Title
A Bayesian Family Factor Model for Multiple Outcomes

Author
Chen, Qiaolin

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Bayesian Family Factor Model for Multiple Outcomes

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Biostatistics

by

Qiaolin Chen

2014

ABSTRACT OF THE DISSERTATION

A Bayesian Family Factor Model for Multiple Outcomes

by

Qiaolin Chen

Doctor of Philosophy in Biostatistics

University of California, Los Angeles, 2014

Professor Robert E. Weiss, Co-chair

Professor Catherine A. Sugar, Co-chair

The UCLA Neurocognitive Family Study (NFS) collected multiple measurements on schizophrenia (SZ) patients and their relatives, as well as control subjects and their relatives, to study heritable vulnerability factors for schizophrenia. Each family has several members enrolled in the study and the same multiple outcomes were measured on each person. The relationship structure is complicated because not only observations on individuals from the same family are correlated, but the multiple outcome measures on the same individuals are also correlated. Traditional familial data analyses model outcomes separately and thus do not provide information about the interrelationships among them. I propose a Bayesian Family Factor Model (BFFM), which extends the classical confirmatory factor analysis (CFA) model to explain the correlations among observed variables using a combination of family-member factors and outcome factors. Traditional methods for fitting CFA models, such as full information maximum likelihood (FIML) estimation using quasi-Newton optimization (QNO) can have convergence problems and Heywood cases caused by empirical under-identification. In contrast, modern Bayesian Markov chain Monte Carlo (MCMC) handles these inference problems easily. Simulations compare the BFFM to FIML-QNO in settings where the true covariance matrix is identified, close to not identified and not identified. For these settings, FIML-QNO fails to fit the data in 85%, 57% and 13% of the cases, re-

spectively, due to non-convergence or invalid estimates, while MCMC provides stable estimates. When both methods successfully fit the data, estimates from the BFFM have smaller variances and comparable mean squared errors. BFFM can test hypotheses of interest easily using Bayes factors computed as the Savage-Dickey ratios. I illustrate the BFFM by analyzing the UCLA NFS data and test hypotheses about differences in means between SZ and control families. Tests of the group mean differences using posterior probabilities suggest that SZ probands perform worse in all 17 neurocogitive measures than control probands, while mothers of SZ subjects do worse than control mothers.

The dissertation of Qiaolin Chen is approved.

Keith H. Nuechterlein

Donatello Telesca

Catherine A. Sugar, Committee Co-chair

Robert E. Weiss, Committee Co-chair

University of California, Los Angeles

2014

This doctoral dissertation is dedicated to my always encouraging parents, Yuwang Chen and Guanlian Zhu, and my sweet little girl, Eunice Gao, with whom I could have spent more time during the past three years. I would like to acknowledge the inspirational instruction and guidance of my advisors, Dr. Robert E. Weiss and Dr. Catherine A. Sugar, who understood and helped me during hard times in my life. Finally, I would like to thank my mathematics and statistics tutor Dr. Wenhua Gao. I could not have changed my career.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The author thanks Dr. Keith H. Nuechterlein and Dr. Robert F. Asarnow for access to the data and for helpful discussions. The author is also deeply grateful to Dr. Weiss and Dr. Sugar for their valuable guidance and help. This work was partially funded by a Dissertation Year Fellowship from the Graduate Division at UCLA and by NIMH Grants MH041953, MH049716, MH045112, MH037705 and MH066286.

VITA

2002-2006	B.S. (Biological Sciences), Peking University, Beijing, China.
2006-2009	Ph.D. Candidate (Microbiology, Immunology and Molecular Genetics), UCLA, Los Angeles, California.
2009-2011	M.S. (Biostatistics), UCLA, Los Angeles, California.
2009–2010	Teaching Assistant, Department of Biostatistics, UCLA.
2010	Clinical Programmer Intern, Baxter Biosciences, Westlake Village, CA.
2010–present	Biostatistics Graduate Research Assistant, Semel Institute Biostatistics Core, UCLA.

PUBLICATIONS AND PRESENTATIONS

Schlosser, D.A., Jacobson, S., **Chen, Q.**, Sugar, C.A., Niendam, T.A., Li, G., Bearden, C.E. and Cannon, T.D. Recovery From an At-Risk State: Clinical and Functional Outcomes of Putatively Prodromal Youth Who Do Not Develop Psychosis. *Schizophrenia Bulletin*, August 2011; 38 (6): 1225-1233.

Miller, M. E., **Chen, Q.**, Elashoff, D., Abemayor, E. and St. John, M. , Positron emission tomography and positron emission tomography-CT evaluation for recurrent papillary thyroid carcinoma: Meta-analysis and literature review. *Head and Neck*, April 2011; 33(4): 562-565.

Miller, M.E., Palla, B., **Chen, Q.**, Elashoff, D.A., Abemayor, E., St John, M.A. and Lai, C.K. A novel classification system for perineural invasion in noncutaneous head and neck squamous cell carcinoma: histologic subcategories and patient outcomes. *American Journal of Otolaryngology*, 2011; 33(2): 212C215.

Miller, M.E., Palla, B., **Chen, Q.**., Elashoff, D.A., Abemayor, E., St John, M.A. and Lai, C.K. A Novel Classification System for Perineural Invasion: Histologic Subcategory Relates to Patient Outcome. *The Laryngoscope*, 120 (S3), S43-S43.

Wei, J., Sun, Z., **Chen, Q.**, and Gu, J. Serum deprivation induced apoptosis in macrophage is mediated by autocrine secretion of type I IFNs. *Apoptosis*, April 2006; 11(4): 545-54.

Bayesian Family Factor Models for Multiple Outcomes. Q Chen*, R Weiss, C Sugar. *Joint Statistical Meetings*, Montreal, Quebec, Canada. **8/3/2013 - 8/8/2013**

Bayesian Family Factor Models for Multiple Outcomes. Q Chen*, R Weiss, C Sugar. *The 20th ASA/IMS Spring Research Conference (SRC 2013) on Statistics in Industry and Technology*, Los Angeles, CA. **6/20/2013 - 6/22/2013**

Longitudinal Analysis of the Impact of Therapist Guidance on Time and Money Spent Gambling. Q Chen*, R Weiss, M Campos, T Fong. *Western North American Region Meetings*, Los Angeles, CA. **6/17/2013 - 6/19/2013**

Bayesian Family Factor Models for Multiple Outcomes. Qiaolin Chen*, R Weiss, C Sugar. *Eastern North American Region Meetings*, Orlando, FL. **3/10/2013 - 3/13/2013**

Bayesian Family Factor Models for Multiple Outcomes in Familial Data. Q Chen*, R

Weiss, C Sugar. *Joint Statistical Meetings*, San Diego, CA. **7/28/2012 - 8/2/2012**

CHAPTER 1

Introduction

Schizophrenia is a severe mental illness which affects a person's ability to differentiate between what is real and what is not, to think logically, to have normal emotional responses and to behave normally in social situations. Schizophrenia patients frequently have strange beliefs or delusions, see or hear things that aren't really there, speak or think in a disorganized way and withdraw from social interactions. The illness is long lasting and highly disabling. About 1 in 100 people will develop schizophrenia over their lifetime (Schultz et al., 2007). Neurocognitive deficits are a key feature of the disease and include reduced attention span, memory problems, difficulties with verbal fluency, executive functions, and rapid perceptual processing (Asarnow et al., 2002). Schizophrenia typically starts in late adolescence or early adulthood (Diagnostic and Statistical Manual of Mental Disorders, DSM-IV-TR, 2000), which is called adult onset. However, it sometimes occurs in children prior to age 12, which is called childhood onset. Vulnerability factors are the non-symptomatic characteristics reflecting an individual's predisposition to schizophrenia. Predisposing genes can cause these non-symptomatic abnormalities, which in turn contribute to schizophrenia. Potential vulnerability factors include abnormalities in neurocognitive functioning and in brain structure.

For complex diseases such as schizophrenia, there are typically multiple important outcome domains. Since these outcomes will be correlated, it is desirable to model them jointly. Separate analyses which ignore the within-subject-across-outcome correlations, both miss important mechanistic and clinical information, and are less pow-

erful. Approaches for joint analysis of multivariate data include linear mixed models (Sammel et al., 1999; McCulloch, 2006), structural equation modeling (SEM) (Bollen, 1998; Byrne, 2009; Kline, 2011) and factor analysis (Rao, 1955; Thompson, 2004; Brown, 2006; Bartholomew et al., 2011).

As with many other psychiatric disorders, genes as well as environmental factors are considered to play an important role in causing schizophrenia. Family studies are often used to identify possible genetic factors involved in a disease (Donner and Koval, 1980; Karlin et al., 1981; Morris, 2009; Wang et al., 2011). In such studies, a proband is an individual who triggers study of other members of the family. Some family studies use a case-control sampling design, collecting data on individuals with a given disorder and matched control subjects, as well as their relatives. Analysis of familial data is complicated by the presence of dependence among observations from the genetically related individuals.

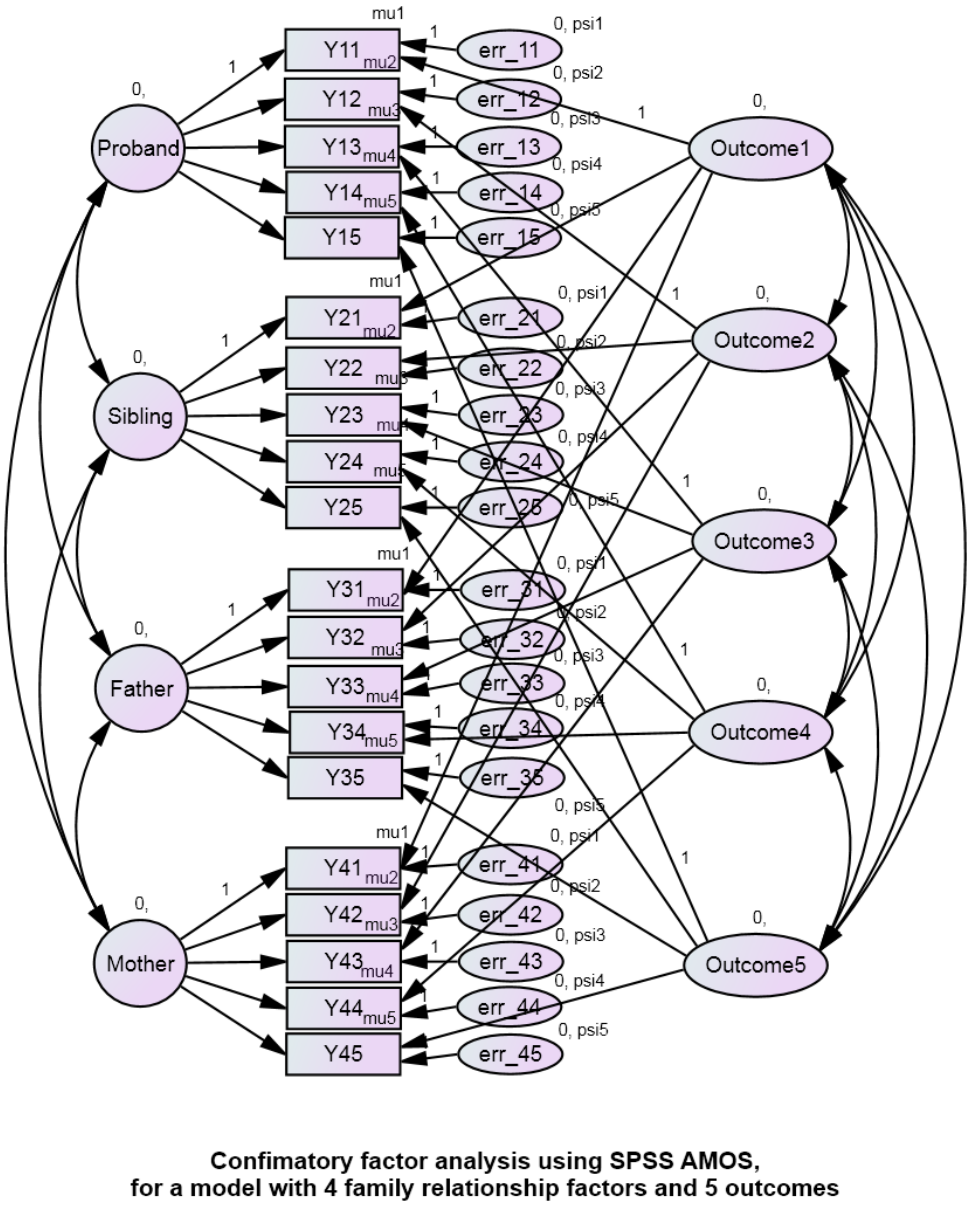
1.1 The UCLA Neurocognitive Family Study Data

The UCLA Neurocognitive Family Study (NFS) (Asarnow et al., 2001, 2002; Nuechterlein et al., 2002) is a cross-sectional case-controlled family study. Multiple cognitive measures were collected on schizophrenia patients and their first-degree relatives, as well as healthy controls and their relatives. The study aimed to investigate potential heritable predisposing or vulnerability factors for the disease, by identifying the features or characteristics which distinguish schizophrenia patients and their families from healthy controls and their families, and by examining how these cognitive deficits are differentially expressed among family members. Our goal here is to find an appropriate model which can address three main objectives: (i) to compare the degree of abnormality between schizophrenia families and control families; (ii) to determine correlations among measurements from first degree relatives and (iii) to identify relationships among the multiple outcome measures.

Modeling data such as that from the UCLA Nuerocognitive Family Study requires a complex covariance structure. Suppose there are K outcome measures for each of J members in a total of N families, so that the observed data for each family, $\boldsymbol{y}_i$, is a JK vector. For example, in a nuclear family, there are $K = 4$ family member types: proband, father, mother and sibling. Both the J family member types and the K outcome types contribute to the variation in $\boldsymbol{y}_i$, which is summarized by a $JK \times JK$ covariance matrix. I assume that the covariances are explained by K unobserved family member factors and J unobserved outcome factors, which induce correlations on the observed measures, both across-family-member within-measure and across-measure within-family-member. As the measurements on individuals from the same family are related, the J family member factors are allowed to be correlated. Similarly, so the K outcome factors are also assumed correlated because outcome measurements within subjects are associated.

The relationships among the JK observed variables and $J + K$ factors can be described using a path diagram (Bollen, 1998; Loehlin, 2004; Brown, 2006). Figure 1.1 shows an example of a path diagram for familial data with $J = 4$ family members and $K = 5$ outcomes drawn using AMOS (Arbuckle, 2011; Blunch, 2012), an add-on to SPSS for structural equation modeling. In a path diagram, all unobserved quantities, including latent factors and residuals are represented by ovals, while all observed variables are represented by rectangles. Bidirectional arrows represent correlations, while the single-headed arrows represent causal effects. In the figure, the observed variables Y_{ijk} labeled Yjk omitting the i (rectangles in the middle), are assumed to be caused by two sets of factors, the correlated family member factors (Proband, Sibling, Father and Mother) and correlated outcome factors (Outcome1, . . . , Outcome5), along with residuals that are unique to each observed variable on each family member, err_jk for $j = 1, \ldots, J$ and $k = 1, \ldots, K$, controlled by variances unique to each observed variable, psi1, . . . , psi5. Means and variance parameters are labeled on the rectangles and ovals/circles, before and after commas, respectively. For example, mu1, . . . , mu5 are

means of Yj1, . . . , Yj5, for $j = 1, \ldots, 4$, and the means of all family member factors, outcome factors and residuals are restricted to 0.



Confimatory factor analysis using SPSS AMOS,
for a model with 4 family relationship factors and 5 outcomes

Figure 1.1: A path diagram for the Bayesian Family Factor Model (BFFM). Responses variables Y_{ijk} labeled Yjk omitting the i (rectangles in the middle), for $j = 1, \ldots, 4$ and $k = 1, \ldots, 5$, are caused by two sets of factors, family member specific factors (circles on the left) and outcome specific factors (ovals on the right), along with a residual error, err_jk, which is unique to each item.

1.2 Current Approaches and Problems

Standard analyses of family studies usually model outcomes separately, which is potentially less efficient and does not provide information about the relationships among outcomes (Hamsten and de Faire, 1987; Harrap et al., 2000; Asarnow et al., 2002). Classical analysis techniques for multiple outcomes are not designed to take into account associations among family members, which is equivalent to omitting the family member factors in Figure 1.1 and only considering the right half of the diagram.

