right) = \frac{n}{2}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= \frac{n}{2}tr\left[\frac{\partial R_{22}}{\partial\alpha}B_{22}\frac{\partial R_{22}}{\partial\alpha}B_{22}\right] \\
I(\phi) &= -E\left(\frac{\partial^2\ell}{\partial\phi^2}\right) = \frac{nt}{2\phi^2} \\
I(\rho,\alpha) &= -E\left(\frac{\partial\ell}{\partial\rho\partial\alpha}\right) = \frac{n}{2}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= n\frac{\partial R_{12}}{\partial\rho}B_{22}\frac{\partial R_{22}}{\partial\alpha}B_{21} \\
I(\rho,\phi) &= -E\left(\frac{\partial\ell}{\partial\rho\partial\phi}\right) = \frac{n}{2\phi}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho}\right] = \frac{n}{\phi}\frac{\partial R_{12}}{\partial\rho}B_{21} \\
&= -\frac{nB_{11}}{2\phi(1-\alpha^2)(1-\rho^2)^2}[(\rho - t\rho^{2t-1} + (t-1)\rho^{2t+1}) \\
&\quad +\alpha^2(2\rho^3 - \rho^5 - (t-1)\rho^{2t-3} + (t-2)\rho^{2t-1}) \\
&\quad -\alpha(3\rho^2 - \rho^4 - (2t-1)\rho^{2t-2} + (2t-3)\rho^{2t})] \\
I(\alpha,\phi) &= -E\left(\frac{\partial\ell}{\partial\alpha\partial\phi}\right) = \frac{n}{2\phi}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] = \frac{n}{2\phi}tr\left[B_{22}\frac{\partial R_{22}}{\partial\alpha}\right] \\
&= -\frac{n\alpha(t-2)}{\phi(1-\alpha^2} + \frac{nB_{11}}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-2}(j)\alpha^{j-1}\left(\rho^{j+2} - \rho^{2t-j}\right) \\
&\quad +\frac{nB_{11}\alpha^2}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-3}(j)\alpha^{j-1}\left(\rho^{j+2} - \rho^{2t-j}\right) \\
&\quad +\frac{nB_{11}\alpha^2}{\phi(1-\alpha^2)^2(1-\rho^2)} + \sum_{j=1}^{t-4}(j)\alpha^{j-1}\left(\rho^{j+4} - \rho^{2t-2-j}\right) \\
&\quad -\frac{nB_{11}\alpha}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-2}(j)\alpha^{j-1}\left(\rho^{j+1} - \rho^{2t+1-j}\right) \\
&\quad -\frac{nB_{11}\alpha}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-3}(j)\alpha^{j-1}\left(\rho^{j+3} - \rho^{2t-1-j}\right) \\
&\quad -\frac{nB_{11}\alpha^3}{\phi(1-\alpha^2)^2(1-\rho^2)}2\sum_{j=1}^{t-3}(j)\alpha^{j-1}\left(\rho^{j+3} - \rho^{2t-1-j}\right) \\
&\quad -\frac{nB_{11}\alpha^3}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-4}(j)\alpha^{j-1}\left(\rho^{j+3} - \rho^{2t-1-j}\right) \\
&\quad -\frac{nB_{11}\alpha^3}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-5}(j)\alpha^{j-1}\left(\rho^{j+5} - \rho^{2t-3-j}\right) \\
&\quad +\frac{nB_{11}\alpha^2}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-2}(j)\alpha^{j-1}\left(\rho^{j+2} - \rho^{2t-j}\right)
\end{aligned}
$$

$$+\frac{nB_{11}\alpha^2}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-3}(j)\alpha^{j-1}\left(\rho^{j+2}-\rho^{2t-j}\right)$$
$$+\frac{nB_{11}\alpha^2}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-4}(j)\alpha^{j-1}\left(\rho^{j+4}-\rho^{2t-2-j}\right)$$
$$+\frac{nB_{11}}{\phi(1-\alpha^2)^2(1-\rho^2)}\sum_{j=1}^{t-4}(j)\alpha^{j-1}\left(\rho^{j+4}-\rho^{2t-2-j}\right)$$

Note that the covariance terms involving β are zero, indicating that $\widehat{\beta}_\ell$ is uncorrelated with the estimators for the other parameters.

II.4.2 Method of Moments

For the MoM method, based on (2.4.1), we have

$$\sqrt{n}(\widehat{\theta}_m-\theta)\sim AMVN\left(0, I_m^{-1}(\theta)M_m(\theta)(I_m^{-1}(\theta))'\right) \tag{2.4.5}$$

where $I_m(\theta)=-\frac{1}{n}\sum_{i=1}^n E\left[\frac{\partial h_{m,i}(\theta)}{\partial\theta'}\right]$, $M_m(\theta)=\frac{1}{n}\sum_{i=1}^n Cov(h_{m,i}(\theta))$ and the $h_{m,i}(\theta)$ are vectors of unbiased estimating equations defined as follows

$$\begin{aligned} h_{m,i}(\theta) &= (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), g_i(\theta))' \\ h_{0i}(\theta) &= X_i'R^{-1}(\lambda)Z_i \\ h_{1i}(\theta) &= Z_i'A(\rho)Z_i = tr(A(\rho)Z_iZ_i') \\ h_{2i}(\theta) &= Z_i'A(\alpha)Z_i = tr(A(\alpha)Z_iZ_i') \\ g_i(\theta) &= Z_i'R^{-1}(\lambda)Z_i - t\phi = tr(R^{-1}(\lambda)Z_iZ_i') - t\phi \end{aligned} \tag{2.4.6}$$

where $A(\rho)$ and $A(\alpha)$ are defined earlier. By taking the negative expectation of the partial derivatives of (2.4.6) with respect to θ and averaging over n we obtain $I_m(\theta)$, and by taking the covariance of (2.4.6) and averaging over n we obtain $M_m(\theta)$. From here it is easy to show that $I_m(\theta)$ has the following elements

$$I_m(\theta)=\begin{pmatrix} I_{11} & 0 & 0 & 0 \\ 0 & I_{22} & 0 & 0 \\ 0 & 0 & I_{33} & 0 \\ 0 & I_{42} & I_{43} & I_{44} \end{pmatrix} \tag{2.4.7}$$

where

$$
\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{0i}(\theta)}{\partial \beta}\right) = \frac{1}{n}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \rho}\right) = -\phi \\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \alpha}\right) = -\phi \\
I_{42} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_i(\theta)}{\partial \rho}\right) = \phi tr\left(R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}\right) \\
&= -\frac{2\phi B_{11}}{1-\alpha^2}\left(\frac{(t-1)\rho^{2t+1} - t\rho^{2t-1} + \rho}{(1-\rho^2)^2}\right) \\
&\quad -\frac{2\phi\alpha^2 B_{11}}{1-\alpha^2}\left(\frac{(t-2)\rho^{2t-1} - (t-1)\rho^{2t-3} - \rho^5 + 2\rho^3}{(1-\rho^2)^2}\right) \\
&\quad +\frac{2\phi\alpha B_{11}}{1-\alpha^2}\left(\frac{(2t-3)\rho^{2t} - (2t-1)\rho^{2t-2} - \rho^4 + 3\rho^2}{(1-\rho^2)^2}\right) \\
I_{43} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_i(\theta)}{\partial \alpha}\right) = \phi tr\left(R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right) \\
&= -\frac{2\alpha(t-2)}{1-\alpha^2} \\
&\quad +\frac{2B_{11}}{(1-\alpha^2)^2}\left(\frac{\sum_{j=1}^{t-2}(j)\alpha^{j-1}(\rho^{j+2} - \rho^{2t-j})}{1-\rho^2}\right) \\
&\quad -\frac{2\alpha B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-2}(j)\alpha^{j-1}(\rho^{j+1} - \rho^{2t+1-j}) + 3\sum_{j=1}^{t-3}(j)\alpha^{j-1}(\rho^{j+3} - \rho^{2t-1-j})\right) \\
&\quad +\frac{2\alpha^2 B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-3}(j)\alpha^{j-1}(\rho^{j+2} - \rho^{2t-j}) + \sum_{j=1}^{t-4}(j)\alpha^{j-1}(\rho^{j+4} - \rho^{2t-2-j})\right) \\
&\quad +\frac{2\alpha^2 B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-2}(j)\alpha^{j-1}(\rho^{j+2} - \rho^{2t-j}) + \sum_{j=1}^{t-3}(j)\alpha^{j-1}(\rho^{j+2} - \rho^{2t-j})\right) \\
&\quad +\frac{4\alpha^2 B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-4}(j)\alpha^{j-1}(\rho^{j+4} - \rho^{2t-2+j})\right) \\
&\quad -\frac{4\alpha^3 B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-3}(j)\alpha^{j-1}(\rho^{j+3} - \rho^{2t+1-j})\right) \\
&\quad -\frac{2\alpha^3 B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-4}(j)\alpha^{j-1}(\rho^{j+3} - \rho^{2t-1-j}) + \sum_{j=1}^{t-5}(j)\alpha^{j-1}(\rho^{j+5} - \rho^{2t-3-j})\right) \\
&\quad \frac{2\alpha^4 B_{11}}{(1-\alpha^2)^2(1-\rho^2)}\left(\sum_{j=1}^{t-4}(j)\alpha^{j-1}(\rho^{j+4} - \rho^{2t-2-j})\right)
\end{aligned}
$$

$$I_{44} = -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_i(\theta)}{\partial \phi}\right) = t.$$

We can also show that $M_m(\theta)$ has the following elements

$$M_m(\theta) = \begin{pmatrix} M_{11} & 0 & 0 & 0 \\ 0 & M_{22} & M_{23} & 0 \\ 0 & M_{23} & M_{33} & 0 \\ 0 & 0 & 0 & M_{44} \end{pmatrix} \tag{2.4.8}$$

where

$$\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{0i}(\theta)) = \frac{\phi}{n}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta)) = 2\phi^2 tr\left[A(\rho)R(\lambda)A(\rho)R(\lambda)\right] \\
&= \frac{2\phi^2\rho^2}{t^2}\left(t + \frac{2(\rho^2-\rho^{2t})}{1-\rho^2} + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right) \\
&\quad -\frac{2\phi^2\rho_p^2}{t}\left(4 + \frac{2(\rho\alpha-(\rho\alpha)^{t-1})}{1-\rho\alpha}\right) + \phi^2(1+\rho^2) \\
M_{23} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), h_{2i}(\theta)) = 2\phi^2 tr\left[A(\rho)R(\lambda)A(\alpha)R(\lambda)\right] \\
&= \frac{2\phi^2\rho\alpha}{t^2}\left(t + \frac{2(\rho^2-\rho^{2t})}{1-\rho^2} + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right) \\
&\quad -\frac{2\phi^2\rho}{t(t-2)}\left(\frac{\rho^3-\rho^{2t-1}}{1-\rho^2} + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j-1}\right) \\
&\quad -\frac{\phi^2\alpha\rho}{t}\left(4 + \frac{2(\rho\alpha-(\rho\alpha)^{t-1})}{1-\rho\alpha}\right) \\
&\quad +\frac{\phi^2}{(t-2)}\left(\frac{\alpha+\rho}{\alpha}\right)\left(\frac{\rho\alpha-(\rho\alpha)^{t-1}}{1-\rho\alpha}\right) \\
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta)) = 2\phi^2 tr\left[A(\alpha)R(\lambda)A(\alpha)R(\lambda)\right] \\
&= \frac{2\phi^2\alpha^2}{t^2}\left(t + \frac{2(\rho^2-\rho^{2t})}{1-\rho^2} + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right) \\
&\quad -\frac{4\phi^2\alpha}{t(t-2)}\left(\frac{\rho^3-\rho^{2t-1}}{1-\rho^2} + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j-1}\right)
\end{aligned}$$

$$+\frac{\phi^2}{2(t-2)^2}\left(2(t-2)+[10(t-3)+2]\alpha^2+\sum_{j=1}^{t-4}(t-3-j)\alpha^{2+2j}\right)$$

$$M_{44} = \frac{1}{n}\sum_{i=1}^{n} Cov(g_i(\theta)) = 2\phi^2 t.$$

Note that based on matrices (2.4.7) and (2.4.8), the covariance terms involving β are zero, indicating that $\widehat{\beta}_m$ is uncorrelated with estimators for the other parameters.

II.4.3 Quasi Least Squares

For the QLS method we note that, based on (2.4.1), we have

$$\sqrt{n}(\widehat{\theta}_q - \theta) \sim AMVN\left(0, I_q^{-1}(\theta)M_q(\theta)(I_q^{-1}(\theta))'\right) \qquad (2.4.9)$$

where $I_q(\theta) = -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{q,i}(\theta)}{\partial \theta'}\right]$, $M_q(\theta) = \frac{1}{n}\sum_{i=1}^{n} Cov(h_{q,i}(\theta))$ and the $h_{q,i}(\theta)$ are vectors of unbiased estimating equations defined as follows

$$h_{q,i}(\theta) = (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), g_i(\theta))' \qquad (2.4.10)$$

$$h_{0i}(\theta) = X_i'(\beta)R^{-1}(\lambda)Z_i$$

$$h_{1i}(\theta) = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\rho}}(Z_iZ_i' - \phi R(\lambda))\right]$$

$$h_{2i}(\theta) = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\alpha}}(Z_iZ_i' - \phi R(\lambda))\right]$$

$$g_i(\theta) = tr\left[R^{-1}(\lambda)Z_iZ_i'\right] - t\phi$$

where $\widetilde{\lambda}$ is the solution to the following equations

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\rho}}R(\lambda)\right] = 0$$

$$\Leftrightarrow \widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\left[\widetilde{B}_{21} + \widetilde{B}_{22}R_{21}\right] + \widetilde{B}_{12}\left[R_{21}\widetilde{B}_{12} + R_{22}\widetilde{B}_{22}\right]\frac{\partial \widetilde{R}_{21}}{\partial \rho} = 0$$

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\alpha}}R(\lambda)\right] = 0$$

$$\Leftrightarrow \widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{21} + tr\left[\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\left(\widetilde{B}_{21}\widetilde{B}_{12} + \widetilde{B}_{22}R_{22}\widetilde{B}_{22}\right)\right] = 0$$

Note that $\lambda = (\rho, \alpha)$ are the "true" values of the correlation parameters. By taking the negative expectation of the partial derivatives for (2.4.10) with respect to θ

and averaging over n we obtain $I_q(\theta)$, and by taking the covariance of (2.4.10) and averaging over n we obtain $M_q(\theta)$. From here it is easy to show that $I_q(\theta)$ has the following elements

$$I_q(\theta) = \begin{pmatrix} I_{11} & 0 & 0 & 0 \\ 0 & I_{22} & I_{23} & 0 \\ 0 & I_{32} & I_{33} & 0 \\ 0 & I_{42} & I_{43} & I_{44} \end{pmatrix} \tag{2.4.11}$$

where

$$\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{0i}(\theta)}{\partial \beta}\right) = -\frac{1}{2}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \rho}\right) = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\rho},\widetilde{\alpha})}{\partial \widetilde{\rho}} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= -2\phi\left[\frac{\partial R_{12}}{\partial \rho}\widetilde{B}_{21}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{21} + \widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{22}\frac{\partial R_{21}}{\partial \rho}\right] \\
I_{23} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \alpha}\right) = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \widetilde{\rho}} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -2\phi\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{22}\frac{\partial R_{22}}{\partial \alpha}\widetilde{B}_{21} \\
I_{32} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \rho}\right) = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \widetilde{\alpha}} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= -2\phi\frac{\partial R_{12}}{\partial \rho}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21} \\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \alpha}\right) = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \widetilde{\alpha}} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -\phi\, tr\left[\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}\frac{\partial R_{22}}{\partial \alpha}\right] \\
I_{42} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_i(\theta)}{\partial \rho}\right) = \phi\, tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}\right] = 2\phi\frac{\partial R_{12}}{\partial \rho}B_{21} \\
I_{43} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_i(\theta)}{\partial \alpha}\right) = \phi\, tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right] = \phi\, tr\left[B_{22}\frac{\partial R_{22}}{\partial \alpha}\right] \\
I_{44} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_i(\theta)}{\partial \phi}\right) = t.
\end{aligned}$$

We can also show that $M_q(\theta)$ has the following elements

$$M_q(\theta) = \begin{pmatrix} M_{11} & 0 & 0 & 0 \\ 0 & M_{22} & M_{23} & 0 \\ 0 & M_{32} & M_{33} & 0 \\ 0 & 0 & 0 & M_{44} \end{pmatrix} \qquad (2.4.12)$$

where

$$\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{0i}(\theta)) = \frac{\phi}{n}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta)) \\
&= 2\phi^2 tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\rho}}R^{-1}(\widetilde{\lambda})R(\lambda)R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\rho}}R^{-1}(\widetilde{\lambda})R(\lambda)\right] \\
M_{23} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), h_{2i}(\theta)) \\
&= 2\phi^2 tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\rho}}R^{-1}(\widetilde{\lambda})R(\lambda)R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\alpha}}R^{-1}(\widetilde{\lambda})R(\lambda)\right] \\
M_{32} &= M_{23} \\
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta)) \\
&= 2\phi^2 tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\alpha}}R^{-1}(\widetilde{\lambda})R(\lambda)R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\alpha}}R^{-1}(\widetilde{\lambda})R(\lambda)\right] \\
M_{44} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_i(\theta)) = 2t\phi^2.
\end{aligned}$$

Note that based on the forms (2.4.11) and (2.4.12), the covariance terms corresponding to β are zero, indicating that $\widehat{\beta}_q$ is uncorrelated with the estimators for the other parameters.

II.4.4 Comparison of Asymptotic Performance

Though all three estimating procedures yield consistent estimates of the correlation parameters, we want to compare their asymptotic performance. To do this we compute asymptotic relative efficiencies (ARE) over the admissible range described in

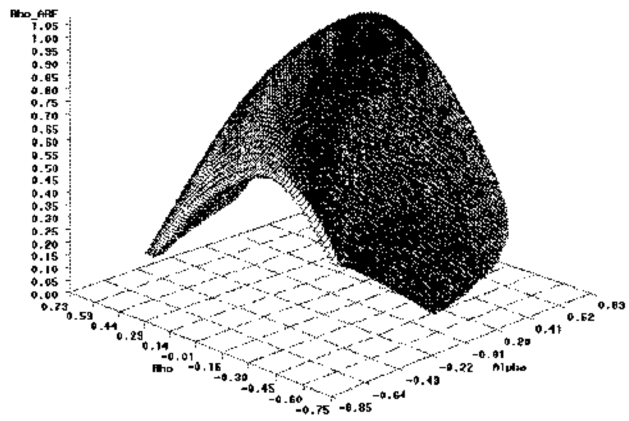

Figure 2.3: ρ ARE for MLE and MOM Methods

Section II.2 for all three estimation procedures. Implicit in this scenario is that observations are drawn from a multi-variate normal distribution.

To begin we set family size at $t = 4$, fix the number of families at $n = 1,000$, set $\phi = 3$ and let both ρ and α vary over the range shown in Figure 2.1. For each pair (ρ, α) we calculate the asymptotic variances derived in Sections II.4.1, II.4.2 and II.4.3 in order to compute ARE. As the more efficient estimator will have the smaller asymptotic variance, and since we are selecting a wide range of correlation values, these plots will show not only which estimating procedure is more efficient but also for which values of ρ and α this is the case.

First we find the ARE for estimators of ρ. We show the ARE for the MLE and MOM methods in Figure 2.3. Here we see the ARE is highest when ρ is close to zero, and the ARE drops sharply as ρ increases in magnitude. Note also that the 'crest' in the ARE plot is weakly slanted in a positive linear fashion. For the MLE and QLS methods, the ARE plot is found in Figure 2.4. Here we note that the ARE is very high over a wide range of ρ and α, indicating that the QLS estimator variance is almost as small as that for the MLE, and only for extreme correlation values close to the positive definite boundary does the efficiency of the QLS estimator decrease with respect to the MLE. Asymptotically, then, we see that QLS is comparable to MLE for most plausible correlation values, though the MLE is slightly better (which

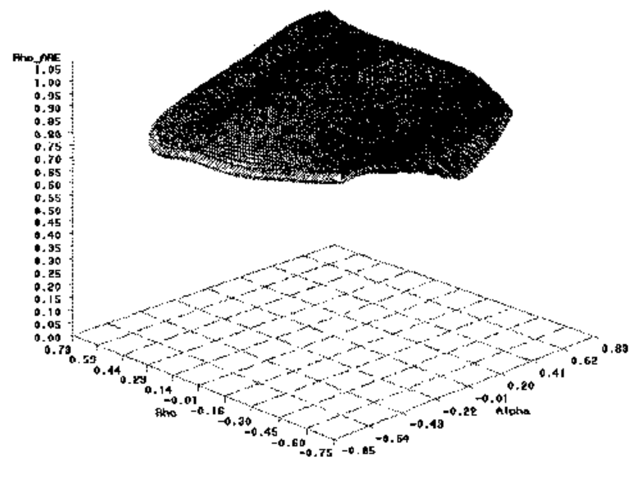

Figure 2.4: ρ ARE for MLE and QLS Methods

must be the case). Lastly we compare the relative efficiences for the MOM and QLS methods, the plot of which is found in Figure 2.5. Here we see a similar pattern to that found in comparing the MLE and MOM methods, that the ARE is highest when ρ is close to zero and falls sharply as you move away from zero. This shows that, like MLE, QLS is asymptotically superior to MOM.

We next compute ARE for the α estimators. Starting with the MLE and MOM methods, we find the ARE plot in Figure 2.6. Here the ARE is highest when α is closest to zero, and the ARE drops quickly as α increases in magnitude. This shows that MLE is superior to MOM. For the MLE and QLS methods, the ARE plot is found in Figure 2.7. In this Figure we see the ARE is highest when α is close to zero, and then slightly decreases as α moves away from zero. As was the case for the ρ estimators, we see that the ARE is high over a wide portion of the admissible range, showing that the variance of the QLS estimator is almost as small as the variance of the MLE. Asymptotically, then, we see that QLS is comparable to MLE for most correlation values. Lastly we compute the relative efficiences for the MOM and QLS methods, the plot of which is found in Figure 2.8. This plot shows a similar pattern to that found in the MLE/MOM case, that the ARE is highest when α is close to zero and falls steadily as you move away from zero. Thus QLS is also superior to MOM.

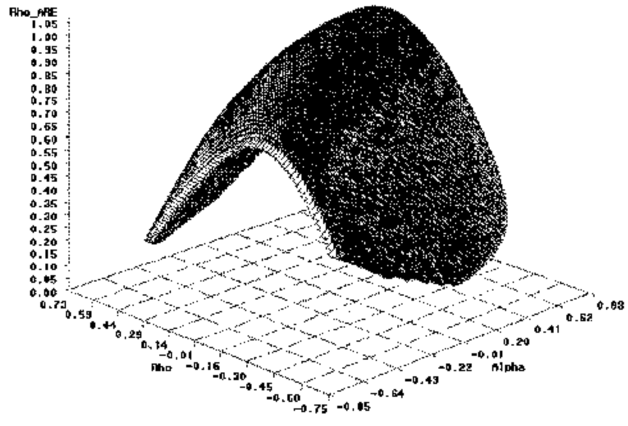

Figure 2.5: ρ ARE for MOM and QLS Methods

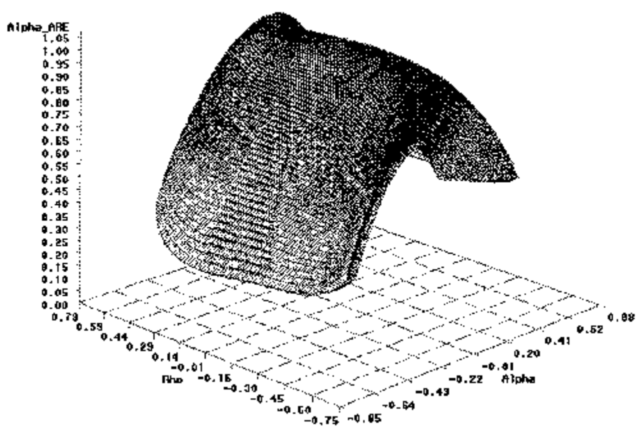

Figure 2.6: α ARE for MLE and MOM Methods

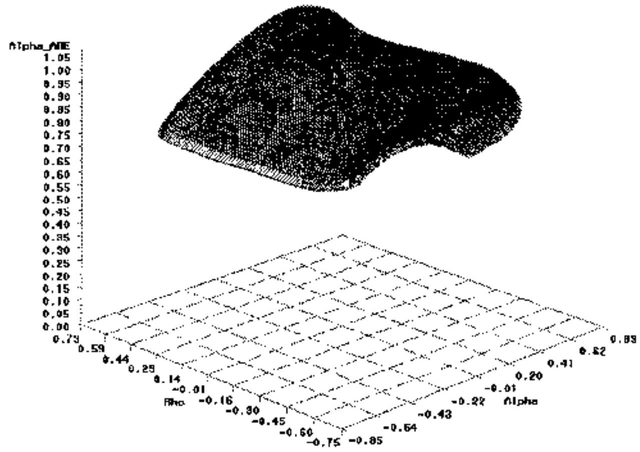

Figure 2.7: α ARE for MLE and QLS Methods

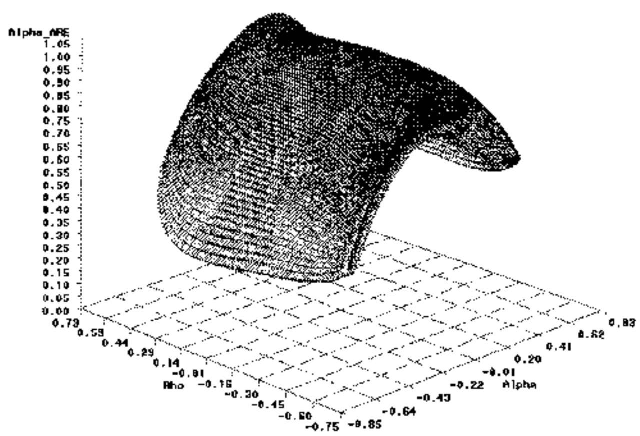

Figure 2.8: α ARE for MOM and QLS Methods

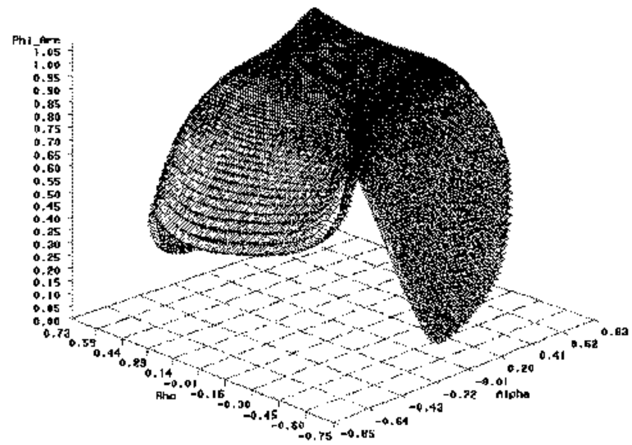

Figure 2.9: ϕ ARE for MLE and MOM Methods

We include ARE plots for the variance parameter as well. Figure 2.9 shows the ARE of ϕ for the MLE and MOM procedures, Figure 2.10 shows the ARE for the MLE and QLS procedures, and Figure 2.11 shows the ARE for the MOM and QLS procedures. These Figures show that the QLS variance estimator is good competitor with the MLE estimator, as the ARE is close to one over most of the admissible range. The MOM variance estimator is a good competitor to both the MLE and QLS estimators over a much narrower region of ρ and α.

II.5 Small-Sample Performance

In the small-sample case, our goal is two-fold. We first gauge the small-sample efficiency for each method under the assumption of normally distributed data, and second we gauge the efficiency when the data are not normally distributed (i.e. when the data come from a skewed or otherwise distinctly non-normal distribution). This later case will shed light not only on efficiency but also on the robustness of each method to departures from normality.

For both cases we fix sample size at $n = 30$, keep family size at $t = 4$ and set $\phi = 3$. We then simulate 1000 such samples for each of many combinations of ρ and α (which vary over their admissible range), and for each sample we estimate the parameters.

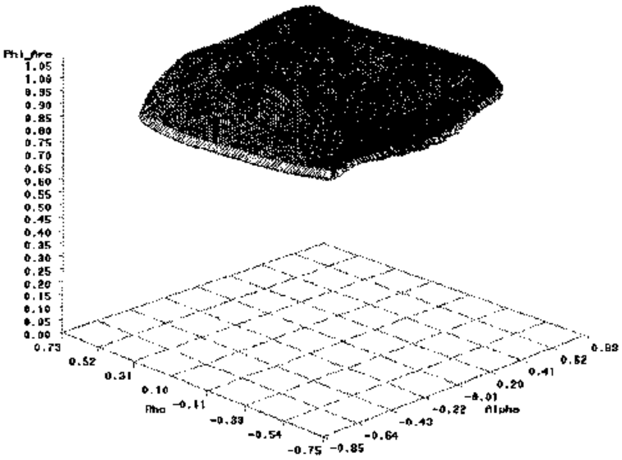

Figure 2.10: ϕ ARE for MLE and QLS Methods

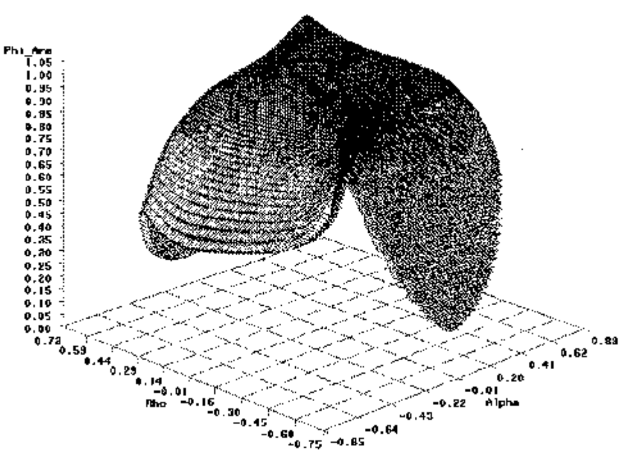

Figure 2.11: ϕ ARE for MOM and QLS Methods

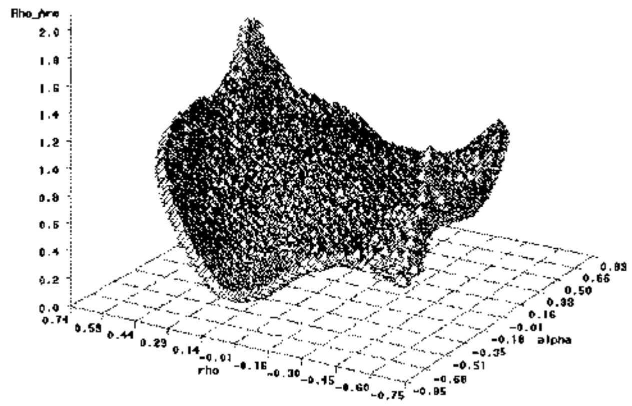

Figure 2.12: ρ RE for MLE and MoM Methods with Normal Data

We then calculate the average squared deviation of the estimated parameter value from the "true" population values. The ratio of the estimated averages for any two estimating procedures is our estimate of small-sample relative efficiency.

II.5.1 Small-Sample Normal Case

We begin with the esimators for ρ. For MLE and MoM procedures, we get the results found in Figure 2.12. Note that the RE is greater than 1 in some places, indicating that for these values, the MoM estimator has smaller estimated variance than the MLE estimator. For most values, however, MLE is still more efficient than MoM. Figure 2.13 shows the relative efficiencies for the MLE and QLS methods. The RE is greater than 1 in some places, notably for large positive and large negative values of ρ and for small values of α. Here we see that QLS is a much better competitor to the MLE. Lastly we compare the QLS and MOM methods, the results of which are found in Figure 2.14. Like the MLE-MoM case, the RE is small for most values of ρ, with the variance for the MoM method being smaller than the variance for the QLS method only for extremely large positive correlation values. Thus, for ρ estimators, we see that QLS is a much better competitor with MLE, and both MLE and QLS are still mostly superior to MoM.

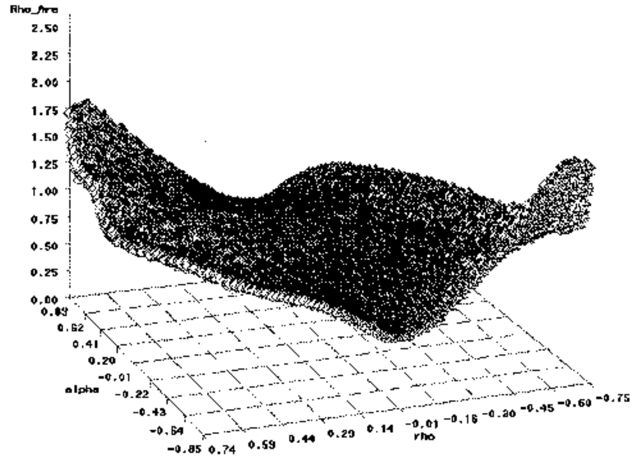

Figure 2.13: ρ RE for MLE and QLS Methods with Normal Data

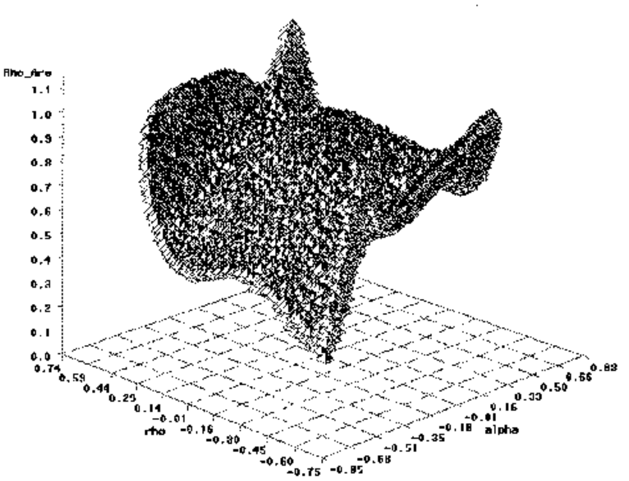

Figure 2.14: ρ RE for QLS and MoM Methods with Normal Data

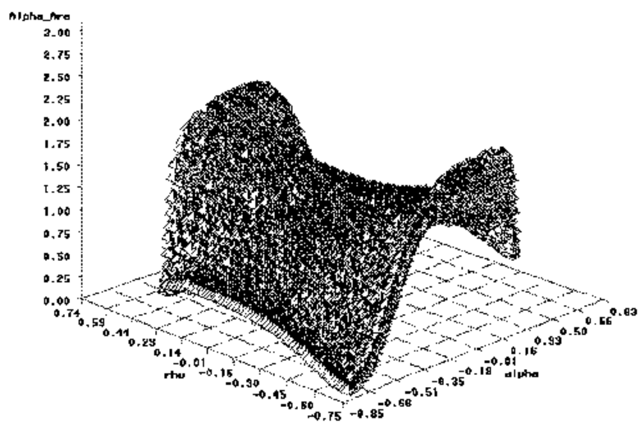

Figure 2.15: α RE for MLE and MoM Methods with Normal Data

We now move on to estimators of α in the small-sample normal case. For the MLE and MoM procedures, we get the results found in Figure 2.15. Here we notice that the RE is greater than 1 for small values of α and is actually high for extreme values of ρ. For some values of ρ, the MoM estimator is more than twice as efficient than the MLE, though this occurs close to the positive definite boundary. We also note that the efficiency of MoM decreases as α increases in magnitude. Figure 2.16 shows the relative efficiencies for the MLE and QLS methods, which resembles the saddle shape found in Figure 2.15. Here the RE is greater than 1 over a wide range of ρ when α is small and for large values of ρ. Only for moderately large values of α is the MLE more efficient than the QLS estimator. Thus we see that QLS is a much better competitor to the MLE in this situation. Lastly we compare the QLS and MOM methods, with the results found in Figure 2.17. This plot is similar to the MLE-MoM plot, noting that the QLS estimator is more efficient than the MoM estimator for most correlation values. Thus, in the small-sample normal case for estimators of α, we see that QLS is a much better competitor with MLE, and both MLE and QLS are better than MoM, though not as much as in the asymptotic case.

Lastly we estimate the small-sample relative efficiencies for estimators of ϕ. Figure 2.18 contains the RE for the MLE and MoM estimators. Here we see the relative efficiency is close to 1 only for very small values of ρ and α, and that the RE quickly

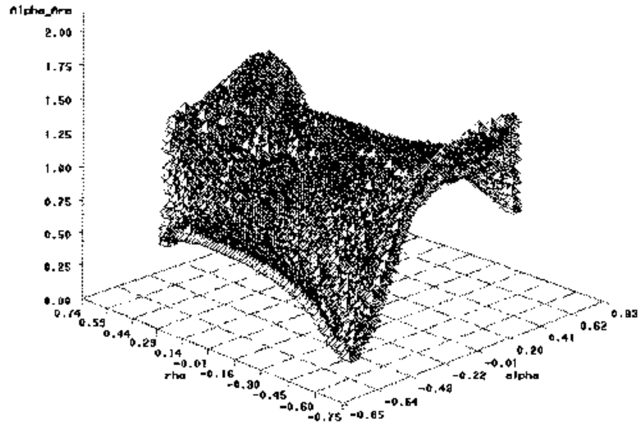

Figure 2.16: α RE for MLE and QLS Methods with Normal Data

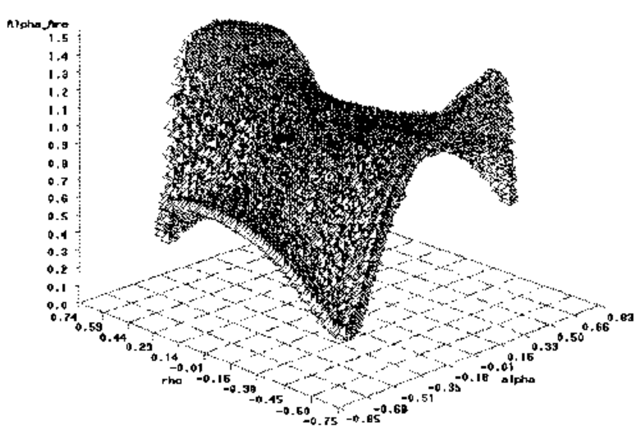

Figure 2.17: α RE for QLS and MoM Methods with Normal Data

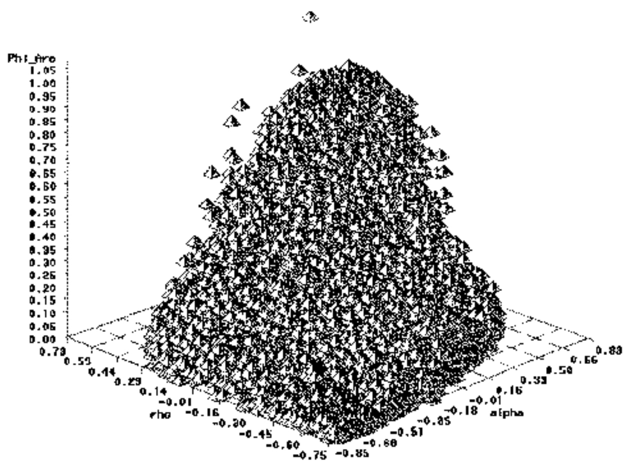

Figure 2.18: ϕ RE for MLE and MoM Methods with Normal Data

decreases as correlation increases in magnitude. This shows that the MLE variance estimator is mostly superior to the MoM estimator. In Figure 2.19 we have the RE for the MLE and QLS estimators. In this plot we see that the estimated relative efficiencies are comparable for most values of ρ and α, and for vlaues close to the positive definite boundary, the QLS estimator is more efficient than the MLE. And finally, Figure 2.20 has the RE for the QLS and MoM estimators. This plot shows that, as was the case for MLE-MoM, the RE is close to 1 only for very small values of ρ and α, and the RE decreases quickly as ρ and α increase in magnitude. Thus, in the small-sample normal case, we see that the QLS variance estimator is at least as good as the MLE and much better than the MoM variance estimator. The MLE variance estimator is also more efficient than the MoM estimator.

Table 2.2 provides estimated infeasibility probabilities, or the probability that each estimating procedure yields correlation estimates outside the positive definite range. Using the same simulation procedure, we compute the estimated probabilities as the number of times the procedure failed to a provide an estimate within the admissible range divided by the total number of simulations (1,000). Note that N/A indicates that those parameter values are outside the positive definite boundary. From this Table it is clear that the QLS procedure has an extremely low probability of producing inadmissible correlation estimates over the entire range of ρ and α. The

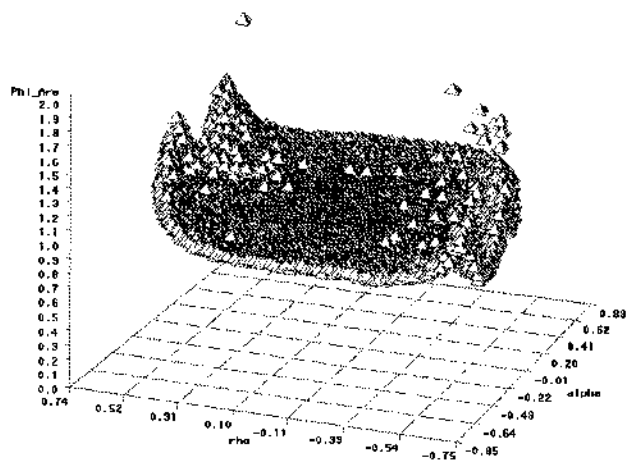

Figure 2.19: ϕ RE for MLE and QLS Methods with Normal Data

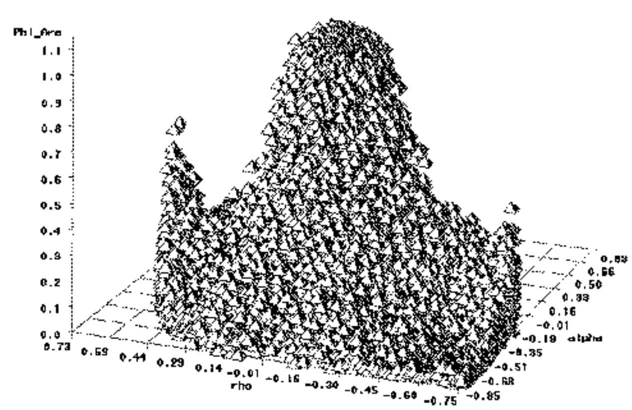

Figure 2.20: ϕ RE for QLS and MoM Methods with Normal Data

Table 2.2: Estimated Infeasibility Probabilities (Normal, Homogeneous Variance Case)

		α				
ρ	Method	-0.80	-0.70	0.3	0.70	0.80
-0.75	MLE	0.541	0.473	N/A	N/A	N/A
	MoM	0.376	0.081	N/A	N/A	N/A
	QLS	0.003	0.001	N/A	N/A	N/A
-0.60	MLE	0.137	0.182	N/A	N/A	N/A
	MoM	0.153	0.061	N/A	N/A	N/A
	QLS	0.008	0.001	N/A	N/A	N/A
0.10	MLE	0.016	0.018	0.019	0.004	0.011
	MoM	0.004	0.002	0.000	0.001	0.000
	QLS	0.002	0.006	0.000	0.003	0.001
0.60	MLE	N/A	N/A	0.195	0.138	0.090
	MoM	N/A	N/A	0.009	0.049	0.171
	QLS	N/A	N/A	0.006	0.000	0.020
0.70	MLE	N/A	N/A	0.462	0.343	0.351
	MoM	N/A	N/A	0.009	0.125	0.307
	QLS	N/A	N/A	0.036	0.001	0.002

MLE and MOM procedures, though competitive for moderate parameter values, have high inadmissible probabilities for large values of ρ and α.

So in the small sample normal case, we see that the QLS procedure is much more competitive with the MLE procedure then they were in the asymptotic case for estimators of both ρ and α. Only for moderate values does the MLE method give the smallest variance in estimating ρ, while for α the QLS method gives the smallest variance when α takes moderate values, while the MLE method gives the smallest variance if α takes more extreme values. For estimators of ϕ, QLS is at least as good as MLE and is better than MoM. Though MoM is inferior to the other methods for all three parameters, it is a better competitor against the other methods for the correlation estimators.

II.5.2 Small-Sample Non-Normal Case

Here the goal is to estimate small-sample efficiencies when the Y_i's are drawn from a non-normally distributed population. This will help us gauge the robustness of the estimating procedures to departures from normality. Following the methodology used in Chaganty and Shi (2004), we simulate random observations from a beta distribution with $\alpha = \beta = \frac{1}{6}$. These parameter values result in a U-shaped pdf and

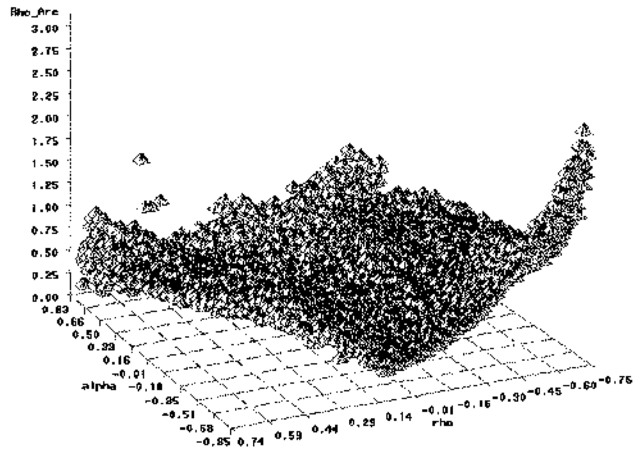

Figure 2.21: ρ RE for MLE and MoM Methods with Non-Normal Data

thus yield non-normal random variates. Since correlation is both scale and location invariant, these simulations pose no problem for ρ and α. However, since variance is not scale invariant, we will not include ϕ in this portion of the analysis. We again fix $t = 4$ and $n = 30$ and simulate $1,000$ such samples for each choice of ρ and α.

We begin with the ρ estimators. Comparing first the MLE and MoM methods we get the RE plot in Figure 2.21. Here we see the MoM procedure is more efficient for extreme values of ρ and α, as well as for a large range of positive α. Elsewhere the MLE is more efficient. We next compare the MLE and QLS methods in Figure 2.22. We see that the two procedures are fairly comparable for some values of ρ and α, with the QLS procedure performing much better for large ρ, and especially for large α where we see a spike in the efficiencies. Here the estimated variance of the MLE is around 4 times as large as the estimated variance for the QLS estimator. Lastly we compare the QLS and MoM procedures in Figure 2.23. Here we see that the QLS method in general has smaller estimated variance than the MoM method except for extreme values of ρ and α.

We now move on to estimators of α. Comparing first the MLE and MoM methods we get the RE plot in Figure 2.24. Here we see that the MLE procedure has smaller relative efficiency when α is large positive and large negative. However, The MoM procedure is comparable when α is close to zero and is better for extremely large ρ.

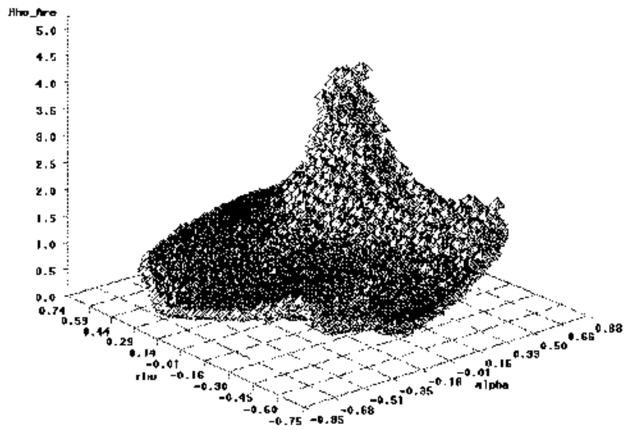

Figure 2.22: ρ RE for MLE and QLS Methods with Non-Normal Data

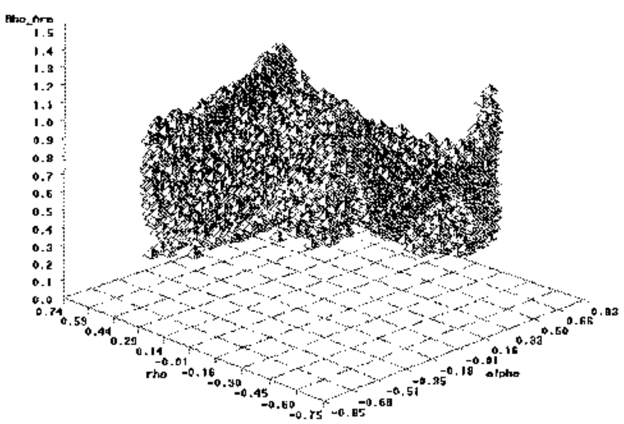

Figure 2.23: ρ RE for QLS and MoM Methods with Non-Normal Data

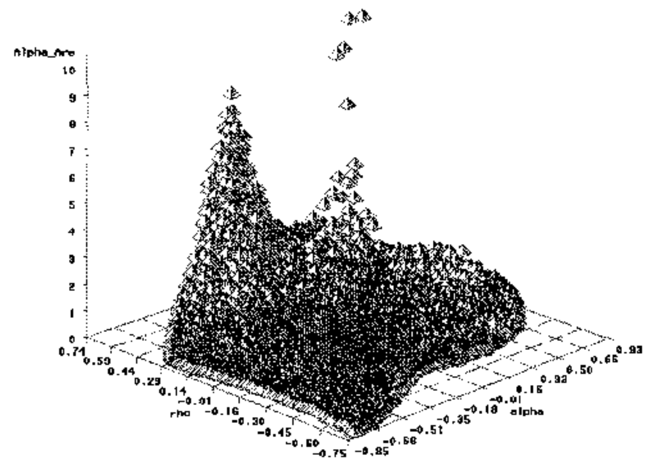

Figure 2.24: α RE for MLE and MoM Methods with Non-Normal Data

We next compare the MLE and QLS methods in Figure 2.25. Here we see that the MLE procedure preforms better for extreme values of α, while the QLS procedure performs better for small to moderate values of α and especially for large ρ, where the QLS estimator vastly outperforms the MLE. Lastly we compare the QLS and MoM procedures in Figure 2.26. Here the two methods are comparable when α is close to zero, QLS is better for moderate and large values of α, and the MoM procedures is better when both α and ρ are large and positive.

Lastly we estimate infeasibility probabilities for each estimation procedure. Table 2.3 shows the estimates of these probabilities over a wide range of ρ and α. Here we see that the QLS procedure has low error probabilities for all values of ρ and α. The MLE procedure is competitive with QLS for small values, yet performs poorly for large values, while the MOM procedure is nowhere competitive.

So in the small-sample non-normal case we see that QLS is now outperforming the MLE procedure for most values of ρ and α, both with regards to estimated efficiency and estimated infeasibility probability. The MoM procedure is also much more competitive with the MLE procedure, though not as much with QLS.

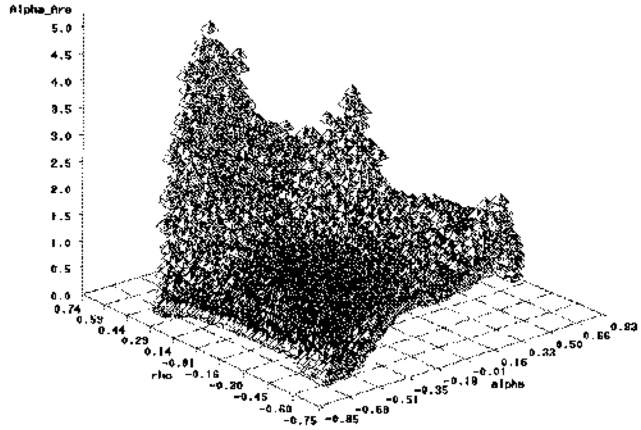

Figure 2.25: α RE for MLE and QLS Methods with Non-Normal Data

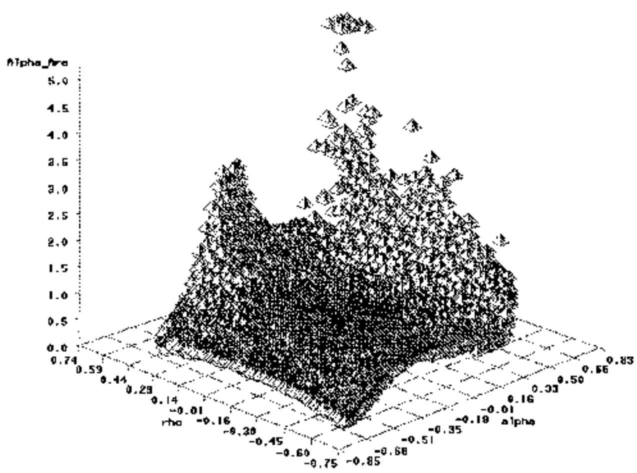

Figure 2.26: α RE for QLS and MoM Methods with Non-Normal Data

Table 2.3: Estimated Infeasibility Probabilities (Non-Normal, Homogeneous Variance Case)

			α			
ρ	Method	-0.80	-0.70	0.3	0.70	0.80
-0.75	MLE	0.783	0.665	N/A	N/A	N/A
	MoM	0.499	0.150	N/A	N/A	N/A
	QLS	0.000	0.004	N/A	N/A	N/A
-0.60	MLE	0.303	0.263	N/A	N/A	N/A
	MoM	0.253	0.082	N/A	N/A	N/A
	QLS	0.000	0.000	N/A	N/A	N/A
0.10	MLE	0.137	0.129	0.046	0.000	0.000
	MoM	0.611	0.462	0.402	0.738	0.578
	QLS	0.002	0.002	0.000	0.000	0.000
0.60	MLE	N/A	N/A	0.114	0.811	0.525
	MoM	N/A	N/A	0.442	1.000	1.000
	QLS	N/A	N/A	0.001	0.000	0.001
0.70	MLE	N/A	N/A	0.679	0.897	0.943
	MoM	N/A	N/A	0.022	1.000	1.000
	QLS	N/A	N/A	0.041	0.001	0.001

II.6 Hypothesis Testing

In this Section we develop hypothesis testing procedures involving the correlation parameters for each estimating method. We develop hypothesis tests for general functions of the correlation parameters, and then concentrate upon specific examples and compare their performance through simulation.

II.6.1 Likelihood Ratio Test

Under maximum likelihood estimation we are assuming the data are normally distributed. Knowledge of the multivariate normal likelihood function allows us to utilize a likelihood ratio test for hypothesis tests regarding the correlation parameters.

Generally, we test a null hypothesis that some function of the correlation parameters ($h(\lambda)$) is equal to some constant, or $H_o : h(\lambda) = c$. To do this, we take the ratio of the likelihood evaluated with the maximum likelihood estimates under H_o (the restricted MLE's) against the likelihood evaluated with the so-called unrestricted maximum likelihood estimates. Let $\widehat{\theta}_o = (\widehat{\beta}_o, \widehat{\lambda}_o, \widehat{\phi}_o)$ be the restricted and

$\widehat{\theta} = (\widehat{\beta}, \widehat{\lambda}, \widehat{\phi})$ the unrestricted MLE's, respectively. Then the likelihood ratio test statistic is

$$\begin{aligned}
\lambda(\theta) &= \frac{\prod_{i=1}^{n} f_i(y_i|\widehat{\theta}_o)}{\prod_{i=1}^{n} f_i(y_i|\widehat{\theta})} \\
&= \frac{(2\pi)^{-\frac{nt}{2}}(\widehat{\phi}_o)^{-\frac{nt}{2}}|R(\widehat{\lambda}_o)|^{-\frac{n}{2}} exp\left(-\frac{1}{2\widehat{\phi}_o}\sum_{i=1}^{n}(Y_i - X_i\widehat{\beta}_o)'R^{-1}(\widehat{\lambda}_o)(Y_i - X_i\widehat{\beta}_o)\right)}{(2\pi)^{-\frac{nt}{2}}(\widehat{\phi})^{-\frac{nt}{2}}|R(\widehat{\lambda})|^{-\frac{n}{2}} exp\left(-\frac{1}{2\widehat{\phi}}\sum_{i=1}^{n}(Y_i - X_i\widehat{\beta})'R^{-1}(\widehat{\lambda})(Y_i - X_i\widehat{\beta})\right)}
\end{aligned}$$

Note that $\widehat{\phi}_o = \frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i\widehat{\beta}_o)'R^{-1}(\widehat{\lambda}_o)(Y_i - X_i\widehat{\beta}_o)$ and $\widehat{\phi} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i\widehat{\beta})'R^{-1}(\widehat{\lambda})(Y_i - X_i\widehat{\beta})$, so that

$$\lambda(\theta) = \left(\frac{\widehat{\phi}_o}{\widehat{\phi}}\right)^{-\frac{nt}{2}} \left(\frac{|R(\widehat{\lambda}_o)|}{|R(\widehat{\lambda})|}\right)^{-\frac{n}{2}}$$

Recall under the central limit theorem that $-2\ln(\lambda(\theta))$ has an asymptotic chi-square distribution with $d = d_{ur} - d_r$ degrees of freedom, where d_{ur} is the number of parameters in the unrestricted model and d_r is the number of parameters under H_o. Thus, the test statistic becomes

$$-2\ln(\lambda(\theta)) = nt\left(\ln(\widehat{\phi}_o) - \ln(\widehat{\phi})\right) + n\left(\ln|R(\widehat{\lambda}_o)| - \ln|R(\widehat{\lambda})|\right). \qquad (2.6.1)$$

The most obvious special cases for the correlation parameters are $H_o : \rho = 0$ and $H_o : \alpha = 0$. For testing $H_o : \rho = 0$, we note that $R(\widehat{\lambda}_o) = R(0, \widehat{\alpha}_o)$. The determinant of the correlation matrix is simply the determinant of a $t-1$ by $t-1$ autoregressive matrix, or $(1-\widehat{\alpha}_o^2)^{t-2}$. Recalling the determinant of $R(\lambda)$ under the full model (2.2.2), we get the following likelihood ratio test statistic for $H_o : \rho = 0$

$$\begin{aligned}
-2\ln(\lambda(\theta)) &= nt\left(\ln(\widehat{\phi}_o) - \ln(\widehat{\phi})\right) + n(t-2)\ln\left(1 - \widehat{\alpha}_o^2\right) \qquad (2.6.2) \\
&- n(t-3)\ln(1-\widehat{\alpha}^2) + n\ln(1-\widehat{\rho}^2) \\
&- n\ln\left((1-\widehat{\alpha}^2)(1-\widehat{\rho}^2) - (\widehat{\rho}^2 - \widehat{\rho}^{2t}) - \widehat{\alpha}^2(\widehat{\rho}^4 - \widehat{\rho}^{2t-2}) + 2\widehat{\alpha}(\widehat{\rho}^3 - \widehat{\rho}^{2t-1})\right).
\end{aligned}$$

Since the difference in the number of parameters between $\widehat{\theta}_o = (\widehat{\beta}_o, 0, \widehat{\alpha}_o, \widehat{\phi}_o)'$ and $\widehat{\theta} = (\widehat{\beta}, \widehat{\rho}, \widehat{\alpha}, \widehat{\phi})'$ is 1, then 2.6.2 is asymptotically χ_1^2.

For testing $H_o : \alpha = 0$, it can be shown that $|R(\rho, 0)| = 1 - \frac{\rho^2 - \rho^{2t}}{1-\rho^2}$. Thus the likelihood ratio test statistic for this null hypothesis is

$$\begin{aligned}
-2\ln(\lambda(\theta)) &= nt\left(\ln(\widehat{\phi}_o) - \ln(\widehat{\phi})\right) + n(t-2)\ln\left(1 - \frac{\widehat{\rho}_o^2 - \widehat{\rho}_o^{2t}}{1-\widehat{\rho}_o^2}\right) \qquad (2.6.3) \\
&- n(t-3)\ln(1-\widehat{\alpha}^2) + n\ln(1-\widehat{\rho}^2) \\
&- n\ln\left((1-\widehat{\alpha}^2)(1-\widehat{\rho}^2) - (\widehat{\rho}^2 - \widehat{\rho}^{2t}) - \widehat{\alpha}^2(\widehat{\rho}^4 - \widehat{\rho}^{2t-2}) + 2\widehat{\alpha}(\widehat{\rho}^3 - \widehat{\rho}^{2t-1})\right)
\end{aligned}$$

which is asymptotically χ^2_1.

II.6.2 Wald's Test

Both the Method of Moment and Quasi-Least Squares procedures employ quasi-log-likelihood functions, as opposed to proper log-likelihood functions, so the likelihood ratio test is not available for these methods. However, we have derived the asymptotic variances for the MOM and QLS estimators, $\widehat{\theta}_m$ and $\widehat{\theta}_q$ respectively ((2.4.7) through (2.4.12)).

Wald's Test states that for testing the null hypothesis $H_o : h(\theta) = 0$, where $h(\theta)$ is a function of θ (possibly vector valued) and $\widehat{\theta}$ is an estimator with asymptotic variance $I^{-1}(\theta)$, $I(\theta)$ being Fisher's Information matrix, the test statistic

$$T = nh(\widehat{\theta})^T \left[H(\widehat{\theta})^T I^{-1}(\widehat{\theta}) H(\widehat{\theta})\right]^{-1} h(\widehat{\theta}) \tag{2.6.4}$$

has a chi-square distribution with $d = rank(h(\theta))$ degrees of freedom. Here $H(\theta) = \frac{\partial h(\theta)}{\partial \theta}$ is a vector (matrix) of partial derivatives of $h(\theta)$ with respect to θ. Recalling that though we cannot calculate $I(\theta)$ for the MoM and QLS cases, we have derived their asymptotic variances $I_n^{-1}(\theta) M_n(\theta) (I_n^{-1}(\theta))^T$, so that the so-called Wald-Type test statistic becomes

$$T = nh(\widehat{\theta})^T \left[H(\widehat{\theta})^T I_n^{-1}(\widehat{\theta}) M_n(\widehat{\theta}) (I_n^{-1}(\widehat{\theta}))^T H(\widehat{\theta})\right]^{-1} h(\widehat{\theta}) \tag{2.6.5}$$

which again has an asymptotic chi-square distribution with $d = rank(h(\theta))$ degrees of freedom.

For instance, if we test $H_o : \rho = 0$, then $h(\theta) = \rho$, $H(\theta) = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}^T$, and our test statistic is

$$T = n\widehat{\rho}^2 \left[\begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} I_n^{-1}(\widehat{\theta}) M_n(\widehat{\theta}) (I_n^{-1}(\widehat{\theta}))^T \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right]^{-1} \tag{2.6.6}$$

which is asymptotically χ^2_1. Likewise, if we were to test $H_o : \alpha = 0$, then $h(\theta) = \alpha$,

$H(\theta) = (\ 0 \quad 0 \quad 1 \quad 0\ )^T$, and our test statistic is

$$T = n\widehat{\alpha}^2 \left[(\ 0 \quad 0 \quad 1 \quad 0\ ) I_n^{-1}(\widehat{\theta}) M_n(\widehat{\theta}) (I_n^{-1}(\widehat{\theta}))^T \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right]^{-1} \tag{2.6.7}$$

which is also asymptotically χ_1^2.

For the MOM estimator recall that we defined the asymptotic variance matrices $I_n(\theta)$ and $M_n(\theta)$ in (2.4.7) and (2.4.8), respectively. If we are testing $H_o : \rho = 0$, then we substitute $\widehat{\theta}_m$, $I_m(\widehat{\theta}_m)$ and $M_m(\widehat{\theta}_m)$ into (2.6.6) to get our estimated Wald-Type test statistic.

$$\widehat{T}_{m,\rho=0} = \frac{n\widehat{\rho}_m^2 \widehat{\phi}_m^2}{\widehat{M}_{22}} = \frac{n\widehat{\rho}_m^2}{2tr\left[A(\widehat{\rho}_m)R(\widehat{\lambda}_m)A(\widehat{\rho}_m)R(\widehat{\lambda}_m)\right]} \tag{2.6.8}$$

where $\widehat{\lambda}_m = (\widehat{\rho}_m, \widehat{\alpha}_m)$ are the MOM estimates of the correlation parameters. Similarly, for testing $H_o : \alpha = 0$, we get

$$\widehat{T}_{m,\alpha=0} = \frac{n\widehat{\alpha}_m^2 \widehat{\phi}_m^2}{\widehat{M}_{33}} = \frac{n\widehat{\alpha}_m^2}{2tr\left[A(\widehat{\alpha}_m)R(\widehat{\lambda}_m)A(\widehat{\alpha}_m)R(\widehat{\lambda}_m)\right]} \tag{2.6.9}$$

Note that both $A(\rho)$ and $A(\alpha)$ are previously defined ((2.3.8) and (2.3.9), respectively).

For the QLS procedure, recall that we defined the asymptotic variance matrices $I_n(\theta)$ and $M_n(\theta)$ in (2.4.11) and (2.4.12), respectively. If we are testing $H_o : \rho = 0$, then we substitute $\widehat{\theta}_q$, $I_q(\widehat{\theta}_q)$ and $M_q(\widehat{\theta}_q)$ into (2.6.6) to get our estimated Wald-Type test statistic.

$$\widehat{T}_{q,\rho=0} = n\widehat{\rho}_q^2 \left[\frac{(\widehat{I}_{22}\widehat{I}_{33} - \widehat{I}_{23}\widehat{I}_{32})^2}{\widehat{M}_{22}\widehat{I}_{33}^2 + \widehat{M}_{33}\widehat{I}_{23}^2} \right] \tag{2.6.10}$$

where I_{ij} and M_{ij} are the ijth elements of (2.4.11) and (2.4.12), respectively, evaluated at $\widehat{\theta}_q$. Similarly, if we are testing $H_o : \alpha = 0$, we get

$$\widehat{T}_{q,\alpha=0} = n\widehat{\alpha}_q^2 \left[\frac{(\widehat{I}_{22}\widehat{I}_{33} - \widehat{I}_{23}\widehat{I}_{32})^2}{\widehat{M}_{22}\widehat{I}_{32}^2 + \widehat{M}_{33}\widehat{I}_{22}^2} \right]. \tag{2.6.11}$$

For both of these hypotheses, the test statistics have an asymptotic chi-square distribution with one degree of freedom.

II.6.3 Estimated Significance Levels

To gauge the performance of the Likelihood Ratio Test for the Maximum Likelihood Estimating procedure and the Wald-Type tests for the Method of Moment and Quasi-Least Squares procedures, we make use of simulations to estimate significance levels. Depending on the hypothesis of interest, we either set $\rho = 0$ or $\alpha = 0$ and fix the other correlation parameter at some admissible value. Note that when $\alpha = 0$, the admissible range for ρ is $(-0.68, 0.74)$, and when $\rho = 0$, the admissible range for α is $(-1, 1)$. For each combination of ρ and α, we simulate $n = 30$ observations of size $t = 4$ with $\phi = 3$ from a multivariate normal distribution. For the likelihood ratio test we calculate both the restricted and unrestricted maximum likelihood estimators, which should be similar as the simulated data reflect the conditions stated in the null hypothesis. For the Wald-Type test, we calculate the the method of moment and quasi-least squares estimators using the data and use these to calculate the asymptotic variances. Since we are simulating data assuming the null hypothesis is true, we expect to not reject the null hypothesis. However, due to randomness there is a chance that the simulated data will yield estimates that will cause us to reject H_o. Recall that for each test we reject H_o if the test statistic is greater than a chi-square critical value χ^2_1 for a particular significance level. If we choose a significance level of 0.05, then the critical value is 3.841. If we repeat these simulations a large number times (5,000) for a particular value of the non-zero correlation parameter, then the estimated significance level of the test is the ratio of the number of times we reject the null hypothesis to the total number of repeated simulations. If we then repeat this procedure over a wide range of values for the non-zero correlation parameter, we get an idea of how the test performs in many scenarios.