Direct product models (Browne, 1984; Cudeck, 1989; Wothke and Browne, 1990; Naik and Rao, 2001; Srivastava et al., 2008) provide a potential method for analyzing familial data with multiple outcomes, which assume that family member factors interact with outcome factors in a multiplicative manner,

$$\text{var}(\boldsymbol{y}_i) = \boldsymbol{\Sigma}_{\text{member}} \otimes \boldsymbol{\Sigma}_{\text{outcome}},$$

where $\boldsymbol{\Sigma}_{\text{member}}$ and $\boldsymbol{\Sigma}_{\text{outcome}}$ are $J \times J$ and $K \times K$ covariance matrices for the two groups of effects, respectively. However, these models are rigid as they assume all outcomes have identical correlation matrices and ratios of variances.

1.2.1 Confirmatory Factor Analysis

Factor analysis (FA) models correlated observed variables using a smaller number of unobservable variables, called latent factors (Rao, 1955; Harman, 1960; Brown, 2006). It is used either for dimension reduction or to improve understanding of the pattern of associations among variables (Rowe, 1998). Variances in the observed variables are explained by both common factors and unique error variances. If some factors are assumed to be independent, the corresponding factor covariances are fixed to zero (Brown, 2006).

The basic structure for a factor model is described below. Suppose K outcome

measures are collected on each of N subjects. The relationship among the K outcomes maybe characterized by a factor analysis model with p factors

$$\boldsymbol{y}_i = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{f}_i + \boldsymbol{\varepsilon}_i, \tag{1.1}$$

where $\boldsymbol{y}_i$ is a $K \times 1$ outcome vector for subject i, $i = 1, \ldots, N$; $\boldsymbol{\mu} = E(\boldsymbol{y}_i)$ is a vector of overall means; $\boldsymbol{\Lambda}$ is a $K \times p$ matrix of factor loadings; $\boldsymbol{f}_i$ is a $p \times 1$ vector of factor scores with mean $\mathbf{0}$ and a $p \times p$ covariance matrix $\boldsymbol{\Phi}$

$$\boldsymbol{f}_i \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi});$$

and $\boldsymbol{\varepsilon}_i$ is a $K \times 1$ vector of unique errors independent of factor scores, with a diagonal variance matrix $\boldsymbol{\Psi} = \text{diag}(\psi_1, \ldots, \psi_K)$

$$\boldsymbol{\varepsilon}_i \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}).$$

The marginal variance-covariance matrix of $\boldsymbol{y}_i$ can be decomposed as the sum of variance and covariances due to the factors and variance due to the unique errors

$$\text{var}(\boldsymbol{y}_i) = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^t + \boldsymbol{\Psi}.$$

Confirmatory factor analysis (CFA) is used to test hypothesized relationships between observed variables and factors (Jöreskog, 1969). Researchers specify the number of factors beforehand and make *a priori* assumptions about which observed variables are related to which factors based on past evidence and theory. The factor loadings specify the pattern of relationships between the observed variables and the factors. Only loadings corresponding to hypothesized relationships between specific observed variables and factors are allowed to be nonzero. All the others, called cross-loadings, are fixed to zero. The scale of the factors can be defined by fixing factor variances to

1, or by setting the scale of a factor to be the same as one of the observed variable to which it contributes. For standard CFA, parameters are estimated using maximum likelihood, EM maximum likelihood or the method of moments (Rao, 1955; Rubin and Thayer, 1982; Basilevsky, 2009; Bartholomew et al., 2011). For standard CFA, parameters are estimated using maximum likelihood, EM maximum likelihood or the method of moments (Rao, 1955; Rubin and Thayer, 1982). Software programs used for performing confirmatory factor analysis include SPSS AMOS (Arbuckle, 2011), LISREL (Jöreskog and Sörbom, 2012), EQS (Byrne, 2013), Mplus (Muthén and Muthén, 1998–2012), SAS CALIS procedue (Hunter, 2005), and sem (Fox, 2006) and lavaan (Rosseel, 2012) packages in R. See Byrne (2001) and Albright and Park (2009) for reviews.

1.2.2 Multitrait-Multimethod (MTMM) Analysis

The structure of familial data with multiple outcomes is similar to that of the multitrait-multimethod (MTMM) data used for studying construct validity: the ability of psychological tests to actually measure the concept being studied (Campbell and Fiske, 1959; Marsh, 1989; Eid et al., 2006; Madans et al., 2011). For MTMM analysis, a certain number of traits (J) are each assessed by several methods (K) for each of N subjects, resulting in a $JK \times JK$ correlation matrix. The path diagram for an MTMM model is similar to Figure 1.1, replacing family members with traits and outcomes with methods.

Despite the similarity in data structure, the focus of MTMM analyses is quite different from analyses of familial data. MTMM analyses only model the correlation matrix not the mean structure, and mainly focuses on estimation and tests of parameters with specific meanings for construct validity. In contrast, in familial data analysis, mean structures may depend on covariates and hypotheses about regression coefficients are of interest. Incomplete data is a significant issue in familial data, as a family may not have all J member types and individual measures may also be missing for a particular subject.

The most popular technique for fitting an MTMM model is confirmatory factor analysis (CFA) using the correlated-trait correlated-method (CTCM) structure, which assumes the inter-related trait factors are independent of the inter-related method factors (Marsh and Hocevar, 1988; Kenny and Kashy, 1992). This model requires at least a total of $J + K \geq 6$ trait and method factors with at least $J \geq 2$ method and $K \geq 2$ trait factors to be identified, and it is not empirically identified when the loading matrix has deficient column rank (Grayson and Marsh, 1994), or when all trait or method factor correlations are equal (Brannick and Spector, 1990). Wothke (1984), Brannick and Spector (1990) and Lance et al. (2002) analyzed 21, 14 and 19 published MTMM matrices, respectively, and reported that in 100%, 94% and 100% of the cases, the algorithm for CFA model failed to converge or gave invalid solutions, such as negative variances or non-positive definite covariance matrices, which are called Heywood cases (Grayson and Marsh, 1994). The algorithm for fitting CFA models to familial data can have the same identification problems, resulting in non-convergence, fits with invalid solutions, improper estimates such as negative loadings, or unstable estimates with extreme standard errors.

1.2.3 Bayesian Factor Analysis

Bayesian factor analysis (BFA) (Press and Shigemasu, 1989; West, 2003; Lopes and West, 2004; Quinn, 2004; Ghosh and Dunson, 2009; Press, 2012) can help to mitigate the identification problem by incorporating available knowledge about parameters in the form of prior distributions based on either expert opinions or previous experiments. Markov chain Monte Carlo (MCMC) methodology has been applied previously in BFA to sample from posterior distributions (Geweke and Zhou, 1996; Press and Shigemasu, 1997; Rowe, 1998; Aguilar and West, 2000; Rowe, 2003).

BFA often makes normality assumptions for the distribution of unique errors. Conditionally conjugate priors for model parameters facilitate straightforward posterior computation by Gibbs sampling (Geman and Geman, 1984). For example, normal

priors are often used for means. Bayesian inference using inverse-gamma priors for unique error variances and inverse-Wishart priors for the covariance matrices avoid the problem of Heywood cases (negative variances and non-positive definite covariance matrices) that occur with maximum likelihood approaches. Normal priors are usually specified for factor loadings. Bayesian methods have not been previously applied to CFA for analyzing familial data with multiple outcomes or for fitting the MTMM models. By incorporating Bayesian techniques, it is possible to solve most problems of standard CFA.

The rest of the paper proceeds as follows: Chapter 2 describes the proposed Bayesian Family Factor Model (BFFM), including the basic model structure, prior specification, a Gibbs algorithm to impute missing data and sample from the posterior distributions. Chapter 3 discusses simulation studies comparing BFFM with the full information maximum likelihood estimation of CFA using quasi-Newton optimization (FIML-QNO) algorithm by the lavaan package in R. In Chapter 4 and 5 I fit BFFM to the motivating UCLA Neurocognitive Family Study (NFS) data. Methods of testing hypotheses and their application to the UCLA NFS data are described in Chapter 6. Implications and possible extensions are discussed in Chapter 7.

CHAPTER 2

The Bayesian Family Factor Model

I propose a Bayesian Family Factor Model (BFFM), which extends the classical confirmatory factor analysis (CFA) model to explain the correlations among observed variables using a combination of family-member factors and outcome factors. This chapter describes the basic structure for a BFFM, the specification of conditionally conjugate priors and a Gibbs sampling algorithm.

2.1 Specification of the BFFM

I propose the following basic structure for a BFFM. Suppose K normally distributed outcomes are collected on each of J members in N families. Let i, j and k index family, member type and outcome, respectively, with $i = 1, \ldots, N$, $j = 1, \ldots, J$ and $k = 1, \ldots, K$. Then y_{ijk} is the k^{th} outcome for the j^{th} member in the i^{th} family, $\boldsymbol{y}_{ij} = (y_{ij1}, \ldots, y_{ijK})^T$ is the $K \times 1$ vector of K outcomes for the j^{th} member in the i^{th} family and $\boldsymbol{y}_i = (y_{i11}, \ldots, y_{i1K}, \ldots, y_{iJ1}, \ldots, y_{iJK})^T$ is the $JK \times 1$ vector of observations for all J members in the i^{th} family. The relationships among the JK observed variables are characterized by a factor analysis model

$$\boldsymbol{y}_i \;=\; \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\Lambda}_A\boldsymbol{f}_{Ai} + \boldsymbol{\Lambda}_B\boldsymbol{f}_{Bi} + \boldsymbol{\varepsilon}_i, \tag{2.1}$$

where $\boldsymbol{X}_{i(JK\times P)}$ is a matrix of known covariates for family i; $\boldsymbol{\beta}_{P\times 1} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_P)$ is a vector of regression coefficients; $\boldsymbol{f}_{Ai(J\times 1)}$ and $\boldsymbol{f}_{Bi(K\times 1)}$ are independent vectors of family member factors and outcome factors, respectively, with corresponding variance

matrices $\boldsymbol{\Phi}_{A(J\times J)}$ and $\boldsymbol{\Phi}_{B(K\times K)}$

$$\boldsymbol{f}_{Ai} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi}_A),$$

$$\boldsymbol{f}_{Bi} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi}_B);$$

$\boldsymbol{\varepsilon}_i$ is a $JK \times 1$ vector of unique errors independent of $\boldsymbol{f}_{Ai}$ and $\boldsymbol{f}_{Bi}$ with diagonal error variance matrix $\boldsymbol{\Psi}_{(JK\times JK)} = \text{diag}(\psi_{11}, \ldots, \psi_{JK})$

$$\boldsymbol{\varepsilon}_i \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi});$$

$\boldsymbol{\Lambda}_{A(JK\times J)} = \text{blockdiag}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_J)$ is a family member factor loading matrix with diagonal blocks of $K \times 1$ vectors $\boldsymbol{\alpha}_j = (1, a_{j2}, \ldots, a_{jK})$; and $\boldsymbol{\Lambda}_{B(JK\times K)} = [\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_K]^T$ is an outcome factor loading matrix, where $\boldsymbol{B}_1 = \boldsymbol{I}_K$, $\boldsymbol{B}_j = \text{diag}(b_{j1}, \ldots, b_{jK})$, for $j = 2, \ldots, J$. Here $\boldsymbol{\alpha}_j$ is a vector of non-zero family factor loadings for the j^{th} family member specific effects and $\boldsymbol{\Lambda}_{Bj}$ is a diagonal matrix of outcome factor loadings for the j^{th} family member.

The variance-covariance matrix of the observed variables, y_i, for the i^{th} family unconditional on the factors is

$$\boldsymbol{\Sigma} = \text{var}(\boldsymbol{y}_i|\boldsymbol{\beta}) = \boldsymbol{\Lambda}_A\boldsymbol{\Phi}_A\boldsymbol{\Lambda}_A^T + \boldsymbol{\Lambda}_B\boldsymbol{\Phi}_B\boldsymbol{\Lambda}_B^T + \boldsymbol{\Psi},$$

while the variance of observed variable y_{ijk} is

$$\text{var}(y_{ijk}|\boldsymbol{\beta}) = a_{jk}^2\phi_{Ajj} + b_{jk}^2\phi_{Bkk} + \psi_{jk},$$

where ϕ_{Ajj} is the j^{th} diagonal element of $\boldsymbol{\Phi}_A$ and ϕ_{Bkk} is the k^{th} diagonal element of $\boldsymbol{\Phi}_B$, a_{jk} is the k^{th} element of $\boldsymbol{\alpha}_j$ and b_{jk} is the k^{th} diagonal element of $\boldsymbol{B}_j$.

The factor loading matrices $\boldsymbol{\Lambda}_A$ and $\boldsymbol{\Lambda}_B$ can be expressed as

$$[\boldsymbol{\Lambda}_A|\boldsymbol{\Lambda}_B] = \left[\begin{array}{ccc|c} \boldsymbol{\alpha}_1 & \cdots & \mathbf{0} & \boldsymbol{B}_1 \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{B}_2 \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \boldsymbol{\alpha}_J & \boldsymbol{B}_J \end{array}\right] = \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & \ldots & 0 \\ a_{12} & 0 & 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1K} & 0 & 0 & 0 & 0 & 0 & \ldots & 1 \\ 0 & 1 & 0 & 0 & b_{21} & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & 0 & 0 & b_{22} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{2K} & 0 & 0 & 0 & 0 & \ldots & b_{2K} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & b_{J1} & 0 & \ldots & 0 \\ 0 & 0 & 0 & a_{J2} & 0 & b_{J2} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{JK} & 0 & 0 & \ldots & b_{JK} \end{array}\right]. \tag{2.2}$$

The model for all K outcomes on K members of all N families is

$$\boldsymbol{Y} = \boldsymbol{Z}(\boldsymbol{\beta} \otimes \boldsymbol{I}_{JK}) + \boldsymbol{F}_A\boldsymbol{\Lambda}_A^T + \boldsymbol{F}_B\boldsymbol{\Lambda}_B^T + \boldsymbol{E},$$

where $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)^T$ is an $N \times JK$ matrix of all observed data for N families, $\boldsymbol{Z} = [\text{vec}(\boldsymbol{X}_1), \ldots, \text{vec}(\boldsymbol{X}_N)]^T$ is an $N \times JKP$ matrix of known covariates, $\boldsymbol{\beta}$ is a $P \times 1$ vector of regression coefficients, $\boldsymbol{F}_A = (\boldsymbol{f}_{A1}, \ldots, \boldsymbol{f}_{AN})^T$ is an $N \times J$ matrix of family member factor scores, $\boldsymbol{F}_B = (\boldsymbol{f}_{B1}, \ldots, \boldsymbol{f}_{BN})^T$ is an $N \times K$ matrix of outcome factor scores, and $\boldsymbol{E} = (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_N)^T$ is an $N \times JK$ matrix of residual errors.

For the k^{th} outcome measured on the first family member (proband, $j = 1$), as

$a_{1k} = 1$, the overall variance var(y_{i1k}) can be decomposed as

$$\text{var}(y_{i1k}) \quad = \quad 1\phi_{A11} + b_{1k}^2\phi_{Bkk} + \psi_{1k},$$

which indicates that the factor variance for probands, ϕ_{A11}, must be smaller than any of the overall variances for probands, var(y_{i1k}). This information helps in setting priors of factor loadings and factor covariance matrix. Therefore, it is better to scale the observed variables to make the overall variances similar, so that ϕ_{A11} will not be forced to be small, which can cause precision problem in computing such as very small values being rounded to 0.

Next, scales for factors and factor loadings are specified. Here scale for observed variables is a combination of both size/magnitude and dispersion. For factor loadings and factor variance matrices, scale is more related to variation or dispersion, as these parameters are used to model the variance-covariance matrix. The scales of all family member factors, f_{Aij}, for $j = 1, \ldots, J$, are set to be the same as the observed variables for the first outcome, y_{ij1}, by fixing the first nonzero loading in each column of $\mathbf{\Lambda}_A$ to 1, $a_{j1} = 1$. The scale of a family member factor loading, a_{jk}, is the ratio of the scale of the k^{th} outcome to that of the first outcome, for $k = 2, \ldots, K$. Factor loading a_{jk} is the amount of change in y_{ijk} associated with a 1 unit increase in f_{Aij} with all else fixed. In addition, because $a_{jk}/a_{j1} = a_{jk}/1 = a_{jk}$, loading a_{jk} is also the ratio of the effect of f_{Aij} on y_{ijk} to its effect on y_{ij1}.

Similarly, the scale for an outcome factor, f_{Bik}, is specified to be the same as that of the observed variable for the first family member (proband), y_{i1k}, by fixing the first nonzero loading in each column of $\mathbf{\Lambda}_B$ to 1, $b_{1k} = 1$. Therefore, the scale of the k^{th} outcome is passed on to the k^{th} outcome factor, f_{Bik}. Similarly, outcome loading b_{jk} is the amount of change in y_{ijk} associated with a 1 unit increase in f_{Bik} and as $b_{jk}/b_{1k} = b_{jk}$ for $j = 2, \ldots, J$, b_{jk} is also the ratio of the effect of f_{Bik} on y_{ijk} to that on y_{i1k}.

The total number of free hyper-parameters in the model is $(3JK + J^2/2 + K^2/2 - J/2 - K/2 + P)$, as there are P regression coefficients, $(J-1)K$ family member factor loadings, $J(K-1)$ outcome factor loadings, $J(J+1)/2$ unique parameters in the family factor variance matrix, $K(K+1)/2$ unique parameters in the family factor variance matrix, and JK unique error variance parameters. Similar to the CFA model for MTMM data, this model requires at least a total $J + K \geq 6$ family member and outcome factors with at least $J \geq 2$ family member and $K \geq 2$ outcome factors to be identified, and it is not empirically identified when the loading matrix has deficient column rank (Grayson and Marsh, 1994), or when all family member or outcome factor correlations are equal (Brannick and Spector, 1990).

There is a one-to-one correspondence between model parameters and lines on the path diagram in Figure 1.1. Factor variances matrices, $\mathbf{\Phi}_A$ and $\mathbf{\Phi}_B$, correspond to bidirectional arrows among the $J = 4$ family member factors on the left and among the $K = 5$ outcome factors on the right, respectively. The non-zero elements of $\mathbf{\Lambda}_A$, namely $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_J$, correspond to unidirectional arrows from family member factors on the left to the JK observed variables, $\boldsymbol{y}_i$. The non-zero elements of $\mathbf{\Lambda}_B$, namely diagonal elements of $\boldsymbol{B}_j$, correspond to unidirectional arrows from family member factors on the right to $\boldsymbol{y}_i$.

2.2 Conditionally Conjugate Priors for BFFM

To complete a Bayesian specification of the model, priors need to be assigned for each unknown parameter. In the absence of strong theoretical or empircal beliefs to the contrary, I specify conditionally conjugate priors for all parameters. In the absence of strong theoretical or empirical beliefs to the contrary, I specify conditionally conjugate priors for all parameters. The prior distributions for the regression coefficients, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$, and free elements a_{jk} and b_{jk} in the factor loading matrices, $\mathbf{\Lambda}_A$ and $\mathbf{\Lambda}_B$,

are independent normal

$$\begin{aligned}
\beta_p &\overset{\text{iid}}{\sim} \mathcal{N}(\beta_{0p}, \sigma^2_{\beta 0p}), \text{ for } p = 1, \ldots, P, \\
a_{jk} &\overset{\text{ind}}{\sim} \mathcal{N}(\mu_{a_{jk}}, \sigma^2_{a_{jk}}), \text{ for } j = 1, \ldots, J, k = 2, \ldots, K, \\
b_{jk} &\overset{\text{ind}}{\sim} \mathcal{N}(\mu_{b_{jk}}, \sigma^2_{b_{jk}}), \text{ for } j = 2, \ldots, J, k = 1, \ldots, K
\end{aligned}$$

The factor variance matrices, $\boldsymbol{\Phi}_A$ and $\boldsymbol{\Phi}_B$, follow independent inverse Wishart distributions

$$\begin{aligned}
\boldsymbol{\Phi}_A &\sim \mathcal{IW}(\boldsymbol{W}_A, \nu_A), \\
\boldsymbol{\Phi}_B &\sim \mathcal{IW}(\boldsymbol{W}_B, \nu_B),
\end{aligned}$$

where ν_A and ν_B are the degrees of freedom parameters, $\boldsymbol{W}_{A(J\times J)} = (\nu_A - J - 1)\boldsymbol{D}_A\boldsymbol{C}_A\boldsymbol{D}_A$ and $\boldsymbol{W}_{B(K\times K)} = (\nu_B - K - 1)\boldsymbol{D}_B\boldsymbol{C}_B\boldsymbol{D}_B$ are location parameters, $\boldsymbol{C}_{A(J\times J)}$ and $\boldsymbol{C}_{B(K\times K)}$ are prior factor correlation matrices, and $\boldsymbol{D}_{A(J\times J)} = \text{diag}(d_{A1}, \ldots, d_{AJ})$ and $\boldsymbol{D}_{B(K\times K)} = \text{diag}(d_{B1}, \ldots, d_{BK})$ are matrices with factor variances as diagonal elements to be specified shortly. Independent inverse-gamma priors are specified for the JK diagonal elements of $\boldsymbol{\Psi}$

$$\psi_{jk} \overset{\text{ind}}{\sim} \mathcal{IG}\left(\frac{\alpha_{\psi_{jk}}}{2}, \frac{\beta_{\psi_{jk}}}{2}\right),$$

for $j = 1, \ldots, J$ and $k = 1, \ldots, K$.

2.3 Specification of Prior Hyper-parameters

This section describes an approach to eliciting prior hyper-parameters based on model interpretation and subject matter knowledge. The basic assumptions are that the variances of the K outcomes are distinct due to scale differences and that the variances across family members of the k^{th} outcome are similar. The first step is to obtain estimated values for the overall variances of the K outcomes, $\widehat{\text{var}}(y_1), \ldots, \widehat{\text{var}}(y_K)$, either from the literature, from previous studies or from expert opinion. When no other infor-

mation is available, $1/4$ of the range of the k^{th} outcome variable in the data set under study is a plausible value of $\widehat{\text{var}}(y_k)^{1/2}$.

To specify priors for factor variance matrices, note that

$$\begin{aligned} \text{var}(y_{ij1}) &= a_{jk}^2\phi_{Ajj} + \phi_{B11} + \psi_{j1}, \\ \text{and} \qquad \text{var}(y_{i1k}) &= \phi_{A11} + b_{jk}^2\phi_{Bkk} + \psi_{1k}, \end{aligned}$$

which implies $\widehat{\text{var}}(y_1)$ can be used as an upper bound up to sampling error of the prior mean of ϕ_{B11}, $d_{B1}^2 = p_{\phi b1}\widehat{\text{var}}(y_1)$, and the minimum of $\widehat{\text{var}}(y_1), \ldots, \widehat{\text{var}}(y_1)$ can be used an upper bound of the prior mean of ϕ_{A11}, $d_{A1}^2 = p_{\phi a1}\text{min}(\widehat{\text{var}}(y_1), \ldots, \widehat{\text{var}}(y_K))$, for scaling constants $0 < p_{\phi a1}, p_{\phi b1} \leq 1$. As the scale of the first outcome is passed on to all family member factors, f_{Aij}, the prior means of factor variances are set to be equal, $d_{A1}^2 = \ldots = d_{AJ}^2$. As the scale of the k^{th} outcome is passed on to the k^{th} outcome factor, f_{Bik}, the prior means of outcome factor variances, ϕ_{Bkk}, are set to be proportional to the estimated overall variances,

$$\frac{d_{B1}^2}{\widehat{\text{var}}(y_1)} = \ldots = \frac{d_{BK}^2}{\widehat{\text{var}}(y_K)} = p_\psi,$$

where the scaling constant $0 < p_\psi \leq 1$.

Information on correlations within outcome across family members and among outcomes within subjects can help to specific $\boldsymbol{C}_A$ and $\boldsymbol{C}_B$, the prior factor correlation matrices. Some information about theoretical associations among family members are available. For example, the genetic correlations between father and mother, between parent and children and between siblings are 0, 0.5 and 0.5, respectively. In addition, some outcome measures are known to be more closely related than others. For example, correlations among sub-scales from the same test will be similar and higher than correlations coming from sub-scales of different tests, which can be reflected in the prior factor correlation matrix $\boldsymbol{C}_B$.

Prior means of factor loadings are elicited as follows. For a particular outcome,

effects of the different family member factors on the observed variables are likely to be similar, so I assume that the prior means of loadings for the same outcome are equal across members,

$$\mu_{a_{1k}} = \mu_{a_{Jk}},$$

for $k = 2, \ldots, K$. As the scale of a_{jk} is the ratio of the scale of the k^{th} outcome to the scale of the first outcome, I set prior means of the loadings proportional to the square root of the estimated overall variances

$$\frac{1}{\widehat{\text{var}}(y_1)^{1/2}} = \frac{\mu_{a_{j2}}}{\widehat{\text{var}}(y_2)^{1/2}} = \ldots = \frac{\mu_{a_{jK}}}{\widehat{\text{var}}(y_K)^{1/2}},$$

for j=1,. . . ,J. For outcome factor loadings, because effects of the same outcome factor on observed variables are likely to be similar across family members, and $b_{11} = \ldots = b_{1K} = 1$, I set prior means of all outcome factor loadings to 1

$$\mu_{b_{jk}} \equiv 1,$$

fo $j = 2, \ldots, J$ and $k = 1, \ldots, K$. To specify prior means for the unique error variance, ψ_{jk}, for $j = 1, \ldots, J$ and $k = 1, \ldots, K$, $\widehat{\text{var}(y_k)}$ can be used as an upper bound, as the total variance the sum of variance due to unique error and variance due to common factors. To specify the priors for regression coefficients, it is necessary to identify plausible values for the covariate effects on each outcome from previous studies or expert opinion. For the special case where covariates are indicators of diagnostic or treatment groups, the estimated means of outcomes in the general population or in patients from earlier studies are useful guides for choosing prior means.

2.3.1 Gibbs Sampling from the Posterior Distribution

Because of the use of conjugate priors, simulation of the posterior distribution proceeds via a Gibbs sampling algorithm where each parameter is sampled from its full

conditional distribution (Geman and Geman, 1984; Gelfand and Smith, 1990; Robert and Casella, 2004). To reduce autocorrelation and improve efficiency, I use a blocked Gibbs sampler to sample the regression coefficients, $\boldsymbol{\beta}$ and the factor scores, $\boldsymbol{f}_i$ from their joint conditional distributions, respectively. Full details of the Gibbs sampler are given in the appendix.

Missing data are imputed at each iteration of the MCMC algorithm with a data augmentation (DA) algorithm, assuming observations are missing at random (MAR) (Little and Rubin, 2002). This approach has the advantage of using BFFM for both imputation and data analysis. Because the missing and observed data are jointly normally distributed, the conditional distribution of the missing data given the observed data is also normal. I implemented Schafer (1997)'s sweep operator algorithm for imputation of multivariate normal data. For details see the appendix.

2.4 Data Likelihood and the Conditional Posterior Distributions

Bayesian inference usually involves specification of priors for model parameters, calculation of data likelihood and calculation of the posterior densities. It is often not possible to obtain the posterior distribution with straightforward analytical solutions, so it is necessary to generate samples from the posterior distribution using sampling methods such as Markov chain Monte Carlo (MCMC). Because of the use of conjugate priors, simulation of the posterior distribution proceeds via a Gibbs sampling algorithm where each parameter is sampled from its full conditional distribution (Geman and Geman, 1984; Gelfand and Smith, 1990; Robert and Casella, 2004). This section describes the computation of data likelihood and the derivation of conditional posterior densities.

Define $\boldsymbol{e}_{i(jk)} = \boldsymbol{y}_i - \boldsymbol{X}_i^T\boldsymbol{\beta} - \boldsymbol{\Lambda}_A\boldsymbol{f}_{Ai} - \boldsymbol{\Lambda}_B\boldsymbol{f}_{Bi}$ and $\boldsymbol{E}_{N\times JK} = \boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B} - \boldsymbol{F}_A\boldsymbol{\Lambda}_A^T - \boldsymbol{F}_B\boldsymbol{\Lambda}_B^T$. The complete data likelihood for all parameters in the model, $\boldsymbol{\Theta} = (\boldsymbol{\beta}, \boldsymbol{F}_A, \boldsymbol{F}_B, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Psi})$, based on K outcomes and J family members for all N fam-

ilies, $\boldsymbol{Y}_{N\times JK} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)^T$, has the following form

$$\begin{aligned} & L(\boldsymbol{\Theta}|\boldsymbol{Y}) \\ = \; & (2\pi)^{-\frac{NJK}{2}} |\boldsymbol{\Psi}|^{-\frac{N}{2}} \exp\left\{ -\tfrac{1}{2}\text{tr}\left[\boldsymbol{\Psi}^{-1} \sum_{i=1}^{N} (\boldsymbol{e}_i)\boldsymbol{e}_i)^T \right] \right\} \\ = \; & (2\pi)^{-\frac{NJK}{2}} |\boldsymbol{\Psi}|^{-\frac{N}{2}} \exp\left\{ -\tfrac{1}{2}\text{tr}\left[\boldsymbol{\Psi}^{-1} \boldsymbol{E}^T \boldsymbol{E} \right] \right\}. \end{aligned}$$

The joint posterior distribution for all parameters is proportional to the complete data likelihood multiplied by the prior density

$$\begin{aligned} & \text{p}(\boldsymbol{\beta}, \boldsymbol{F}_A, \boldsymbol{F}_B, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Psi}|\boldsymbol{Y}) \\ \propto \; & \text{p}(\boldsymbol{Y}|\boldsymbol{\beta}, \boldsymbol{F}_A, \boldsymbol{F}_B, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Psi}) \\ & \qquad \times \text{p}(\boldsymbol{\beta})\text{p}(\boldsymbol{\Lambda}_A)\text{p}(\boldsymbol{\Lambda}_B)\text{p}(\boldsymbol{F}_A|\boldsymbol{\Phi}_A)\text{p}(\boldsymbol{F}_B|\boldsymbol{\Phi}_B)\text{p}(\boldsymbol{\Phi})\text{p}(\boldsymbol{\Psi}). \end{aligned}$$

Because it is easier to compute conditional posterior of the parameters given the complete data, I use the data augmentation (DA) algorithm which treats missing data as unknown parameters and impute them as a step in the MCMC algorithm (Little and Rubin, 2002; Schafer, 1997). The rest of this section presents the computation of the conditional posterior densities for all of the model parameters, including the missing data, regression coefficients, factor loadings, factor scores, unique error variances and factor variance matrices.