Based on the Law of Large Numbers, we expect the estimated significance level to be close the chosen level (0.05) if the estimating procedure is providing accurate estimates. The variance of the estimators increases with the absolute size of the correlation parameters, which means that we should on average reject the null hypothesis more often for large values of the correlation parameters than we would for small values. Thus, the estimated significance level should be small for small correlation values and larger for large correlation values.

We begin by analyzing $H_o : \alpha = 0$. The results for the likelihood ratio test

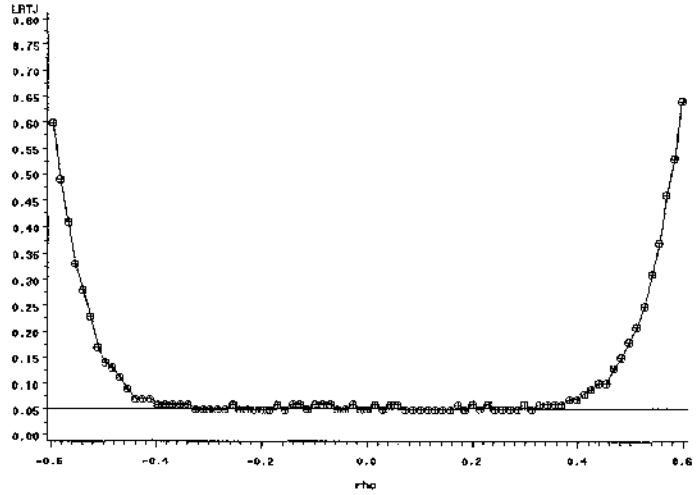

Figure 2.27: $H_o : \alpha = 0$ – Estimated Significance Level for LRT (MLE)

(LRT) are seen in Figure 2.27. From this we clearly see that the LRT yields an estimated significance level close to the actual significance level over a wide range of ρ. However, as ρ approaches its positive definite boundary the estimated significance level increases dramatically, as expected. This shows that for testing $H_o : \alpha = 0$, the likelihood ratio test works well for small to moderate values of ρ, but not for large values. Figure 2.28 shows the estimated significance level for MoM using the Wald-Type test. Here we see the same general pattern shown for the LRT, with an estimated significance level approximately equal to 0.05 over a wide range of ρ. Again, the level increases as the magnitude of ρ increases, but not as much as in the LRT. Lastly, Figure 2.29 shows the estimated significance level for QLS using the Wald-Type test. It is clear from this plot that the estimated significance level is close to 0.05 for a wide range of ρ, and then increases as the magnitude of ρ increases for moderately large values. Note that plots for the MLE and QLS are very similar. Thus, for testing $H_o : \alpha = 0$, all three tests (LRT, and Wald-Type test for both MoM and QLS) perform similarly.

Now we concentrate on the $H_o : \rho = 0$. Figure 2.30 shows the results of the LRT, and we see that the estimated significance level is vaguely U-shaped, centered at small values of α. However, noting the range of the estimates, we see that the estimated significance levels are close to 0.05 for most α values. This shows that

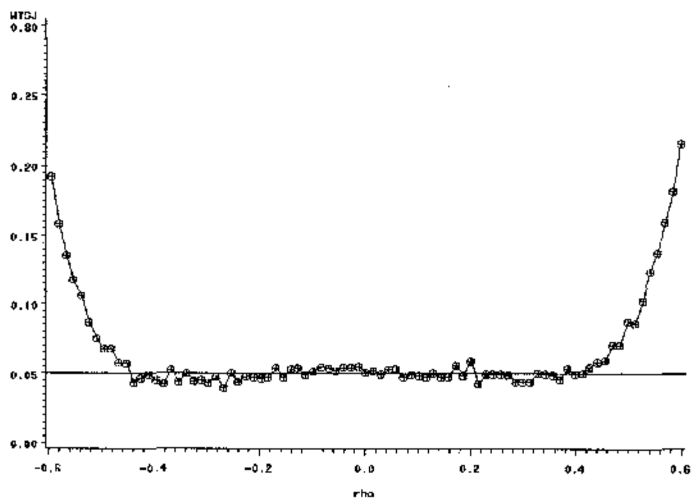

Figure 2.28: $H_o : \alpha = 0$ – *Estimated Significance Level for Wald-Type test (MoM)*

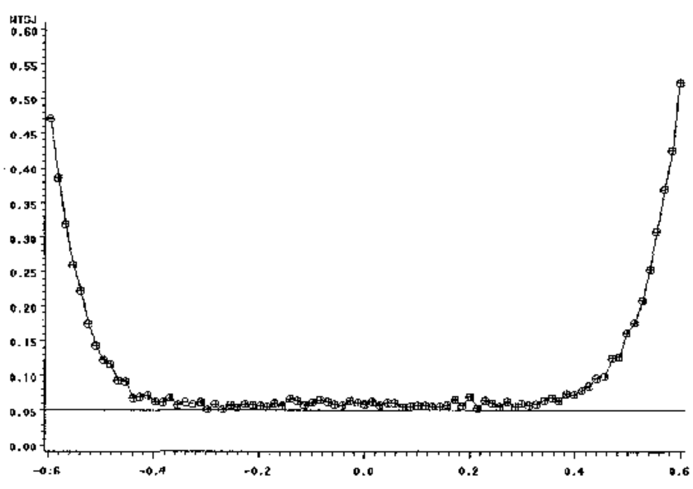

Figure 2.29: $H_o : \alpha = 0$ – *Estimated Significance Level for Wald-Type test (QLS)*

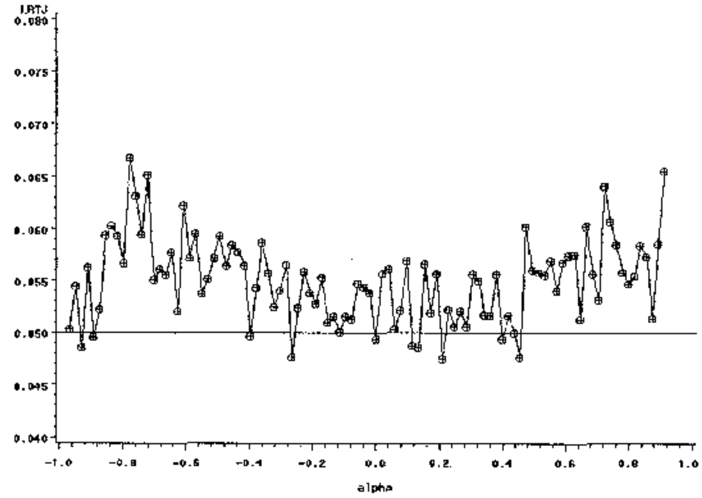

Figure 2.30: $H_o : \rho = 0$ – Estimated Significance Level for LRT (MLE)

the LRT is a very strong test for this hypothesis. Figure 2.31 shows the results for MoM using the Wald-Type test. Here we see that the estimated significance level is accurate for small values of α, but it then increases as the magnitude of α increases. Finally, Figure 2.32 gives the estimated significance levels for the QLS procedure using the Wald-Type test. Here we see that the estimated significance levels are high for small levels of α (around 0.10), but then decrease as the magnitude of α increases. Thus, it is clear that for testing $H_o : \rho = 0$, the LRT is much better than the Wald-Type test.

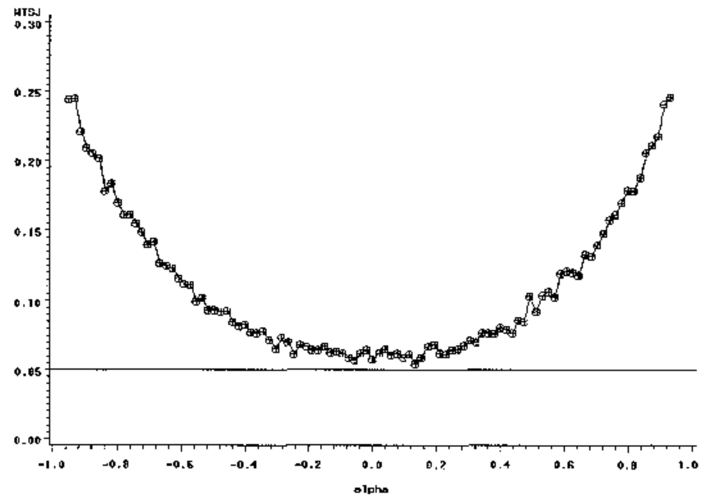

Figure 2.31: $H_o : \rho = 0$ – *Estimated Significance Level for Wald-Type test (MoM)*

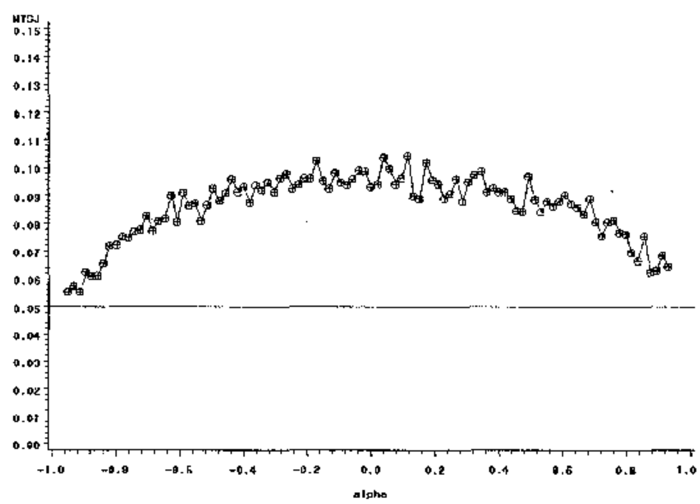

Figure 2.32: $H_o : \rho = 0$ – *Estimated Significance Level for Wald-Type test (QLS)*

CHAPTER III

AR(1) STRUCTURE WITH HETEROGENEOUS VARIANCE

III.1 Introduction

In this Chapter we again focus on the autoregressive familial correlation structure, where we now model heterogeneous intra-class variances. This variance-covariance matrix is of the following form.

$$\begin{aligned}
\Sigma(\Phi,\lambda) &= D(\Phi)R(\lambda)D(\Phi) \\
&= \begin{pmatrix}
\phi_p & \sqrt{\phi_p\phi_s}\rho & \sqrt{\phi_p\phi_s}\rho^2 & \sqrt{\phi_p\phi_s}\rho^3 & & \sqrt{\phi_p\phi_s}\rho^{t-1} \\
\sqrt{\phi_p\phi_s}\rho & \phi_s & \phi_s\alpha & \phi_s\alpha^2 & \cdots & \phi_s\alpha^{t-1} \\
\sqrt{\phi_p\phi_s}\rho^2 & \phi_s\alpha & \phi_s & \phi_s\alpha & & \phi_s\alpha^{t-2} \\
 & \vdots & & & \ddots & \vdots \\
\sqrt{\phi_p\phi_s}\rho^{t-1} & \phi_s\alpha^{t-2} & \phi_s\alpha^{t-3} & \phi_s\alpha^{t-4} & \cdots & \phi_s
\end{pmatrix}
\end{aligned} \tag{3.1.1}$$

Recall that ϕ_p is the parental variance and ϕ_s is the child variance. Important to note here is that neither of the variance parameters ($\Phi = (\phi_p, \phi_s)$) factor out of the variance-covariance matrix in scalar form. Thus parameter estimation will differ from that in Chapter II.

This chapter is organized as follows. In Section III.2 we discuss important properties of the correlation structure (3.1.1), specifically the inverse, determinant and positive definite range of the correlation parameters. We briefly present the three estimation procedures in Section III.3 and apply them to (3.1.1). In Section III.4 we find asymptotic variances of the estimators and compare their asymptotic performance. Section III.5 provides small-sample properties for each estimation procedure in both the normal and non-normal cases.

III.2 Properties of Correlation Matrix

To find the determinant of (3.1.1) we make use of the property that the determinant of a matrix product is the product of matrix determinants, or

$$|\Sigma(\lambda,\Phi)| = |D(\Phi)R(\lambda)D(\Phi)| = |D(\Phi)||R(\lambda)||D(\Phi)| = |D(\Phi)|^2|R(\lambda)|.$$

Further, the determinant of a diagonal matrix with non-zero elements is the product of those elements (here, $D(\Phi) = [diag(\phi_p, \phi_s, \cdots, \phi_s)]^{1/2}$). Thus,

$$|\Sigma(\lambda,\Phi)| = \phi_p\phi_s^{t-1}|R(\lambda)|.$$

We have already shown in Chapter II that

$$|R(\lambda)| = \frac{(1-\alpha^2)^{t-3}}{1-\rho^2} \times \left[(1-\alpha^2)(1-\rho^2) - (\rho^2-\rho^{2t}) - \alpha^2(\rho^4-\rho^{2t-2}) + 2\alpha(\rho^3-\rho^{2t-1})\right] \quad (3.2.1)$$

so that we get

$$|\Sigma(\lambda,\Phi)| = \phi_p\phi_s^{t-1}\left(\frac{(1-\alpha^2)^{t-3}}{1-\rho^2}\right) \times \left[(1-\alpha^2)(1-\rho^2) - (\rho^2-\rho^{2t}) - \alpha^2(\rho^4-\rho^{2t-2}) + 2\alpha(\rho^3-\rho^{2t-1})\right]. \quad (3.2.2)$$

To find the positive-definite range of the variance (Φ) and correlation (λ) parameters, we create an inequality by setting the determinant of the variance-covariance matrix (3.2.2) greater than zero, or $|\mathbf{\Sigma}(\lambda,\Phi)| > 0$. Since both ϕ_p and ϕ_s must be greater than zero, we are left with $|R(\lambda)| > 0$, as defined in (3.2.1). Thus the positive definite range is the same as that found in Section II.2 for the homogeneous variance case.

Recall that the inverse of a product of symmetric matrices is equal to the product of matrix inverses. Thus,

$$\Sigma^{-1}(\lambda,\Phi) = D^{-1}(\Phi)R^{-1}(\lambda)D^{-1}(\Phi)$$

where $D^{-1}(\Phi) = [diag(1/\phi_p, 1/\phi_s, \cdots, 1/\phi_s)]^{1/2}$, and $R^{-1}(\lambda)$ is the same as Chapter II. More formally, we find $\Sigma^{-1}(\lambda,\Phi)$ by partitioning the matrices $D(\Phi)$ and $R(\lambda)$ as follows. Let $\Gamma_{11} = (1/\phi_p)^{1/2}$ and $\Gamma_{22} = (1/\phi_s)^{1/2}I_{t-1}$ represent the 1×1 and

$(t-1)\times(t-1)$ non-zero partitions of $D^{-1}(\Phi)$, respectively. Partitioning $R^{-1}(\lambda)$ as in Chapter II, we get

$$\begin{aligned}\Sigma^{-1}(\lambda,\Phi) &= \begin{bmatrix}\Gamma_{11} & 0\\ 0 & \Gamma_{22}\end{bmatrix}\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}\begin{bmatrix}\Gamma_{11} & 0\\ 0 & \Gamma_{22}\end{bmatrix}\\ &= \begin{bmatrix}\Gamma_{11}B_{11}\Gamma_{11} & \Gamma_{11}B_{12}\Gamma_{22}\\ \Gamma_{22}B_{21}\Gamma_{11} & \Gamma_{22}B_{22}\Gamma_{22}\end{bmatrix}\end{aligned}$$

with

$$\begin{aligned}\Gamma_{11}B_{11}\Gamma_{11} &= \Gamma_{11}^2 B_{11} = \frac{1}{\phi_p}B_{11}\\ &= \frac{1}{\phi_p}\left[\frac{(1-\alpha^2)(1-\rho^2)}{(1-\alpha^2)(1-\rho^2)-(\rho^2-\rho^{2t})-\alpha^2(\rho^4-\rho^{2t-2})+2\alpha(\rho^3-\rho^{2t-1})}\right]\\ \Gamma_{11}B_{12}\Gamma_{22} &= \frac{1}{\sqrt{\phi_p\phi_s}}B_{12}I_{t-1} = \frac{1}{\sqrt{\phi_p\phi_s}}B_{12}\\ &= -\frac{B_{11}}{\sqrt{\phi_p\phi_s}(1-\alpha^2)}\left[\begin{pmatrix}\rho\\ \rho^2\\ \vdots\\ \rho^{t-1}\end{pmatrix}' + \alpha^2\begin{pmatrix}0\\ \rho^2\\ \vdots\\ \rho^{t-2}\\ 0\end{pmatrix}' - \alpha\begin{pmatrix}\rho^2\\ \rho+\rho^3\\ \vdots\\ \rho^{t-3}+\rho^{t-1}\\ \rho^{t-2}\end{pmatrix}'\right]\\ \Gamma_{22}B_{21}\Gamma_{11} &= (\Gamma_{11}B_{12}\Gamma_{22})'\\ \Gamma_{22}B_{22}\Gamma_{22} &= \frac{1}{\phi_s}I_{t-1}B_{22}I_{t-1} = \frac{1}{\phi_s}B_{22}\\ &= \frac{1}{\phi_s(1-\alpha^2)}\left(I_{t-1}+\alpha^2C_2-\alpha C_1\right)\\ &\quad\times\left[I_{t-1}+\frac{B_{11}}{1-\alpha^2}\begin{pmatrix}\rho^2 & \rho^3 & \cdots & \rho^t\\ \rho^3 & \rho^4 & \cdots & \rho^{t+1}\\ \vdots & \vdots & \ddots & \vdots\\ \rho^t & \rho^{t+1} & \cdots & \rho^{2t-2}\end{pmatrix}\left(I_{t-1}+\alpha^2C_2-\alpha C_1\right)\right].\end{aligned}$$

A list of partial and second derivatives of the variance-covariance matrix appear in Appendix A.2.

III.3 Parameter Estimation

In this section we derive parameter estimators for the maximum likelihood (MLE), method of moments (MoM) and quasi-least squares procedures. For each procedure we use the same estimator for the regression parameter

$$\widehat{\beta} = \left[\sum_{i=1}^{n} X_i'\Sigma^{-1}(\widehat{\lambda}, \widehat{\Phi})X_i\right]^{-1} \sum_{i=1}^{n} X_i'\Sigma^{-1}(\widehat{\lambda}, \widehat{\Phi})Y_i \tag{3.3.1}$$

where $\widehat{\lambda} = (\widehat{\rho}, \widehat{\alpha})$ is the vector of correlation parameter estimators and $\widehat{\Phi} = (\widehat{\phi}_p, \widehat{\phi}_s)$ is the vector of variance parameter estimators. To avoid certain problems explained in the next subsection, we use moment estimators for the two variance parameters as shown in Elston (1975). The estimators are found using the following unbiased estimating equations.

$$\sum_{i=1}^{n} [Z_i'A(\phi_p)Z_i - \phi_p] = 0 \tag{3.3.2}$$

$$\Leftrightarrow A(\rho) = \begin{pmatrix} 1 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}$$

$$\sum_{i=1}^{n} [Z_i'A(\phi_s)Z_i - (t-1)\phi_s] = 0 \tag{3.3.3}$$

$$\Leftrightarrow A(\alpha) = \begin{pmatrix} 0 & \underline{0} \\ \underline{0} & I_{t-1} \end{pmatrix}$$

Solving (3.3.2) and (3.3.3) for ϕ_p and ϕ_s, respectively, yields

$$\widehat{\phi}_p = \frac{1}{n}\sum_{i=1}^{n} \widehat{z}_{i1}^2 \tag{3.3.4}$$

$$\widehat{\phi}_s = \frac{1}{n(t-1)}\sum_{i=1}^{n}\sum_{j=2}^{t} \widehat{z}_{ij}^2 \tag{3.3.5}$$

where $\widehat{z}_{ij} = y_{ij} - x_{ij}\widehat{\beta}$ is the residual of the jth member of the ith family. Note that $\widehat{\phi}_p$ uses only the squared parent residual (z_{i1}), as it is the only residual with expectation ϕ_p, and $\widehat{\phi}_s$ uses the squared residuals of all $(t-1)$ children, as their expectation is ϕ_s. Also, $Z_n = \sum_{i=1}^{n}(Y_i - X_i\beta)(Y_i - X_i\beta)'$ is the residual matrix, which is partitioned as follows

$$Z_n = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$$

where

$$\begin{aligned} Z_{11} &= \sum_{i=1}^{n} z_{i1}^2 \\ Z_{12} &= \left(\sum_{i=1}^n z_{i1}z_{i2} \;\; \sum_{i=1}^n z_{i1}z_{i3} \;\; \ldots \;\; \sum_{i=1}^n z_{i1}z_{it} \right) \\ Z_{21} &= Z_{12}' \\ Z_{22} &= \begin{pmatrix} \sum_{i=1}^n z_{i2}^2 & \sum_{i=1}^n z_{i2}z_{i3} & & \sum_{i=1}^n z_{i2}z_{it} \\ \sum_{i=1}^n z_{i2}z_{i3} & \sum_{i=1}^n z_{i3}^2 & \cdots & \sum_{i=1}^n z_{i3}z_{it} \\ & \vdots & \ddots & \vdots \\ \sum_{i=1}^n z_{i2}z_{it} & \sum_{i=1}^n z_{i3}z_{it} & \cdots & \sum_{i=1}^n z_{it}^2 \end{pmatrix}. \end{aligned}$$

III.3.1 Maximum Likelihood

Maximum likelihood estimation in the heterogeneous variance case is similar to that found in the homogeneous variance case, except now the variance parameters are more deeply embedded in the likelihood function. Nevertheless, we attempt to find estimators by maximizing the log-likelihood function with respect to each parameter. The log-likelihood function is now

$$\ell = -\frac{nt}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma(\lambda,\Phi)| - \frac{1}{2}\sum_{i=1}^{n}(Y_i - X_i\beta)'\Sigma^{-1}(\lambda,\Phi)(Y_i - X_i\beta).$$

Recall that $\Sigma(\lambda,\Phi) = D(\Phi)R(\lambda)D(\Phi)$, and the log-likelihood becomes

$$\begin{aligned} \ell &= -\frac{nt}{2}\ln(2\pi) - \frac{n}{2}\ln(\phi_p) - \frac{n(t-1)}{2}\ln(\phi_s) - \frac{n}{2}|R(\lambda)| \\ &\quad -\frac{1}{2}\sum_{i=1}^{n}(Y_i - X_i\beta)'D^{-1}(\Phi)R^{-1}(\lambda)D^{-1}(\Phi)(Y_i - X_i\beta). \qquad (3.3.6) \end{aligned}$$

Note that we will sometimes express the last term in (3.3.6) as $-\frac{1}{2}tr\left[\Sigma^{-1}(\lambda,\Phi)Z_n\right]$.

Since Φ is embedded quadratically into (3.3.6), we will not get a closed-form variance estimator as we did in the homogeneous case. We could solve for the two variance parameters ($\Phi = (\phi_p, \phi_s)$) simultaneously by taking the derivative of (3.3.6) with respect to Φ to get the two estimating equations for the variance parameters as found in (3.3.7).

$$\begin{aligned}
\frac{\partial \ell}{\partial \phi_p} &\propto \frac{1}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \phi_p} \Sigma^{-1}(\lambda, \Phi) \widehat{Z}_n \right] \qquad (3.3.7) \\
&\quad - \frac{n}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \phi_p} \right] = 0 \\
&\leftrightarrow \phi_p^{-1} B_{11} \widehat{Z}_{11} + \phi_p^{-\frac{1}{2}} \phi_s^{-\frac{1}{2}} B_{12} \widehat{Z}_{21} - n = 0 \\
\frac{\partial \ell}{\partial \phi_s} &\propto \frac{1}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \phi_s} \Sigma^{-1}(\lambda, \Phi) \widehat{Z}_n \right] \\
&\quad - \frac{n}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \phi_s} \right] = 0 \\
&\leftrightarrow \phi_s^{-1} tr(B_{22} \widehat{Z}_{22}) + \phi_p^{-\frac{1}{2}} \phi_s^{-\frac{1}{2}} B_{12} \widehat{Z}_{21} - n(t-1) = 0.
\end{aligned}$$

However, these estimating equations do not yield closed form solutions and have problems with convergence if solved for numerically using Newton-Raphson or some other iterative technique. Thus, we use moment estimators (3.3.4) and (3.3.5) for the variance parameters.

The correlation parameters may be solved for simultaneously using the Newton-Raphson method with estimating equations (3.3.8) and (3.3.9).

$$\begin{aligned}
\frac{\partial \ell}{\partial \rho} &\propto \frac{1}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \rho} \Sigma^{-1}(\lambda, \Phi) \widehat{Z} \right] \qquad (3.3.8) \\
&\quad - \frac{n}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \rho} \right] = 0 \\
&\Leftrightarrow tr \left[D^{-1}(\Phi) R^{-1}(\lambda) \frac{\partial R(\lambda)}{\partial \rho} R^{-1}(\lambda) D^{-1}(\Phi) \widehat{Z} \right] - ntr \left[R^{-1}(\lambda) \frac{\partial R(\lambda)}{\partial \rho} \right] = 0 \\
&\Leftrightarrow \widehat{\phi}_p^{-1} B_{11} \frac{\partial R_{12}}{\partial \rho} B_{21} \widehat{Z}_{11} + \widehat{\phi}_p^{-\frac{1}{2}} \widehat{\phi}_s^{-\frac{1}{2}} \left[B_{11} \frac{\partial R_{12}}{\partial \rho} B_{22} + B_{12} \frac{\partial R_{21}}{\partial \rho} B_{12} \right] \widehat{Z}_{21} \\
&\quad \widehat{\phi}_s^{-1} B_{12} \widehat{Z}_{22} B_{22} \frac{\partial R_{21}}{\partial \rho} - n B_{12} \frac{\partial R_{21}}{\partial \rho} = 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &\propto \frac{1}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \alpha} \Sigma^{-1}(\lambda, \Phi) \widehat{Z} \right] \qquad (3.3.9) \\
&\quad - \frac{n}{2} tr \left[\Sigma^{-1}(\lambda, \Phi) \frac{\partial \Sigma(\lambda, \Phi)}{\partial \alpha} \right] = 0 \\
&\Leftrightarrow tr \left[D^{-1}(\Phi) R^{-1}(\lambda) \frac{\partial R(\lambda)}{\partial \alpha} R^{-1}(\lambda) D^{-1}(\Phi) \widehat{Z} \right] - ntr \left[R^{-1}(\lambda) \frac{\partial R(\lambda)}{\partial \alpha} \right] = 0 \\
&\Leftrightarrow \widehat{\phi}_p^{-1} B_{12} \frac{\partial R_{22}}{\partial \alpha} B_{21} \widehat{Z}_{11} + \widehat{\phi}_p^{-\frac{1}{2}} \widehat{\phi}_s^{-\frac{1}{2}} B_{12} \frac{\partial R_{22}}{\partial \alpha} B_{22} \widehat{Z}_{21} \\
&\quad \widehat{\phi}_s^{-1} tr \left[B_{22} \frac{\partial R_{22}}{\partial \alpha} B_{22} \widehat{Z}_{22} \right] - ntr \left[B_{22} \frac{\partial R_{22}}{\partial \alpha} \right] = 0
\end{aligned}$$

Thus, we find the MLE of θ by iterating between estimation of β, Φ, and λ. More specifically, we start with initial values λ_o and Φ_o of the correlation and variance parameters and use these to estimate β. We then use $\widehat{\beta}$ to calculate $\widehat{Z}_n$, and then estimate Φ. Using $\widehat{Z}_n$ and $\widehat{\Phi}$, we then estimate λ and use $\widehat{\Phi}$ and $\widehat{\lambda}$ to re-estimate β. Repeating in this manner until convergence, we arrive at the maximum likelihood estimate, $\widehat{\theta}_\ell = (\widehat{\beta}_\ell, \widehat{\lambda}_\ell, \widehat{\Phi}_\ell)$.

III.3.2 Method of Moments

We begin the method of moment procedure by using (3.3.1), (3.3.4) and (3.3.5 to estimate β and Φ, respectively. To find estimators for the elements of λ we use the following two estimating equations

$$\sum_{i=1}^{n} Z_i' A(\rho) Z_i = 0 \tag{3.3.10}$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} Z_i' \left[\frac{\rho}{\sqrt{\phi_p}} \begin{pmatrix} 1 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} - \frac{1}{2\sqrt{\phi_s}} \begin{pmatrix} C_3 & \underline{0} \\ \underline{0} & 0 \end{pmatrix} \right] Z_i = 0$$

and

$$\sum_{i=1}^{n} Z_i' A(\alpha) Z_i = 0 \tag{3.3.11}$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} Z_i' \left[\frac{\alpha}{t-1} \begin{pmatrix} 0 & \underline{0} \\ \underline{0} & I_{t-1} \end{pmatrix} - \frac{1}{2(t-2)} \begin{pmatrix} 0 & \underline{0} \\ \underline{0} & C_1 \end{pmatrix} \right] Z_i = 0$$

where $C_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and C_1 is defined in Chapter II. Solving (3.3.10) and (3.3.11) for ρ and α, respectively, gives the following moment estimators

$$\widehat{\rho}_m = \frac{\widehat{\phi}_{p,m}^{\frac{1}{2}} \sum_{i=1}^{n} \widehat{z}_{i1}\widehat{z}_{i2}}{\widehat{\phi}_{s,m}^{\frac{1}{2}} \sum_{i=1}^{n} \widehat{z}_{i1}^2} \tag{3.3.12}$$

$$\widehat{\alpha}_m = \frac{(t-1)\sum_{i=1}^{n}\sum_{i=2}^{t-1} \widehat{z}_{i,j}\widehat{z}_{i,j+1}}{(t-2)\sum_{i=1}^{n}\sum_{i=2}^{t} \widehat{z}_{ij}^2}. \tag{3.3.13}$$

Note that the numerators of these estimators are practically the same as those used in Chapter II.

Procedurally, then, we first assume that $\Phi = (1, 1)$ and $\lambda = (0, 0)$ (i.e. response variables of all family members are independent) or use sample statistics and solve for β using (3.3.1). We then use $\widehat{\beta}$ to compute $\widehat{Z}_n$, and then estimate $\widehat{\Phi} = (\widehat{\phi}_p, \widehat{\phi}_s)$ with (3.3.4) and (3.3.5) and $\widehat{\lambda} = (\widehat{\rho}, \widehat{\alpha})$ with (3.3.12) and (3.3.13). We then use these to re-estimate β and continue in this manner until convergence. These estimators, $\widehat{\theta}_m = (\widehat{\beta}_m, \widehat{\lambda}_m, \widehat{\Phi}_m)$, are then the moment estimators for θ.

III.3.3 Quasi-Least Squares

In this case we note that the quasi-log-likelihood function is defined as

$$\begin{aligned} S(\theta) &= \sum_{i=1}^{n}(Y_i - X_i\beta)'\Sigma^{-1}(\lambda, \Phi)(Y_i - X_i\beta) \\ &= \sum_{i=1}^{n} Z_i'\Sigma^{-1}(\lambda, \Phi)Z_i = tr\left[\Sigma^{-1}(\lambda, \Phi)Z_n\right]. \end{aligned} \tag{3.3.14}$$

Recall that $\Sigma(\Phi, \lambda) = D(\Phi)R(\lambda)D(\Phi)$. For the *Step 1* estimators of β and λ we maximize (3.3.14) with respect to those parameters, however we must account for Φ, the vector of variance parameters that do not factor out of the quasi-log-likelihood. Thus we propose including simple moment estimators of the variance parameters in *Step 1* so that estimators of the other parameters may be obtained, recalling that in the homogeneous case there was no need to estimate variance parameters in the first step.

The QLS procedure outlined here contains two steps. In *Step 1* we begin by selecting initial values $\Phi_o = (\phi_{p,o}, \phi_{s,o})$ and $\lambda_o = (\rho_o, \alpha_o)$, which we take as either sample statistics or $(1,1)$ and $(0,0)$, respectively. We then find $\widetilde{\beta}$, the *Step 1* estimator of β using (3.3.1) evaluated at the initial values Φ_o and λ_o. We then use $\widetilde{\beta}$ to update the residual matrix Z_n and estimate Φ with (3.3.4) and (3.3.5). We now use $\widetilde{\Phi} = (\widetilde{\phi}_p, \widetilde{\phi}_s)$ and $\widetilde{Z}_n$ to estimate the correlation parameters, which is done by maximizing the quasi-log-likelihood function with respect to ρ and α.

$$\begin{aligned}
\frac{\partial S(\theta)}{\partial \rho} &= \frac{\partial}{\partial \rho}\left[tr\left(\Sigma^{-1}(\lambda_o,\widetilde{\Phi})\widetilde{Z}_n\right)\right] = \left[tr\left(\frac{\partial \Sigma^{-1}(\lambda_o,\widetilde{\Phi})}{\partial \rho}\widetilde{Z}_n\right)\right] = 0 \\
&\leftrightarrow -tr\left[\Sigma^{-1}(\lambda_o,\widetilde{\Phi})\frac{\partial \Sigma(\lambda_o,\widetilde{\Phi})}{\partial \rho}\Sigma^{-1}(\lambda_o,\widetilde{\Phi})\widetilde{Z}_n\right] = 0 \qquad (3.3.15) \\
&\leftrightarrow \widetilde{\phi}_p^{-1}B_{11,o}\frac{\partial R_{12,o}}{\partial \rho}B_{21,o}\widetilde{Z}_{11} + \widetilde{\phi}_s^{-1}B_{12,o}\widetilde{Z}_{22}B_{22,o}\frac{\partial R_{21,o}}{\partial \rho} \\
&\quad +\widetilde{\phi}_p^{-1/2}\widetilde{\phi}_s^{-1/2}\left[B_{12,o}\frac{\partial R_{21,o}}{\partial \rho}B_{12,o} + B_{11,o}\frac{\partial R_{12,o}}{\partial \rho}B_{22,o}\right]\widetilde{Z}_{21} = 0 \\
\frac{\partial S(\theta)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[tr\left(\Sigma^{-1}(\lambda_o,\widetilde{\Phi})\widetilde{Z}_n\right)\right] = \left[tr\left(\frac{\partial \Sigma^{-1}(\lambda_o,\widetilde{\Phi})}{\partial \alpha}\widetilde{Z}_n\right)\right] = 0 \\
&\leftrightarrow -tr\left[\Sigma^{-1}(\lambda_o,\widetilde{\Phi})\frac{\partial \Sigma(\lambda_o,\widetilde{\Phi})}{\partial \alpha}\Sigma^{-1}(\lambda_o,\widetilde{\Phi})\widetilde{Z}_n\right] = 0 \qquad (3.3.16) \\
&\leftrightarrow \widetilde{\phi}_p^{-1}B_{11,o}\frac{\partial R_{22,o}}{\partial \alpha}B_{21,o}\widetilde{Z}_{11} + \widetilde{\phi}_s^{-1}tr\left[B_{22,o}\frac{\partial R_{22,o}}{\partial \alpha}B_{22,o}\widetilde{Z}_{22}\right] \\
&\quad +2\widetilde{\phi}_p^{-1/2}\widetilde{\phi}_s^{-1/2}B_{12,o}\frac{\partial R_{22,o}}{\partial \alpha}B_{22,o}\widetilde{Z}_{21} = 0
\end{aligned}$$

We must solve for ρ and α simultaneously using Newton-Raphson. Now we use $\widetilde{\Phi}$ and $\widetilde{\lambda}$ to re-estimate β, and iterate in this manner until convergence. Then $\widetilde{\lambda} = (\widetilde{\rho}, \widetilde{\alpha})$ is the *Step 1* estimate of λ.

As shown in Shults and Chaganty (1998), the *Step 1* estimates of the correlation parameters are biased, as can be seen by taking their expectation.

$$\begin{aligned}
E\left[\frac{\partial S(\theta)}{\partial \rho}\right] &= tr\left[\frac{\partial \Sigma^{-1}(\lambda_o,\widetilde{\Phi})}{\partial \rho}E(\widetilde{Z}_n)\right] \neq 0 \qquad (3.3.17) \\
&\propto tr\left[\frac{\partial \Sigma^{-1}(\lambda_o,\widetilde{\Phi})}{\partial \rho}\Sigma(\lambda,\Phi)\right] \neq 0 \\
E\left[\frac{\partial S(\theta)}{\partial \alpha}\right] &= tr\left[\frac{\partial \Sigma^{-1}(\lambda_o,\widetilde{\Phi})}{\partial \alpha}E(\widetilde{Z}_n)\right] \neq 0 \\
&\propto tr\left[\frac{\partial \Sigma^{-1}(\lambda_o,\widetilde{\Phi})}{\partial \alpha}\Sigma(\lambda,\Phi)\right] \neq 0
\end{aligned}$$

To eliminate this bias we turn (3.3.17) into a set of estimating equations ((3.3.18) and (3.3.19)) that we solve for ρ and α, respectively. Here we again set $\lambda_o = (\rho_o, \alpha_o)$, either sample statistics or (0,0), and fix $\widetilde{\lambda} = (\widetilde{\rho}, \widetilde{\alpha})$ as the *Step 1* estimators of the correlation parameters. Note that we do not need values for Φ as they cancel out of the equations. We then get the following *Step 2* estimating equations for λ

$$tr\left[\Sigma^{-1}(\widetilde{\lambda},\widetilde{\Phi})\frac{\partial\Sigma(\widetilde{\lambda},\widetilde{\Phi})}{\partial\widetilde{\rho}}\Sigma^{-1}(\widetilde{\lambda},\widetilde{\Phi})\Sigma(\lambda_o,\widetilde{\Phi})\right]=0 \tag{3.3.18}$$

$$\Leftrightarrow\quad tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\rho}}R^{-1}(\widetilde{\lambda})R(\lambda_o)\right]=0$$

$$\Leftrightarrow\quad \widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\widetilde{\rho}}\left[\widetilde{B}_{21}+\widetilde{B}_{22}R_{21,o}\right]+\widetilde{B}_{12}\left[R_{21,o}\widetilde{B}_{12}+\widetilde{B}_{22}R_{22,o}\right]\frac{\partial\widetilde{R}_{21}}{\partial\widetilde{\rho}}=0$$

$$tr\left[\Sigma^{-1}(\widetilde{\lambda},\widetilde{\Phi})\frac{\partial\Sigma(\widetilde{\lambda},\widetilde{\Phi})}{\partial\widetilde{\alpha}}\Sigma^{-1}(\widetilde{\lambda},\widetilde{\Phi})\Sigma(\lambda_o,\widetilde{\Phi})\right]=0 \tag{3.3.19}$$

$$\Leftrightarrow\quad tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\widetilde{\alpha}}R^{-1}(\widetilde{\lambda})R(\lambda_o)\right]=0$$

$$\Leftrightarrow\quad \widetilde{B}_{12}\frac{\partial\widetilde{R}_{22}}{\partial\widetilde{\alpha}}\widetilde{B}_{21}+2\widetilde{B}_{12}\frac{\partial\widetilde{R}_{22}}{\partial\widetilde{\alpha}}\widetilde{B}_{22}R_{21,o}+tr\left[\widetilde{B}_{22}\frac{\partial\widetilde{R}_{22}}{\partial\widetilde{\alpha}}\widetilde{B}_{22}R_{22,o}\right]=0.$$

The two estimating equations (3.3.18) and (3.3.19) are solved simultaneously using Newton-Raphson. These values, $\widehat{\lambda}_q$, are the *Step 2* estimators for the correlation parameters ρ and α. We then use these estimators to update β and Φ as before. The QLS estimator of θ is $\widehat{\theta}_q=(\widehat{\beta}_q,\widehat{\lambda}_q,\widehat{\Phi}_q)'$.

III.4 Asymptotic Variance and Performance

In this section we derive the asymptotic variance covariance matrices for the MLE, MOM and QLS estimators described in Section III.3. In each case we make use of the Theorem by Joe (2.4.1) to find the asymptotic variances for the estimators.

III.4.1 Maximum Likelihood

Typically for the MLE ($\widehat{\theta}_\ell$) we use Fisher's Information matrix to find the asymptotic variance, knowing that

$$\sqrt{n}(\widehat{\theta}_\ell-\theta)\sim AMVN(\underline{0},I^{-1}(\theta)) \tag{3.4.1}$$

where $I(\theta)=-E\left(\partial^2\ell/\partial\theta\partial\theta'\right)$. However, since we are using moment estimators for the variance parameters, we cannot use the Fisher Information. Thus, we make use of the Theorem described in (2.4.1) to find a more general form of the asymptotic variance and show that

$$\sqrt{n}\left(\widehat{\theta}_\ell - \theta\right) \sim AMVN\left(\underline{0}, I_\ell^{-1}(\theta)M_\ell(\theta)(I_\ell^{-1}(\theta))'\right) \tag{3.4.2}$$

where $I_\ell(\theta) = -\frac{1}{n}\sum_{i=1}^n(\partial h_{\ell,i}(\theta)/\partial\theta)$, $M_\ell(\theta) = frac1n\sum_{i=1}^n Cov(h_{\ell,i}(\theta))$ and $h_{\ell,i}(\theta)$ is a vector of the unbiased estimating equations that lead to the maximum likelihood estimators for θ. We define $h_{\ell,i}(\theta)$ as

$$\begin{aligned}
h_{\ell,i}(\theta) &= (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), g_{1i}(\theta), g_{2i}(\theta))' \qquad (3.4.3)\\
h_{0i} &= X_i'\Sigma^{-1}(\lambda,\Phi)Z_i \\
h_{1i} &= Z_i'\frac{\partial\Sigma^{-1}(\lambda,\Phi)}{\partial\rho}Z_i - tr\left[\frac{\partial\Sigma^{-1}(\lambda,\Phi)}{\partial\rho}\Sigma(\lambda,\Phi)\right] \\
&= tr\left[\frac{\partial\Sigma^{-1}(\lambda,\Phi)}{\partial\rho}(Z_iZ_i' - \Sigma(\lambda,\Phi))\right] \\
h_{2i} &= Z_i'\frac{\partial\Sigma^{-1}(\lambda,\Phi)}{\partial\alpha}Z_i - tr\left[\frac{\partial\Sigma^{-1}(\lambda,\Phi)}{\partial\alpha}\Sigma(\lambda,\Phi)\right] \\
&= tr\left[\frac{\partial\Sigma^{-1}(\lambda,\Phi)}{\partial\alpha}(Z_iZ_i' - \Sigma(\lambda,\Phi))\right] \\
g_{1i} &= Z_i'A(\phi_p)Z_i - \phi_p = Z_i'\begin{bmatrix} 1 & \underline{0} \\ \underline{0} & \underline{0}\end{bmatrix}Z_i - \phi_p \\
&= tr\left[A(\phi_p)Z_iZ_i'\right] - \phi_p
\end{aligned}$$

$$\begin{aligned}
g_{2i} &= Z_i'A(\phi_s)Z_i - (t-1)\phi_s = Z_i'\begin{bmatrix} 0 & \underline{0} \\ \underline{0} & I_{t-1}\end{bmatrix}Z_i - (t-1)\phi_s \\
&= tr\left[A(\phi_s)Z_iZ_i'\right] - (t-1)\phi_s.
\end{aligned}$$

By taking the negative expectation of the derivative of (3.4.3) with respect to θ and averaging over n we obtain $I_\ell(\theta)$, and by taking the covariance of (3.4.3) and averaging over n we obtain $M_\ell(\theta)$. From here it is easy to show that $I_\ell(\theta)$ has the following elements

$$I_\ell(\theta) = \begin{pmatrix} I_{11} & 0 & 0 & 0 & 0 \\ 0 & I_{22} & I_{23} & I_{24} & I_{25} \\ 0 & I_{23} & I_{33} & 0 & I_{35} \\ 0 & 0 & 0 & I_{44} & 0 \\ 0 & 0 & 0 & 0 & I_{55}\end{pmatrix} \tag{3.4.4}$$

where

$$
\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{0i}(\theta)}{\partial \beta}\right) = \frac{1}{n}\sum_{i=1}^{n} X_i'\Sigma^{-1}(\lambda,\Phi)X_i \\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \rho}\right) = -tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= -2B_{11}\frac{\partial R_{12}}{\partial \rho}B_{22}\frac{\partial R_{21}}{\partial \rho} - 2B_{12}\frac{\partial R_{21}}{\partial \rho}B_{12}\frac{\partial R_{21}}{\partial \rho} \\
I_{23} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \alpha}\right) = -tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= -2B_{12}\frac{\partial R_{22}}{\partial \alpha}B_{22}\frac{\partial R_{21}}{\partial \rho} \\
I_{24} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \phi_p}\right) = -2tr\left[\frac{\partial R(\lambda)}{\partial \rho}R^{-1}(\lambda)D^{-1}(\Phi)\frac{\partial D(\Phi)}{\partial \phi_p}\right] \\
&= -\phi_p^{-1}\frac{\partial R_{12}}{\partial \rho}B_{21} \\
I_{25} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \phi_s}\right) = -2tr\left[\frac{\partial R(\lambda)}{\partial \rho}R^{-1}(\lambda)D^{-1}(\Phi)\frac{\partial D(\Phi)}{\partial \phi_s}\right] \\
&= -\phi_s^{-1}B_{12}\frac{\partial R_{21}}{\partial \rho} \\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \alpha}\right) = -tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -tr\left[B_{22}\frac{\partial R_{22}}{\partial \alpha}B_{22}\frac{\partial R_{22}}{\partial \alpha}\right]
\end{aligned}
$$

$$
\begin{aligned}
I_{35} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \phi_s}\right) = -2tr\left[\frac{\partial R(\lambda)}{\partial \alpha}R^{-1}(\lambda)D^{-1}(\Phi)\frac{\partial D(\Phi)}{\partial \phi_s}\right] \\
&= -\phi_s^{-1}tr\left[\frac{\partial R_{22}}{\partial \alpha}B_{22}\right] \\
I_{44} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_{1i}(\theta)}{\partial \phi_p}\right) = 1 \\
I_{55} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_{2i}(\theta)}{\partial \phi_s}\right) = t-1.
\end{aligned}
$$

We can also show that $M_\ell(\theta)$ has the following elements

$$M_\ell(\theta) = \begin{pmatrix} M_{11} & 0 & 0 & 0 & 0 \\ 0 & M_{22} & M_{23} & 0 & 0 \\ 0 & M_{23} & M_{33} & 0 & 0 \\ 0 & 0 & 0 & M_{44} & M_{45} \\ 0 & 0 & 0 & M_{45} & M_{55} \end{pmatrix} \tag{3.4.5}$$

where

$$\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{0i}(\theta)) = \frac{1}{n}\sum_{i=1}^{n} X_i'\Sigma^{-1}(\lambda,\Phi)X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta)) = 2tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho}\right] \\
&= 2B_{12}\frac{\partial R_{21}}{\partial\rho}B_{12}\frac{\partial R_{21}}{\partial\rho} + 2B_{11}\frac{\partial R_{12}}{\partial\rho}B_{22}\frac{\partial R_{21}}{\partial\rho} \\
M_{23} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), h_{2i}(\theta)) = 2tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= 4B_{12}\frac{\partial R_{22}}{\partial\alpha}B_{22}\frac{\partial R_{21}}{\partial\rho} \\
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta)) = 2tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= 2tr\left[B_{22}\frac{\partial R_{22}}{\partial\alpha}B_{22}\frac{\partial R_{22}}{\partial\alpha}\right] \\
M_{44} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_{1i}(\theta)) = 2tr\left[A(\phi_p)\Sigma(\lambda,\Phi)A(\phi_p)\Sigma(\lambda,\Phi)\right] = 2\phi_p^2 \\
M_{45} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_{1i}(\theta), g_{2i}(\theta)) = 2tr\left[A(\phi_p)\Sigma(\lambda,\Phi)A(\phi_s)\Sigma(\lambda,\Phi)\right] = \frac{2\phi_p\phi_s(\rho^2-\rho^{2t})}{1-\rho^2}
\end{aligned}$$

$$\begin{aligned}
M_{55} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_{2i}(\theta)) = 2tr\left[A(\phi_s)\Sigma(\lambda,\Phi)A(\phi_s)\Sigma(\lambda,\Phi)\right] \\
&= 2\phi_s^2\left[(t-1) + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right].
\end{aligned}$$

By taking the inverse of $I_\ell(\theta)$ and pre- and post-multiplying upon $M_\ell(\theta)$ we obtain the asymptotic variance of the MLE estimators. Note that, based on matrices (3.4.4) and (3.4.5), the estimator for β is uncorrelated with the estimators for the other parameters.

III.4.2 Method of Moments

For the Method of Moments (MOM) estimator ($\widehat{\theta}_m$) we have

$$\sqrt{n}\left(\widehat{\theta}_m - \theta\right) \sim AN\left(\underline{0}, I_m^{-1}(\theta)M_m(\theta)(I_m^{-1}(\theta))'\right) \tag{3.4.6}$$

where $I_m(\theta) = -\frac{1}{n}\sum_{i=1}^{n}(\partial h_{m,i}(\theta)/\partial\theta)$, $M_m(\theta) = \frac{1}{n}\sum_{i=1}^{n} Cov(h_{m,i}(\theta))$ and $h_{m,i}(\theta)$ is a vector of the unbiased estimating equations that lead to the MOM estimators for θ. For any $i = 1, ..., n$, $h_{m,i}(\theta)$ is defined as follows

$$\begin{aligned}
h_{m,i}(\theta) &= (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), g_{1i}(\theta), g_{2i}(\theta))' \qquad (3.4.7)\\
h_{0i} &= X_i'\Sigma^{-1}(\lambda, \Phi)Z_i\\
h_{1i} &= Z_i'A(\rho)Z_i = tr\left[A(\rho)Z_iZ_i'\right]\\
&= Z_i'\left[\frac{\rho}{\sqrt{\phi_p}}\begin{pmatrix} 1 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} - \frac{1}{2\sqrt{\phi_s}}\begin{pmatrix} C_3 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}\right] Z_i\\
h_{2i} &= Z_i'A(\alpha)Z_i = tr\left[A(\alpha)Z_iZ_i'\right]\\
&= Z_i'\left[\frac{\alpha}{t-1}\begin{pmatrix} 0 & \underline{0} \\ \underline{0} & I_{t-1} \end{pmatrix} - \frac{1}{2(t-2)}\begin{pmatrix} \underline{0} & \underline{0} \\ \underline{0} & C_1 \end{pmatrix}\right] Z_i\\
g_{1i} &= Z_i'A(\phi_p)Z_i - \phi_p = Z_i'\begin{bmatrix} 1 & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} Z_i - \phi_p\\
&= tr\left[A(\phi_p)Z_iZ_i'\right] - \phi_p\\
g_{2i} &= Z_i'A(\phi_s)Z_i - (t-1)\phi_s = Z_i'\begin{bmatrix} 0 & \underline{0} \\ \underline{0} & I_{t-1} \end{bmatrix} Z_i - (t-1)\phi_s\\
&= tr\left[A(\phi_s)Z_iZ_i'\right] - (t-1)\phi_s.
\end{aligned}$$

By taking the negative expectation of the derivative of (3.4.7) with respect to θ and averaging over n we obtain $I_m(\theta)$, and by taking the covariance of (3.4.7) and averaging over n we obtain $M_m(\theta)$. From here it is easy to show that $I_m(\theta)$ has the following elements

$$I_m(\theta) = \begin{pmatrix} I_{11} & 0 & 0 & 0 & 0 \\ 0 & I_{22} & 0 & I_{24} & I_{25} \\ 0 & 0 & I_{33} & 0 & 0 \\ 0 & 0 & 0 & I_{44} & 0 \\ 0 & 0 & 0 & 0 & I_{55} \end{pmatrix} \tag{3.4.8}$$

where

$$
\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{0i}(\theta)}{\partial \beta}\right) = \frac{1}{n}\sum_{i=1}^{n} X_i'\Sigma^{-1}(\lambda,\Phi)X_i \\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \rho}\right) = -\phi_p^{1/2} \\
I_{24} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \phi_p}\right) = \frac{\rho}{2\phi_p^{1/2}} \\
I_{25} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \phi_s}\right) = \frac{\rho\phi_p^{1/2}}{2\phi_s} \\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \alpha}\right) = -\phi_s \\
I_{44} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{3i}(\theta)}{\partial \phi_p}\right) = 1 \\
I_{55} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{4i}(\theta)}{\partial \phi_s}\right) = t-1.
\end{aligned}
$$

We can also show that $M_m(\theta)$ has the following elements

$$
M_m(\theta) = \begin{pmatrix}
M_{11} & 0 & 0 & 0 & 0 \\
0 & M_{22} & M_{23} & 0 & M_{25} \\
0 & M_{23} & M_{33} & M_{34} & M_{35} \\
0 & 0 & M_{34} & M_{44} & M_{45} \\
0 & M_{25} & M_{35} & M_{45} & M_{55}
\end{pmatrix} \tag{3.4.9}
$$

where

$$
\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{0i}(\theta)) = \frac{1}{n}\sum_{i=1}^{n} X_i'\Sigma^{-1}(\lambda,\Phi)X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta)) = 2tr\left[A(\rho)\Sigma(\lambda,\Phi)A(\rho)\Sigma(\lambda,\Phi)\right] = \phi_p(1-\rho^2)
\end{aligned}
$$

$$
\begin{aligned}
M_{23} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), h_{2i}(\theta)) = 2tr\left[A(\rho)\Sigma(\lambda,\Phi)A(\alpha)\Sigma(\lambda,\Phi)\right] \\
&= \frac{2\alpha\rho\phi_p^{1/2}\phi_s(\rho^2-\rho^{2t})}{(t-1)(1-\rho^2)} - \frac{2\rho\phi_p^{1/2}\phi_s(\rho^3-\rho^{2t-1})}{(t-2)(1-\rho^2)} - \frac{2\phi_p^{1/2}\phi_s(\alpha\rho-(\alpha\rho)^t)}{(t-1)(1-\alpha\rho)} \\
&\quad + \frac{\phi_p^{1/2}\phi_s(\alpha+\rho)(\alpha\rho-(\alpha\rho)^{t-1})}{\alpha(t-2)(1-\alpha\rho)} \\
M_{25} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), h_{4i}(\theta)) = 2tr\left[A(\rho)\Sigma(\lambda,\Phi)A(\phi_s)\Sigma(\lambda,\Phi)\right] \\
&= 2\rho\phi_p^{1/2}\phi_s\left(\frac{\rho^2-\rho^{2t}}{1-\rho^2}\right) - \frac{2\phi_p^{1/2}\phi_s(\alpha\rho-(\alpha\rho)^t)}{\alpha(1-\alpha\rho)} \\
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta)) = 2tr\left[A(\alpha)\Sigma(\lambda,\Phi)A(\alpha)\Sigma(\lambda,\Phi)\right] \\
&= \frac{2\alpha^2\phi_s^2}{(t-1)^2}\left[(t-1)+2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right] - \frac{8\alpha\phi_s^2}{(t-1)(t-2)}\left[\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j-1}\right] \\
&\quad + \frac{\phi_s^2}{(t-2)^2}\left[(t-2)(1+\alpha^2)+4\sum_{j=1}^{t-3}(t-2-j)\alpha^{2j}\right] \\
M_{34} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta), h_{3i}(\theta)) = 2tr\left[A(\alpha)\Sigma(\lambda,\Phi)A(\phi_p)\Sigma(\lambda,\Phi)\right] \\
&= \frac{2\phi_p\phi_s}{1-\rho^2}\left[\frac{\alpha(\rho^2-\rho^{2t})}{t-1} - \frac{\rho^3-\rho^{2t-1}}{t-2}\right] \\
M_{35} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta), h_{4i}(\theta)) = 2tr\left[A(\alpha)\Sigma(\lambda,\Phi)A(\phi_s)\Sigma(\lambda,\Phi)\right] \\
&= \frac{2\alpha\phi_s^2}{(t-1)}\left[(t-1)+2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right] - \frac{4\phi_s^2}{(t-2)}\left[\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j-1}\right] \\
M_{44} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{3i}(\theta)) = 2tr\left[A(\phi_p)\Sigma(\lambda,\Phi)A(\phi_p)\Sigma(\lambda,\Phi)\right] = 2\phi_p^2 \\
M_{45} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{3i}(\theta), h_{4i}(\theta)) = 2tr\left[A(\phi_p)\Sigma(\lambda,\Phi)A(\phi_s)\Sigma(\lambda,\Phi)\right] \\
&= 2\phi_p\phi_s\frac{\rho^2-\rho^{2t}}{1-\rho^2} \\
M_{55} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{4i}(\theta)) = 2tr\left[A(\phi_s)\Sigma(\lambda,\Phi)A(\phi_s)\Sigma(\lambda,\Phi)\right] \\
&= 2\phi_s^2\left[(t-1)+2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right].
\end{aligned}
$$

By taking the inverse of $I_m(\theta)$ and pre- and post-multiplying upon $M_m(\theta)$ we obtain the asymptotic variance of the MOM estimators. Note that, based on matrices (3.4.8) and (3.4.9), the estimator for β is uncorrelated with the estimators for the other parameters.

III.4.3 Quasi Least Squares

For the Quasi-Least Squares (QLS) estimator, $\widehat{\theta}_q$, we have

$$\sqrt{n}\left(\widehat{\theta}_q - \theta\right) \sim AMVN\left(\underline{0}, I_q^{-1}(\theta)M_q(\theta)(I_q^{-1}(\theta))'\right) \tag{3.4.10}$$

where $I_q(\theta) = -frac1n\sum_{i=1}^{n}(\partial h_{q,i}(\theta)/\partial\theta)$, $M_q(\theta) = (1/n)\sum_{i=1}^{n} Cov(h_{q,i}(\theta)$ and $h_{q,i}(\theta)$ is a vector of the unbiased estimating equations that lead to the QLS estimators for θ. For any $i = 1, ..., n$, $h_{q,i}(\theta)$ is defined as follows,

$$\begin{aligned}
h_{q,i}(\theta) &= (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), g_{1i}(\theta), g_{2i}(\theta))' \qquad (3.4.11)\\
h_{0i} &= X_i'\Sigma^{-1}(\lambda, \Phi)Z_i\\
h_{1i} &= Z_i'\frac{\partial\Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial\rho}Z_i - tr\left[\frac{\partial\Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial\rho}\Sigma(\lambda, \Phi)\right]\\
&= tr\left[\frac{\partial\Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial\rho}(Z_iZ_i' - \Sigma(\lambda, \Phi))\right]\\
h_{2i} &= Z_i'\frac{\partial\Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial\alpha}Z_i - tr\left[\frac{\partial\Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial\alpha}\Sigma(\lambda, \Phi)\right]\\
&= tr\left[\frac{\partial\Sigma^{-1}(\widetilde{\lambda}, \Phi)}{\partial\alpha}(Z_iZ_i' - \Sigma(\lambda, \Phi))\right]\\
g_{1i} &= Z_i'A(\phi_p)Z_i - \phi_p = Z_i'\begin{bmatrix} 1 & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix}Z_i - \phi_p\\
&= tr\left[A(\phi_p)Z_iZ_i'\right] - \phi_p\\
g_{2i} &= Z_i'A(\phi_s)Z_i - (t-1)\phi_s = Z_i'\begin{bmatrix} 0 & \underline{0} \\ \underline{0} & I_{t-1} \end{bmatrix}Z_i - (t-1)\phi_s\\
&= tr\left[A(\phi_s)Z_iZ_i'\right] - (t-1)\phi_s
\end{aligned}$$

where $\widetilde{\lambda}$ is the simultaneous solution to

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\rho}} R(\lambda)\right] = 0$$

$$\Leftrightarrow \widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \widetilde{\rho}}\left[\widetilde{B}_{21} + \widetilde{B}_{22}R_{21}\right] + \widetilde{B}_{12}\left[R_{21}\widetilde{B}_{12} + R_{22}\widetilde{B}_{22}\right]\frac{\partial \widetilde{R}_{21}}{\partial \widetilde{\rho}} = 0$$

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \widetilde{\alpha}} R(\lambda)\right] = 0$$

$$\Leftrightarrow \widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \widetilde{\alpha}}\widetilde{B}_{21} + 2\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \widetilde{\alpha}}\widetilde{B}_{22}R_{21} + tr\left[\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \widetilde{\alpha}}\widetilde{B}_{22}R_{22}\right].$$

By taking the negative expectation of the derivative of (3.4.11) with respect to θ and averaging over n we obtain $I_q(\theta)$, and by taking the covariance of (3.4.11) and averaging over n we obtain $M_q(\theta)$. From here it is easy to show that $I_q(\theta)$ has the following elements

$$I_q(\theta) = \begin{pmatrix} I_{11} & 0 & 0 & 0 & 0 \\ 0 & I_{22} & I_{23} & I_{24} & I_{25} \\ 0 & I_{32} & I_{33} & I_{34} & I_{35} \\ 0 & 0 & 0 & I_{44} & 0 \\ 0 & 0 & 0 & 0 & I_{55} \end{pmatrix} \tag{3.4.12}$$

where

$$\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{0i}(\theta)}{\partial \beta}\right) = \frac{1}{n}\sum_{i=1}^{n} X_i'\Sigma^{-1}(\lambda,\Phi)X_i \\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \rho}\right) = tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho}R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= -2\widehat{B}_{12}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{12}\frac{\partial R_{21}}{\partial \rho} - 2\widetilde{B}_{11}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{22}\frac{\partial R_{21}}{\partial \rho} \\
I_{23} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \alpha}\right) = tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho}R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -2\widetilde{B}_{12}\frac{\partial R_{22}}{\partial \alpha}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{21}}{\partial \rho} \\
I_{24} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \phi_p}\right) = -2tr\left[\frac{\partial D^{\frac{1}{2}}(\Phi)}{\partial \phi_p}R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho}R^{-1}(\widetilde{\lambda})R(\lambda)D^{\frac{1}{2}}(\Phi)\right] \\
&= -2\phi_p^{-1}\widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{21} - \phi_p^{-1}\widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{22}R_{21} - \phi_p^{-1}\widetilde{B}_{12}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{12}R_{21}
\end{aligned}$$

$$
\begin{aligned}
I_{25} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{1i}(\theta)}{\partial \phi_s}\right) = -2tr\left[\frac{\partial D^{\frac{1}{2}}(\Phi)}{\partial \phi_s} R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho} R^{-1}(\widetilde{\lambda}) R(\lambda) D^{\frac{1}{2}}(\Phi)\right] \\
&= -2\phi_s^{-1}\widetilde{B}_{12}R_{22}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{21}}{\partial \rho} - \phi_s^{-1}\widetilde{B}_{12}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{12}R_{21} - \phi_s^{-1}\widetilde{B}_{11}R_{12}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{21}}{\partial \rho} \\
I_{32} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \rho}\right) = tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \alpha} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho}\right] \\
&= -2\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}\frac{\partial R_{21}}{\partial \rho} \\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \alpha}\right) = -tr\left[\widehat{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}\frac{\partial R_{22}}{\partial \alpha}\right] \\
I_{34} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \phi_p}\right) = -2tr\left[\frac{\partial D^{\frac{1}{2}}(\Phi)}{\partial \phi_p} R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \alpha} R^{-1}(\widetilde{\lambda}) R(\lambda) D^{\frac{1}{2}}(\Phi)\right] \\
&= -\phi_p^{-1}\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21} - \phi_p^{-1}\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{21} \\
I_{35} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial h_{2i}(\theta)}{\partial \phi_s}\right) = -2tr\left[\frac{\partial D^{\frac{1}{2}}(\Phi)}{\partial \phi_s} R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \alpha} R^{-1}(\widetilde{\lambda}) R(\lambda) D^{\frac{1}{2}}(\Phi)\right] \\
&= -\phi_s^{-1}R_{12}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21} - \phi_s^{-1}tr\left[\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}\right] \\
I_{44} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_{1i}(\theta)}{\partial \phi_p}\right) = 1 \\
I_{55} &= -\frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\partial g_{2i}(\theta)}{\partial \phi_s}\right) = t-1.
\end{aligned}
$$

We can also show that $M_q(\theta)$ has the following elements

$$
M_q(\theta) = \begin{pmatrix} M_{11} & 0 & 0 & 0 & 0 \\ 0 & M_{22} & M_{23} & M_{24} & M_{25} \\ 0 & M_{23} & M_{33} & M_{34} & M_{35} \\ 0 & M_{24} & M_{34} & M_{44} & M_{45} \\ 0 & M_{25} & M_{35} & M_{45} & M_{55} \end{pmatrix} \tag{3.4.13}
$$

where

$$
\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{0i}(\theta)) = \frac{1}{n}\sum_{i=1}^{n} X_i'\Sigma^{-1}(\lambda,\rho)X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta))
\end{aligned}
$$

$$
\begin{aligned}
&= 2tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\rho}R^{-1}(\widetilde{\lambda})R(\lambda)R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\rho}R^{-1}(\widetilde{\lambda})R(\lambda)\right] \\
&= 2\left[\widetilde{B}_{12}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}+\widetilde{B}_{12}R_{21})\right]^2 + 2\left[\widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\right]^2 \\
&\quad +4\widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\widetilde{B}_{12}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}+\widetilde{B}_{12}R_{21}) \\
&\quad 4\widetilde{B}_{12}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}R_{12}+\widetilde{B}_{12}R_{22})\left[\widetilde{B}_{22}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}+\widetilde{B}_{12}R_{21})\right] \\
&\quad +4\widetilde{B}_{12}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}R_{12}+\widetilde{B}_{12}R_{22})\left[\widetilde{B}_{21}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\right] \\
&\quad +4\widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\left[\widetilde{B}_{22}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}+\widetilde{B}_{12}R_{21})\right] \\
&\quad 4\widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\left[\widetilde{B}_{21}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\right] \\
&\quad +2\left[(\widetilde{B}_{11}R_{12}+\widetilde{B}_{12}R_{22})\widetilde{B}_{22}\frac{\partial\widetilde{R}_{21}}{\partial\rho}\right]^2 \\
&\quad +4\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{22}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}R_{12}+\widetilde{B}_{12}R_{21})\widetilde{B}_{21} \\
&\quad +2\left[\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{21}\right]^2
\end{aligned}
$$

$$
\begin{aligned}
M_{23} &= \frac{1}{n}\sum_{i=1}^{n}Cov(h_{1i}(\theta),h_{2i}(\theta)) \\
&= 2tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\rho}R^{-1}(\widetilde{\lambda})R(\lambda)R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial\alpha}R^{-1}(\widetilde{\lambda})R(\lambda)\right] \\
&= 2\left(\widetilde{B}_{12}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}+\widetilde{B}_{12}R_{21})\right)\widetilde{B}_{12}\frac{\partial\widetilde{R}_{22}}{\partial\alpha}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21}) \\
&\quad 2\left(\widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\right)\widetilde{B}_{12}\frac{\partial\widetilde{R}_{22}}{\partial\alpha}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21}) \\
&\quad 2\left(\widetilde{B}_{12}\frac{\partial\widetilde{R}_{21}}{\partial\rho}(\widetilde{B}_{11}R_{12}+\widetilde{B}_{12}R_{22})\right)\widetilde{B}_{22}\frac{\partial\widetilde{R}_{22}}{\partial\alpha}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21}) \\
&\quad 2\left(\widetilde{B}_{11}\frac{\partial\widetilde{R}_{12}}{\partial\rho}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\right)\widetilde{B}_{22}\frac{\partial\widetilde{R}_{22}}{\partial\alpha}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21}) \\
&\quad 2\left((\widetilde{B}_{11}+\widetilde{B}_{12}R_{21})\widetilde{B}_{12}\right)\frac{\partial\widetilde{R}_{22}}{\partial\alpha}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{22}\frac{\partial\widetilde{R}_{21}}{\partial\rho} \\
&\quad 2\left((\widetilde{B}_{11}R_{12}+\widetilde{B}_{12}R_{22})\widetilde{B}_{22}\right)\frac{\partial\widetilde{R}_{22}}{\partial\alpha}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{22}\frac{\partial\widetilde{R}_{21}}{\partial\rho}
\end{aligned}
$$

$$2\frac{\partial \widetilde{R}_{12}}{\partial \rho}\left((\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\widetilde{B}_{12}\right)\frac{\partial \widetilde{R}_{22}}{\partial \alpha}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{21}$$

$$2\frac{\partial \widetilde{R}_{12}}{\partial \rho}\left((\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{22}\right)\frac{\partial \widetilde{R}_{22}}{\partial \alpha}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{21}$$

$$\begin{aligned}
M_{24} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), g_{1i}(\theta)) \\
&= -4\phi_p\left[\widetilde{B}_{12}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{11}+\widetilde{B}_{12}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{12}R_{21}\right] \\
&\quad -4\phi_p\left[\widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{22}R_{21}+\widetilde{B}_{12}R_{21}R_{12}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\right] \\
M_{25} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{1i}(\theta), g_{2i}(\theta)) \\
&= -4\phi_s\left[\widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{21}R_{12}R_{21}+\widetilde{B}_{12}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\widetilde{B}_{12}R_{22}R_{21}\right] \\
&\quad -4\phi_s\left[\widetilde{B}_{11}\frac{\partial \widetilde{R}_{12}}{\partial \rho}\widetilde{B}_{22}R_{22}R_{21}+\widetilde{B}_{12}R_{22}R_{22}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{21}}{\partial \rho}\right] \\
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta)) \\
&= 2tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \alpha}R^{-1}(\widetilde{\lambda})R(\lambda)R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \alpha}R^{-1}(\widetilde{\lambda})R(\lambda)\right] \\
&= 2\left[\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21})\right]^2+2\left[R_{12}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21}\right]^2 \\
&\quad +4\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}(\widetilde{B}_{21}R_{12}+\widetilde{B}_{22}R_{22})\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}(\widetilde{B}_{21}+\widetilde{B}_{22}R_{21}) \\
&\quad +4R_{12}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21} \\
&\quad +2tr\left[\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}\right] \\
M_{34} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta), g_{1i}(\theta)) \\
&= -2\phi_p\left[\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21}+2\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{21}+R_{12}\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{21}\right] \\
M_{35} &= \frac{1}{n}\sum_{i=1}^{n} Cov(h_{2i}(\theta), g_{2i}(\theta))
\end{aligned}$$

$$
\begin{aligned}
&= -2\phi_s \left[\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{21}R_{12}R_{21} + 2\widetilde{B}_{12}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}R_{21}\right] \\
&\quad -2\phi_s \left[tr\left(\widetilde{B}_{22}\frac{\partial \widetilde{R}_{22}}{\partial \alpha}\widetilde{B}_{22}R_{22}R_{22}\right)\right] \\
M_{44} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_{1i}(\theta)) = 2\phi_p^2 \\
M_{45} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_{1i}(\theta), g_{2i}(\theta)) = 2\phi_p\phi_s\left(\frac{\rho^2-\rho^{2t}}{1-\rho^2}\right) \\
M_{55} &= \frac{1}{n}\sum_{i=1}^{n} Cov(g_{2i}(\theta)) = 2\phi_s^2\left[(t-1) + 2\sum_{j=1}^{t-2}(t-1-j)\alpha^{2j}\right].
\end{aligned}
$$

By taking the inverse of $I_q(\theta)$ and pre- and post-multiplying upon $M_q(\theta)$ we obtain the asymptotic variance of the QLS estimators. Note that, based on (3.4.12) and (3.4.13), the estimator for β is uncorrelated with the estimators for the other parameters. It is also important to note that (3.4.3) and (3.4.11) differ only by the value of $\widetilde{\lambda}$. Thus, if $\widetilde{\lambda}$ is close to the population value of λ, the asymptotic variances of the MLE and QLS estimators will also be close.