2.4.1 Missing Data Imputation

Missing data are handled using a data augmentation (DA) algorithm, which sequentially imputes missing data and samples from a complete-data Bayesian model via MCMC (Little and Rubin, 2002), assuming observations are missing at random (MAR). This approach has the advantage of using BFFM for both imputation and data analysis. Because the missing and observed data are jointly normally distributed, the conditional distribution of the missing data given the observed data is also normal. Let $\boldsymbol{y}_{i,o}$ and $\boldsymbol{y}_{i,m}$ denote the observed and missing parts of $\boldsymbol{y}_i$ by respectively. At iteration l with

current parameter value $\boldsymbol{\Theta}^{(l)}$, sample

$$\boldsymbol{y}_{i,m}^{(l+1)} \sim p(\boldsymbol{y}_{i,m}|\boldsymbol{\Theta}^{(l)}, \boldsymbol{y}_{i,o}),$$

for $i = 1, \ldots, N$, where $\boldsymbol{\Theta}^{(l)}$ does not include factors $\boldsymbol{f}_{Ai}$ nor $\boldsymbol{f}_{Bi}$. Define the mean vector and variance matrix of $\boldsymbol{y}_i$ as $\boldsymbol{\mu}_i = \boldsymbol{X}_i\boldsymbol{\beta}$ and $\boldsymbol{\Sigma} = \text{var}(\boldsymbol{y}_i|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}) = \boldsymbol{\Lambda}_A\boldsymbol{\Phi}_A\boldsymbol{\Lambda}_A^T + \boldsymbol{\Lambda}_B\boldsymbol{\Phi}_B\boldsymbol{\Lambda}_B^T + \boldsymbol{\Psi}$. After grouping $\boldsymbol{y}_{i(JK\times 1)}$ in the order of observed and missing parts, the missing and observed data are jointly normally distributed

$$\boldsymbol{y}_{i(JK\times 1)}^{*} = \begin{bmatrix} \boldsymbol{y}_{i,obs} \\ \boldsymbol{y}_{i,miss} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_{i1} \\ \boldsymbol{\mu}_{i2} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11i} & \boldsymbol{\Sigma}_{12i} \\ \boldsymbol{\Sigma}_{12i}^T & \boldsymbol{\Sigma}_{22i} \end{bmatrix} \right),$$

where the normal mean vector and variance matrix are obtained by permuting $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}$ in the order of $\boldsymbol{y}_{i,o}$ and $\boldsymbol{y}_{i,m}$. Therefore conditioned on the observed data and all parameters, the missing data are also normally distributed

$$\boldsymbol{y}_{i,m}|\boldsymbol{y}_{i,o} = \boldsymbol{y}_{i,1} \sim \mathcal{N}\left(\boldsymbol{\mu}_{i2} + \boldsymbol{\Sigma}_{21i}\boldsymbol{\Sigma}_{11i}^{-1}(\boldsymbol{y}_{i,1} - \boldsymbol{\mu}_{i1}), \boldsymbol{\Sigma}_{22i} - \boldsymbol{\Sigma}_{21i}\boldsymbol{\Sigma}_{11i}^{-1}\boldsymbol{\Sigma}_{12i}\right).$$

Then the algorithm proposed by Schafer (1997) is implemented, which organizes the mean vector $\boldsymbol{\mu}$ and the variance-covariance matrix $\boldsymbol{\Sigma}$ into a $(JK + 1) \times (JK + 1)$ parameter matrix $\boldsymbol{\Omega}$,

$$\boldsymbol{\Omega} = \begin{bmatrix} -1 & \boldsymbol{\mu}^T \\ \boldsymbol{\mu} & \boldsymbol{\Sigma} \end{bmatrix}$$

The sweep operation is a function transforming an $M \times M$ matrix $\boldsymbol{\Omega}$ to an $M \times M$ matrix $\boldsymbol{H}$. A sweep on position m, $m = 1, \ldots, M$ is defined as

$$\begin{aligned}
(1) \quad h_{mm} &= -\tfrac{1}{\omega_{mm}} \\
(2) \quad h_{jm} &= h_{mj} = \tfrac{\omega_{jm}}{\omega_{mm}}, \text{ for } j \neq m \\
(3) \quad h_{jl} &= h_{lj} = \omega_{jl} - \tfrac{\omega_{jm}\omega_{lm}}{\omega_{mm}} \text{ for } j \neq m \text{ and } l \neq m.
\end{aligned}$$

For each family, the conditional mean and variance matrix of missing data given observed data can be obtained by sweeping the rows and columns of $\mathbf{\Omega}$ on positions of the observed variables. The missing data can then be simulated from a normal distribution with this mean and variance.

For example, if $\boldsymbol{y}_{i(JK\times 1)}$ is permuted so that the first p_1 elements are observed and the rest $p_2 = JK - p_1$ elements are missing, $\boldsymbol{y}_i^* = (\boldsymbol{y}_{i,o}, \boldsymbol{y}_{i,m})$, then sweeping

$$\mathbf{\Omega} = \begin{bmatrix} -1 & \boldsymbol{\mu}_{i1}^T & \boldsymbol{\mu}_{i2}^T \\ \boldsymbol{\mu}_{i1} & \mathbf{\Sigma}_{11i} & \mathbf{\Sigma}_{12i} \\ \boldsymbol{\mu}_{i2} & \mathbf{\Sigma}_{12i}^T & \mathbf{\Sigma}_{22i} \end{bmatrix}$$

on positions of observed values, $2, \ldots, p_1 + 1$, would yield

$$\begin{bmatrix} -1 - \boldsymbol{\mu}_{i1}^T \mathbf{\Sigma}_{11i}^{-1} \boldsymbol{\mu}_{i1} & \boldsymbol{\mu}_{i1}^T \mathbf{\Sigma}_{11i}^{-1} & (\boldsymbol{\mu}_{i2} - \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1}\mathbf{\Sigma}_{12i})^T \\ (\boldsymbol{\mu}_{i1}^T \mathbf{\Sigma}_{11i}^{-1})^T & \mathbf{\Sigma}_{11i}^{-1} & (\mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1})^T \\ \boldsymbol{\mu}_{i2} - \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1}\mathbf{\Sigma}_{12i} & \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1} & \mathbf{\Sigma}_{22i} - \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1}\mathbf{\Sigma}_{12i} \end{bmatrix},$$

where

$$\mathrm{E}(\boldsymbol{y}_{i,o}|\boldsymbol{\mu}_i, \mathbf{\Sigma}, \boldsymbol{y}_{i,m} = \boldsymbol{y}_{i,1}) = \boldsymbol{\mu}_{i2} - \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1}\mathbf{\Sigma}_{12i} + \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1}\boldsymbol{y}_{i,1}$$

and

$$\mathrm{var}(\boldsymbol{y}_{i,o}|\boldsymbol{\mu}_i, \mathbf{\Sigma}, \boldsymbol{y}_{i,m} = \boldsymbol{y}_{i,1}) = \mathbf{\Sigma}_{22i} - \mathbf{\Sigma}_{21i}\mathbf{\Sigma}_{11i}^{-1}\mathbf{\Sigma}_{12i}.$$

Unlike Schafer (1997) which assumed equal means for all families, for this analysis the mean, $\boldsymbol{\mu}_i = \boldsymbol{X}_i\boldsymbol{\beta}$ can be different for different families. Therefore, the observed data are not grouped by the missing pattern and for each new i it start from $\mathbf{\Omega}$.

2.4.2 Regression Coefficients, $\boldsymbol{\beta}$

As the prior distribution of the regression coefficients, $\boldsymbol{\beta}$, is multivariate normal

$$\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\mu}_{\beta 0}, \boldsymbol{\Sigma}_{\beta 0}),$$

where $\boldsymbol{\mu}_{\beta 0} = (\beta_{01}, \ldots, \beta_{0P})^T$ and $\boldsymbol{\Sigma}_{\beta 0} = \text{diag}(\sigma^2_{\beta 01}, \ldots, \sigma^2_{\beta 0P})$, conditional on all data $\boldsymbol{Y}$ and variance matrix $\boldsymbol{\Sigma}$, the posterior distribution of $\boldsymbol{\beta}$ is also multivariate normal

$$(\boldsymbol{\beta}|\boldsymbol{\Sigma}, \boldsymbol{Y}) \sim \mathcal{N}(\boldsymbol{\beta}_p, \boldsymbol{\Sigma}_{\beta p}),$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_{\beta p} &= (\boldsymbol{\Sigma}^{-1}_{\beta 0} + \sum_{i=1}^{N} \boldsymbol{X}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{X}_i)^{-1}, \\ \boldsymbol{\beta}_p &= \boldsymbol{\Sigma}_{\beta p}(\boldsymbol{\Sigma}^{-1}_{\beta 0} \boldsymbol{\mu}_{\beta 0} + \sum_{i=1}^{N} \boldsymbol{X}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y}_i). \end{aligned} \tag{2.3}$$

2.4.3 Factor Loading Matrices, $\boldsymbol{\Lambda}_A$ and $\boldsymbol{\Lambda}_B$

Factor loadings $\boldsymbol{\Lambda}_A$ and $\boldsymbol{\Lambda}_B$ represent the effects of factors $\boldsymbol{f}_{Ai}$ and $\boldsymbol{f}_{Bi}$ for predicting observed variables, $\boldsymbol{y}_i$. Define the N-vector of the k^{th} outcome on the j^{th} family member for all N families as $\boldsymbol{y}_{jk(N\times 1)} = (y_{1jk}, \ldots, y_{Njk})$, the family member factor loading scores of the j^{th} member for all N families as $\boldsymbol{f}_{Aj(N\times 1)} = (f_{A1j}, \ldots, f_{ANj})^T$, the outcome factor loading scores of the k^{th} outcome for all N families as $\boldsymbol{f}_{Bk(N\times 1)} = (f_{B1k}, \ldots, f_{BNk})^T$, and the covariates for the k^{th} outcome and the j^{th} family member for family i as $\boldsymbol{x}_{ijk(P\times 1)} = (x_{1ijk}, \ldots, x_{Pijk})^T$.

For the family member factor loadings, when $k = 1$, $a_{jk} \equiv 1$ for $j = 1, \ldots, J$; when $k \neq 1$, the conditional posterior distribution of a non-zero element, a_{jk}, in the family member factor loading matrix, $\boldsymbol{\Lambda}_A$, is normal

$$(a_{jk}|\boldsymbol{f}_{Aj}, \psi_k, \boldsymbol{\beta}, b_{jk}, \boldsymbol{y}_{jk}) \sim \mathcal{N}(\mu_{a_{jk}p}, \sigma^2_{a_{jk}p})$$

with

$$\begin{aligned}\sigma^2_{a_{jkp}} &= \left(\frac{1}{\sigma^2_{a_{jk}}} + \frac{\boldsymbol{f}^T_{Aj}\boldsymbol{f}_{Aj}}{\psi_k}\right)^{-1}, \\ \mu_{a_{jkp}} &= \sigma^2_{a_{jkp}}\left[\frac{\mu_{a_{jk}}}{\sigma^2_{a_{jk}}} + \frac{1}{\psi_k}\boldsymbol{f}^T_{Aj}(\boldsymbol{y}_{jk} - \boldsymbol{x}_{ijk}\boldsymbol{\beta} - b_{jk}\boldsymbol{f}_{Bk})\right].\end{aligned}$$

For the outcome factor loadings, when $j = 1$, $b_{jk} \equiv 1$ for $k = 1, \ldots, K$; when $j \neq 1$, the conditional posterior density of a non-zero elements, b_{jk}, in the outcome factor loading matrix, $\boldsymbol{\Lambda}_B$, is normal

$$(b_{jk}|\boldsymbol{f}_{Bk}, \psi_k, \boldsymbol{\beta}, a_{jk}, \boldsymbol{y}_{jk}) \sim \mathcal{N}(\mu_{b_{jkp}}, \sigma^2_{b_{jkp}}),$$

where

$$\begin{aligned}\sigma^2_{b_{jkp}} &= \left(\frac{1}{\sigma^2_{b_{jk}}} + \frac{\boldsymbol{f}^T_{Bk}\boldsymbol{f}_{Bk}}{\psi_k}\right)^{-1}, \\ \mu_{b_{jkp}} &= \sigma^2_{b_{jkp}}\left[\frac{\mu_{b_{jk}}}{\sigma^2_{b_{jk}}} + \frac{1}{\psi_k}\boldsymbol{f}^T_{Bk}(\boldsymbol{y}_{jk} - \boldsymbol{X}\boldsymbol{\beta}_{jk} - a_{jk}\boldsymbol{f}_{Aj})\right].\end{aligned}$$

2.4.4 Unique Error Variances, ψ_k

For the k^{th} outcome measure, denote the complete data on this outcome for all N families as

$$\boldsymbol{Y}_{..k(N\times J)} = \begin{bmatrix} y_{11k} & \ldots & y_{1Jk} \\ \vdots & \ddots & \vdots \\ y_{N1k} & \ldots & y_{NJk} \end{bmatrix},$$

the covariates for this outcome for family i as

$$\boldsymbol{X}_{i.k(J\times P)} = \begin{bmatrix} x_{i1k1} & \ldots & x_{i1kP} \\ \vdots & \ddots & \vdots \\ x_{iJk1} & \ldots & x_{iJkP} \end{bmatrix},$$

and the covariates of this outcome for all N families as

$$\boldsymbol{Z}_{..k(N\times JP)} = [\text{vec}(\boldsymbol{X}_{1.k}), \ldots, \text{vec}(\boldsymbol{X}_{N.k})]^T,$$

the k^{th} outcome factor scores for all N families as

$$\boldsymbol{f}_{Bk(N\times 1)} = (f_{B1k}, \ldots, f_{BNk})^T,$$

the family member factor loading corresponding to the k^{th} outcome as

$$\boldsymbol{A}_{k(J\times J)} = \text{diag}(a_{1k}, \ldots, a_{Jk}),$$

and the non-zero outcome factor loading corresponding to the k^{th} outcome as

$$\boldsymbol{b}_{k(J\times 1)} = (b_{1k}, \ldots, b_{Jk})^T.$$

Given the complete data only depends on $\boldsymbol{Y}_{..k}$ and also given the other parameters, $\boldsymbol{F}_A$, $\boldsymbol{f}_{Bk}$, $\boldsymbol{A}_k$, $\boldsymbol{b}_k$, the conditional posterior distribution of ψ_k is inverse-gamma

$$(\psi_k|\boldsymbol{F}_A, \boldsymbol{f}_{Bk}, \boldsymbol{A}_k, \boldsymbol{b}_k, \boldsymbol{Y}_{..k}) \sim \mathcal{IG}\left(\frac{\alpha_{\psi_{kp}}}{2}, \frac{\beta_{\psi_{kp}}}{2}\right),$$

where

$$\begin{aligned} \alpha_{\psi_{kp}} &= \alpha_{\psi_k} + NJ \\ \text{and} \quad \beta_{\psi_{kp}} &= \beta_{\psi_k} + \text{tr}[(\boldsymbol{Y}_{..k} - \boldsymbol{X}\boldsymbol{B}_{.k} - \boldsymbol{F}_A\boldsymbol{A}_k^T - \boldsymbol{f}_{Bk}\boldsymbol{b}_k^T)^T \\ &\qquad (\boldsymbol{Y}_{..k} - \boldsymbol{X}\boldsymbol{B}_{.k} - \boldsymbol{F}_A\boldsymbol{A}_k^T - \boldsymbol{f}_{Bk}\boldsymbol{b}_k^T)]. \end{aligned} \tag{2.4}$$

2.4.5 Factor Variance Matrices, $\boldsymbol{\Phi}_A$ and $\boldsymbol{\Phi}_B$

Conditional on the data and $\boldsymbol{F}_{A(N\times J)} = (\boldsymbol{f}_{A1}^T, \ldots, \boldsymbol{f}_{AN}^T)^T$, the posterior distribution of $\boldsymbol{\Phi}_{A(J\times J)}$ is an inverse Wishart distribution

$$(\boldsymbol{\Phi}_A|\cdot) \sim \mathcal{IW}(\boldsymbol{W}_A + \boldsymbol{F}_A\boldsymbol{F}_A^T, \nu_A + N).$$

Similarly, conditional on the data and $\boldsymbol{F}_{B(K\times K)} = (\boldsymbol{f}_{B1}^T, \ldots, \boldsymbol{f}_{BN}^T)^T$, the posterior distribution of $\boldsymbol{\Phi}_{B(K\times K)}$ is an inverse Wishart distribution

$$(\boldsymbol{\Phi}_B|\cdot) \sim \mathcal{IW}(\boldsymbol{W}_B + \boldsymbol{F}_B\boldsymbol{F}_B^T, \nu_B + N).$$

2.4.6 Factor Scores, $\boldsymbol{f}_{Ai}$ and $\boldsymbol{f}_{Bi}$

Conditional on complete data $\boldsymbol{Y}$ and all other parameters, $\boldsymbol{\Omega}_A = (\boldsymbol{\Phi}_A, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Psi}, \boldsymbol{\beta}, \boldsymbol{f}_{Bi})$, the posterior distribution of $\boldsymbol{f}_{Ai}$ is a multivariate normal distribution

$$\boldsymbol{f}_{Ai}|\boldsymbol{\Omega}_A, \boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}_{fAi}, \boldsymbol{\Sigma}_{fA})$$

$$\begin{array}{lll} \text{where} & \boldsymbol{\Sigma}_{fA(J\times J)} & = (\boldsymbol{\Phi}_A^{-1} + \boldsymbol{\Lambda}_A^T\boldsymbol{\Psi}^{-1}\boldsymbol{\Lambda}_A)^{-1} \\ \text{and} & \boldsymbol{\mu}_{fAi(J\times 1)} & = \boldsymbol{\Sigma}_{fA}\boldsymbol{\Lambda}_A^T\boldsymbol{\Psi}^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{\Lambda}_B\boldsymbol{f}_{Bi}). \end{array}$$

Similarly, Conditional on complete data $\boldsymbol{Y}$ and all other parameters, $\boldsymbol{\Omega}_B = \boldsymbol{\Phi}_B, \boldsymbol{\Lambda}_B, \boldsymbol{\Lambda}_A, \boldsymbol{\Psi}, \boldsymbol{\beta}, \boldsymbol{f}_{Ai})$, the posterior distribution of $\boldsymbol{f}_{Ai}$ is a multivariate normal distribution

$$\boldsymbol{f}_{Bi}|\boldsymbol{\Omega}_B, \boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}_{fBi}, \boldsymbol{\Sigma}_{fB})$$

$$\begin{array}{lll} \text{where} & \boldsymbol{\Sigma}_{fB(K\times K)} & = (\boldsymbol{\Phi}_B^{-1} + \boldsymbol{\Lambda}_B^T\boldsymbol{\Psi}^{-1}\boldsymbol{\Lambda}_B)^{-1} \\ \text{and} & \boldsymbol{\mu}_{fBi(J\times 1)} & = \boldsymbol{\Sigma}_{fB}\boldsymbol{\Lambda}_B^T\boldsymbol{\Psi}^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{\Lambda}_A\boldsymbol{f}_{Ai}). \end{array}$$