III.4.4 Comparison of Asymptotic Performance

In this section we compute the asymptotic relative efficiency (ARE) of the variance and correlation parameters for the MLE, MOM and QLS procedures. In each of the three cases we set $t = 4$ and $n = 1,000$ and compute the asymptotic variances of the estimators derived in III.4 at specific values of Φ and λ. The ratio of the asymptotic variances for the same estimator, then , is the ARE. By varying ρ and α over their admissible range, we get an idea not only of the large-sample efficiency of one estimating procedure with regards to another but also how the efficiency changes with the parameter values. For our purposes, we have selected $\phi_p = 49$ and $\phi_s = 16$.

We start with estimators of ρ. The ARE plot for the MLE and MOM procedures is found in Figure 3.1. This plot shows that the asymptotic variances are comparable only for a small region when ρ is close to zero. The ARE is low elsewhere. Figure 3.2 shows ARE for the MLE and QLS procedures. Here we see that the variances are comparable over a wide range of admissible values. The ARE is low only when α is extremely large (both positive and negative). Finally, Figure 3.3 shows the ARE

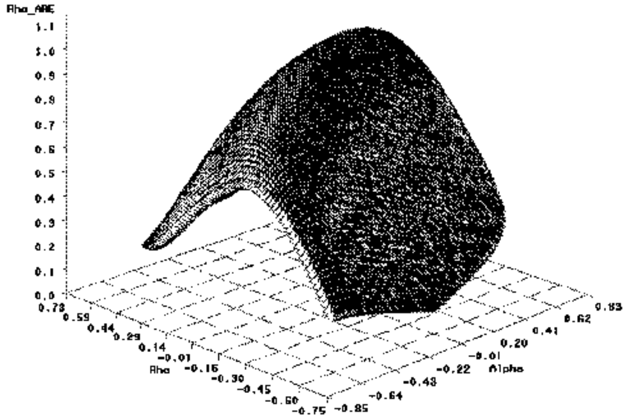

Figure 3.1: ρ ARE for MLE and MOM Methods

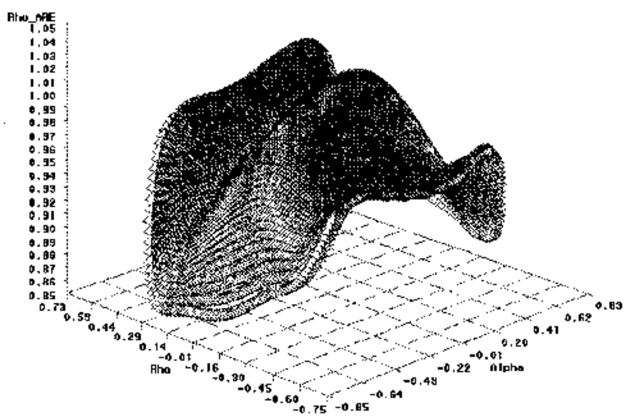

Figure 3.2: ρ ARE for MLE and QLS Methods

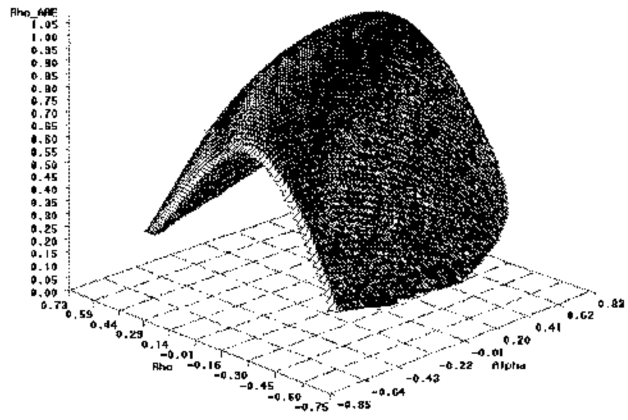

Figure 3.3: ρ ARE for QLS and MOM Methods

between the QLS and MOM procedures. Here we see that the ARE is comparable only for a small area when ρ is close to zero. These results imply that the the MLE and QLS estimators of ρ are highly competitive asymptotically, whereas both methods are asymptotically superior to the MOM correlation estimators.

We now focus on α estimators. The ARE plot for the MLE and MOM procedures is found in Figure 3.4. This plot shows the asymptotic variances are comparable for small values of α, and the efficiency of the MLE increases with respect to the MoM estimator as α increases in magnitude. Figure 3.5 shows the ARE for the MLE and QLS procedures. Here we see that ARE is comparable over a wide range of admissible values and is low only when ρ and α are extremely large (both positive and negative). Finally, Figure 3.6 shows the ARE for the QLS and MOM procedures. Here we see that the ARE is comparable over an area corresponding to small values of α, and the ARE decreases as α increases in magnitude. These results imply that the QLS correlation estimator of α is highly competitive with the MLE asymptotically, whereas both methods are asymptotically superior to the MOM correlation estimator of α.

Lastly we analyze the variance parameters. Recall that we used the same estimators for Φ in all three methods, and thus we would expect that the ARE be close between each procedure. In fact, we see that this is indeed the case. Figures 3.7

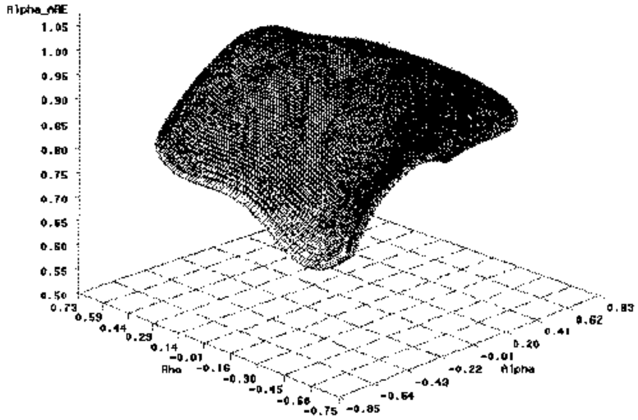

Figure 3.4: α ARE for MLE and MOM Methods

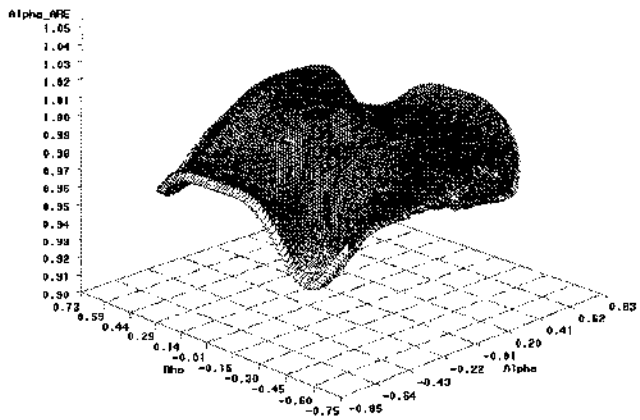

Figure 3.5: α ARE for MLE and QLS Methods

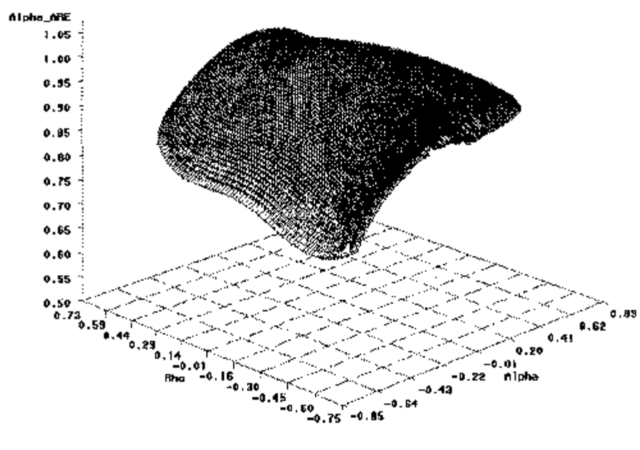

Figure 3.6: α ARE for QLS and MOM Methods

and 3.8 give the ARE for all three comparisons: the MLE and MOM procedures, the MLE and QLS procedures, and the QLS and MoM procedures, for ϕ_p and ϕ_s, respectively. Based on these plots, we see that (asymptotically) all three procedures estimate $\Phi = (\phi_p, \phi_s)$ with the same precision.

III.5 Small-Sample Performance

In this Section we estimate the small-sample variance of the correlation parameter estimators through use of simulated data. To do this we fix Φ, with $\phi_p = 49$, $\phi_s = 16$, and select a pair of values for ρ and α within their positive definite range. With these parameter values, we simulate $n = 30$ observations from a multivariate normal distribution with $t = 4$ and calculate the ML, MoM and QLS estimators. We then repeat this procedure 1,000 times for the same values of λ. We estimate the variance of the correlation parameter estimator by summing the squared deviations of the estimate from the "true" correlation parameter value and divide by the number of times the estimating procedure yielded feasible estimates. We then repeat this procedure for other values of ρ and α so that we can see how the estimated variance of the correlation parameter estimators changes as the correlation parameter values themselves change.

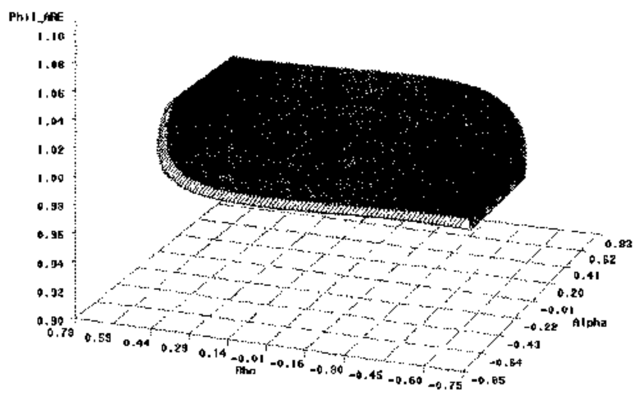

Figure 3.7: ϕ_p ARE for MLE and MOM, MLE and QLS, and QLS and Mom Methods

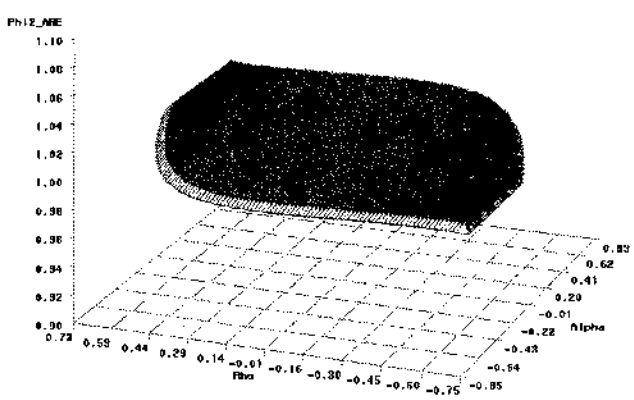

Figure 3.8: ϕ_s ARE for MLE and MOM, MLE and QLS, and QLS and Mom Methods

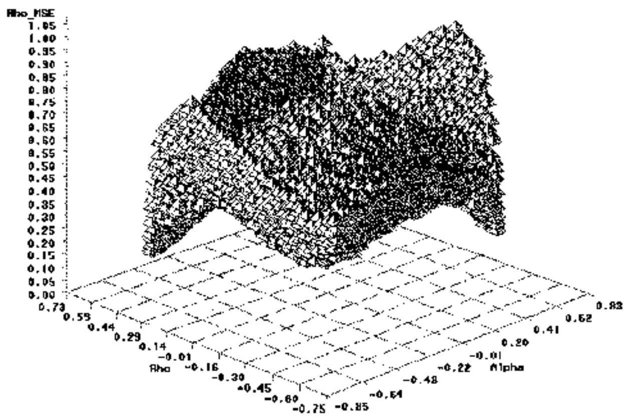

Figure 3.9: ρ RE for MLE and MOM Methods with Normal Data

We also estimate the small-sample variances of the correlation parameters when simulating data from a non-normal distribution. This allows us to gauge the robustness of each estimating procedure to departures from normality. Specifically, we simulate data from a beta distribution with both parameters equal to 1/6, as this gives a u-shaped pdf, which is distinctly non-normal.

To compare the small-sample performance of the estimating procedures, we use the small-sample estimated variances to calculate relative efficiencies. These ratios allow us to determine which estimating procedure has the smallest estimated variance for the correlation parameter estimators, and for which values of ρ and α that this is the case. Note that since the asymptotic relative efficiencies for the variance parameters ϕ_p and ϕ_s everywhere equal to one, we will not include the small-sample efficiencies for those parameters here. However, they were found to be close to one for most values of ρ and α away from the positive definite boundary.

III.5.1 Small-Sample Normal Case

We first study the case of normally distributed simulated data, and begin with estimators of ρ. Figure 3.9 gives the estimated efficiencies between the MLE and MoM procedures. In this Figure we see that the efficiencies are below one everywhere,

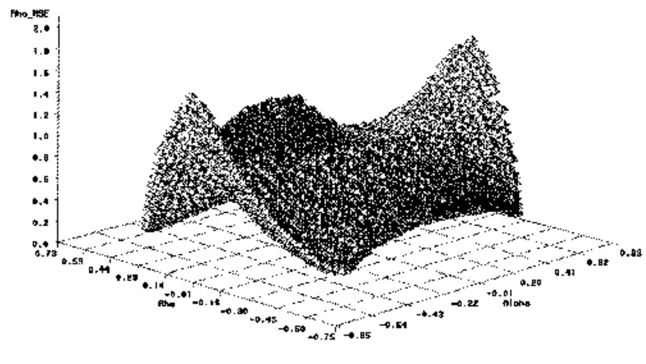

Figure 3.10: ρ RE for MLE and QLS Methods with Normal Data

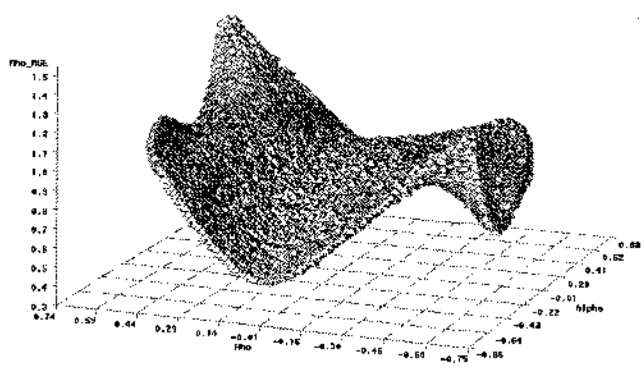

Figure 3.11: ρ RE for QLS and MOM Methods with Normal Data

meaning that the MLE has smaller estimated variance than the MoM estimator for all correlation parameter values. Note that for small values of ρ and α the estimated variance of the MoM estimator is comparable to that of the MLE, and this is especially the case for large values of α. For large values of ρ the MLE has much smaller estimated variance than the MoM estimator. The efficiencies for the MLE and QLS procedures are found in Figure 3.10. Here we note that the estimated variance for the QLS estimator is comparable to that of the MLE for small and moderate values of ρ, and is smaller for large values of α. For large values of ρ, the MLE has smaller

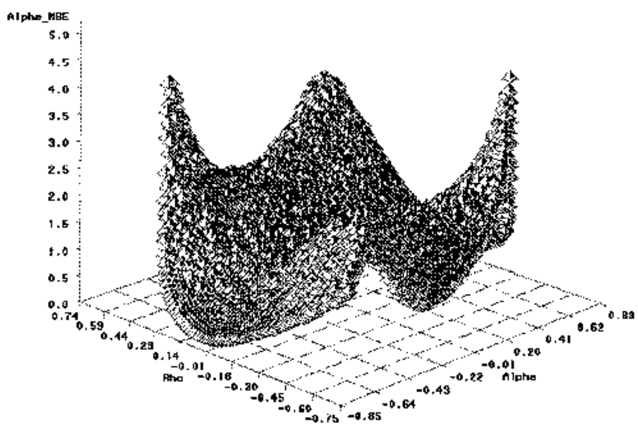

Figure 3.12: α RE for MLE and MOM Methods with Normal Data

estimated variance. Lastly, the efficiencies for the QLS and MoM procedures are found in Figure 3.11. In this Figure we see that for small ρ and α the QLS estimator has smaller estimated variance than the MoM estimator. Only for extreme values of the correlation parameters does the MoM estimator have smaller estimated variance than the QLS estimator. So among estimators of ρ, both the MLE and QLS procedures outperform the MoM procedure in the small-sample normal-data case, and the QLS procedure is comparable to the MLE for most values of the correlation parameters.

We now move on to estimators of α. Figure 3.12 shows the estimated efficiencies for the MLE and MoM procedures. Here we see that for all but extreme values of α the MoM estimator has smaller estimated variance than the MLE. This is especially the case for extreme values of ρ and α. Figure 3.13 gives the estimated efficiencies for the MLE and QLS procedures. In this Figure we see that the estimated variance of the QLS estimator, like that for MoM, is smaller than that for the MLE almost everywhere, especially for extreme values of ρ and α. Lastly, Figure 3.14 gives the estimated efficiencies for the QLS and MoM procedures. Here we see that for small and moderate values of α, the estimated variances for the QLS and MoM estimators are roughly the same. For large values of α, the QLS estimator has smaller estimated variance, and for large values of ρ, the MoM estimator has smaller estimated variance.

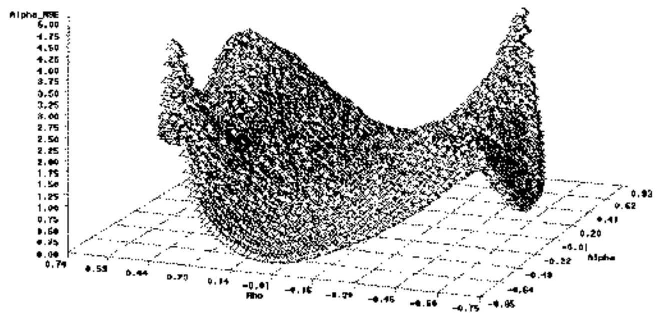

Figure 3.13: α RE for MLE and QLS Methods with Normal Data

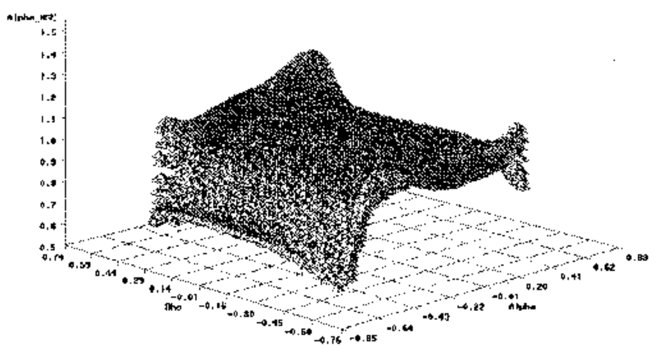

Figure 3.14: α RE for QLS and MOM Methods with Normal Data

Table 3.1: Estimated Infeasibility Probabilities (Normal, Heterogeneous Variance Case)

			α			
ρ	Method	-0.80	-0.70	0.3	0.70	0.80
-0.75	MLE	0.175	0.590	N/A	N/A	N/A
	MoM	0.000	0.000	N/A	N/A	N/A
	QLS	0.000	0.026	N/A	N/A	N/A
-0.60	MLE	0.001	0.065	N/A	N/A	N/A
	MoM	0.000	0.000	N/A	N/A	N/A
	QLS	0.000	0.000	N/A	N/A	N/A
0.10	MLE	0.000	0.002	0.030	0.001	0.000
	MoM	0.002	0.001	0.000	0.001	0.001
	QLS	0.000	0.000	0.000	0.000	0.000
0.60	MLE	N/A	N/A	0.436	0.068	0.002
	MoM	N/A	N/A	0.000	0.000	0.001
	QLS	N/A	N/A	0.012	0.000	0.000
0.70	MLE	N/A	N/A	0.518	0.307	0.042
	MoM	N/A	N/A	0.000	0.000	0.000
	QLS	N/A	N/A	0.031	0.007	0.000

For estimators of α, then, we see that both the QLS and MoM estimators perform better than the MLE in the small-sample normal-data case, with the QLS and MoM procedures performing equally well.

Along with the estimated variances and efficiencies, we also estimate the probability of infeasibility for each procedure. Using the same simulation procedure (and simulations) that generated the estimated variances, we estimate the infeasibility probability as the number of times the estimating procedure gave correlation parameter estimates that were outside the positive definite boundary, divided by the total number of simulations $(1,000)$. The estimated infeasibility probabilites for select values or ρ and α are given in Table 3.1. Most strikingly we see that both the MoM and QLS procedures have almost negligible estimated infeasibility probabilities for all values of ρ and α listed in the Table. Recall that in Table 2.2 in Chapter II, the MoM procedure had large estimated probabilities for extreme values of ρ and α. This essentially means that these two procedures produce correlation parameter estimators within the positive definite boundary nearly all the time. Note that the estimated probabilities are high for the MLE procedure for extreme values of the correlation parameters.

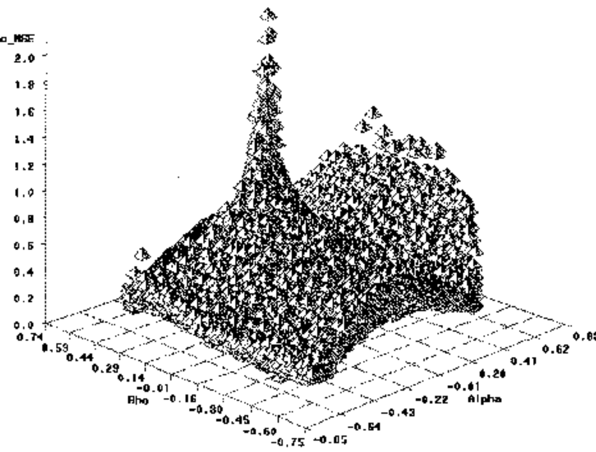

Figure 3.15: ρ RE for MLE and MOM Methods with Non-Normal Data

III.5.2 Small-Sample Non-Normal Case

Now we study the case where our data are simulated from a non-normal distribution. Beginning with estimators of ρ, the plot of small-sample estimated efficiencies for the MLE and MoM procedures is found in Figure 3.15. Here we see that for most values of ρ and α the estimated efficiency is below one, indicating that the estimated variance for the MLE is smaller than that for the MoM estimator. Notably, the estimated variance of the moment estimator is close to that of the MLE for small values of α, and its efficiency with respect to the MLE decreases as α increases in magnitude. Figure 3.16 gives the estimated relative efficiency for the MLE and QLS procedures. Here we see that only for extremely large values of α is the variance of the QLS estimator smaller than that of the MLE. However, over a wide range of small to moderately large values of α the estimated variance of the QLS estimator is comparable to that of the MLE. Lastly, Figure 3.17 gives the estimated relative efficiency for the QLS and MoM procedures. Here we see that for most values of ρ and α the QLS procedure has smaller estimated variance than the MoM procedure. For small values of ρ we see that this is especially the case. For small correlation values, the estimated variances of both procedures are more or less equal. In the small-sample non-normal case with regards to estimators of ρ, we see that both the MLE and QLS procedures have smaller estimated variances than the MoM procedure,

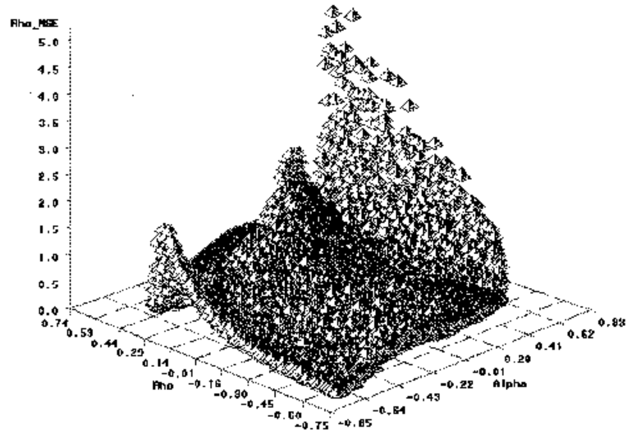

Figure 3.16: ρ RE for MLE and QLS Methods with Non-Normal Data

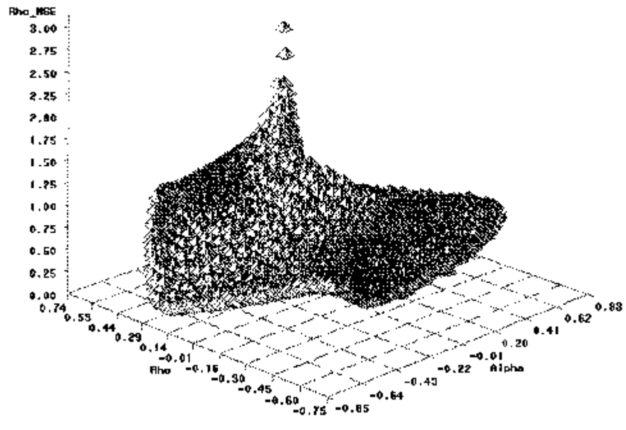

Figure 3.17: ρ RE for QLS and MOM Methods with Non-Normal Data

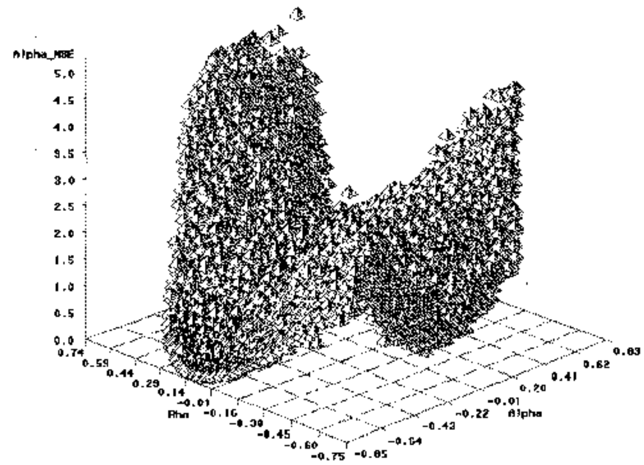

Figure 3.18: α RE for MLE and MOM Methods with Non-Normal Data

and the QLS and MLE estimators are comparable for most correlation values.

We now move on to estimators of α. Figure 3.18 gives the estimated relative efficiency for the MLE and MoM procedures. Here we see that for all but very large α, the estimated relative efficiency is larger than one, indicating that the variance of the MoM estimator is smaller than that for the MLE. This is especially the case for large values of ρ. Figure 3.19 gives the estimated relative efficiency for the MLE and QLS procedures. Like the MLE and MoM case, we see here that the estimated efficiency is greater than one for almost all correlation parameter values, indicating that the estimated variance of the QLS estimator is smaller than that of the MLE. Notice in some places the estimated efficiency is as high as 8. Finally, the estimated relative efficiency for the QLS and MoM procedures is found in Figure 3.20. In this plot we see that for small values of α, the estimated efficiency is close to one, indicating that the estimated variances for the parameter estimators are close in value. However, as α increases in magnitude, the variance of the moment estimator increases with respect to the QLS estimator. Only for very large values of ρ close to the positive definite boundary does the MoM estimator have smaller estimated variance. For estimators of α, then, both the QLS and MoM estimators have smaller estimated variance than the MLE, while the estimated variance for the QLS estimator is smaller than that for the MoM estimator.

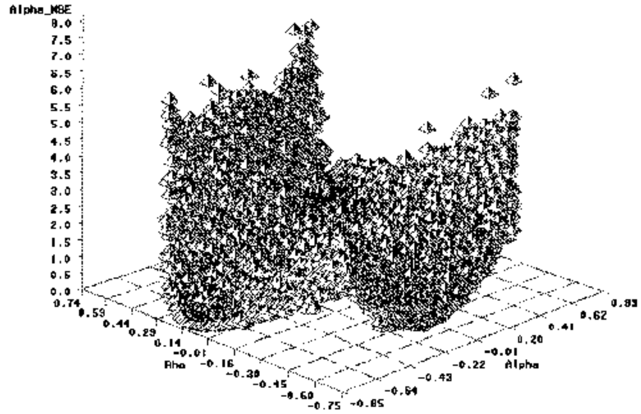

Figure 3.19: α RE for MLE and QLS Methods with Non-Normal Data

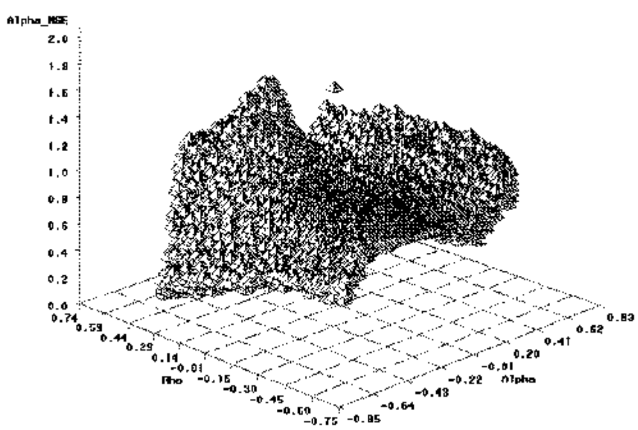

Figure 3.20: α RE for QLS and MOM Methods with Non-Normal Data

Table 3.2: Estimated Infeasibility Probabilities (Non-Normal, Heterogeneous Variance Case)

			α			
ρ	Method	-0.80	-0.70	0.3	0.70	0.80
-0.75	MLE	0.052	0.016	N/A	N/A	N/A
	MoM	0.000	0.000	N/A	N/A	N/A
	QLS	0.000	0.002	N/A	N/A	N/A
-0.60	MLE	0.505	0.015	N/A	N/A	N/A
	MoM	0.001	0.000	N/A	N/A	N/A
	QLS	0.000	0.000	N/A	N/A	N/A
0.10	MLE	0.169	0.050	0.004	0.001	0.000
	MoM	0.155	0.090	0.000	0.017	0.112
	QLS	0.000	0.000	0.000	0.000	0.000
0.60	MLE	N/A	N/A	0.836	0.337	0.122
	MoM	N/A	N/A	0.000	0.442	0.832
	QLS	N/A	N/A	0.014	0.000	0.000
0.70	MLE	N/A	N/A	0.172	0.836	0.338
	MoM	N/A	N/A	0.001	0.006	0.718
	QLS	N/A	N/A	0.011	0.000	0.000

We have also estimated infeasibility probabilities for the estimating procedures using the same simulated data used to estimate the small sample variances in the non-normal case. These estimates are found in Table 3.2. Note that the MLE procedure has high estimated infeasibility probabilities for large values of ρ and α, while the QLS procedure has very small estimated probabilities for all correlation values. The MoM procedure has very small estimated infeasibility probabilities for all but large positive values of ρ.

CHAPTER IV

EQUICORRELATED STRUCTURE FOR A NUCLEAR FAMILY

IV.1 Introduction

In this Chapter we focus on the nuclear family, consisting of two parents and $t-2$ children, where the dependencies exhibited between parents and children, as well as dependencies between children, are equicorrelated. Here we assume that the $(t \times 1)$ response vector Y_i has mean vector $X_i\beta$ and variance-covariance matrix $\Sigma(\lambda, \phi) = \phi R(\lambda)$, where X_i is the $(t \times p)$ matrix of covariates for the ith family, β is a $(p \times 1)$ vector of regression coefficients, ϕ is the variance parameter, and $\lambda = (\gamma, \rho_1, \rho_2, \alpha)$ is the vector of correlation parameters. The correlation matrix $R(\lambda)$ is of the following form.

$$R(\lambda) = \begin{pmatrix} 1 & \gamma & \rho_1 & \rho_1 & & \rho_1 \\ \gamma & 1 & \rho_2 & \rho_2 & \cdots & \rho_2 \\ \rho_1 & \rho_2 & 1 & \alpha & & \alpha \\ \rho_1 & \rho_2 & \alpha & 1 & & \alpha \\ & \vdots & & & \ddots & \vdots \\ \rho_1 & \rho_2 & \alpha & \alpha & \cdots & 1 \end{pmatrix} \tag{4.1.1}$$

For correlation structure (4.1.1), note that γ is the correlation between parents, ρ_1 is the correlation between the first parent and the children, ρ_2 is the correlation between the second parent and the children, and α is the correlation between children. Also recall from Chapter I that (4.1.1) is the same correlation structure used in Shoukri and Ward (1989) where the authors modeled heterogeneous variances. Note that we are using a homogeneous intra-class variance structure. Though this correlation structure is not new, we do introduce its application to the quasi-least squares estimating procedure.

For the one-parent case of the equicorrelated structure, Srivastava (1984) showed that a simple transformation simplifies both the correlation matrix and estimation of the correlation parameters. In a similar fashion, we extend that transformation to the nuclear family case. Define Γ as the following transformation matrix

$$\Gamma = \begin{pmatrix} I_2 & \underline{0} \\ \underline{0} & H \end{pmatrix} \tag{4.1.2}$$

where I_2 is a (2×2) identity matrix, and H is a $(t-2) \times (t-2)$ Helmert matrix of the following form.

$$H = \begin{pmatrix} \frac{1}{t-2} & \frac{1}{t-2} & \frac{1}{t-2} & \frac{1}{t-2} & & \frac{1}{t-2} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & \cdots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ \frac{1}{\sqrt{(t-2)(t-3)}} & \frac{1}{\sqrt{(t-2)(t-3)}} & \cdots & & \frac{1}{\sqrt{(t-2)(t-3)}} & \frac{-(t-3)}{\sqrt{(t-2)(t-3)}} \end{pmatrix} \tag{4.1.3}$$

Based on this Helmert matrix, we have

$$HH' = \begin{pmatrix} \frac{1}{t-2} & \underline{0} \\ \underline{0} & I_{t-3} \end{pmatrix}$$

and

$$He = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e'H' = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$$

where e is a $(t-2)\times 1$ vector of 1's. If $Z_i = (Y_i - X_i\beta)$ has variance-covariance matrix $\Sigma(\lambda, \phi) = \phi R(\lambda)$, then ΓZ_i has variance-covariance matrix $\Gamma\Sigma(\lambda,\phi)\Gamma' = \phi\Gamma R(\lambda)\Gamma'$. If we partition $R(\lambda)$ as follows

$$R(\lambda) = \left(\begin{array}{cc|cccc} 1 & \gamma & \rho_1 & \rho_1 & \cdots & \rho_1 \\ \gamma & 1 & \rho_2 & \rho_2 & \cdots & \rho_2 \\ \hline \rho_1 & \rho_2 & 1 & \alpha & \cdots & \alpha \\ \rho_1 & \rho_2 & \alpha & 1 & \cdots & \alpha \\ & \vdots & & & \ddots & \vdots \\ \rho_1 & \rho_2 & \alpha & 1 & \cdots & 1\alpha \end{array}\right) = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

then the transformed correlation matrix becomes

$$\begin{aligned} \Gamma R(\lambda)\Gamma' &= \begin{pmatrix} I & \underline{0} \\ \underline{0} & H \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} I & \underline{0} \\ \underline{0} & H' \end{pmatrix} \\ &= \begin{pmatrix} R_{11} & R_{12}H' \\ HR_{21} & HR_{22}H' \end{pmatrix} \end{aligned}$$

where

$$
\begin{aligned}
R_{12}H' &= \begin{pmatrix} \rho_1 e'H' \\ \rho_2 e'H' \end{pmatrix} = \begin{pmatrix} \rho_1 & 0 & \cdots & 0 \\ \rho_2 & 0 & \cdots & 0 \end{pmatrix} \\
HR_{21} &= (R_{12}H')' \\
HR_{22}H' &= H\left[(1-\alpha)I_{t-2} + \alpha ee'\right]H' \\
&= (1-\alpha)HH' + \alpha Hee'H' \\
&= \begin{pmatrix} \frac{1+(t-3)\alpha}{t-2} & 0 & 0 & \cdots & 0 \\ 0 & 1-\alpha & 0 & \cdots & 0 \\ 0 & 0 & 1-\alpha & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1-\alpha \end{pmatrix}.
\end{aligned}
$$

Thus, the fully transformed variance-covariance matrix becomes

$$
\Gamma\Sigma(\lambda,\phi)\Gamma' = \phi \begin{pmatrix} 1 & \gamma & \rho_1 & 0 & \cdots & 0 \\ \gamma & 1 & \rho_2 & 0 & \cdots & 0 \\ \rho_1 & \rho_2 & \frac{1+(t-3)\alpha}{t-2} & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1-\alpha & \cdots & 0 \\ & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1-\alpha \end{pmatrix}. \tag{4.1.4}
$$

For simplicity, we refer to ΓZ_i as Z_i and $\Gamma R(\lambda)\Gamma'$ as $R(\lambda)$ for the remainder of this Chapter. To avoid confusion, we will not refer to the untransformed varieties unless specified.

The rest of this Chapter is outlined as follows. In Section IV.2 we derive the determinant and inverse of (4.1.4) and also find the positive definite range of the parameters. In Section IV.3 we derive parameter estimators using the maximum likelihood, method of moment and quasi-least squares procedures, and in Section IV.4 we find the asymptotic variances of those estimators and compare their asymptotic performance.

IV.2 Properties of Correlation Matrix

To find the determinant of (4.1.4) it helps to partition the correlation matrix as follows

$$R(\lambda) = \begin{pmatrix} R_{11} & \underline{0} \\ \underline{0} & R_{22} \end{pmatrix}$$

where

$$\begin{aligned} R_{11} &= \begin{pmatrix} 1 & \gamma & \rho_1 \\ \gamma & 1 & \rho_2 \\ \rho_1 & \rho_2 & c \end{pmatrix} \\ R_{22} &= (1-\alpha)I_{t-3} \end{aligned}$$

and $c = \frac{1+(t-3)\alpha}{t-2}$. Then we have

$$|R(\lambda)| = \begin{vmatrix} R_{11} & \underline{0} \\ \underline{0} & R_{22} \end{vmatrix} = |R_{11}||R_{22}|.$$

Using properties of the determinant of a partitioned matrix, we have

$$\begin{aligned} |R_{11}| &= \begin{vmatrix} 1 & \gamma & \rho_1 \\ \gamma & 1 & \rho_2 \\ \rho_1 & \rho_2 & c \end{vmatrix} = \begin{vmatrix} 1 & \gamma \\ \gamma & 1 \end{vmatrix} \left(c - \begin{pmatrix} \rho_1 & \rho_2 \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ \gamma & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \right) \\ &= c(1-\gamma^2) - \left(\rho_1^2 + \rho_2^2 - 2\gamma\rho_1\rho_2\right) \end{aligned}$$

and by recalling that the determinant of a diagonal matrix is the product of those diagonal elements, we have

$$|R_{22}| = |(1-\alpha)I_{t-3}| = (1-\alpha)^{t-3}.$$

Putting these together, then

$$|R(\lambda)| = (1-\alpha)^{t-3}\left[c(1-\gamma^2) - \left(\rho_1^2 + \rho_2^2 - 2\gamma\rho_1\rho_2\right)\right]. \tag{4.2.1}$$

To find the inverse of (4.1.4) we again make use of the partitioned form to get

$$R^{-1}(\lambda) = \begin{pmatrix} R_{11} & \underline{0} \\ \underline{0} & R_{22} \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{4.2.2}$$

where

$$
\begin{aligned}
B_{11} &= \left(R_{11} - \underline{0}R_{22}^{-1}\underline{0}'\right)^{-1} = R_{11}^{-1} = \begin{pmatrix} 1 & \gamma & \rho_1 \\ \gamma & 1 & \rho_2 \\ \rho_1 & \rho_2 & c \end{pmatrix}^{-1} \\
&= \frac{1}{(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2} \begin{pmatrix} c-\rho_2^2 & \rho_1\rho_2 - \gamma c & \gamma\rho_2 - \rho_1 \\ \rho_1\rho_2 - \gamma c & c - \rho_1^2 & \gamma\rho_1 - \rho_2 \\ \gamma\rho_2 - \rho_1 & \gamma\rho_1 - \rho_2 & 1-\gamma^2 \end{pmatrix} \\
B_{12} &= -B_{11}\underline{0}R_{22}^{-1} = \underline{0} \\
B_{21} &= B_{12}' = \underline{0} \\
B_{22} &= R_{22}^{-1} + R_{22}^{-1}\underline{0}B_{11}\underline{0}R_{22}^{-1} = R_{22}^{-1} = \left((1-\alpha)I_{t-3}\right)^{-1} = \frac{1}{1-\alpha}I_{t-3}.
\end{aligned}
$$

To find the positive definite range of the correlation parameters ($\lambda = (\gamma, \rho_1, \rho_2, \alpha)$) we set the determinants of the leading minors of (4.1.4) greater than zero and solve for parameter values that satisfy the inequality, the last of which is

$$(1-\alpha)^{t-3}\left[c(1-\gamma^2) - \left(\rho_1^2 + \rho_2^2 - 2\gamma\rho_1\rho_2\right)\right] > 0. \tag{4.2.3}$$

We begin with γ, noting that we only have to use the last principle minor (i.e, the determinant (4.2.1)) as the first $(t-1)$ do not include γ. So we start with the following expression

$$c(1-\gamma^2) - \left(\rho_1^2 + \rho_2^2 - 2\gamma\rho_1\rho_2\right) > 0 \tag{4.2.4}$$

which is a quadratic expression in terms of γ. Thus we find the positive definite range by solving for γ using the quadratic formula. Doing so gives the following bounds for γ.

$$\frac{\rho_1\rho_2 - \sqrt{\rho_1^2\rho_2^2 - c(\rho_1^2+\rho_2^2) + c^2}}{c} < \gamma < \frac{\rho_1\rho_2 + \sqrt{\rho_1^2\rho_2^2 - c(\rho_1^2+\rho_2^2) + c^2}}{c}$$

In a similar fashion (4.2.4) is also quadratic in terms of both ρ_1 and ρ_2. Solving for both parameters using the quadratic formula gives the following positive definite bounds.

$$
\begin{aligned}
-\gamma\rho_2 - \sqrt{(c-\rho_2^2)(1-\gamma^2)} &< \rho_1 < -\gamma\rho_2 + \sqrt{(c-\rho_2^2)(1-\gamma^2)} \\
-\gamma\rho_1 - \sqrt{(c-\rho_1^2)(1-\gamma^2)} &< \rho_2 < -\gamma\rho_1 + \sqrt{(c-\rho_1^2)(1-\gamma^2)}
\end{aligned}
$$

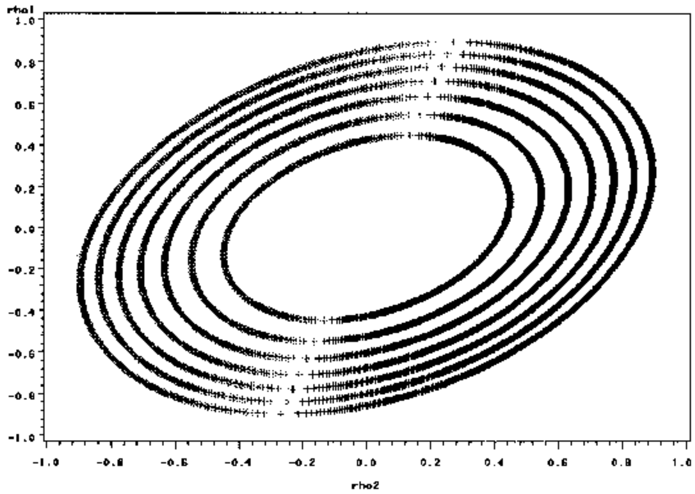

Figure 4.1: P.D. Range Contour Plots of α vs. (ρ_1, ρ_2) with $\gamma = -0.3$

Solving for α, we note that the first $(t-3)$ principle minors yield $(1-\alpha)^i > 0$, which simplifies to $\alpha < 1$. Finally, using, (4.2.4), we see that the expression is linear in terms of α (via 'c'), and we get the following positive definite bounds for α.

$$\frac{(t-2)(\rho_1^2 + \rho_2^2 - 2\gamma\rho_1\rho_2) - (1-\gamma^2)}{(t-3)(1-\gamma^2)} < \alpha < 1$$

To find exact bounds for any of these parameters, we select values of the other parameters and enter those into the positive definite range expressions. Figures 4.1 through 4.4 show the positive definite ranges for ρ_1 and ρ_2 for select values of γ for $t = 5$. Each Figure is a contour plot with each ellipse representing a particular value of α. The values of α are $(\pm 0.6, \pm 0.4, \pm 0.2, 0.0)$, with $\alpha = -0.6$ corresponding to the smallest contour in each Figure and $\alpha = 0.6$ corresponding to the largest.

Lastly, partial derivatives of (4.1.4) are listed in Appendix A.3

IV.3 Parameter Estimation

In this section we derive estimators using the Maximum Likelihood, Method of Moment, and Quasi-Least Squares procedures. For each we use the following estimators

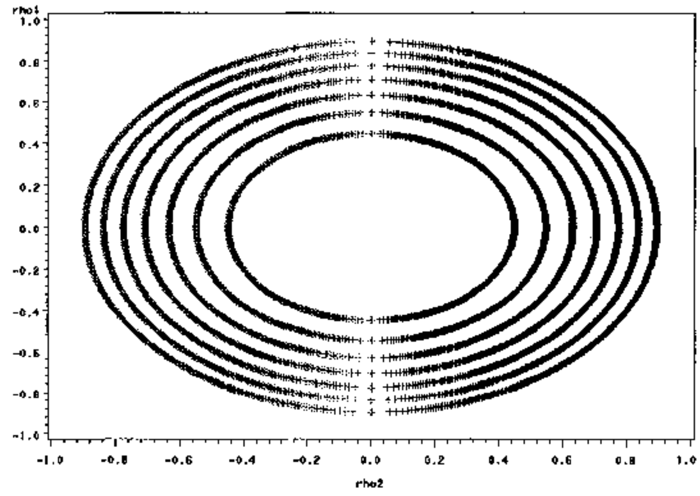

Figure 4.2: P.D. Range Contour Plots of α vs. (ρ_1, ρ_2) with $\gamma = 0.0$

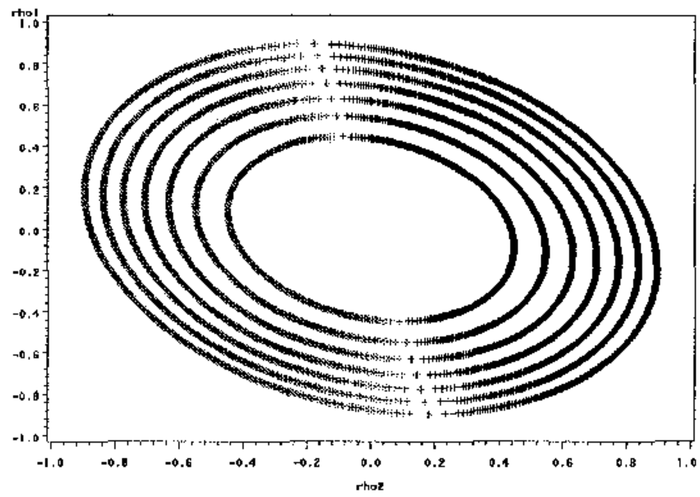

Figure 4.3: P.D. Range Contour Plots of α vs. (ρ_1, ρ_2) with $\gamma = 0.2$

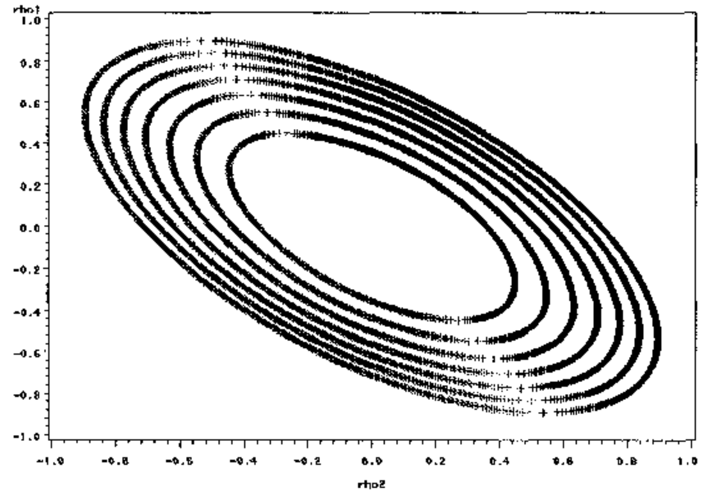

Figure 4.4: P.D. Range Contour Plots of α vs. (ρ_1, ρ_2) with $\gamma = 0.6$

of β and ϕ, respectively,

$$\hat{\beta} = \left[\sum_{i=1}^{n} X_i' R^{-1}(\widehat{\lambda}) X_i\right]^{-1} \sum_{i=1}^{n} X_i' R^{-1}(\widehat{\lambda}) Y_i \tag{4.3.1}$$

$$\widehat{\phi} = \frac{1}{nt} tr\left[R^{-1}(\widehat{\lambda})\widehat{Z}_n\right] = \frac{1}{nt}\sum_{i=1}^{n} \widehat{Z}_i' R^{-1}(\widehat{\lambda})\widehat{Z}_i \tag{4.3.2}$$

where $\widehat{\lambda}$ is the vector of correlation parameter estimators and

$$Z_n = \sum_{i=1}^{n} Z_i Z_i' = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$$

where

$$Z_{11} = \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{12} & z_{22} & z_{23} \\ z_{13} & z_{23} & z_{33} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} z_{i1}^2 & \sum_{i=1}^{n} z_{i1}z_{i2} & \sum_{i=1}^{n} z_{i1}z_{i3} \\ \sum_{i=1}^{n} z_{i1}z_{i2} & \sum_{i=1}^{n} z_{i2}^2 & \sum_{i=1}^{n} z_{i2}z_{i3} \\ \sum_{i=1}^{n} z_{i1}z_{i3} & \sum_{i=1}^{n} z_{i2}z_{i3} & \sum_{i=1}^{n} z_{i3}^2 \end{pmatrix}$$

$$Z_{22} = \begin{pmatrix} z_{44} & z_{45} & \cdots & z_{4t} \\ z_{54} & z_{55} & \cdots & z_{5t} \\ & \vdots & \ddots & \vdots \\ z_{t4} & z_{t5} & \cdots & z_{tt} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} z_{i4}^2 & \sum_{i=1}^{n} z_{i4}z_{i5} & \cdots & \sum_{i=1}^{n} z_{i4}z_{it} \\ \sum_{i=1}^{n} z_{i4}z_{i5} & \sum_{i=1}^{n} z_{i5}^2 & \cdots & \sum_{i=1}^{n} z_{i5}z_{it} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} z_{i4}z_{it} & \sum_{i=1}^{n} z_{i5}z_{it} & \cdots & \sum_{i=1}^{n} z_{it}^2 \end{pmatrix}$$

Note that Z_{12} and $Z_{21} = Z_{12}'$ are defined analogously.

IV.3.1 Maximum Likelihood

For the Maximum Likelihood (MLE) procedure the log-likelihood function is

$$\begin{aligned}\ell &= -\frac{nt}{2}\ln(2\pi) - \frac{nt}{2}\ln(\phi) - \frac{n}{2}\ln|R(\lambda)| - \frac{1}{2\phi}\sum_{i=1}^{n} Z_i' R^{-1}(\lambda) Z_i \qquad (4.3.3)\\ &= -\frac{nt}{2}\ln(2\pi) - \frac{nt}{2}\ln(\phi) - \frac{n}{2}\ln|R(\lambda)| - \frac{1}{2\phi} tr\left[R^{-1}(\lambda) Z_n\right]\end{aligned}$$

where Z_i is the transformed family response-vector for the ith family and $R(\lambda)$ is the transformed correlation structure described in (4.1.4). To find the maximum likelihood estimators we set the first derivative of (4.3.3) with respect to θ equal to zero and solve for the given parameter.

For the correlation parameters, we get the following estimating equations.

$$\begin{aligned}\frac{\partial \ell}{\partial \gamma} &= \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma}\right] - \frac{1}{2\phi} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma} R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.4)\\ \Leftrightarrow\ & n\, tr\left[B_{11}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}\right] - \frac{1}{\hat{\phi}} tr\left[B_{11}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix} B_{11}\widehat{Z}_{11}\right] = 0\\ \Leftrightarrow\ & n(\rho_1\rho_2 - c\gamma)\left((c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2\right)\\ & -\frac{1}{\hat{\phi}}(c-\rho_2^2)(\rho_1\rho_2 - c\gamma)z_{11} - \frac{1}{\hat{\phi}}(\rho_1\rho_2 - c\gamma)(c-\rho_1^2)z_{22}\\ & -\frac{1}{\hat{\phi}}(\gamma\rho_2 - \rho_1)(\gamma\rho_1 - \rho_2)z_{33} - \frac{1}{\hat{\phi}}\left[(c-\rho_2^2)(c-\rho_1^2) + (\rho_1\rho_2 - c\gamma)^2\right] z_{12}\\ & -\frac{1}{\hat{\phi}}\left[(c-\rho_2^2)(\gamma\rho_1 - \rho_2) + (\rho_1\rho_2 - c\gamma)(\gamma\rho_2 - \rho_1)\right] z_{13}\\ & -\frac{1}{\hat{\phi}}\left[(\rho_1\rho_2 - c\gamma)(\gamma\rho_1 - \rho_2) + (\gamma\rho_2 - \rho_1)(c-\rho_1^2)\right] z_{23} = 0\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell}{\partial \rho_1} &= \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1}\right] - \frac{1}{2\phi} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.5)\\
&\Leftrightarrow\ n\,tr\left[B_{11}\begin{pmatrix} 0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}\right] - \frac{1}{\widehat{\phi}} tr\left[B_{11}\begin{pmatrix} 0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}B_{11}\widehat{Z}_{11}\right] = 0\\
&\Leftrightarrow\ n(\gamma\rho_2 - \rho_1)\left((c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2\right)\\
&\quad -\frac{1}{\widehat{\phi}}(c-\rho_2^2)(\gamma\rho_2-\rho_1)z_{11} - \frac{1}{\widehat{\phi}}(\rho_1\rho_2 - c\gamma)(\gamma\rho_1-\rho_2)z_{22}\\
&\quad -\frac{1}{\widehat{\phi}}(\gamma\rho_2-\rho_1)(1-\gamma^2)z_{33}\\
&\quad -\frac{1}{\widehat{\phi}}\left[(c-\rho_2^2)(\gamma\rho_1-\rho_2) + (\rho_1\rho_2 - c\gamma)(\gamma\rho_2-\rho_1)\right]z_{12}\\
&\quad -\frac{1}{\widehat{\phi}}\left[(c-\rho_2^2)(1-\gamma^2) + (\gamma\rho_2-\rho_1)^2\right]z_{13}\\
&\quad -\frac{1}{\widehat{\phi}}\left[(\rho_1\rho_2 - c\gamma)(1-\gamma^2) + (\gamma\rho_2-\rho_1)(\gamma\rho_1-\rho_2)\right]z_{23} = 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell}{\partial \rho_2} &= \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2}\right] - \frac{1}{2\phi} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.6)\\
&\Leftrightarrow\ n\,tr\left[B_{11}\begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\right] - \frac{1}{\widehat{\phi}} tr\left[B_{11}\begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}B_{11}\widehat{Z}_{11}\right] = 0\\
&\Leftrightarrow\ n(\gamma\rho_1 - \rho_2)\left((c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2\right)\\
&\quad -\frac{1}{\widehat{\phi}}(\rho_1\rho_2 - c\gamma)(\gamma\rho_2-\rho_1)z_{11} - \frac{1}{\widehat{\phi}}(c-\rho_1^2)(\gamma\rho_1-\rho_2)z_{22}\\
&\quad -\frac{1}{\widehat{\phi}}(\gamma\rho_1-\rho_2)(1-\gamma^2)z_{33}\\
&\quad -\frac{1}{\widehat{\phi}}\left[(\rho_1\rho_2 - c\gamma)(\gamma\rho_1-\rho_2) + (\gamma\rho_2-\rho_1)(c-\rho_1^2)\right]z_{12}\\
&\quad -\frac{1}{\widehat{\phi}}\left[(\rho_1\rho_2 - c\gamma)(1-\gamma^2) + (\gamma\rho_2-\rho_1)(\gamma\rho_1-\rho_2\right]z_{13}\\
&\quad -\frac{1}{\widehat{\phi}}\left[(c-\rho_1^2)(1-\gamma^2) + (\gamma\rho_1-\rho_2)^2\right]z_{23} = 0
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &= \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right] - \frac{1}{2\phi} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha} R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.7) \\
\Leftrightarrow & \ \frac{n(t-3)}{(t-2)} tr\left[B_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] - ntr\left[B_{22}\right] \\
& -\frac{(t-3)}{\widehat{\phi}(t-2)} tr\left[B_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} B_{11}\widehat{Z}_{11}\right] + \frac{1}{(1-\alpha)^2} tr\left[Z_{22}\right] = 0 \\
\Leftrightarrow & \ \frac{n(t-3)(1-\gamma^2)}{(t-2)\left((c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right)} - \frac{n(t-3)}{1-\alpha} + \frac{1}{(1-\alpha)^2}\sum_{j=4}^{t} z_{jj}^2 \\
& -\frac{(t-3)\left((\gamma\rho_2-\rho_1)^2 z_{11} + (\gamma\rho_1-\rho_2)^2 z_{22}\right)}{\widehat{\phi}(t-2)\left((c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right)^2} \\
& -\frac{(t-3)\left((1-\gamma^2)^2 z_{33} + 2(\gamma\rho_2-\rho_1)(\gamma\rho_1-\rho_2)z_{12}\right)}{\widehat{\phi}(t-2)\left((c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right)^2} \\
& -\frac{2(t-3)\left((\gamma\rho_2-\rho_1)(1-\gamma^2)z_{13} + (\gamma\rho_1-\rho_2)(1-\gamma^2)z_{23}\right)}{\widehat{\phi}(t-2)\left((c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right)^2} = 0
\end{aligned}
$$

Here, note that $\widehat{\phi}$ is the MLE of ϕ. These four estimating equations ((4.3.4), (4.3.5), (4.3.6) and (4.3.7)) are used to find the MLE's for λ. Of course, these estimators are not in closed-form and are solved simultaneously using the Newton-Raphson method. So the Helmert transformation does not achieve the objective of obtaining closed-form solutions of the correlation parameters, though it does simplify the estimating equations considerably.

To find the MLE's we start with trial values of the correlation parameters (λ_0), and use them to obtain an initial estimate of β using (4.3.1). We then use this estimate to update the residuals ($\widehat{Z}_n$) and estimate ϕ using (4.3.2). Then $\widehat{\phi}$ and $\widehat{Z}_n$ are used to estimate the correlation parameters using (4.3.4), (4.3.5), (4.3.6) and (4.3.7). The estimates of the correlation parameters ($\widehat{\lambda}$), are then used to re-estimate β, and the process is continued until convergence. These estimates, then, are the MLE's of θ, specifically $\widehat{\theta}_\ell = (\widehat{\beta}_\ell, \widehat{\lambda}_\ell, \widehat{\phi}_\ell)'$.

IV.3.2 Method of Moments

For the Method of Moments (MOM) we find unbiased moment estimators for each of the correlation parameters. For γ, we get the following estimator

$$\widehat{\gamma}_m = \frac{2\sum_{i=1}^n \widehat{z}_{i1}\widehat{z}_{i2}}{\sum_{i=1}^n \sum_{j=1}^2 \widehat{z}_{ij}^2} \tag{4.3.8}$$

which is based on the estimating equation

$$\begin{aligned} \sum_{i=1}^n Z_i' A(\gamma) Z_i &= 0 \\ \Leftrightarrow \quad A(\gamma) &= \frac{\gamma}{2}\begin{pmatrix} I_2 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}. \end{aligned}$$

For ρ_1 we get the following estimator

$$\widehat{\rho}_{1,m} = \frac{2\sum_{i=1}^n \widehat{z}_{i1}\widehat{z}_{i3}}{\sum_{i=1}^n \sum_{j=1}^2 \widehat{z}_{ij}^2} \tag{4.3.9}$$

which is based on the estimating equation

$$\begin{aligned} \sum_{i=1}^n Z_i' A(\rho_1) Z_i &= 0 \\ \Leftrightarrow \quad A(\rho_1) &= \frac{\rho_1}{2}\begin{pmatrix} I_2 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}. \end{aligned}$$

For ρ_2 we get the following estimator

$$\widehat{\rho}_{2,m} = \frac{2\sum_{i=1}^n \widehat{z}_{i2}\widehat{z}_{i3}}{\sum_{i=1}^n \sum_{j=1}^2 \widehat{z}_{ij}^2} \tag{4.3.10}$$

which is based on the estimating equation

$$\begin{aligned} \sum_{i=1}^n Z_i' A(\rho_2) Z_i &= 0 \\ \Leftrightarrow \quad A(\rho_2) &= \frac{\rho_2}{2}\begin{pmatrix} I_2 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}. \end{aligned}$$

Finally, for α we get the following estimator

$$\widehat{\alpha}_m = \frac{1+(t-3)(t-2)}{(t-3)^2} - \frac{2(t-2)\sum_{i=1}^{n}\sum_{j=3}^{t}\widehat{z}_{ij}^2}{(t-3)^2\sum_{i=1}^{n}\sum_{j=1}^{2}\widehat{z}_{ij}^2} \qquad (4.3.11)$$

which is based on the estimating equation

$$\begin{aligned}\sum_{i=1}^{n} Z_i' A(\alpha) Z_i &= 0 \\ \Leftrightarrow \quad A(\alpha) &= \frac{\alpha(t-3)^2 - (1+(t-3)(t-2))}{2}\begin{pmatrix} I_2 & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} \\ &\quad +(t-2)\begin{pmatrix} \underline{0} & \underline{0} \\ \underline{0} & I_{t-3} \end{pmatrix}.\end{aligned}$$

To find the MOM estimators, we select initial values for λ (either, all zeros or sample statistics) and estimate β using (4.3.1). We then use $\widehat{\beta}$ to update the residuals ($\widehat{Z}_n$) and then estimate ϕ using (4.3.2). We then use $\widehat{Z}_n$ to estimate λ, which we in turn use to re-estimate β. We continue in this manner until convergence. Those estimators are then the MOM estimators, specifically $\widehat{\theta}_m = (\widehat{\beta}_m, \widehat{\lambda}_m, \widehat{\phi}_m)'$.

IV.3.3 Quasi-Least Squares

For the Quasi-Least Squares Method (QLS) we begin with the following quasi-log-likelihood function

$$\begin{aligned} S(\theta) &= \sum_{i=1}^{n}(Y_i - X_i\beta)' R^{-1}(\lambda)(Y_i - X_i\beta) \qquad (4.3.12) \\ &= tr\left[R^{-1}(\lambda) Z_n\right]. \end{aligned}$$

We can find estimators for β and λ by differentiating (4.3.12) with respect to each parameter, setting the resulting expression equal to zero and solving for that parameter.