2.5 A Gibbs Sampling Algorithm

To reduce autocorrelation and improve efficiency, I use a blocked Gibbs sampler to sample the regression coefficients, $\boldsymbol{\beta}$ and the factor scores, $\boldsymbol{f}_i$ from their joint conditional distributions. Denote all missing observations as $\boldsymbol{Y}_{miss}$, denote the matrices of all family member factor scores and all outcome factor scores as $\boldsymbol{F}_{A(N\times J)} = [\boldsymbol{f}_{A1}, \ldots, \boldsymbol{f}_{AN}]^T$ and $\boldsymbol{F}_{B(N\times K)} = [\boldsymbol{f}_{B1}, \ldots, \boldsymbol{f}_{BN}]^T$, then MCMC samples from the joint posterior density $p(\boldsymbol{Y}_{miss}, \boldsymbol{F}_A, \boldsymbol{F}_B, \boldsymbol{\beta}, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Phi}_A, \boldsymbol{\Phi}_B, \boldsymbol{\Psi}_\varepsilon|\boldsymbol{Y}_{obs})$ proceeds as

follows: At the $(l+1)^{th}$ iteration with current values of $(\boldsymbol{Y}_{miss}^{(l)}, \boldsymbol{F}_A^{(l)}, \boldsymbol{F}_B^{(l)}, \boldsymbol{\beta}^{(l)}, \boldsymbol{\Lambda}_A^{(l)}, \boldsymbol{\Lambda}_B^{(l)}, \boldsymbol{\Phi}_A^{(l)}, \boldsymbol{\Phi}_B^{(l)}, \boldsymbol{\Psi}_\varepsilon^{(l)})$,

1. Simulate $\boldsymbol{Y}_{miss}^{(l+1)}$ from
$$p(\boldsymbol{Y}_{miss}|\boldsymbol{\Sigma}^{(l)}, \boldsymbol{\beta}^{(l)}, \boldsymbol{Y}_{Obs}),$$
where $\boldsymbol{\Sigma}^{(l)} = \boldsymbol{\Lambda}^{(l)}\boldsymbol{\Phi}^{(l)}\boldsymbol{\Lambda}^{(l)T} + \boldsymbol{\Psi}^{(l)}$;
2. Simulate $\boldsymbol{\beta}^{(l+1)}$ from $p(\boldsymbol{\beta}|\boldsymbol{\Sigma}^{(l)}, \boldsymbol{Y}_{miss}^{(l+1)}, \boldsymbol{Y}_{obs})$;
3. Simulate $\boldsymbol{F}_A^{(l+1)}$ from $p(\boldsymbol{F}_A|\boldsymbol{\Phi}_A^{(l)}, \boldsymbol{\Lambda}^{(l)}, \boldsymbol{\Psi}^{(l)}, \boldsymbol{\beta}^{(l+1)}, \boldsymbol{Y}_{miss}^{(l+1)}, \boldsymbol{Y}_{obs})$;
4. Simulate $\boldsymbol{F}_B^{(l+1)}$ from $p(\boldsymbol{F}_B|\boldsymbol{\Phi}_B^{(l)}, \boldsymbol{\Lambda}^{(l)}, \boldsymbol{\Psi}^{(l)}, \boldsymbol{B}^{(l+1)}, \boldsymbol{Y}_{miss}^{(l+1)}, \boldsymbol{Y}_{obs})$;
5. Simulate $\boldsymbol{\Phi}_A^{(l+1)}$ from $p(\boldsymbol{\Phi}_A|\boldsymbol{F}_A^{(l+1)})$;
6. Simulate $\boldsymbol{\Phi}_B^{(l+1)}$ from $p(\boldsymbol{\Phi}_B|\boldsymbol{F}_B^{(l+1)})$;
7. Simulate $\boldsymbol{\Lambda}^{(l+1)}$ from $p(\boldsymbol{\Lambda}|\boldsymbol{\Psi}^{(l)}, \boldsymbol{F}^{(l+1)}, \boldsymbol{\beta}^{(l+1)}, \boldsymbol{Y}_{miss}^{(l+1)}, \boldsymbol{Y}_{obs})$;
8. Simulate $\boldsymbol{\Psi}^{(l+1)}$ from $p(\boldsymbol{\Psi}|\boldsymbol{F}^{(l+1)}, \boldsymbol{\beta}^{(l+1)}, \boldsymbol{\Lambda}^{(l+1)}, \boldsymbol{Y}_{miss}^{(l+1)}, \boldsymbol{Y}_{obs})$.

CHAPTER 3

Analysis of Simulated Data

To assess the performance of the Bayesian Family Factor Model (BFFM) in different scenarios, simulation studies are used to compare BFFM with CFA estimated by full information maximum likelihood (FIML) using quasi-Newton optimization (QNO), on the basis of ability to fit the data, as well as examining mean squared errors (MSE), squared biases and variances of parameters estimated by the two methods.

3.1 Generating Data Sets for Simulation Studies

Grayson and Marsh (1994) proved that a CFA model is not identified when the true factor loading matrix, $\mathbf{\Lambda}$, is not full rank. One sufficient condition for deficient column rank is $\mathbf{\Lambda} = [\boldsymbol{C} \otimes \boldsymbol{a}_0 | \boldsymbol{d} \otimes \boldsymbol{B}_0]$, where $\boldsymbol{C}_{(J\times J)}$ and $\boldsymbol{B}_{0(K\times K)}$ are diagonal full rank matrices, and $\boldsymbol{a}_0$ and $\boldsymbol{d}$ are $K \times 1$ and $J \times 1$ vectors, respectively (Grayson and Marsh, 1994). I generated 3 scenarios where the true covariance matrices are identified, close to not identified and not identified, by specifying different true factor loading matrices, $\mathbf{\Lambda}$, which were far from equal to, almost equal to, and equal to $[\boldsymbol{C} \otimes \boldsymbol{a}_0 | \boldsymbol{d} \otimes \boldsymbol{B}_0]$.

Two hundred data sets were simulated from each scenario. Each data set has $N = 200$ families, $K = 5$ outcomes and $J = 4$ members: proband, sibling, father and mother. True regression coefficients as well as true unique error variances are set to be equal for 3 scenarios. True parameters are specified as follows: unique error variances of the same outcomes are assumed equal across family members, so K distinct unique error variances are $\phi_k = \phi_{1k} = \ldots = \phi_{Jk}$, for $k = 1, \ldots, K$. For the UCLA NFS, I am

interest in differences of mean outcome measures between control and schizophrenia (SZ) families and across family members, so the matrix of covariates is specified as

$$\boldsymbol{X}_{i(JK\times 2JK)} = \left[\begin{array}{cc} d_i \boldsymbol{I}_{JK} & (1-d_i)\boldsymbol{I}_{JK} \end{array} \right],$$

where $d_i = 0, 1$ for control and SZ families respectively, and $\boldsymbol{I}_{JK}$ is a $JK \times JK$ identity matrix. The corresponding regression coefficient vector is $\boldsymbol{\beta}_{(2JK\times 1)} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)^T$, where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are $JK \times 1$ vectors of means of all K outcomes on the J family members in the control and SZ families. For $J = 4$ and $K = 5$, the total number of parameters is 101. The observations are set to be missing completely at random with probability $p = 0.15$ and the missingness pattern is the same across all 200 data sets in each scenario. As psychological measures usually have different ranges and scales, I assumed that there were different scales associated with different outcomes, and the ratio was $1 : 2 : 5 : 8 : 10$. Family member factor loadings for for different outcomes had about the same ratio, with some random variation added in. The ratio of true error variances for different outcomes and the ratio of outcome factor variances were also $1 : 4 : 25 : 64 : 100$. The variance-covariance matrix of family member factors were chosen to be close to 1. True values for all true parameters are listed in Tables A.1, A.2 and A.3 in the Appendix.

3.2 Comparing BFFM and CFA: Producing Valid Solutions

Standard non-Bayesian CFA models are fit to simulated data using the lavaan package in R (Rosseel, 2012), which uses full information maximum likelihood (FIML) estimation to handle missing data and uses a quasi-Newton optimization algorithm to estimate parameters. FIML estimation maximizes the likelihood function for each family based on the observed variables y_{ijk} that are not missing so that all the available data are used. Full information maximum likelihood estimation with quasi-Newton optimization (FIML-QNO) is defined as successful in fitting the data if the algorithm converges

and provides valid solutions (e.g. having positive-definite covariance matrices and positive variances). In many cases, FIML-QNO fails to find a fit to the data due to empirical under-identification. The percentages of data sets for which FIML-QNO was successful in fitting in the 3 scenarios are 85%, 57% and 13%, respectively (Figure 3.1). When CFA model using FIML-QNO was fit to the same simulated data but with no missing observation, the percentages increase slightly to 21%, 50.5% and 92.5%, respectively, suggesting that the missing data was not the major cause of the failure of FIML-QNO.

Next, a BFFM is fit to 200 data sets in each scenario, with $10,000$ iterations after an initial burn-in of 1000 iterations. Priors are chosen to be partially informative and centered at true values with large dispersions. The trace plots, density plots and autocorrelation plots show no obvious evidence of bad mixing, non-convergence or high autocorrelations. BFFM successfully fit all 600 data sets and the resulting posterior means were always valid solutions (i.e. positive variances and positive definite covariance matrices).

3.3 Comparing the Performance of BFFM and FIML-QNO When FIML-QNO Was Successful

Besides the ability to fit data, I also want to compare the performance characteristics of BFFM and FIML-QNO, when FIML-QNO was successful in fitting the data sets. The mean squared error (MSE) of an estimator $\widehat{\theta}$ for a parameter θ, $\text{MSE}_{\widehat{\theta}} = \text{E}(\widehat{\theta}-\theta)^2$, measures the average squared distance between the estimator $\widehat{\theta}$ and the true parameter value θ. The MSE can be decomposed as the sum of the variance of the estimator, $\text{Var}(\widehat{\theta})$, which measures the uncertainty of $\widehat{\theta}$, and the squared bias, $[\text{E}(\widehat{\theta}) - \theta]^2$, which measures accuracy. Denote $\widehat{\theta}_l$ as the posterior mean of θ from the MCMC outputs of the l^{th} data set, for $l = 1, \ldots, 200$, then $\widehat{\text{var}}(\widehat{\theta}) = \frac{1}{200}\sum_{l=1}^{200}\left(\widehat{\theta}_l - \frac{1}{200}\sum_{l=1}^{200}\widehat{\theta}_l\right)^2$. The relative

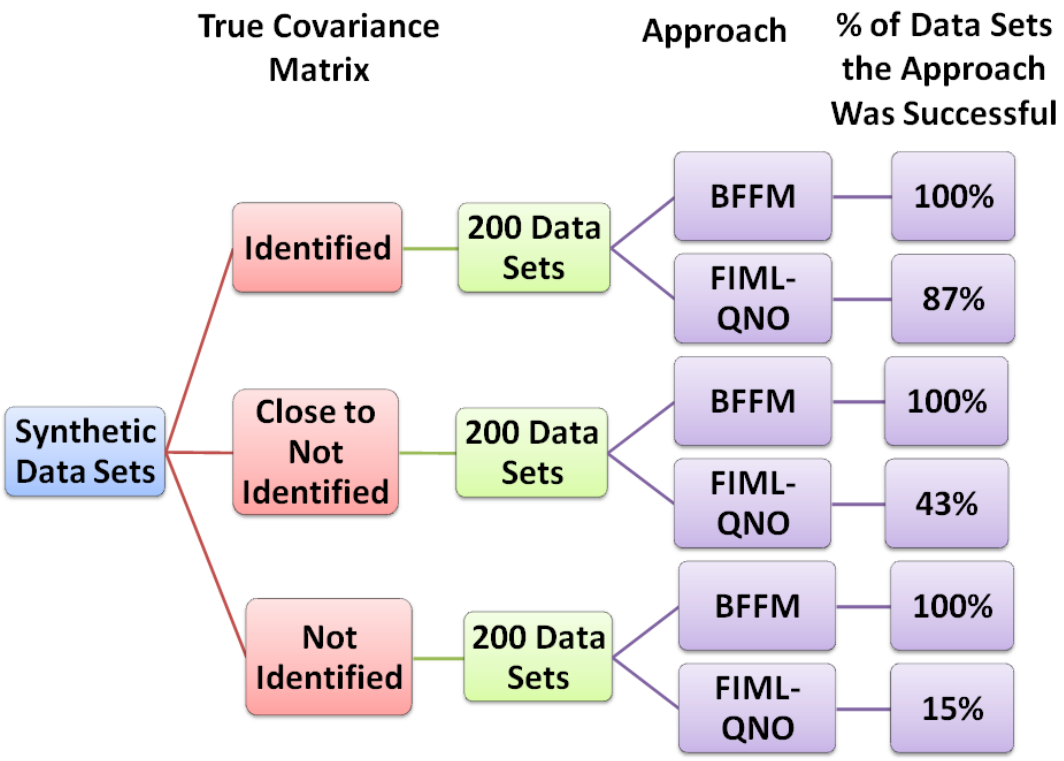


Figure 3.1: The percent of data sets for which Full Information Maximum Likelihood (FIML) estimation with quasi-Newton optimization (QNO) converges and gives valid estimates, in scenarios where the true covariance matrices are identified, close to under-identified and under-identified.

MSE, relative variance and relative squared bias are estimated as $\frac{1}{200}\sum_{l=1}^{200}\left[(\widehat{\theta}_l-\theta)/\theta\right]^2$, $\widehat{\text{var}}(\widehat{\theta})/\theta^2$ and $\left[\frac{1}{200}\sum_{l=1}^{200}\widehat{\theta}_l-\theta\right]^2$, respectively.

In the scenario where the true covariance matrix is close to not identified, I compare relative mean squared errors (RMSE), relative variances and relative squared biases of all parameters estimated by fitting BFFM and FIML-QNO to the 43% of the data sets which FIML-QNO was successful in fitting (Tables B.7, B.2, B.9, B.10, B.11 and B.12 in Appendix B). Overall, parameter estimates from BFFM and FIML-QNO are similar and are close to the true values. Figure 3.2(a) plots on a log-log scale the relative

MSEs of parameters estimated by BFFM vs. those of parameters estimated by FIML-QNO. There are 101 dots representing all parameters. Different symbols represent different groups of parameters (factor loadings, factor variance-covariance parameters, regression coefficients and unique error variances). For a given parameter, if RMSEs estimated by two models are the same, the dot will lie on a diagonal line with slope 1; when the RMSE estimated by BFFM is smaller, the dot will lie above the diagonal line; and when the RMSE estimated by FIML-QNO is smaller, the dot will lie below the diagonal line. For more than 60% of the parameters, the RMSEs estimated by BFFM are smaller. For most parameters, the RMSEs estimated by both methods are small (RMSE < 0.1, dots in the lower left corner). However, for some factor loadings and factor variance-covariances, the RMSEs estimated by FIML-QNO are much larger than those estimated by BFFM (dots in the upper half).

Figure 3.2(b) plots relative variances of parameters estimated by BFFM vs. those estimated by FIML-QNO. Almost all dots lie above the diagonal line, where BFFM has smaller relative variances for almost all parameters. Similarly, Figure 3.2(c) plots the relative squared biases ($[\mathrm{E}(\widehat{\theta}) - \theta]^2/\theta^2$) of parameters estimate by BFFM and FIML-QNO. For about 40% of the parameters, the relative squared biases estimated by BFFM are smaller, but the FIML-QNO has smaller relative squared biases when both methods perform well (relative squared biases < 0.1). However, as with the relative MSEs, for some factor variance-covariances and factor loadings, the squared biases estimated by FIML-QNO are much larger than those estimated by BFFM.

It is important to check whether BFFM will also perform worse when the FIML-QNO failed. In the scenario where the true covariance matrix is close to not identified, I compare the MSEs, squared biases and variances for BFFM on the 43% of the data sets for which FIML-QNO did not fail to those on the 57% data sets for which FIML-QNO failed to converge or provide admissible solutions. The plot of the relative MSEs in Figure 3.2(d) shows that almost all dots are close to the diagonal line with slope 1, indicating that BFFM works equally well for both kinds of data sets. The plots of

relative squared biases and relative variances are very similar.

The relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in 3 scenarios where the true covariance matrix is identified, close to not identified and not identified are summarized in tables in Appendix B).

In summary, simulation studies show that FIML-QNO failed to fit the CFA model to the data in many cases, especially when the true covariance matrix is not identified or close to not identified, due to non-convergence or invalid solutions, while BFFM fit all 600 data sets in the 3 scenarios and gives estimates of similar consistency. When FIML-QNO is successful, variances estimated by BFFM are smaller for almost all parameters and is competitive in MSE. Although FIML-QNO produces smaller squared biases in some cases, the MSEs and variances from some parameters are very large, suggesting these estimates are unstable. The BFFM is overall superior, providing stability and much broader applicability, in exchange for (in some cases) a small bias.

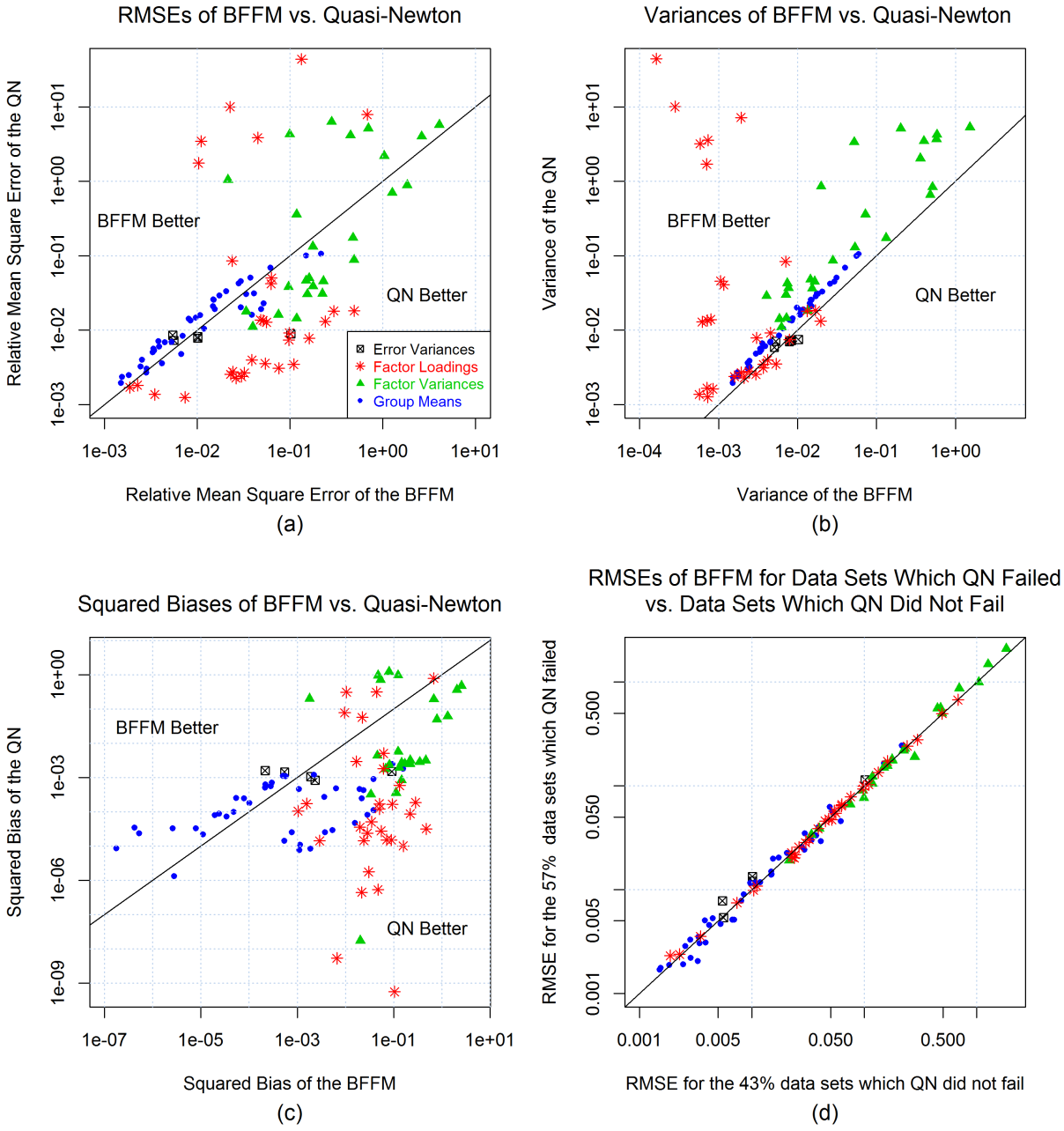


Figure 3.2: Plots of relative mean squared errors (RMSE, a), relative variances (b) and relative squared biases (c) for parameters estimated by BFFM against those estimated by FIML-QNO, and plot of the relative mean squared errors by BFFM for the 43% of the data sets which FIML-QN failed vs for the 57% of the data sets which FIML-QN was successful (d), in the scenario where the true covariance matrix is close to not identified, on a log-log scale.

3.4 Impact of Missing Data on Failure of FIML-QNO

In the previous section, I described simulation studies for comparing BFFM and FIML-QNO using data sets generated under 3 scenarios with different degrees of identification problem. Here I want to further identify potential causes for the failure of FIML-QNO to fit some data sets. I examine whether the missing data contributes to non-convergence and invalid estimates of FIML-QNO. Missing data in the previous analyses were generated by randomly setting observations of the 200 data sets in each scenario to missing, at a target missing rate $r = 15\%$. In this analysis, I examine the ability FIML-QNO to fit the corresponding complete data sets. Table 3.1 presents the percentages of complete and 15%-missing data sets which FIML-QNO is successful in fitting, in 3 scenariors. For the identified, close to under-identified and under-identified scenarios, the percentages of data sets for which FIML-QNO did not fail increases from $87\%, 41\%$ and 14% for data with 15% missing, to $92.5\%, 50.5\%$ and 21% for complete data, respectively, suggesting that missing data contributes partly to the failure of FIML-QNO.

Scenario	Complete Data	15% Missing
Identified	92.5	87
Close to Under-identified	50.5	41
Under-identified	21	14

Table 3.1: Percent of complete and 15% missing data sets which FILM-QNO was successful to fit. The data sets with missing are generated from complete data sets by setting observations to missing at $p = 0.15$.

To further evaluate effects of missing data on the ability of FIML-QNO to fit data sets, I generated new data sets by setting 5%, 30% and 40% of the observations to missing for the 200 data sets, in the scenario where the true covariance matrix is close to not identified. The percent of data sets which FIML-QNO is successful in fitting are compared in Table 3.2 and plotted in Figure 3.3. As missingness percentage increases

from 0% to 40%, the percentages of data sets which FIML-QNO can fit decreases from 52.5% to 16.5%, while the proportions of data sets with non-convergence and invalid estimates increase.

Missing Rate	0%	5%	15%	30%	40%
Converge	52.5	48	41.5	26	16.5
Invalid	34.5	38	38	51.5	56.5
Not converge	13	14	20.5	22.5	27

Table 3.2: Percent of data sets which FIML-QNO did not fail, in the close to Under--identified scenario. Invalid, not converge and converge refer to the situations where FIML-QNO gives invalid estimates, fails to converge, and neither of the above.

Figure 3.3: Percent of data sets which FIML-QNO is successful in fitting, in the scenario where the true covariance matrix is close to not identified. Invalid Est, non–convergence and convergence refer to the situations where FIML-QNO gives invalid estimates, fails to converge, and neither of the above.

CHAPTER 4

Application of the Bayesian Family Factor Model to the UCLA Family Study Data with Five Outcomes

In this Chapter, I illustrate the Bayesian Family Factor Model (BFFM) by analyzing the data on $K = 5$ primary outcome variables from the UCLA Neurocognitive Family Study (NFS). I compute descriptive statistics, elicit prior hyper-parameters, implement the Gibbs sampling algorithm using R and summarize the posterior distributions.

These five measures are analyzed because they are the most representative measures of each major cognitive domain of interest, and have been successfully used in assessing schizophrenia related cognitive deficits. Furthermore, it is reasonable to start constructing the model with a smaller number of outcomes, so that the algorithm runs faster and interpretation of results is easier, as the total number of parameters is smaller. This data structure with $K = 5$ outcomes and $J = 4$ family member types was used as a template for designing simulation studies in Section 3 to test the algorithm and to assess the performance of the model for all parameters.

4.1 The UCLA Neurocognitive Family Study Data

The UCLA Neurocognitive Family Study (NFS) is a cross-sectional case-control study collecting multiple outcomes on schizophrenia subjects and their relatives, as well as community control subjects and their relatives. There are two parallel studies, one for adult onset and one for childhood onset, for which the data were collected by Dr. K.H. Nuechterlein and Dr. R.F. Asarnow. The $J = 4$ family member types, proband, sibling,

father and mother are indexed by $j = 1, \ldots, 4$, respectively. Table 4.1 presents presents counts of families and individuals. There are a total of 210 families and 635 subjects in the study, about half in the adult onset arm and half in the childhood onset arm. The number of schizophrenia probands and their relatives are roughly the same as the number of community control probands and families. Fifty two percent of the subjects are male. To study most cognitive measures, age is a critical factor. The mean age for all subjects is 33 (Std Dev $= 7$), with a minimum of 7 and a maximum of 85. The average age of subjects in the adult onset group is older than that of the childhood onset group, as expected. For now, age is not included in the current analyses as I want to keep the model as simple as possible to begin with. In future analyses, I can include age in the mean structure to adjust for age effects.

The various batteries of cognitive tests assessed include the Wechsler Abbreviated Scale of Intelligence (WASI), the Test of Memory and Learning (TOMAL), the digital span subset from Wechsler Memory Scale (WMS-III), the Maintenance and Manipulation Test (MNM), the Minnesota Multiphasic Personality Inventory (MMPI), the California Verbal Learning Test C Children's Version (CVLT-C) and the California Verbal Learning Test C Second Edition (CVLTCII). Figure 4.1 organized the major neurocognitive outcome measures by these tests.

The seven cognitive tests of primary interest are described in details below.

1. Memory-Load Continuous Performance Test (3-7 CPT):

   In the conventional continuous performance test (CPT), a random series of single numbers or letters are presented on a computer monitor. Subjects are asked to indicate that they have detected a target event by pressing a response button and to avoid responding to distracting stimuli. Outcome measures of this test include the level of signal/noise discrimination, d prime (CPT37D), the hit rate (Hitr37) and the false alarm rate (Falr37).

2. Degraded Stimulus Continuous Performance Test (DS CPT):

In this version of the CPT, the image of numerals presented to the subject are degraded, that is, the numerals appear extremely blurred and indistinct. Similar to 3-7 CPT, subjects are asked to indicate that they have detected a target event. Outcome variables of this test include the level of signal/noise discrimination, d prime (CPTDSD), the hit rate (HitrDS) and the false alarm rate (FalrDS);

3. Forced choice Span of Apprehension (SPAN):

   In this test, either a T or F will be flashed briefly on the computer screen along with other irrelevant letters in an array of 1, 5, and l0 letters. The subjects were instructed to press one button when a T was present and another button when an F was present. The primary dependent variables are the number of correct target detections for 1-letter, 5-letter, and l0-letter arrays (SPAN1, SPAN5 and SPAN10, repectively).

4. Trail making test (TRAILS):

   The Trail Making Test from the Halstead-Reitan Neuropsychological Battery (Springate and Fein, 2013) requires subjects to connect numbers (1-25) in part A or alternating numbers (1-13) and letters (A-L) in part B (i.e., 1-A-2-B-3-C, etc.) in sequence as rapidly as possible. The subject's scores are the number of seconds required to complete Part A (logTRLAA) and Part B (logTRLBA).

5. Facial Recognition:

   The Benton test of Facial Recognition (BFRT), which consists of a short form requiring 27 responses and a long form requiring 54 responses. On each item, subjects are presented with a target face above six test faces, and they are asked to indicate which of the six images match the target face (Benton, 1994). An outcome measure of interest is short form score (NCFRSFSC).

6. Verbal Fluency:

   In the Controlled Oral Word Association Test for verbal fluency, participants were asked to generate as many words as possible beginning with the letters "F," "A,"

and "S", each for 60 seconds. The combined score of "F," "A," and "S" (VFFAS) is the outcome variable of interest.

7. Maintenance and Manipulation (MNM):

   In the Maintenance and Manipulation test, an array of 4 objects first appeared on the computer screen for 2 seconds. In the maintenance only condition, subjects were then asked to decide whether the new array was the same as the previous one. In the maintenance plus manipulation conditions, there was a delay period when the subjects are told to reorganize the array held in memory. Outcome measures of this test include the main (hold) trials mean accuracy (MAINacc) and reaction time (MAINrt), as well as the manipulate (flip) trials accuracy (Manipa) and reaction time (MANIPrt).

The $K = 5$ primary outcomes analyzed in this Chapter are Maintenance and Manipulation Test (MnM Test) manipulation accuracy, degraded stimulus CPT (DS-CPT) block sum d prime, memory-load CPT (3-7 CPT) block sum d prime, forced-choice Span of Apprehension (Span) 10-letter accuracy and Trail Making Test b time in seconds, corresponding to $k = 1, \ldots, 5$, respectively. In the next chapter, I will include 12 additional variables in the model, develop a way to incorporate knowledge on clustering structure of outcomes into the priors and compare consistency of posterior across the five- and seventeen-outcome models.

It is desirable for the ease of interpretation to have higher scores mean better test performance for all outcomes. Therefore, the sign of Trail Making Test b has been reversed. In addition, the factor variance for probands, ϕ_{A11}, must be smaller than any of the overall variances for probands, var(y_{i1k}), as described in Section 2.1, therefore, it is useful to scale the observed variables to make the overall variances similar, so that ϕ_{A11} will not be forced to be small, which causes precision problem, as small values may be rounded to zero in computation. Scaling and transformation of these five outcomes are described in Table 4.2.

Variable	Value	Control ($n = 321$)	Sz ($n = 314$)	Total ($n = 635$)
	Proband	84	116	200
Family	Sibling	115	79	194
member	Father	45	38	83
	Mother	77	81	158
Gender	Female	163	152	315
	Male	158	162	320

Table 4.1: Frequency tables of family member and gender by schizophrenia and control families.

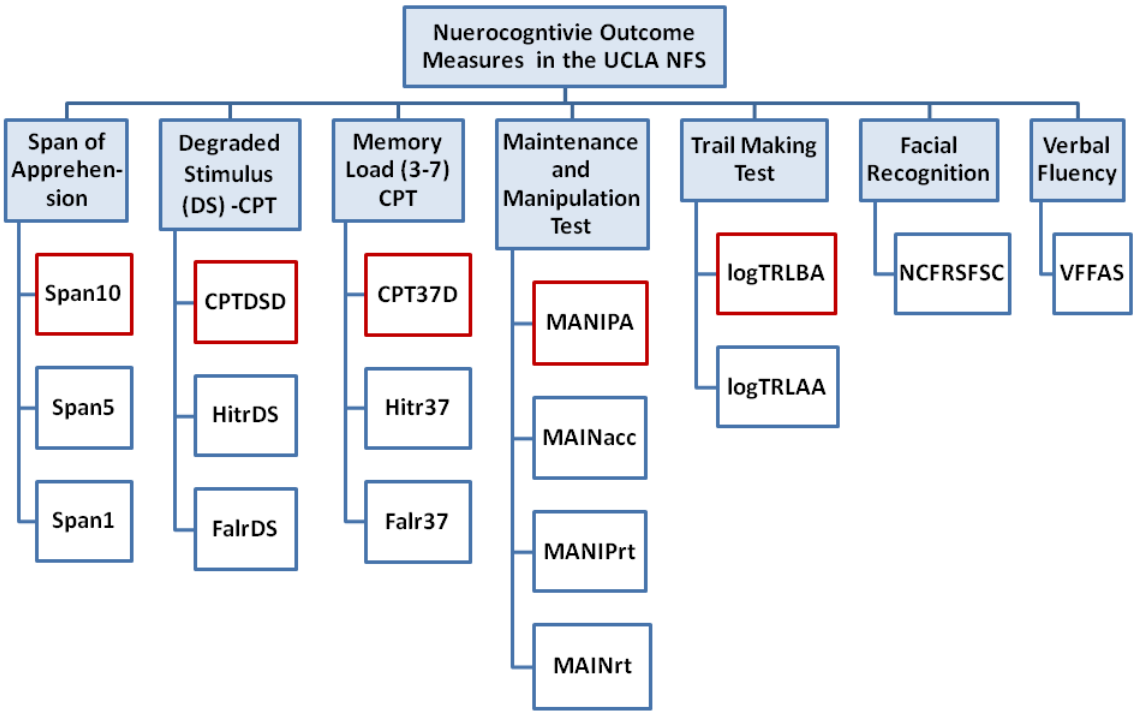


Figure 4.1: Neurocognitive performance measures collected in the UCLA Neurocognitive Family Study.

k	Variable	Description	Transformation
1	MANIPA	Maintenance and Manipulation (MnM) test accuracy during manipulation	$100 * y$
2	CPTDSD	Degraded Stimulus Continuous Performance Test (DS-CPT) block sum d prime	$10 * y$
3	CPT37D	Memory-load Continuous Performance Test (37-CPT) block sum d prime	$10 * y$
4	SPAN10	the Forced-choice Span of Apprehension test 10-letter accuracy	$100 * y$
5	logTRLBA	The trail making test b time in seconds	$-100 * \log_{10}(y)$

Table 4.2: Variable descriptions and transformations for five neurocognitive measurements of primary interest. The first four outcomes are scaled while the logTRLBA with a skewed distribution and negatively correlated with other outcomes are log-transformed and has its sign reversed.

Descriptive statistics are used to summarize the data. Table C.1 in Appendix C presents the raw group means and standard deviations of the $K = 5$ outcomes measured on probands, siblings, fathers and mothers for the schizophrenia (SZ) and control families. Table C.2 presents the complete correlation matrix of all 20 combinations of 4 family member and 5 outcomes. In particular, the block diagonal matrices are within-family-member across outcome correlations, which range from 0.2 to 0.4. The across-member within-outcome correlations of are the diagonal elements of off-diagonal blocks in Table C.2, which extracted and summarized in Table C.3. For all five outcomes, the correlation between observed variables measured on proband and on sibling is the highest (about 0.2).

4.2 Prior Specification

To fit the BFFM, partially informative priors are specified using priors described in Section 2.2 and methods from Section 2.3. First, estimated values for the overall variances of the K outcomes, $\widehat{\text{var}}(y_1), \ldots, \widehat{\text{var}}(y_K)$ are obtained from a previous study and from the literature. Phase 1 of the UCLA Family Study (Asarnow et al., 2002; Nuechterlein et al., 2002) collected four of these five outcomes (CPTDSD, CPT37D, SPAN10 and logTRLBA); furthermore, these measures have been analyzed in previous studies (Kim et al., 2004; Nuechterlein et al., 2011; Koide et al., 2012). For outcomes for which which phase 1 data are not available, estimates from the literature are used. The estimates of overall means and variances from all those various sources for all 17 outcome measures are summarized in Table 4.3. Summaries of correlations are given in Tables 4.4 and 4.5. For now, only prior information of the 5 primary outcome measures (highlighted in bold) are used.

Using the informtion in Tables 4.3, 4.4 and 4.5, prior hyper-parameters for factor loadings, factor variance matrix, unique error variances and regression coefficients can be specified using the methods described in Section 2.3. What follows illustrate this process by selecting specific values appropriate to this data.

First, information from the Phase 1 study can help to specify prior correlation matrices $\boldsymbol{D}_A$ and $\boldsymbol{D}_B$, for outcome factors and family member factors, respectively. For Phase 1 study data, correlations across four outcomes (CPTDSD, CPT37D, SPAN10 and logTRLBA) ignoring the family structure are all between 0.30 and 0.48 (Table 4.4). Therefore, I choose a compound symmetric correlation structure for $\boldsymbol{D}_B$, by setting all the prior factor correlations to be 0.35. The within-outcome across-family-member correlations in the Phase 1 data are summarized in Table 4.5. Most of these correlations are positive, as observations on individuals from the same family are expected to be positively associated; there are some negative correlations, though none are significantly different from 0. Furthermore, the correlations between observations from all pairs of

family members due to pure genetic effects are all 0.5, except the 0 correlation between father and mother. In practice, due to various kinds of noise and environmental factors outside of the family, the real correlations can be lower than 0.5. Based on the Phase 1 correlations and the theoretical correlations among family members, the prior means of correlations among family members are set to be 0.15 between father and mother and 0.2 otherwise, as father and mother are not genetically associated.

For prior specification of factor variances, it is necessary to choose the degree of freedom parameterfor the inverse Wishart priors, ν_A and ν_A, which are inversely proportional to dispersions of the factor variance matrices. The priors are less informative when ν_A and ν_B are smaller. Furthermore, it is necessary to have $\nu_A > J + 1$ and $\nu_B > K + 1$ for the inverse Wishart distributions to center at $\boldsymbol{W}_A/(\nu_A - K - 1)$ and $\boldsymbol{W}_B/(\nu_B - K - 1)$, respectively. The degrees of freedom were set to be $\nu_A = 9$ and $\nu_A = 10$.

Next, the covariate of primary interest for the UCLA NSF is the indicator of whether the person is in a SZ or control family, so the regression coefficients are the means of each outcome by family member type in two groups, referred to as group means later. Without strong belief to the contrary, priors for group means of a particular outcome are assumed to be the same for both groups across family members, that is, there are 5 distinct priors for group means, one for each outcome. In this case, any differences in posterior means across family members or between groups will be driven by the data, not the prior. Means and variances of independent normal priors are set to values in Table 4.3.

The total variance for the observed variable of the 1^{st} outcome (MANIPA) measured on the 1^{st} family member (proband) can be decomposed as

$$\text{var}(y_{i11}) = \phi_{A11} + \phi_{B11} + \psi_{11},$$

so $\widehat{\text{var}}(y_k)$ can be used as an upper bound for these 3 components. To estimate the

fraction of total variance contributed by each component, it would be natural to fit the CFA model using the FIML-QNO to the Phase 1 data. However, it failed to converge when the model includes both $J = 4$ family member factors and $K = 4$ outcome factors. When "half-models" with only family member factors or with only outcome factors are fit to the Phase 1 data, about 40% of the total variance is explained by family member factors, or by outcome factors, respectively. Therefore, the prior means of both ϕ_{A11} and ϕ_{B11}, are set the be 40% of the estimated overall variance for the first outcome, $\widehat{\text{var}}(y_1)$, as listed in Table 4.3. Specification of other factor variances and the overall factor variance matrix proceeds as described in Section 2.3. Finally, the prior mean of the unique error variance of the k^{th} outcome, ψ_k, are set to be 20% of $\widehat{\text{var}}(y_1)$, while the degrees of freedom for these inverse gamma prior distributions are all set to be 15.

	Mean	SD	Variance	SD ratio
CPTDSD	28	11	121	1.00
DShitr	63	23	529	2.09
DSfalr	7	15	225	1.36
SPAN10	50	5	25	0.46
SPAN1	60	4.50	20.3	0.41
SPAN5	60	7.70	59.3	0.70
CPT37D	41	9	81	0.82
Hitr37	95	19	361	1.73
Falr37	0.6	4	16	0.36
logTRLBA	-140	20	40	1.82
logTRLAA	-140	17	289	1.55
VFFAS	38	10	100	0.91
NCFRSFSC	23	4.20	17.6	0.38
Manipa	70	13	169	1.18
MAINacc	70	13	169	1.18
MANIPrt	12	3	9	0.27
MAINrt	12	3	9	0.27

Table 4.3: Summary of estimates of overall means and variances for the $K = 17$ outcome measures, obtained from Phase 1 data of the UCLA Family Study and from previous literature (Kim et al., 2004; Nuechterlein et al., 2011; Koide et al., 2012). SD denotes standard deviation. SD ratio is the ratio of SD of an outcome to the SD of the first outcome.

	CPTDSD	CPT37D	SPAN10	logTRLBA
CPTDSD	1.00			
CPT37D	0.48	1.00		
SPAN10	0.30	0.33	1.00	
logTRLBA	0.34	0.45	0.41	1.00

Table 4.4: Correlations across outcomes ignoring the family structure from the Phase 1 data.

		Proband	Sibling	Father	Mother
CPTDSD	Proband	1.00			
	Sibling	0.22	1.00		
	Father	0.20	0.14	1.00	
	Mother	0.25	0.21	0.26	1.00
CPT37D	Proband	1.00			
	Sibling	0.25	1.00		
	Father	0.04	0.02	1.00	
	Mother	0.15	-0.09	0.15	1.00
SPAN10	Proband	1.00			
	Sibling	0.13	1.00		
	Father	0.21	-0.04	1.00	
	Mother	0.14	0.14	0.13	1.00
logTRLBA	Proband	1.00			
	Sibling	0.46	1.00		
	Father	0.14	-0.11	1.00	
	Mother	0.17	-0.01	0.27	1.00

Table 4.5: The within-outcome across-family-member correlations for the four outcomes in Phase 1 data.

4.3 Summary of Posterior Distributions

When the classic CFA model using full information likelihood estimation and the full information maximum likelihood estimation using quasi-Newton optimization (FIML-QNO) is fit to the UCLA Neurocognitve Family Study data using the lavaan package in R, the algorithm fails to converge.

The BFFM estimation procedure using a Gibbs sampling algorithm is implemented in R, with a total of $100,000$ iterations after excluding $10,000$ initial burn-in iterations. Trace plots, density plots and autocorrelation plots show no obvious evidence of bad mixing, non-convergence or high autocorrelation.

Tables D.3, D.1 and D.2 in Appendix D present summaries of the posterior distributions for all 101 parameter estimates, which include means, standard deviations (SD), and posterior probabilities $p(\theta < 0|\boldsymbol{Y})$. The posterior means are summarized and organized in separate tables as discussed below.

Table 4.6 presents posterior means of factor variances, factor correlations and factor loadings for family member and outcome factors. The posterior means of all family member factor correlations are positive and vary from a low of 0.034 between mother and sibling to a high of 0.390 between proband and sibling. Similarly, the posterior means of all outcome factor correlations are positive and range from 0.29 to 0.61. In addition, The posterior means of all factor loadings are all positive, suggesting the observed variables are positively associated with the factors they load on.

Table 4.7 lists the $2JK = 40$ posterior means of regression coefficients, β_{pjk} (top), which are the means of the k^{th} outcomes for the j^{th} family member in the control families ($p = 1$) and the SZ families ($p = 2$), for $j = 1, \ldots 4$ corresponding to prbands, siblings, fathers and mothers, and $k = 1, \ldots, 5$ corresponding to MANIPA, CPTDSD, CPT37D, SPAN10 and logTRLBA, and posterior means of difference in group means, control minus SZ, $\beta_{1jk} - \beta_{2jk}$ (bottom). These results suggest that SZ probands performed worse than the control probands for all five outcomes, while the sign of the

differences in mean outcomes between siblings of the two groups are not well determined by looking at the posterior probabilities $p(\theta < 0|\boldsymbol{Y})$. Parents of schizophrenia probands did worse in span of apprehension and trails B than control parents.

Besides posterior means, posterior distributions of model parameters are also of interest. Figure 4.2 plots the posterior distribution of group means of CPT37D for probands, siblings, fathers and mothers in the control and SZ families (left) and the differences between two groups. These plots show that the means of CPT37D for SZ probands are much smaller than those for control probands, while there are no obvious differences in means between the two groups for fathers and siblings. Mothers in the control families have larger means CPT37D than all others, including the SZ mothers. Additional plots of posterior distributions for group means and factor loadings are provided in Appendix D.

Family Member		Proband	Sibling	Father	Mother	
Factor Variance		50.57	50.81	33.31	37.14	
	Proband	1.000				
Factor	Sibling	0.390 *	1.000			
Correlations	Father	0.173	0.070	1.000		
	Mother	0.081	0.034	0.104	1.000	
	MANIPA	1	1	1	1	
Factor	CPTDSD	0.88 *	0.61 *	0.92 *	0.77 *	
Loadings	CPT37D	1.01 *	0.99 *	0.73 *	0.90 *	
	SPAN10	0.61 *	0.46 *	0.47 *	0.52 *	
	logTRLBA	2.13 *	2.05 *	1.77 *	2.36 *	
Outcome		MANIPA	CPTDSD	CPT37D	SPAN10	logTRLBA
Factor Variance		50.20	24.07	16.70	4.67	97.69
	MANIPA	1.000				
Factor	CPTDSD	0.594 *	1.000			
Correlations	CPT37D	0.493 *	0.609 *	1.000		
	SPAN10	0.349 *	0.354 *	0.291	1.000	
	logTRLBA	0.590	0.526	0.529	0.470	1.000
	Proband	1	1	1	1	1
Factor	Sibling	1.18 *	1.23 *	0.84 *	1.52 *	0.81 *
Loadings	Father	0.18	0.66 *	0.60 *	1.08 *	0.84 *
	Mother	0.54 *	0.65 *	0.65 *	0.98 *	0.76 *
Unique Error		MANIPA	CPTDSD	CPT37D	SPAN10	logTRLBA
Variances		150.17	65.41	28.57	15.77	229.25

Table 4.6: Posterior means of factor variances, factor correlations, factor loadings and unique error variances estimated by BFFM. For a parameter with *, the posterior probability being smaller than zero, is smaller than 0.05, $p(\theta < 0|\boldsymbol{Y}) < 0.05$.

Group	Control				SZ			
Outcome	Proband	Sibling	Father	Mother	Proband	Sibling	Father	Mother
MANIPA	76.01	70.25	77.00	74.39	66.56	72.37	75.59	71.84
CPTDSD	25.27	22.85	20.83	26.07	21.38	23.44	24.02	23.25
CPT37D	44.63	43.36	45.01	48.55	38.44	43.77	44.36	42.90
SPAN10	56.05	54.79	53.88	53.75	53.32	55.33	51.13	51.68
logTRLBA	-139.4	-143.0	-139.5	-140.8	-151.1	-145.1	-150.0	-151.8
	Difference (Control − SZ)							
Outcome	Proband	Sibling	Father	Mother				
MANIPA	9.46 *	-2.12	1.41	2.55				
CPTDSD	3.89 *	-0.59	-3.19	2.82 *				
CPT37D	6.19 *	-0.42	0.65	5.65 *				
SPAN10	2.73 *	-0.54	2.75 *	2.06 *				
logTRLBA	11.8 *	2.1	10.5 *	11.0 *				

Table 4.7: Posterior means of regression coefficients for $J = 4$ family members and $K = 5$ outcomes per family in the control and schizophrenia (SZ) families (top), and the differences between groups (bottom). For a parameter with *, the posterior probability being smaller than zero, is smaller than 0.05, $p(\theta < 0|\boldsymbol{Y}) < 0.05$.

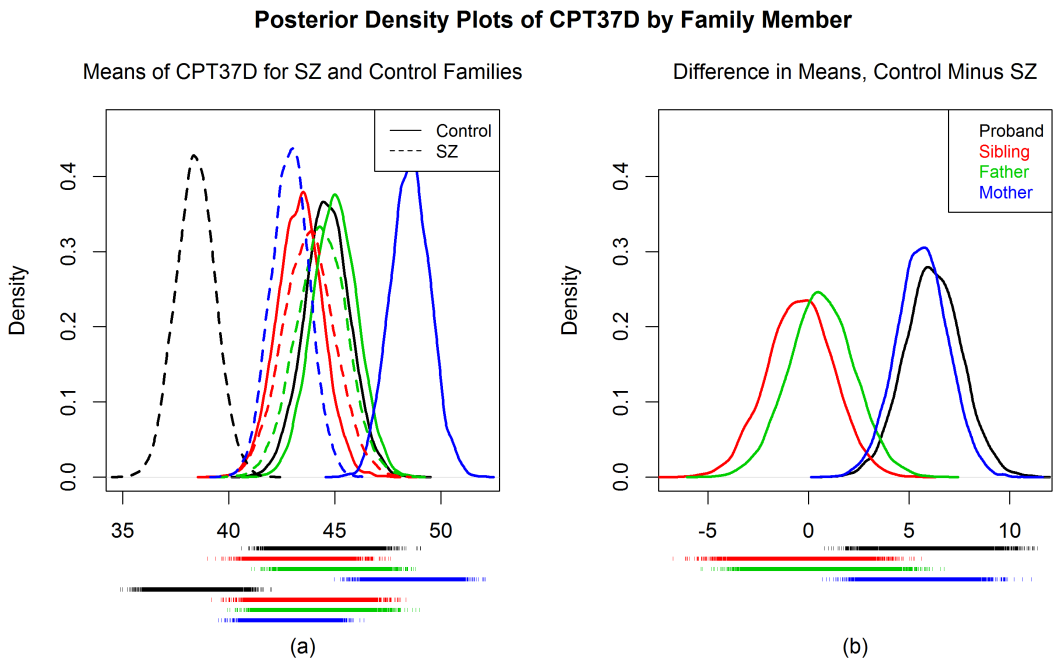


Figure 4.2: (a) Posterior density of means of CPT37D for probands, siblings, fathers and mothers in the control (black) and SZ (grey) families. The 8 1-dimensional density plots at the bottom represent locations of posterior samples for probands, siblings, fathers and mothers in control and SZ families, from top to bottom. (b) Posterior densities for differences in means of CPT37D between two groups, control minus SZ. The 4 1-dimensional density plots at the bottom represent locations of posterior samples for probands, siblings, fathers and mothers, from top to bottom.