Using (4.3.12), we obtain the following estimating equations for the correlation parameters.

$$\begin{aligned}\frac{\partial S(\theta)}{\partial \gamma} &= tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.13)\\ &= tr\left[B_{11}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}B_{11}Z_{11}\right]=0\\ &= a_\gamma\gamma^2+b_\gamma\gamma+c_\gamma=0,\end{aligned}$$

where

$$\begin{aligned}a_\gamma &= c^2z_{11}-c\rho_2z_{13}-c\rho_1z_{23}+\rho_1\rho_2z_{33}\\ b_\gamma &= c(\rho_2^2-c)z_{11}-2c\rho_1\rho_2z_{12}+2c\rho_1z_{13}\\ &\quad +c(\rho_1^2-c)z_{22}+2c\rho_2z_{23}-(\rho_1^2+\rho_2^2)z_{33}\\ c_\gamma &= \rho_1\rho_2(c-\rho_2)z_{11}+\rho_1\rho_2(c-\rho_1)z_{22}+\rho_1\rho_2z_{33}\\ &\quad +\left[2\rho_1^2\rho_2^2+c^2-c(\rho_1^2+\rho_2^2)\right]z_{12}+\left[\rho_2^3-\rho_2(c+\rho_1^2)\right]z_{13}+\left[\rho_1^3-\rho_1(c+\rho_2^2)\right]z_{23}.\end{aligned}$$

$$\begin{aligned}\frac{\partial S(\theta)}{\partial \rho_1} &= tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.14)\\ &= tr\left[B_{11}\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}B_{11}Z_{11}\right]=0\\ &= a_{\rho_1}\rho_1^2+b_{\rho_1}\rho_1+c_{\rho_1}=0\end{aligned}$$

where

$$\begin{aligned}a_{\rho_1} &= \gamma\rho_2z_{22}+z_{13}-\rho_2z_{12}-\gamma z_{23}\\ b_{\rho_1} &= -(\gamma^2+\rho_2^2)z_{22}-(1-\gamma^2)z_{33}-2\gamma\rho_2z_{13}+2c\gamma z_{12}+2\rho_2z_{23}\\ c_{\rho_1} &= \gamma\rho_2(c-\rho_2^2)z_{11}+c\gamma\rho_2z_{22}+\gamma\rho_2(1-\gamma^2)z_{33}\\ &\quad \left[\gamma^2\rho_2^2+(c-\rho_2^2)(1-\gamma^2)\right]z_{13}+\rho_2\left[\rho_2^2-c(1-\gamma^2)\right]z_{12}-\gamma\left[\rho_2^2+c(1-\gamma^2)\right]z_{23}.\end{aligned}$$

$$\begin{aligned}\frac{\partial S(\theta)}{\partial \rho_2} &= tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.15)\\ &= tr\left[B_{11}\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}B_{11}Z_{11}\right]=0\\ &= a_{\rho_2}\rho_2^2+b_{\rho_2}\rho_2+c_{\rho_2}=0\end{aligned}$$

where

$$\begin{aligned}
a_{\rho_2} &= \gamma\rho_1 z_{11} + (\rho_1^2 - c)z_{22} - \rho_1 z_{12} - \gamma z_{13} + z_{23} \\
b_{\rho_2} &= -(\rho_1^2 + c\gamma^2)z_{11} + (\rho_1^2 - c)z_{22} - (1-\gamma^2)z_{33} \\
&\quad 2\rho_1 z_{13} + 2c\gamma z_{12} - 2\gamma\rho_1 z_{23} \\
c_{\rho_2} &= c\gamma\rho_1 z_{11} + \gamma\rho_1(c - \rho_1^2)z_{22} + \gamma\rho_1(1-\gamma^2)z_{33} \\
&\quad \gamma\left[\rho_1^2 - c(1-\gamma^2)\right] z_{13} + \rho_1\left[\rho_1^2 - c(1-\gamma^2)\right] z_{12} + \left[\gamma^2\rho_1^2 + (c-\rho_1^2)(1-\gamma^2)\right] z_{23}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial S(\theta)}{\partial \alpha} &= tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}R^{-1}(\lambda)\widehat{Z}_n\right] = 0 \qquad (4.3.16)\\
&= tr\left[B_{11}\frac{\partial R_{11}}{\partial \alpha}B_{11}Z_{11}\right] + tr\left[B_{22}\frac{\partial R_{22}}{\partial \alpha}B_{22}Z_{22}\right] = 0 \\
&= a_\alpha\alpha^2 + b_\alpha\alpha + c_\alpha = 0
\end{aligned}$$

where

$$\begin{aligned}
a_\alpha &= f_1(\gamma,\rho_1,\rho_2,Z_n) - \frac{(t-3)}{(t-2)}(1-\gamma^2)^2\sum_{j=4}^{t} z_{jj} \\
b_\alpha &= -2f_1(\gamma,\rho_1,\rho_2,Z_n) - \frac{2(1-\gamma^2)^2}{(t-2)}\sum_{j=4}^{t} z_{jj} \\
&\quad +2(1-\gamma^2)\left[\rho_1(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2\right]\sum_{j=4}^{t} z_{jj} \\
c_\alpha &= f_1(\gamma,\rho_1,\rho_2,Z_n) - \frac{(1-\gamma^2)^2}{(t-2)(t-3)}\sum_{j=4}^{t} z_{jj} \\
&\quad +2\frac{(1-\gamma^2)}{(t-3)}\left[\rho_1^2(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2\right]\sum_{j=4}^{t} z_{jj} \\
f_1(\gamma,\rho_1,\rho_2,Z_n) &= (\gamma\rho_2 - \rho_1)^2 z_{11} + (\gamma\rho_1 - \rho_2)^2 z_{22} + (1-\gamma^2)^2 z_{33} \\
&\quad +2(\gamma\rho_1 - \rho_2)(\gamma\rho_2 - \rho_1)z_{12} + 2(\gamma\rho_2 - \rho_1)(1-\gamma^2)z_{13} \\
&\quad +2(\gamma\rho_1 - \rho_2)(1-\gamma^2)z_{23}.
\end{aligned}$$

Solving these four estimating equations iteratively gives $\widetilde{\lambda}$, the *Step 1* estimator of the correlation vector. Note that we must iterate between estimating $\widetilde{\beta}$ (with (4.3.1)) and $\widetilde{\lambda}$ until convergence to obtain the *Step 1* estimates of those parameters.

However, as we have seen in Chapters II and III, $\widetilde{\lambda}$ is a biased estimator of λ. This is shown by taking the expectation of each estimating equation listed in (4.3.14) through (4.3.17).

$$E\left[\frac{\partial S(\theta)}{\partial \gamma}\right] = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \gamma}E(Z_n)\right] \propto tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \gamma}R(\lambda)\right] \neq 0 \quad (4.3.17)$$

$$E\left[\frac{\partial S(\theta)}{\partial \rho_1}\right] = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1}E(Z_n)\right] \propto tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1}R(\lambda)\right] \neq 0 \quad (4.3.18)$$

$$E\left[\frac{\partial S(\theta)}{\partial \rho_2}\right] = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2}E(Z_n)\right] \propto tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2}R(\lambda)\right] \neq 0 \quad (4.3.19)$$

$$E\left[\frac{\partial S(\theta)}{\partial \alpha}\right] = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha}E(Z_n)\right] \propto tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha}R(\lambda)\right] \neq 0 \quad (4.3.20)$$

To find asymptotically unbiased estimators we make equations (4.3.17) through (4.3.20) our *Step 2* estimating equations by setting them equal to zero and solving for the respective correlation parameter. This gives us the following.

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \gamma}R(\lambda)\right] \propto tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] = 0 \quad (4.3.21)$$

$$\Leftrightarrow \quad \widehat{\gamma}_q = -\left[\frac{\widetilde{b}_{11}\widetilde{b}_{12} + \widetilde{b}_{12}\widetilde{b}_{22} + c\widetilde{b}_{13}\widetilde{b}_{23} + \rho_1(\widetilde{b}_{11}\widetilde{b}_{23} + \widetilde{b}_{12}\widetilde{b}_{13}) + \rho_2(\widetilde{b}_{12}\widetilde{b}_{23} + \widetilde{b}_{13}\widetilde{b}_{22})}{\left(\widetilde{b}_{11}\widetilde{b}_{22} + \widetilde{b}_{12}^2\right)}\right]$$

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1}R(\lambda)\right] \propto tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] = 0 \quad (4.3.22)$$

$$\Leftrightarrow \quad \widehat{\rho}_{1,q} = -\left[\frac{\widetilde{b}_{11}\widetilde{b}_{13} + \widetilde{b}_{12}\widetilde{b}_{23} + c\widetilde{b}_{13}\widetilde{b}_{33} + \gamma(\widetilde{b}_{11}\widetilde{b}_{23} + \widetilde{b}_{12}\widetilde{b}_{13}) + \rho_2(\widetilde{b}_{12}\widetilde{b}_{33} + \widetilde{b}_{13}\widetilde{b}_{23})}{\left(\widetilde{b}_{11}\widetilde{b}_{33} + \widetilde{b}_{13}^2\right)}\right]$$

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2}R(\lambda)\right] \propto tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] = 0 \quad (4.3.23)$$

$$\Leftrightarrow \quad \widehat{\rho}_{2,q} = -\left[\frac{\widetilde{b}_{12}\widetilde{b}_{13} + \widetilde{b}_{22}\widetilde{b}_{23} + c\widetilde{b}_{23}\widetilde{b}_{33} + \gamma(\widetilde{b}_{12}\widetilde{b}_{23} + \widetilde{b}_{13}\widetilde{b}_{22}) + \rho_1(\widetilde{b}_{12}\widetilde{b}_{33} + \widetilde{b}_{13}\widetilde{b}_{23})}{\left(\widetilde{b}_{22}\widetilde{b}_{33} + \widetilde{b}_{23}^2\right)}\right]$$

$$
\begin{aligned}
& tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\alpha}R(\lambda)\right] \\
& \propto tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] - \left(\frac{(1-\alpha)(t-2)}{(1-\widetilde{\alpha})^2}\right) = 0 \qquad (4.3.24) \\
\Leftrightarrow\quad & \widehat{\alpha}_q = \frac{(t-3) - (1-\widetilde{\alpha})^2\left[\widetilde{b}_{13}^2 + \widetilde{b}_{23}^2 + \frac{\widetilde{b}_{33}^2}{t-2} + 2\left(\gamma\widetilde{b}_{13}\widetilde{b}_{23} + \rho_1\widetilde{b}_{13}\widetilde{b}_{33} + \rho_2\widetilde{b}_{23}\widetilde{b}_{33}\right)\right]}{(t-3)\left[(1-\widetilde{\alpha})^2\widetilde{b}_{33}^2 + 1\right]}
\end{aligned}
$$

where $\widetilde{b}_{ij}$ is the ijth element of $\widetilde{B}_{11}$, $i, j = 1, 2, 3$. Note that here we have achieved the goal of closed-form estimators. The resulting estimators are then $\widehat{\lambda}_q$, the *Step 2* estimators of λ.

Once we have our *Step 2* estimates of the correlation parameters, we can substitute those values into (4.3.1) to obtain $\widehat{\beta}_q$, which we use to update the residual matrix $\widehat{Z}_n$. This, along with $\widehat{\lambda}$ can then be used to estimate the variance parameter using (4.3.2). Thus the QLS estimators are $\widehat{\theta}_q = (\widehat{\beta}_q, \widehat{\lambda}_q, \widehat{\phi}_q)'$.

IV.4 Asymptotic Variance and Performance

IV.4.1 Maximum Likelihood

For the maximum likelihood estimators of Section IV.3.1, we find the asymptotic variance by finding the inverse of Fisher's Information matrix, as we see in the following relation

$$\sqrt{n}(\widehat{\theta}_\ell - \theta) \sim AMVN\left(\underline{0}, I^{-1}(\theta)\right). \qquad (4.4.1)$$

To find this we take the negative expectation of the second derivative of the likelihood function with respect to θ. From here, it is straightforward to show that the information matrix $I_\ell(\theta)$ is of the following form.

$$I(\theta) = \begin{pmatrix} I(\beta) & 0 & 0 & 0 & 0 & 0 \\ 0 & I(\gamma) & I(\gamma,\rho_1) & I(\gamma,\rho_2) & I(\gamma,\alpha) & I(\gamma,\phi) \\ 0 & I(\gamma,\rho_1) & I(\rho_1) & I(\rho_1,\rho_2) & I(\rho_1,\alpha) & I(\rho_1,\phi) \\ 0 & I(\gamma,\rho_2) & I(\rho_1,\rho_2) & I(\rho_2) & I(\rho_2,\alpha) & I(\rho_2,\phi) \\ 0 & I(\gamma,\alpha) & I(\rho_1,\alpha) & I(\rho_2,\alpha) & I(\alpha) & I(\alpha,\phi) \\ 0 & I(\gamma,\phi) & I(\rho_1,\phi) & I(\rho_2,\phi) & I(\alpha,\phi) & I(\phi) \end{pmatrix} \tag{4.4.2}$$

where

$$\begin{aligned}
I(\beta) &= -E\left[\frac{\partial^2 \ell}{\partial \beta}\right] = \frac{1}{\phi}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
I(\gamma) &= -E\left[\frac{\partial^2 \ell}{\partial \gamma}\right] = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma}\right] \\
&= \frac{n\left[(c-\rho_2^2)(c-\rho_1^2) + (\rho_1\rho_2 - c\gamma)^2\right]}{\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\rho_1) &= -E\left[\frac{\partial^2 \ell}{\partial \rho_1}\right] = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1}\right] \\
&= \frac{n\left[(c-\rho_2^2)(1-\gamma^2) + (\gamma\rho_2 - \rho_1)^2\right]}{\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\rho_2) &= -E\left[\frac{\partial^2 \ell}{\partial \rho_2}\right] = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2}\right] \\
&= \frac{n\left[(c-\rho_1^2)(c-\rho_1^2) + (\gamma\rho_1 - \rho_2)^2\right]}{\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\alpha) &= -E\left[\frac{\partial^2 \ell}{\partial \alpha}\right] = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= \frac{n}{n}\left[\frac{(t-3)(1-\gamma^2)}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1-\rho_2)^2\right]}\right]^2 + \frac{n(t-3)}{2(1-\alpha)^2} \\
I(\phi) &= -E\left[\frac{\partial^2 \ell}{\partial}\right] = \frac{nt}{2\phi^2} \\
I(\gamma,\rho_1) &= -E\left[\frac{\partial^2 \ell}{\partial\gamma\partial\rho_1}\right] = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma}\right] \\
&= \frac{n\left[(c-\rho_2^2)(\gamma\rho_1-\rho_2) + (\rho_1\rho_2 - c\gamma)(\gamma\rho_2-\rho_1)\right]}{\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\gamma,\rho_2) &= -E\left[\frac{\partial^2 \ell}{\partial\gamma\partial\rho_2}\right] = \frac{n}{2} tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2} R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma}\right] \\
&= \frac{n\left[(\rho_1\rho_2 - c\gamma)(\gamma\rho_1-\rho_2) + (\gamma\rho_2-\rho_1)(c-\rho_1^2)\right]}{\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1-\rho_2)^2\right]^2}
\end{aligned}$$

$$
\begin{aligned}
I(\gamma,\alpha) &= -E\left[\frac{\partial^2\ell}{\partial\gamma\partial\alpha}\right] = \frac{n}{2}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\gamma}\right] \\
&= \frac{n(t-3)(\gamma\rho_2-\rho_1)(\gamma\rho_1-\rho_2)}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\rho_1,\rho_2) &= -E\left[\frac{\partial^2\ell}{\partial\rho_1\partial\rho_2}\right] = \frac{n}{2}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho_1}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho_2}\right] \\
&= \frac{n\left[(\rho_1\rho_2-c\gamma)(1-\gamma^2)+(\gamma\rho_2-\rho_1)(\gamma\rho_1-\rho_2)\right]}{\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\rho_1,\alpha) &= -E\left[\frac{\partial^2\ell}{\partial\rho_1\partial\alpha}\right] = \frac{n}{2}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho_1}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= \frac{n(t-3)(\gamma\rho_2-\rho_1)(1-\gamma^2)}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\rho_2,\alpha) &= -E\left[\frac{\partial^2\ell}{\partial\rho_2\partial\alpha}\right] = \frac{n}{2}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho_2}R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= \frac{n(t-3)(\gamma\rho_1-\rho_2)(1-\gamma^2)}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]^2} \\
I(\gamma,\phi) &= -E\left[\frac{\partial^2\ell}{\partial\gamma\partial\phi}\right] = \frac{n}{2\phi}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\gamma}\right] \\
&= \frac{n(\rho_1\rho_2-c\gamma)}{\phi\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I(\rho_1,\phi) &= -E\left[\frac{\partial^2\ell}{\partial\rho_1\partial\phi}\right] = \frac{n}{2\phi}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho_1}\right] \\
&= \frac{n(\gamma\rho_2-\rho_1)}{\phi\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I(\rho_2,\phi) &= -E\left[\frac{\partial^2\ell}{\partial\rho_2\partial\phi}\right] = \frac{n}{2\phi}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\rho_2}\right] \\
&= \frac{n(\gamma\rho_1-\rho_2)}{\phi\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I(\alpha,\phi) &= -E\left[\frac{\partial^2\ell}{\partial\alpha\partial\phi}\right] = \frac{n}{2\phi}tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial\alpha}\right] \\
&= \frac{n(t-3)}{2\phi}\left[\frac{1-\gamma^2}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]}-\frac{1}{1-\alpha}\right].
\end{aligned}
$$

IV.4.2 Method of Moments

For the MoM method we again make use of the theorem described in Chapter II.4. Under regularity conditions, we have

$$
\sqrt{n}(\widehat{\theta}_m-\theta) \sim AMVN\left(\underline{0}, I_m^{-1}(\theta)M_m(\theta)(I_m^{-1}(\theta))'\right) \tag{4.4.3}
$$

where $I_m(\theta) = -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{m,i}(\theta)}{\partial \theta'}\right]$, $M_m(\theta) = \frac{1}{n}\sum_{i=1}^{n} Cov(h_{m,i}(\theta))$ and the $h_{m,i}(\theta)$ are vectors of unbiased estimating equations. For any i, let $h_{m,i}(\theta)$ be defined as follows

$$\begin{aligned}
h_{m,i}(\theta) &= (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), h_{3i}(\theta), h_{4i}(\theta), g_i(\theta))' \qquad (4.4.4)\\
h_{0i}(\theta) &= X_i' R^{-1}(\lambda) Z_i\\
h_{1i}(\theta) &= Z_i' A(\gamma) Z_i = tr(A_i(\gamma) Z_i Z_i')\\
h_{2i}(\theta) &= Z_i' A(\rho_1) Z_i = tr(A_i(\rho_1) Z_i Z_i')\\
h_{3i}(\theta) &= Z_i' A(\rho_2) Z_i = tr(A_i(\rho_2) Z_i Z_i')\\
h_{4i}(\theta) &= Z_i' A(\alpha) Z_i = tr(A_i(\alpha) Z_i Z_i')\\
g_i(\theta) &= Z_i' R^{-1}(\lambda) Z_i - t\phi = tr(R^{-1}(\lambda) Z_i Z_i') - t\phi
\end{aligned}$$

where $A(\gamma)$, $A(\rho_1)$, $A(\rho_2)$ and $A(\alpha)$ are defined earlier. By taking the negative expectation of the partial derivatives of (4.4.4) with respect to θ and averaging over n we obtain $I_m(\theta)$, and by taking the covariance of (4.4.4) and averaging over n we obtain $M_m(\theta)$. From here it is easy to show that $I_m(\theta)$ has the following elements

$$I_m(\theta) = \begin{pmatrix} I_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{55} & 0 \\ 0 & I_{62} & I_{63} & I_{64} & I_{65} & I_{66} \end{pmatrix} \qquad (4.4.5)$$

where

$$\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{0i}(\theta)}{\partial \beta}\right] = \frac{1}{n}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i\\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{1i}(\theta)}{\partial \gamma}\right] = -\phi\\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{2i}(\theta)}{\partial \rho_1}\right] = -\phi\\
I_{44} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{3i}(\theta)}{\partial \rho_2}\right] = -\phi\\
I_{55} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{4i}(\theta)}{\partial \alpha}\right] = -\phi(t-3)^2
\end{aligned}$$

$$
\begin{aligned}
I_{62} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \gamma}\right] = \frac{2\phi(\rho_1\rho_2 - c\gamma)}{[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2]} \\
I_{63} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \rho_1}\right] = \frac{2\phi(\gamma\rho_2 - \rho_1)}{[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2]} \\
I_{64} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \rho_2}\right] = \frac{2\phi(\gamma\rho_1 - \rho_2)}{[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2]} \\
I_{65} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \alpha}\right] = \frac{\phi(t-3)(1-\gamma^2)}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2) - (\gamma\rho_1 - \rho_2)^2\right]} - \frac{\phi(t-3)}{(1-\alpha)} \\
I_{66} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \phi}\right] = t.
\end{aligned}
$$

We can also show that $M_m(\theta)$ has the following elements

$$
M_m(\theta) = \begin{pmatrix}
M_{11} & 0 & 0 & 0 & 0 & 0 \\
0 & M_{22} & M_{23} & M_{24} & M_{25} & 0 \\
0 & M_{23} & M_{33} & M_{34} & M_{35} & 0 \\
0 & M_{24} & M_{34} & M_{44} & M_{45} & 0 \\
0 & M_{25} & M_{35} & M_{45} & M_{55} & 0 \\
0 & 0 & 0 & 0 & 0 & M_{66}
\end{pmatrix} \tag{4.4.6}
$$

where

$$
\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{0i}(\theta)\right] = \frac{\phi}{n}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta)\right] = \phi^2(1+\gamma^2)^2 - 4\phi^2\gamma^2 \\
M_{23} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta), h_{2i}(\theta)\right] = \phi^2(\rho_2 - \gamma\rho_1)(1-\gamma^2) \\
M_{24} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta), h_{3i}(\theta)\right] = \phi^2(\rho_1 - \gamma\rho_2)(1-\gamma^2) \\
M_{25} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta), h_{4i}(\theta)\right] = \phi^2(t-2)\left[\gamma(\rho_1^2+\rho_2^2) - 2\rho_1\rho_2\right] \\
&\quad -\phi^2\gamma(1-\gamma^2)\left[\alpha(t-3)^2 - (1+(t-3)(t-2))\right] \\
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{2i}(\theta)\right] = \phi^2\left[\gamma\rho_1(\gamma\rho_1 - 2\rho_2) + \frac{1+(t-3)\alpha}{(t-2)}\right] \\
M_{34} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{2i}(\theta), h_{3i}(\theta)\right] = \phi^2\gamma\left[\gamma\rho_1\rho_2 - \rho_1^2 - \rho_2^2 + \frac{1+(t-3)\alpha}{(t-2)}\right]
\end{aligned}
$$

$$M_{35} = \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{2i}(\theta), h_{4i}(\theta)\right] = \phi^2\gamma(\gamma\rho_1 - \rho_2)\left[\alpha(t-3)^2 - (1+(t-3)(t-2))\right]$$
$$+\phi^2\rho_1(t-2)\left[\rho_1^2 + \rho_2^2 - 2\left(\frac{1+(t-3)\alpha}{(t-2)}\right)\right]$$
$$M_{44} = \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{3i}(\theta)\right] = \phi^2\left[\gamma\rho_2(\gamma\rho_2 - 2\rho_1) + \frac{1+(t-3)\alpha}{(t-2)}\right]$$
$$M_{45} = \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{3i}(\theta), h_{4i}(\theta)\right] = \phi^2\gamma(\gamma\rho_2 - \rho_1)\left[\alpha(t-3)^2 - (1+(t-3)(t-2))\right]$$
$$+\phi^2\rho_2(t-2)\left[\rho_1^2 + \rho_2^2 - 2\left(\frac{1+(t-3)\alpha}{(t-2)}\right)\right]$$
$$M_{55} = \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{4i}(\theta)\right] = \phi^2(1+\gamma^2)\left[\alpha(t-3)^2 - (1+(t-3)(t-2))\right]^2$$
$$+2\phi^2(t-2)\left[\alpha(t-3)^2 - (1+(t-3)(t-2))\right] + 2\phi^2(1+(t-3)\alpha)^2$$
$$+2\phi^2(t-2)^2(t-3)(1-\alpha)^2$$
$$M_{66} = \frac{1}{n}\sum_{i=1}^{n} Cov\left[g_i(\theta)\right] = 2\phi^2 t$$

IV.4.3 Quasi-Least Squares

For the QLS method we have

$$\sqrt{n}(\widehat{\theta}_q - \theta) \sim AMVN\left(\underline{0}, I_q^{-1}(\theta)M_q(\theta)(I_q^{-1}(\theta))'\right) \qquad (4.4.7)$$

where $I_q(\theta) = -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{q,i}(\theta)}{\partial\theta'}\right]$, $M_q(\theta) = \frac{1}{n}\sum_{i=1}^{n} Cov(h_{q,i}(\theta))$ and the $h_{q,i}(\theta)$ are vectors of unbiased estimating equations for the QLS method. For any i, let $h_{q,i}(\theta)$ be defined as follows.

$$h_{q,i}(\theta) = (h_{0i}(\theta), h_{1i}(\theta), h_{2i}(\theta), h_{3i}(\theta), h_{4i}(\theta), g_i(\theta))' \qquad (4.4.8)$$
$$h_{0i}(\theta) = X_i'(\beta)R^{-1}(\lambda)Z_i$$
$$h_{1i}(\theta) = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\gamma}}(Z_iZ_i' - \phi R(\lambda))\right]$$
$$h_{2i}(\theta) = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\rho_1}}(Z_iZ_i' - \phi R(\lambda))\right]$$
$$h_{3i}(\theta) = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\rho_2}}(Z_iZ_i' - \phi R(\lambda))\right]$$

$$h_{4i}(\theta) = tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\alpha}}(Z_iZ_i' - \phi R(\lambda))\right]$$
$$g_i(\theta) = tr\left[R^{-1}(\lambda)Z_iZ_i'\right] - t\phi$$

where $\widetilde{\lambda}$ is the solution to the following equations

$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\gamma}}R(\lambda)\right] = tr\left[\widetilde{B}_{11}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}\widetilde{B}_{11}R_{11}\right] = 0$$
$$\Leftrightarrow\quad \widetilde{b}_{12}(\widetilde{b}_{11}+\widetilde{b}_{22}) + \gamma\left(\widetilde{b}_{12}^2 + \widetilde{b}_{11}\widetilde{b}_{22}\right) + \rho_1\left(\widetilde{b}_{11}\widetilde{b}_{23} + \widetilde{b}_{12}\widetilde{b}_{13}\right)$$
$$+\rho_2\left(\widetilde{b}_{12}\widetilde{b}_{23} + \widetilde{b}_{13}\widetilde{b}_{22}\right) + c\widetilde{b}_{13}\widetilde{b}_{23} = 0$$
$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\rho}_1}R(\lambda)\right] = tr\left[\widetilde{B}_{11}\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}\widetilde{B}_{11}R_{11}\right] = 0$$
$$\Leftrightarrow\quad \widetilde{b}_{11}\widetilde{b}_{13} + \widetilde{b}_{12}\widetilde{b}_{23} + \gamma\left(\widetilde{b}_{12}\widetilde{b}_{13} + \widetilde{b}_{11}\widetilde{b}_{23}\right) + \rho_1\left(\widetilde{b}_{13}^2 + \widetilde{b}_{11}\widetilde{b}_{33}\right)$$
$$+\rho_2\left(\widetilde{b}_{12}\widetilde{b}_{33} + \widetilde{b}_{13}\widetilde{b}_{23}\right) + c\widetilde{b}_{13}\widetilde{b}_{33}$$
$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\rho}_2}R(\lambda)\right] = tr\left[\widetilde{B}_{11}\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}\widetilde{B}_{11}R_{11}\right] = 0$$
$$\Leftrightarrow\quad \widetilde{b}_{12}\widetilde{b}_{13} + \widetilde{b}_{22}\widetilde{b}_{23} + \gamma\left(\widetilde{b}_{12}\widetilde{b}_{23} + \widetilde{b}_{13}\widetilde{b}_{22}\right) + \rho_1\left(\widetilde{b}_{12}\widetilde{b}_{33} + \widetilde{b}_{13}\widetilde{b}_{23}\right)$$
$$+\rho_2\left(\widetilde{b}_{23}^2 + \widetilde{b}_{22}\widetilde{b}_{33}\right) + c\widetilde{b}_{23}\widetilde{b}_{33}$$
$$tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\widetilde{\alpha}}R(\lambda)\right] \propto tr\left[\widetilde{B}_{11}\begin{pmatrix}0&0&0\\0&0&0\\0&0&1\end{pmatrix}\widetilde{B}_{11}R_{11}\right] - \frac{(1-\alpha)(t-2)}{(1-\widetilde{\alpha})^2} = 0$$
$$\Leftrightarrow\quad \widetilde{b}_{13}^2 + \widetilde{b}_{23}^2 + c\widetilde{b}_{33}^2 + 2\left(\gamma\widetilde{b}_{13}\widetilde{b}_{23} + \rho_1\widetilde{b}_{13}\widetilde{b}_{33} + \rho_2\widetilde{b}_{23}\widetilde{b}_{33}\right) - \frac{(1-\alpha)(t-2)}{(1-\widetilde{\alpha})^2} = 0.$$

Note that $\lambda = (\gamma, \rho_1, \rho_2, \alpha)$ are the population values of the correlation parameters. By taking the expectation of the partial derivatives for (4.4.8) with respect to θ and averaging over n we obtain $I_q(\theta)$, and by taking the covariance of (4.4.8) and averaging over n we obtain $M_q(\theta)$. From here it is easy to show that $I_q(\theta)$ has the following elements

$$I_q(\theta) = \begin{pmatrix} I_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{22} & I_{23} & I_{24} & I_{25} & 0 \\ 0 & I_{23} & I_{33} & I_{34} & I_{35} & 0 \\ 0 & I_{24} & I_{34} & I_{44} & I_{45} & 0 \\ 0 & I_{25} & I_{35} & I_{45} & I_{55} & 0 \\ 0 & I_{62} & I_{63} & I_{64} & I_{65} & I_{66} \end{pmatrix} \tag{4.4.9}$$

where

$$\begin{aligned}
I_{11} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{0i}(\theta)}{\partial \beta}\right] = \frac{1}{n}\sum_{i=1}^{n} X_i' R^{-1} X_i \\
I_{22} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{1i}(\theta)}{\partial \gamma}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \gamma} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \gamma}\right] \\
&= -\frac{2\phi\left[(\widetilde{c}-\widetilde{\rho}_2^2)(\widetilde{c}-\widetilde{\rho}_1^2)+(\widetilde{\rho}_1\widetilde{\rho}_2-\widetilde{c}\widetilde{\gamma})^2\right]}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{23} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{1i}(\theta)}{\partial \rho_1}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \gamma} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho_1}\right] \\
&= -\frac{2\phi\left[(\widetilde{c}-\widetilde{\rho}_2^2)(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)+(\widetilde{\rho}_1\widetilde{\rho}_2-\widetilde{c}\widetilde{\gamma})(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)\right]}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{24} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{1i}(\theta)}{\partial \rho_2}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \gamma} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho_2}\right] \\
&= -\frac{2\phi\left[(\widetilde{\rho}_1\widetilde{\rho}_2-\widetilde{c}\widetilde{\gamma})(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)+(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)(\widetilde{c}-\widetilde{\rho}_1^2)\right]}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{25} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{1i}(\theta)}{\partial \alpha}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \gamma} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -\frac{2\phi\frac{(t-3)}{(t-2)}(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{33} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{2i}(\theta)}{\partial \rho_1}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho_1} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho_1}\right] \\
&= -\frac{2\phi\left[(\widetilde{c}-\widetilde{\rho}_2^2)(1-\widetilde{\gamma}^2)+(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)^2\right]}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{34} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{2i}(\theta)}{\partial \rho_2}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho_1} R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho_2}\right] \\
&= -\frac{2\phi\left[(\widetilde{\rho}_1\widetilde{\rho}_2-\widetilde{c}\widetilde{\gamma})(1-\widetilde{\gamma}^2)+(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)\right]}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2}
\end{aligned}$$

$$
\begin{aligned}
I_{35} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{2i}(\theta)}{\partial \alpha}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho_1}R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -\frac{2\phi\frac{(t-3)}{(t-2)}(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)(1-\widetilde{\gamma}^2)}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{44} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{3i}(\theta)}{\partial \rho_2}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho_2}R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \rho_2}\right] \\
&= -\frac{2\phi\left[(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)\left[(\widetilde{\gamma}\widetilde{\rho}_2-\widetilde{\rho}_1)+(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)\right]+(1-\widetilde{\gamma}^2)\left[(\widetilde{\rho}_1\widetilde{\rho}_2-\widetilde{c}\widetilde{\gamma})+(\widetilde{c}-\widetilde{\rho}_1^2)\right]\right]}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{45} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{3i}(\theta)}{\partial \alpha}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \rho_2}R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -\frac{2\phi\frac{(t-3)}{(t-2)}(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)(1-\widetilde{\gamma}^2)}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2} \\
I_{55} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial h_{4i}(\theta)}{\partial \alpha}\right] = -\phi\, tr\left[R^{-1}(\widetilde{\lambda})\frac{\partial R(\widetilde{\lambda})}{\partial \alpha}R^{-1}(\widetilde{\lambda})\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -\frac{2\phi\frac{(t-3)}{(t-2)}(1-\widetilde{\gamma}^2)^2}{\left[(\widetilde{c}-\widetilde{\rho}_1^2)(1-\widetilde{\gamma}^2)-(\widetilde{\gamma}\widetilde{\rho}_1-\widetilde{\rho}_2)^2\right]^2}-\frac{\phi(t-3)}{(1-\widetilde{\alpha}^2)^2} \\
I_{62} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \gamma}\right] = \phi\, tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \gamma}\right] \\
&= \frac{2\phi(\rho_1\rho_2-c\gamma)}{\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I_{63} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \rho_1}\right] = \phi\, tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_1}\right] \\
&= \frac{2\phi(\gamma\rho_2-\rho_1)}{\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I_{64} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \rho_2}\right] = \phi\, tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \rho_2}\right] \\
&= \frac{2\phi(\gamma\rho_1-\rho_2)}{\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I_{65} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \alpha}\right] = \phi\, tr\left[R^{-1}(\lambda)\frac{\partial R(\lambda)}{\partial \alpha}\right] \\
&= -\frac{(t-3)}{(1-\alpha)}+\frac{(t-3)(1-\gamma^2)}{(t-2)\left[(c-\rho_1^2)(1-\gamma^2)-(\gamma\rho_1-\rho_2)^2\right]} \\
I_{66} &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial g_i(\theta)}{\partial \phi}\right] = t.
\end{aligned}
$$

We can also show that $M_q(\theta)$ has the following elements

$$M_q(\theta) = \begin{pmatrix} M_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & M_{22} & M_{23} & M_{24} & M_{25} & 0 \\ 0 & M_{23} & M_{33} & M_{34} & M_{35} & 0 \\ 0 & M_{24} & M_{34} & M_{44} & M_{45} & 0 \\ 0 & M_{25} & M_{35} & M_{45} & M_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & M_{66} \end{pmatrix} \tag{4.4.10}$$

where

$$\begin{aligned}
M_{11} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{0i}(\theta)\right] = \frac{\phi}{n}\sum_{i=1}^{n} X_i' R^{-1}(\lambda) X_i \\
M_{22} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\gamma} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\gamma} R(\lambda)\right] \\
&= 2\phi^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{23} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta), h_{2i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\gamma} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\rho_1} R(\lambda)\right] \\
&= 2\phi^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{24} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta), h_{3i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\gamma} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\rho_2} R(\lambda)\right] \\
&= 2\phi^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{25} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{1i}(\theta), h_{4i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\gamma} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial\alpha} R(\lambda)\right] \\
&= \frac{2\phi^2(t-3)}{(t-2)} tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\widetilde{B}_{11}R_{11}\right]
\end{aligned}$$

$$
\begin{aligned}
M_{33} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{2i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1} R(\lambda)\right] \\
&= 2\phi^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{34} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{2i}(\theta), h_{3i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2} R(\lambda)\right] \\
&= 2\phi^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{35} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{2i}(\theta), h_{4i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_1} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha} R(\lambda)\right] \\
&= \frac{2\phi^2(t-3)}{(t-2)} tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{44} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{3i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2} R(\lambda)\right] \\
&= 2\phi^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{45} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{3i}(\theta), h_{4i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \rho_2} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha} R(\lambda)\right] \\
&= \frac{2\phi^2(t-3)}{(t-2)} tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
M_{55} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[h_{4i}(\theta)\right] = 2\phi^2 tr\left[\frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha} R(\lambda) \frac{\partial R^{-1}(\widetilde{\lambda})}{\partial \alpha} R(\lambda)\right] \\
&= 2\left(\frac{\phi(t-3)}{(t-2)}\right)^2 tr\left[\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\widetilde{B}_{11}R_{11}\widetilde{B}_{11}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\widetilde{B}_{11}R_{11}\right] \\
&\quad + \frac{2\phi^2(1-\alpha)^2(t-3)}{(1-\widetilde{\alpha})^4} \\
M_{66} &= \frac{1}{n}\sum_{i=1}^{n} Cov\left[g_i(\theta)\right] = 2\phi^2 t.
\end{aligned}
$$

Table 4.1: γ ARE of MLE vs. MoM (MLE vs. QLS)

	γ	0.0			0.2			0.4		
ρ_1	ρ_2 / α	0.0	0.2	0.4	0.0	0.2	0.4	0.0	0.2	0.4
0.0	0.0	1.000	1.000	1.000	0.988	0.989	0.990	0.955	0.957	0.960
		(1.000)	(1.000)	(1.000)	(0.988)	(0.989)	(0.990)	(0.955)	(0.957)	(0.960)
	0.2	1.000	1.000	1.000	0.987	0.987	0.988	0.948	0.950	0.954
		(1.000)	(1.000)	(1.000)	(0.988)	(0.988)	(0.989)	(0.953)	(0.954)	(0.957)
	0.4	1.000	1.000	1.000	0.980	0.981	0.983	0.920	0.923	0.930
		(1.000)	(1.000)	(1.000)	(0.986)	(0.985)	(0.986)	(0.948)	(0.943)	(0.947)
0.2	0.0	1.000	1.000	1.000	0.987	0.987	0.988	0.948	0.950	0.954
		(1.000)	(1.000)	(1.000)	(0.988)	(0.988)	(0.989)	(0.953)	(0.954)	(0.957)
	0.2	0.998	0.998	0.998	0.994	0.994	0.995	0.964	0.965	0.968
		(1.000)	(1.000)	(1.000)	(0.990)	(0.991)	(0.992)	(0.957)	(0.960)	(0.963)
	0.4	0.990	0.990	0.991	0.998	0.998	0.998	0.966	0.968	0.971
		(0.999)	(0.998)	(0.998)	(0.992)	(0.993)	(0.995)	(0.956)	(0.960)	(0.964)
0.4	0.0	1.000	1.000	1.000	0.980	0.981	0.983	0.920	0.923	0.930
		(1.000)	(1.000)	(1.000)	(0.986)	(0.985)	(0.986)	(0.948)	(0.943)	(0.947)
	0.2	0.990	0.990	0.991	0.998	0.998	0.998	0.966	0.968	0.971
		(0.999)	(0.998)	(0.998)	(0.992)	(0.993)	(0.995)	(0.956)	(0.960)	(0.964)
	0.4	0.949	0.951	0.958	0.998	0.998	0.998	0.988	0.988	0.989
		(1.000)	(0.990)	(0.988)	(0.993)	(0.998)	(0.999)	(0.959)	(0.969)	(0.975)

IV.4.4 Comparison of Asymptotic Performance

We now compare the asymptotic performance of each estimating procedure discussed in Section IV.3 by computing asymptotic relative efficiencies (ARE). This is done by calculating the asymptotic variances derived in Section IV.4 for particular values of the correlation parameters. For our purposes, we assume that $t = 5$ (i.e. a family consists of two parents and three children), $n = 5,000$ and $\phi = 3$. As there are four correlation parameters, it is impractical for us to display ARE as done in Chapters II and III. Thus we make use of tables, choosing values of 0.0, 0.2 and 0.4 for each correlation parameter.

We begin with estimators of γ, the ARE's for which are found in Table 4.1. Here we see that the ARE is close to one for both the MLE v. MoM and MLE v. QLS comparisons, implying that the asymptotic variances for estimators of γ for all three estimating procedures are very similar. Specifically, note that the ARE is exactly one or extremely close when $\gamma = 0.0$. Only for $\gamma = 0.4$ do any of the ARE's drop below 0.95, and nowhere are they below 0.9. Thus, for estimators of γ, we see that

Table 4.2: ρ_1 ARE of MLE vs. MoM (MLE vs. QLS)

	γ	0.0			0.2			0.4		
ρ_1	ρ_2 / α	0.0	0.2	0.4	0.0	0.2	0.4	0.0	0.2	0.4
0.0	0.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	0.2	1.000	1.000	1.000	0.996	0.997	0.998	0.982	0.988	0.991
		(1.000)	(1.000)	(1.000)	(0.998)	(0.999)	(0.999)	(0.992)	(0.994)	(0.995)
	0.4	1.000	1.000	1.000	0.980	0.986	0.990	0.914	0.940	0.956
		(1.000)	(1.000)	(1.000)	(0.991)	(0.993)	(0.995)	(0.961)	(0.969)	(0.976)
0.2	0.0	0.924	0.936	0.942	0.920	0.933	0.939	0.907	0.923	0.931
		(0.983)	(0.987)	(0.989)	(0.982)	(0.986)	(0.988)	(0.978)	(0.983)	(0.985)
	0.2	0.923	0.935	0.941	0.950	0.956	0.958	0.967	0.970	0.970
		(0.984)	(0.987)	(0.988)	(0.987)	(0.992)	(0.993)	(0.987)	(0.993)	(0.996)
	0.4	0.920	0.931	0.936	0.973	0.976	0.976	0.995	0.997	0.995
		(0.987)	(0.986)	(0.986)	(0.988)	(0.994)	(0.996)	(0.976)	(0.991)	(0.997)
0.4	0.0	0.690	0.751	0.774	0.678	0.740	0.765	0.633	0.704	0.736
		(0.914)	(0.945)	(0.955)	(0.911)	(0.941)	(0.951)	(0.895)	(0.925)	(0.938)
	0.2	0.668	0.747	0.770	0.726	0.780	0.799	0.741	0.794	0.811
		(0.922)	(0.946)	(0.953)	(0.921)	(0.952)	(0.962)	(0.918)	(0.953)	(0.964)
	0.4	0.681	0.733	0.753	0.771	0.817	0.827	0.835	0.873	0.876
		(0.976)	(0.951)	(0.947)	(0.942)	(0.964)	(0.969)	(0.924)	(0.967)	(0.979)

all three procedures perform similarly.

The ARE for estimators of ρ_1 are found in Table 4.2. Here we see that the ARE for both comparisons are high for values of ρ_1 (0.0 and 0.2), as in this region most ARE values are close to 1.0 and none are less than 0.9. However, for $\rho_1 = 0.4$ we see that the ARE for the MLE and MoM procedures is everywhere below 0.9 and in some cases below 0.65, implying that the MLE has smaller variance than the MoM estimator. The ARE for the MLE and QLS procedures are still high in this region and nowhere lower than 0.895. Though the asymptotic variances for the MoM and QLS ρ_1 estimators are comparable to that of the MLE for small to moderate values of ρ_1, only the QLS estimator has comparable asymptotic variance with the MLE for large values of ρ_1.

The ARE for estimators of ρ_2 are found in Table 4.3. Here we see that for ρ_2 equal to 0.0 and 0.2, the ARE values for both comparisons are everywhere greater than 0.9, and for small γ and ρ_1 we see that the ARE is close to one. However, for $\rho_2 = 0.4$, we see the the ARE for the MLE and MoM procedures is everywhere less than 0.9 and in some cases lower than 0.65. The ARE for the MLE and QLS procedures, however, is

Table 4.3: ρ_2 ARE of MLE vs. MoM (MLE vs. QLS)

	γ	0.0			0.2			0.4		
ρ_1	ρ_2 / α	0.0	0.2	0.4	0.0	0.2	0.4	0.0	0.2	0.4
0.0	0.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	0.2	0.924	0.936	0.942	0.920	0.933	0.939	0.907	0.923	0.931
		(0.983)	(0.987)	(0.989)	(0.982)	(0.986)	(0.988)	(0.978)	(0.983)	(0.985)
	0.4	0.690	0.751	0.774	0.678	0.740	0.765	0.633	0.704	0.736
		(0.914)	(0.945)	(0.955)	(0.911)	(0.941)	(0.951)	(0.895)	(0.925)	(0.938)
0.2	0.0	1.000	1.000	1.000	0.996	0.997	0.998	0.982	0.988	0.991
		(1.000)	(1.000)	(1.000)	(0.998)	(0.999)	(0.999)	(0.992)	(0.994)	(0.995)
	0.2	0.923	0.935	0.941	0.950	0.956	0.958	0.967	0.970	0.970
		(0.984)	(0.987)	(0.988)	(0.987)	(0.992)	(0.993)	(0.987)	(0.993)	(0.996)
	0.4	0.688	0.747	0.770	0.726	0.780	0.799	0.741	0.794	0.811
		(0.922)	(0.946)	(0.953)	(0.921)	(0.952)	(0.962)	(0.918)	(0.953)	(0.964)
0.4	0.0	1.000	1.000	1.000	0.980	0.986	0.990	0.914	0.940	0.956
		(1.000)	(1.000)	(1.000)	(0.991)	(0.993)	(0.995)	(0.961)	(0.969)	(0.976)
	0.2	0.920	0.931	0.936	0.973	0.976	0.975	0.995	0.997	0.995
		(0.987)	(0.986)	(0.986)	(0.988)	(0.994)	(0.996)	(0.976)	(0.991)	(0.997)
	0.4	0.681	0.733	0.753	0.771	0.817	0.827	0.835	0.873	0.876
		(0.976)	(0.951)	(0.947)	(0.942)	(0.964)	(0.969)	(0.924)	(0.967)	(0.979)

nowhere less than 0.9. We also see that, based on the comparisons between the MLE and MoM procedures, the asymptotic variance of the QLS estimator is everywhere at least as small as the MoM estimator. For estimators of ρ_2, then, we see that the QLS estimator is a good competitor with the MLE, while for large values of ρ_2, both the MLE and QLS estimators have smaller asymptotic variances than the MoM estimator.

Lastly, the ARE for estimators of α are found Table 4.4. Here we see that for all values of the correlation parameters the ARE is less than or equal to 0.4 for MLE and MoM comparison, implying that the variance of the MoM estimator is much larger than that of the MLE. Alternatively, we see that the efficiencies for the MLE and QLS procedures are high for all values of the correlation parameters, with no value less than 0.94, and many close to 1.0. For estimating α, then, we see that both the MLE and QLS procedures are far superior to MoM, and QLS is highly comparable to MLE.

Table 4.4: α ARE of MLE vs. MoM (MLE vs. QLS)

	γ	0.0			0.2			0.4		
ρ_1	ρ_2 / α	0.0	0.2	0.4	0.0	0.2	0.4	0.0	0.2	0.4
0.0	0.0	0.248	0.315	0.337	0.245	0.311	0.333	0.237	0.301	0.323
		(1.000)	(0.984)	(0.942)	(1.000)	(0.985)	(0.943)	(1.000)	(0.986)	(0.946)
	0.2	0.249	0.320	0.344	0.246	0.316	0.340	0.238	0.306	0.330
		(0.998)	(0.987)	(0.946)	(0.998)	(0.987)	(0.947)	(0.998)	(0.988)	(0.950)
	0.4	0.237	0.329	0.364	0.234	0.324	0.359	0.225	0.312	0.346
		(0.962)	(0.986)	(0.955)	(0.962)	(0.985)	(0.955)	(0.961)	(0.983)	(0.956)
0.2	0.0	0.249	0.320	0.344	0.246	0.316	0.340	0.238	0.306	0.330
		(0.998)	(0.987)	(0.946)	(0.998)	(0.987)	(0.977)	(0.998)	(0.988)	(0..950)
	0.2	0.251	0.326	0.352	0.248	0.322	0.348	0.240	0.311	0.336
		(0.998)	(0.991)	(0.950)	(0.998)	(0.992)	(0.952)	(0.998)	(0.992)	(0.955)
	0.4	0.240	0.336	0.373	0.237	0.331	0.368	0.229	0.319	0.354
		(0.968)	(0.992)	(0.961)	(0.965)	(0.992)	(0.962)	(0.963)	(0.991)	(0.964)
0.4	0.0	0.237	0.329	0.364	0.234	0.324	0.359	0.225	0.312	0.346
		(0.962)	(0.986)	(0.955)	(0.962)	(0.985)	(0.955)	(0.961)	(0.983)	(0.956)
	0.2	0.240	0.336	0.373	0.237	0.331	0.368	0.229	0.319	0.354
		(0.968)	(0.992)	(0.961)	(0.965)	(0.992)	(0.962)	(0.963)	(0.991)	(0.964)
	0.4	0.225	0.349	0.400	0.224	0.344	0.394	0.220	0.331	0.378
		(0.969)	(1.000)	(0.974)	(0.941)	(0.999)	(0.977)	(0.941)	(0.999)	(0.979)

CHAPTER V

CONCLUSION

V.1 Summary

In Chapter II we analyzed the Autoregressive Familial correlation structure with regards to the maximum likelihood, method of moment and quasi-least squares procedures, finding unbiased estimators and their asymptotic variances. Asymptotically, we found that quasi-least squares correlation estimators are good competitors with the maximum likelihood estimators, and both are superior to the moment estimators. In the small sample case, we estimated small-sample efficiencies and found that the quasi-least squares estimators are much more competitive against the maximum likelihood estimators, especially in the presence of non-normally distributed data. We also proposed a likelihood ratio test for the maximum likelihood estimators and Wald's Tests for the moment and quasi-least squares estimators.

In Chapter III we analyzed the Autoregressive Familial correlation structure in the case of heterogeneous intra-class variances. The main procedural difference between the estimation methods in this Chapter and those in Chapter II is that here we used moment estimators for the variance parameters in each procedure. Estimation of the correlation parameters, however, was similar. Asymptotically, we again saw that the QLS estimator has comparably small variance with the MLE, and both the MLE and QLS correlation parameter estimators are more efficient than the MoM. In the small-sample case, we simulated data from a normal distribution, and found that for estimating ρ the QLS procedure is comparable with the MLE procedure with regards to estimated small-sample variance, and both the QLS and MoM procedures outperform the MLE procedure for estimators of α. We saw similar results in the small-sample case with data simulated from a non-normal distribution.

Finally, in Chapter IV we analyzed the Equicorrelated Nuclear Familial structure. Making use of a canonical transformation we simplified the correlation structure into a more manageable form, which simplified the process of finding estimators and asymptotic variances. Asymptotically, we found that the QLS estimators for each correlation parameter has comparably small variance with the MLE for all values of the correlation parameters, while both the MLE and QLS estimators have much smaller asymptotic variances than the MoM estimator for large correlation values.

V.2 Future Research

The first extension of the work provided in this thesis would be to analyze the unbalanced case, or to account for data sets with families of various sizes. Allowing $t_i, i = 1, \ldots, n$ to vary between families, provided the family dependence structures are the same, should not be too arduous.

One natural progression from the autoregessive familial correlation structure is instead to incorporate age differences into the modeling. This correlation structure for a family of size j could look as follows.

$$\begin{pmatrix} 1 & \rho^{|a_1-a_2|} & \rho^{|a_1-a_3|} & \rho^{|a_1-a_4|} & & \rho^{|a_1-a_j|} \\ \rho^{|a_2-a_1|} & 1 & \alpha^{|a_2-a_3|} & \alpha^{|a_2-a_4|} & \cdots & \alpha^{|a_2-a_j|} \\ \rho^{|a_3-a_1|} & \alpha^{|a_3-a_2|} & 1 & \alpha^{|a_3-a_4|} & & \alpha^{|a_3-a_j|} \\ & \vdots & & & \ddots & \vdots \\ \rho^{|a_j-a_1|} & \alpha^{|a_j-a_2|} & \alpha^{|a_j-a_3|} & \alpha^{|a_j-a_4|} & \cdots & 1 \end{pmatrix}$$

where a_i is the age of the ith family member. Accounting for actual differences between family members in this manner would be more accurate than simply reducing correlation by a power. However, using age differences to reduce correlation could also dilute existing dependencies too much. Another, yet more complicated alternative is the generalized Markov model, which is given by

$$\begin{pmatrix} 1 & \eta^{e_2} & \eta^{e_2+e_3} & \eta^{e_2+e_3+e_4} & & \eta^{e_2+e_3+\cdots+e_j} \\ \eta^{e_2} & 1 & \xi^{o_3} & \xi^{o_3+o_4} & \cdots & \xi^{o_3+o_4+\cdots+o_j} \\ \eta^{e_2+e_3} & \xi^{o_3} & 1 & \xi^{o_3} & & \xi^{o_4+o_5+\cdots+o_j} \\ & \vdots & & & \ddots & \vdots \\ \eta^{e_2+e_3+\cdots+e_j} & \xi^{o_3+o_4+\cdots+o_j} & \xi^{o_5+o_6+\cdots+o_j} & \xi^{o_6+o_7+\cdots+o_j} & \cdots & 1 \end{pmatrix}$$

where η is a par-sib correlation parameter, ξ is a sib-sib correlation parameter, the e_i's are functions of the parameter λ and the o_i's are functions of the parameter γ

defined as

$$e_i(\lambda) = \begin{cases} \frac{[a_i^\lambda - a_{i-1}^\lambda]}{\lambda} & if \lambda \neq 0, \\ \log(a_i) - \log(a_{i-1}) & if \lambda = 0, \end{cases}$$

$$o_i(\gamma) = \begin{cases} \frac{[a_i^\gamma - a_{i-1}^\gamma]}{\gamma} & if \gamma \neq 0, \\ \log(a_i) - \log(a_{i-1}) & if \gamma = 0, \end{cases}$$

where $2 \leq i \leq j$. This structure allows us to adjust the dampening parameter (via λ or γ) to more accurately model the existing correlation. Naturally, with the increase in parameters this structure will also be increasingly intractable algebraically.

With regards to the nuclear family model discussed in Chapter IV, we would first like to analyze the small-sample case, as was done in Chapters II and III. This would give a much better picture of the performance of the three estimating procedures. Another natural extension for the nuclear model is to add further family members (grandparents, step-parents, adopted children, etc.). Modeling the dependence for this family might best be served with an unstructured model given by

$$\begin{pmatrix} 1 & \rho_{1,2} & \rho_{1,3} & \rho_{1,4} & \cdots \\ \rho_{1,2} & 1 & \alpha_{2,3} & \alpha_{2,4} & \cdots \\ \rho_{1,3} & \alpha_{2,3} & 1 & \gamma_{3,4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where each parameter corresponds to a specific family member and the subscripts correspond to which two members the parameter applies. Note however, that we cannot apply the same canonical reduction that was applied in Chapter IV. More generally, however, we could model k arbitrary classes of family members (as in Elston (1975)) with

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1k} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{k1} & \Sigma_{k2} & \cdots & \Sigma_{kk} \end{pmatrix}$$

where Σ_{jj} is the $(t_j \times t_j)$ intra-class variance-covariance structure for the jth class, and Σ_{ij} is the $(t_i \times t_j)$ inter-class variance-covariance structure between the ith and jth classes. If we assume that the parameters within each class follow equicorrelated structures, then we could theoretically find a canonical reduction to simplify the structure into a more manageable form.

REFERENCES

Chaganty, N. R. (1997). An alternative approach to the analysis of longitudinal data via generalized estimating equations. *Journal of Statistical Planning and Inference* 63, 39-54.

Chaganty, N. R. and Shi, G. (2004). A Note on the Estimation of Autocorrelation in Repeated Measurements. *Communications in Statistics: Theory and Methods.* 33 (5), 1157-1170.

Chaganty, N. R. and Shults, J. (1999). On eliminating the asymptotic bias in the quasi-least squares estimate of the correlation parameter. *Journal of Statistical Planning and Inference* 76, 145-161.

Donner, A. and Koval, J. J. (1980). The Estimation of Intraclass Correlation in the Analysis of Family Data. *Biometrics* 36, 19-25.

Elston, R. C. (1975) On the Correlation Between Correlations. *Biometrika* 62 (1), 133-140.

Fieller, E. C. and Smith, C. A. B. (1951) Note on the Analysis of Variance and Intraclass Correlation. *Annals of Eugenics* 16, 97-104.

Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* 52, 399-433.

Fisher, R. A. (1925). *Statistical Methods for Research Workers*, Oliver & Boyd, London.

Guo, G. and Wang, J. (2002) The mixed or multilivel model for behavior genetics *Behavior Genetics* 32, 37-49.

Hand, D. and Crowder, M. (1996) *Practical Longitudinal Data Analysis*, Chapman & Hall, London.

Hardin, J. W. and Hilbe, J. M. (2003). *Generalized Estimating Equations*, New York: Chapman & Hall.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*, New York: Chapman & Hall/CRC.

Khattree, R. and Naik, D. (1994). Estimation of interclass correlation under circular covariance *Biometrika* 81, 612-617.

Magnus, P., Gjessing, H. K., Skrondal, A. and Skjaerven, R. (2001). Paternal contribution to birthweight *Journal of Epidemiology and Community Health* 55, 873-877.

Mak, T. K. and Ng, K. W. (1981). Analysis of Familial Data: Linear-Model Approach. *Biometrika* 68 (2), 457-461.

McArdle, J.J. and Prescott, C.A. (2005). Mixed-effects variance components models for biometric family analyses *Behavior Genetics* 35, 631-652.

Munoz, A., Rosner, B. and Carey, V. (1986). Regression Analysis in the Presence of Heterogeneous Intraclass Correlations *Biometrics* 42, 653-658.

Paul, S. R. (1990). Maximum Likelihood Estimation of Intraclass Correlation in the Analysis of Familial Data: Estimating Equation Approach. *Biometrika* 77 (3), 549-555.

Pawitan, Y., Reilly, M., Nilsson, E., Cnattingius, S. and lichtenstein, P. (2004). Estimation of genetic and enviornmental factors for binary traits using family data *Statistics in Medicine* 23, 449-465.

Rabe-Hesketh, S., Skrondal, A. and Gjessing, H.K. (2005). Biometrical modeling of twin and fmaily data using standard mixed model software *Biometrics* in press.

Rosner, B, Donner, A and Hennekens, C. H. (1977). Estimation of Interclass Correlation from Familial Data *Applied Statistics* 26 (2), 179-187.

Rosner, B. (1979). Maximum Likelihood Estimation of Interclass Correlations. *Biometrika* 66 (3), 533-538.

Shoukri, M.M. and Ward, R. H. (1989). Use of linear models to estimate genetic parameters and measures of familial resemblance in man. *Applied Statistics* 38, 467-479.

Shults, J. and Chaganty, N. R. (1998). Analysis of Serially Correlated Data Using Quasi-Least Squares. *Biometrics* 54, 1622-1630.

Smith, C. A. B. (1980). Estimating Genetic Correlations. *Annals of Human Genetics* 43, 265-284.

Srivastava, M. S. (1984). Estimation of Interclass Correlations in Familial Data. *Biometrika* 71 (1), 177-185.

Stanish, W. M. and Taylor, N. (1983). Estimation of the Intraclass Correlation Coefficient for Analysis of Covariance Model. *The American Statistician* 37 (3), 221-224.

Wedderburn,R. W. M. (1974). Quasi-likelihood functions, generalised linear models, and the Gauss-Newton method. *Biometrika* 61, 439-447.

APPENDIX

LISTS OF PARTIAL DERIVATIVES

A.1 List of partial derivatives from Chapter 2

$$\frac{\partial R(\lambda)}{\partial \rho} = \begin{pmatrix} 0 & 1 & 2\rho & 3\rho^2 & & (t-1)\rho^{t-2} \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 2\rho & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ (t-1)\rho^{t-2} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial R(\lambda)}{\partial \alpha} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 2\alpha & \cdots & (t-2)\alpha^{t-3} \\ 0 & 1 & 0 & 1 & \cdots & (t-3)\alpha^{t-4} \\ 0 & 2\alpha & 1 & 0 & \cdots & (t-4)\alpha^{t-5} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & (t-2)\alpha^{t-3} & (t-3)\alpha^{t-4} & (t-4)\alpha^{t-5} & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial^2 R(\lambda)}{\partial \rho^2} = \begin{pmatrix} 0 & 0 & 2 & 6\rho & & (t-1)(t-2)\rho^{t-3} \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 2 & 0 & 0 & 0 & & 0 \\ 6\rho & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ (t-1)(t-2)\rho^{t-3} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial^2 R(\lambda)}{\partial \alpha^2} = \begin{pmatrix} 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & \cdots & (t-2)(t-3)\alpha^{t-4} \\ 0 & 0 & 0 & \cdots & (t-3)(t-4)\alpha^{t-5} \\ 0 & 2 & 0 & & (t-4)(t-5)\alpha^{t-6} \\ & \vdots & & \ddots & \vdots \\ 0 & (t-2)(t-3)\alpha^{t-4} & (t-3)(t-4)\alpha^{t-5} & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial R(\lambda)}{\partial\rho\partial\alpha} = \underline{0}$$

A.2 List of partial derivatives from Chapter 3

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_p} = \begin{pmatrix} 1 & \frac{\phi_s^{\frac{1}{2}}\rho}{2\phi_p^{\frac{1}{2}}} & \frac{\phi_s^{\frac{1}{2}}\rho^2}{2\phi_p^{\frac{1}{2}}} & & \frac{\phi_s^{\frac{1}{2}}\rho^{t-1}}{2\phi_p^{\frac{1}{2}}} \\ \frac{\phi_s^{\frac{1}{2}}\rho}{2\phi_p^{\frac{1}{2}}} & 0 & 0 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ \frac{\phi_s^{\frac{1}{2}}\rho^{t-1}}{2\phi_p^{\frac{1}{2}}} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial^2\Sigma(\lambda,\Phi)}{\partial\phi_p^2} = -\frac{\phi_s^{\frac{1}{2}}}{4\phi_p^{\frac{3}{2}}} \begin{pmatrix} 0 & \rho & \rho^2 & & \rho^3 \\ \rho & 0 & 0 & \cdots & 0 \\ \rho^2 & 0 & 0 & & 0 \\ & \vdots & & \ddots & \vdots \\ \rho^3 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_s} = \begin{pmatrix} 0 & \frac{\phi_p^{\frac{1}{2}}\rho}{2\phi_s^{\frac{1}{2}}} & \frac{\phi_p^{\frac{1}{2}}\rho^2}{2\phi_s^{\frac{1}{2}}} & \frac{\phi_p^{\frac{1}{2}}\rho^3}{2\phi_s^{\frac{1}{2}}} & & \frac{\phi_p^{\frac{1}{2}}\rho^{t-1}}{2\phi_s^{\frac{1}{2}}} \\ \frac{\phi_p^{\frac{1}{2}}\rho}{2\phi_s^{\frac{1}{2}}} & 1 & \alpha & \alpha^2 & \cdots & \alpha^{t-2} \\ \frac{\phi_p^{\frac{1}{2}}\rho^2}{2\phi_s^{\frac{1}{2}}} & \alpha & 1 & \alpha & & \alpha^{t-3} \\ & \vdots & & & \ddots & \vdots \\ \frac{\phi_p^{\frac{1}{2}}\rho^{t-1}}{2\phi_s^{\frac{1}{2}}} & \alpha^{t-2} & \alpha^{t-3} & \alpha^{t-4} & \cdots & 1 \end{pmatrix}$$

$$\frac{\partial^2\Sigma(\lambda,\Phi)}{\partial\phi_s^2} = -\frac{\phi_p^{\frac{1}{2}}}{4\phi_s^{\frac{3}{2}}} \begin{pmatrix} 0 & \rho & \rho^2 & & \rho^{t-1} \\ \rho & 0 & 0 & \cdots & 0 \\ \rho^2 & 0 & 0 & & 0 \\ & \vdots & & \ddots & \vdots \\ \rho^{t-1} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\rho} = \phi_p^{\frac{1}{2}}\phi_s^{\frac{1}{2}}\begin{pmatrix} 0 & 1 & 2\rho & 3\rho^2 & & (t-1)\rho^{t-2} \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 2\rho & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ (t-1)\rho^{t-2} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial^2\Sigma(\lambda,\Phi)}{\partial\rho^2} = \phi_p^{\frac{1}{2}}\phi_s^{\frac{1}{2}}\begin{pmatrix} 0 & 0 & 2 & 6\rho & & (t-1)(t-2)\rho^{t-3} \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 2 & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ (t-1)(t-2)\rho^{t-3} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\alpha} = \phi_s\begin{pmatrix} 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 2\alpha & \cdots & (t-2)\alpha^{t-3} \\ 0 & 1 & 0 & 1 & & (t-3)\alpha^{t-4} \\ & \vdots & & & \ddots & \vdots \\ 0 & (t-2)\alpha^{t-3} & (t-3)\alpha^{t-4} & (t-4)\alpha^{t-5} & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial^2\Sigma(\lambda,\Phi)}{\partial\alpha^2} = \phi_s\begin{pmatrix} 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & \cdots & (t-2)(t-3)\alpha^{t-4} \\ 0 & 0 & 0 & & (t-3)(t-4)\alpha^{t-5} \\ & \vdots & & \cdots & \vdots \\ 0 & (t-2)(t-3)\alpha^{t-4} & (t-3)(t-4)\alpha^{t-5} & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_p\partial\phi_s} = \frac{1}{4\phi_p^{\frac{1}{2}}\phi_s^{\frac{1}{2}}}\begin{pmatrix} 0 & \rho & \rho^2 & & \rho^{t-1} \\ \rho & 0 & 0 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ \rho^{t-1} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_p\partial\rho} = \frac{\phi_s^{\frac{1}{2}}}{2\phi_p^{\frac{1}{2}}}\begin{pmatrix} 0 & 1 & 2\rho & 3\rho^2 & & (t-1)\rho^{t-2} \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 2\rho & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ (t-1)\rho^{t-2} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_p\partial\alpha} = \underline{\mathbf{0}}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_s\partial\rho} = \frac{\phi_p^{\frac{1}{2}}}{2\phi_s^{\frac{1}{2}}}\begin{pmatrix} 0 & 1 & 2\rho & 3\rho^2 & & (t-1)\rho^{t-2} \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 2\rho & 0 & 0 & 0 & & 0 \\ & & \vdots & & \ddots & \vdots \\ (t-1)\rho^{t-2} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\phi_s\partial\alpha} = \begin{pmatrix} 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 2\alpha & \cdots & (t-2)\alpha^{t-3} \\ 0 & 1 & 0 & 1 & & (t-3)\alpha^{t-4} \\ & & \vdots & & \ddots & \vdots \\ 0 & (t-2)\alpha^{t-3} & (t-3)\alpha^{t-4} & (t-4)\alpha^{t-5} & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial\Sigma(\lambda,\Phi)}{\partial\rho\partial\alpha} = \underline{\mathbf{0}}$$

A.3 List of partial derivatives from Chapter 4

$$\frac{\partial R(\lambda)}{\partial\gamma} = \begin{pmatrix} 0 & 1 & 0 & & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial R(\lambda)}{\partial\rho_1} = \begin{pmatrix} 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial R(\lambda)}{\partial \rho_2} = \begin{pmatrix} 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & & 0 \\ & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\frac{\partial R(\lambda)}{\partial \alpha} = \begin{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{t-3}{t-2} \end{pmatrix} & \underline{0} \\ \underline{0} & -I_{t-3} \end{pmatrix}$$

$$\frac{\partial R(\lambda)}{\partial \gamma \partial \rho_1} = \frac{\partial R(\lambda)}{\partial \gamma \partial \rho_2} = \frac{\partial R(\lambda)}{\partial \gamma \partial \alpha} = \underline{0}$$

$$\frac{\partial R(\lambda)}{\partial \rho_1 \partial \rho_2} = \frac{\partial R(\lambda)}{\partial \rho_1 \partial \alpha} = \frac{\partial R(\lambda)}{\partial \rho_2 \partial \alpha} = \underline{0}$$

VITA

Roy Sabo
Department of Mathematics and Statistics
Old Dominion University
Norfolk, VA 23529

Education

Ph.D.	Old Dominion University, Norfolk, VA. (December 2007) Major: Computational and Applied Mathematics (Statistics)
BS	Hamilton College, Clinton, New York, United States. (May 2000) Major: Economics.

Experience

Biostatistics Graduate Assistant (08/2006 - 06/2007)
Old Dominion University, Norfolk, VA

Teaching Assistant (06/2004 - 08/2006 & 8/2007 - 12/2007)
Old Dominion University, Norfolk, VA

Publications

Chaganty, N. R. and **Sabo, R.**, *"Estimation Methods for an Autoregressive Familial Correlation Structure with Homogeneous Variance"*, under preparation.
Chaganty, N. R. and **Sabo, R.**, *"Estimation Methods for Nuclear Family Correlation Matrix Under Transformation"*, under preparation.

Typeset using LaTeX.