CHAPTER 5

Application of the Bayesian Family Factor Model to the UCLA Family Study Data with Seventeen Outcomes

In Chapter 4, I illustrated the Bayesian Family Factor Model by analyzing the UCLA Neurocognitive Family Study (NSF) data with 5 primary outcomes. Twelve additional measures in the same domains are also collected as secondary outcomes. When more outcomes are considered, the relational structure becomes more complex and the compound symmetric prior correlation model for outcome factors used in Chapter 4 may be too restrictive. Moreover, there may be substantive prior information about the relationships among outcome measures. For example, associations among sub-scales from the same test will be similar and higher than correlations between sub-scales from different tests. Furthermore, outcomes which are designed to measure similar concepts will be more highly correlated than those from different domains. It is desirable to accommodate this substantive information about the clustering of the outcomes in the prior specification for the BFFM.

This chapter describes a way to elicit priors for correlations among outcome factors in familial data with a large number of outcomes, and illustrates the approach by fitting the BFFM to the UCLA NSF data with 17 outcomes. These outcomes are listed in Table 5.1, along with the transformations to make the scales similar and all correlations positive. Except for two stand-alone tests, facial recgnition (NCFRSFSC) and verbal fluency (VFFAS), all of the outcome measures belong to five tests: the Maintenance and Manipulation (MNM), the Degraded Stimulus-Continuous Performance Test (DS-CPT), the memory-load Continuous Performance Test (3-7 CPT), the Forced-

Choiced Span of Apprehension (SPAN) and the Trail Making Test Adolescent Version (TRAILS). Each of these tests include one of the 5 primary outcomes analyzed in Chapter 4, highlighted in bold.

Outcome	Description	Transformation
CPTDSD	Degraded Stimulus(DS)-CPT: block sum d prime	$\times 10$
HitrDS	DS-CPT: block sum hit rate	$\times 100\%$
FalrDS	DS-CPT: negative block sum false alarm rate	$\times(-100\%)$
SPAN10	Span of apprehension: number correct matrix size 10	
SPAN1	Span of apprehension: number correct matrix size 1	
SPAN5	Span of apprehension: number correct matrix size 5	
CPT37D	Memory load 3-7 CPT block sum d prime	$\times 10$
Hitr37	3-7 CPT block sum hit rate	$\times 100\%$
Falr37	Negative 3-7 CPT block sum false alarm rate	$\times(-100\%)$
logTRLBA	Trail making test B, Adolescent Version: time (sec)	$-100\log_{10}(y)$
logTRLAA	Trail making test A, Adolescent Version: time (sec)	$-100\log_{10}(y)$
VFFAS	Verbal Fluency: sum of f, a and s total scores	
NCFRSFSC	Facial recognition: short form score	
MANIPA	MNM (Maintenance and Manipulation Test): manipulate (flip) trials mean accuracy	$\times 100\%$
MAINacc	MNM: main (hold) trials mean accuracy	$\times 100\%$
MANIPrt	MNM: manipulate trials mean reaction time (sec)	$/100$
MAINrt	MNM: main trials mean reaction time	$/100$

Table 5.1: Descriptions and transformations or scalings of 17 outcome measures. The signs of variables DS-CPT block sum false alarm rate, 3-7 CPT block sum false alarm rate, Trails A Adolescent Version time and Trails B Adolescent Version time are reversed so that for all outcome measures, a larger value means better performance.

5.1 Descriptive Statistics

In this section, descriptive statistics are calculated to investigate the raw means and variance-covariance structure of the 17 outcomes from the UCLA NFS. Table 5.2 summarizes these values, by schizophrenia (SZ) and control families.

Sample correlations among the 17 outcomes, ignoring the family structure, are listed in Table E.1 in Appendix E and the corresponding heat map is presented in Figure 5.1. A heat map is a scale colour image for representing values in two dimensions. Overall, correlations among measures from the same test are higher than those of measures from different tests.

To further explore relationships among outcome measures, I perform a cluster analysis on the 17 outcomes using the VARCLUS procedure (Nelson, 2001) in SAS. The resulting dendrogram in Figure 5.2 shows that the measures from the same test are tightly grouped as expected. Based on the test domains and the results of the clustering, these 17 outcomes are organized into 5 groups, as listed in Table 5.3.

	Mean		StdDev	
	Control	SZ	Control	SZ
CPTDSD	24.2	22.7	9.9	11.1
HitrDS	73.6	70.0	18.8	21.6
FalrDS	-6.8	-8.4	6.1	8.5
SPAN10	54.80	52.94	5.47	6.36
SPAN1	63.12	62.71	1.36	2.21
SPAN5	61.57	59.75	2.83	4.93
CPT37D	45.4	41.6	8.2	9.8
Hitr37	94.5	89.8	7.7	12.8
Falr37	-0.59	-0.89	1.43	1.75
logTRLBA	-141	-150	22	23
logTRLAA	-112	-119	18	19
VFFAS	39.71	33.79	12.65	12.54
NCFRSFSC	22.99	22.62	2.15	2.58
MANIPA	74.1	70.9	13.5	16.7
MAINacc	80.5	75.9	12.0	16.7
MANIPrt	12.8	12.4	2.3	2.3
MAINrt	11.6	11.3	2.4	2.3

Table 5.2: Means and standard deviations of 17 outcomes by schizophrenia (SZ) and control family.

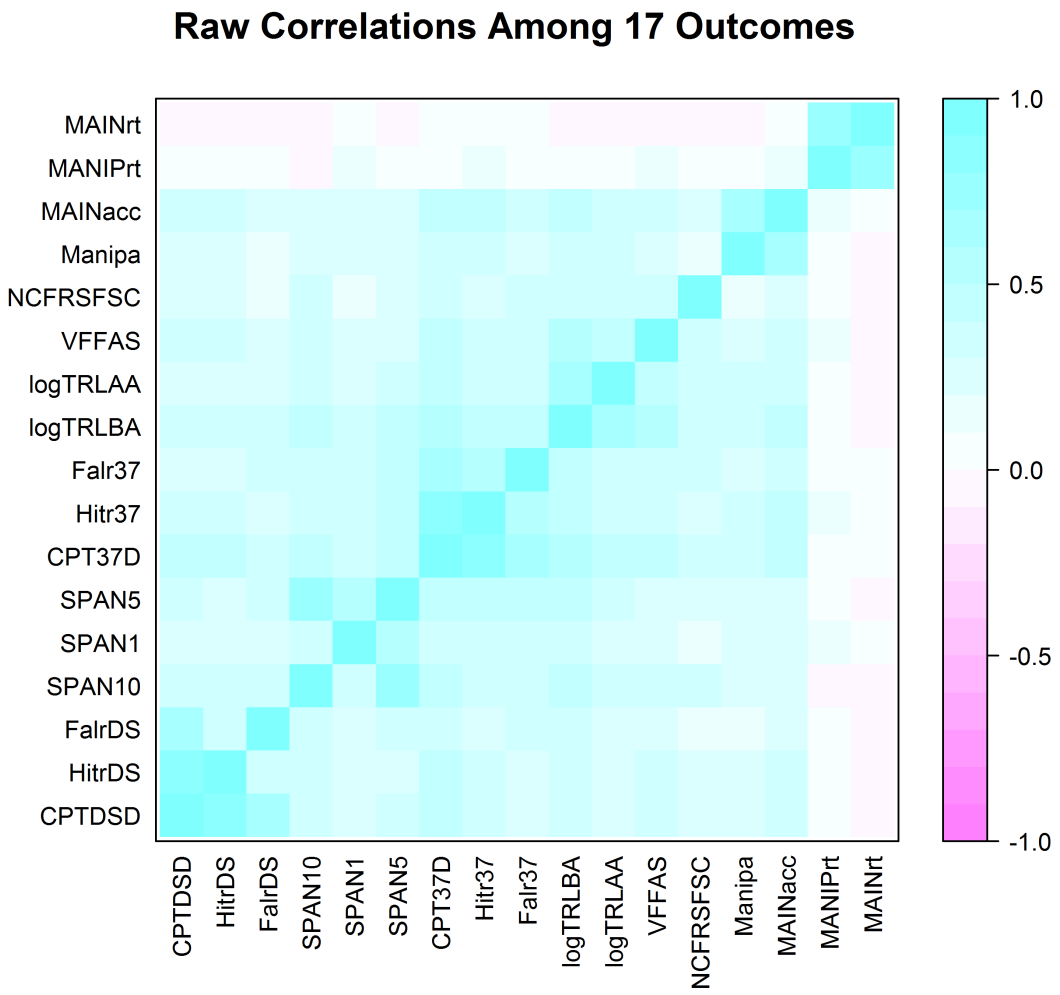


Figure 5.1: A heat map of raw pair-wise correlations among observations of the 17 outcome measures, ignoring the within family correlations. Cyan and pink represent positive and negative values.

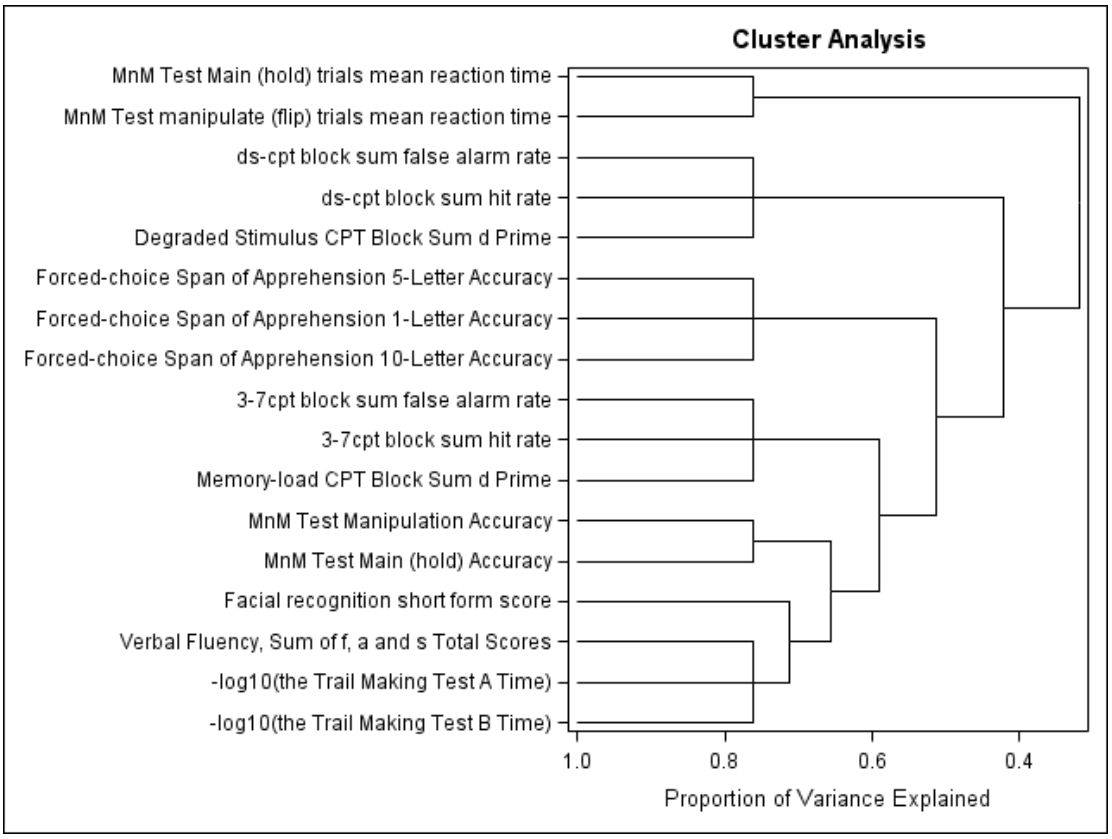


Figure 5.2: A dendrogram corresponding to clustering of the 17 outcome measures using VARCLUS in SAS, ignoring the family structure.

Cluster	Outcomes
Degraded Stimulus-CPT	CPTDSD, HitrDS, FalrDS
Memory-Load CPT	CPT37D, Hitr37, Falr37
Span of Apprehension (SPAN)	SPAN10, SPAN1, SPAN5
Trail making test (TRAILS) and others	logTRLBA, logTRLAA, VFFAS, NCFRSFSC
Maintenance and Manipulation (MNM)	MAINrt, MANIPrt, MAINacc, Manipa

Table 5.3: Grouping the 17 outcome measures into 5 clusters, based on the sets of tests and results of variable clustering.

5.2 Prior Specification

To elicit hyper-parameters for all priors, I obtain estimated values for the overall variances of the $K = 17$ outcomes, $\widehat{\text{var}}(y_1), \ldots, \widehat{\text{var}}(y_K)$ from (i) Phase 1 of the UCLA Family Study (4 measures) and (ii) the means and standard deviations reported in previous literature (13 measures) (Kim et al., 2004; Kopelowicz et al., 2005; Nuechterlein et al., 2011; Koide et al., 2012). The estimated values are listed in Table 4.3. The priors for other parameters, including regression coefficients, $\boldsymbol{\beta}$, factor loadings, $\boldsymbol{\Lambda}_A$ and $\boldsymbol{\Lambda}_B$, family factor variance matrix, $\boldsymbol{\Phi}_A$, and unique error variance matrix, $\boldsymbol{\Psi}$, can be specified using methods described in Section 2.3. Tables E.2 and E.3 in the appendix list all prior hyper-parameters specified for fitting the BFFM.

For outcome factor variances,

$$\boldsymbol{\Phi}_B \sim \mathcal{IW}(\boldsymbol{W}_B, \nu_B),$$

where $\boldsymbol{W}_{B(K\times K)} = (\nu_B - K - 1)\boldsymbol{D}_B\boldsymbol{C}_B\boldsymbol{D}_B$, ν_B and $\boldsymbol{C}_{B(K\times K)} = \text{diag}(d_{B1}, \ldots, d_{BK})$

are specified according to Section 2.3. When the number of outcomes, K, is larger, more meaningful prior outcome factor correlations, $\boldsymbol{D}_B$, can be specified, as follows.

First, the K outcome measures are grouped into a smaller number (s) of clusters, based on the nature of the tests. Sub-scales from the same test are grouped together. Measures designed to assess similar concepts are also assigned to the same cluster. Computation of sample correlations and variable clustering are useful for confirming that the theoretically selected groups are cohesive and for assigning outcome measures that do not belong a priori to a particular group. Without strong belief to the contrary, the within-cluster prior correlations are set to be all equal and higher than the cross-cluster prior correlations, which are also set to be all equal.

For the UCLA NFS, the $K = 17$ outcomes are grouped into $s = 5$ clusters (Table 5.3), as discussed in Section 5.1. For the outcome factors, all with-in cluster prior correlations are set to be 0.4, while all across-cluster correlations are set to be 0.2, because correlations among the four outcomes from the Phase 1 Study range from 0.3 to 0.5. Table E.3 in Appendix E presents the prior correlation matrix for outcome factors and Figure 5.3 shows the corresponding heat map. The prior mean of the covariance matrix can be calculated from the variances and the correlation matrix. Parallelling Section 4.2, I choose $\nu_A = 9$ and $\nu_B = 20$ (Note that it is necessary to have $\nu_A > J+1 = 5$ and $\nu_B > K+1 = 18$ for the inverse Wishart distributions to center at $\boldsymbol{W}_A/(\nu_A - J - 1)$ and $\boldsymbol{W}_B/(\nu_B - K - 1)$).

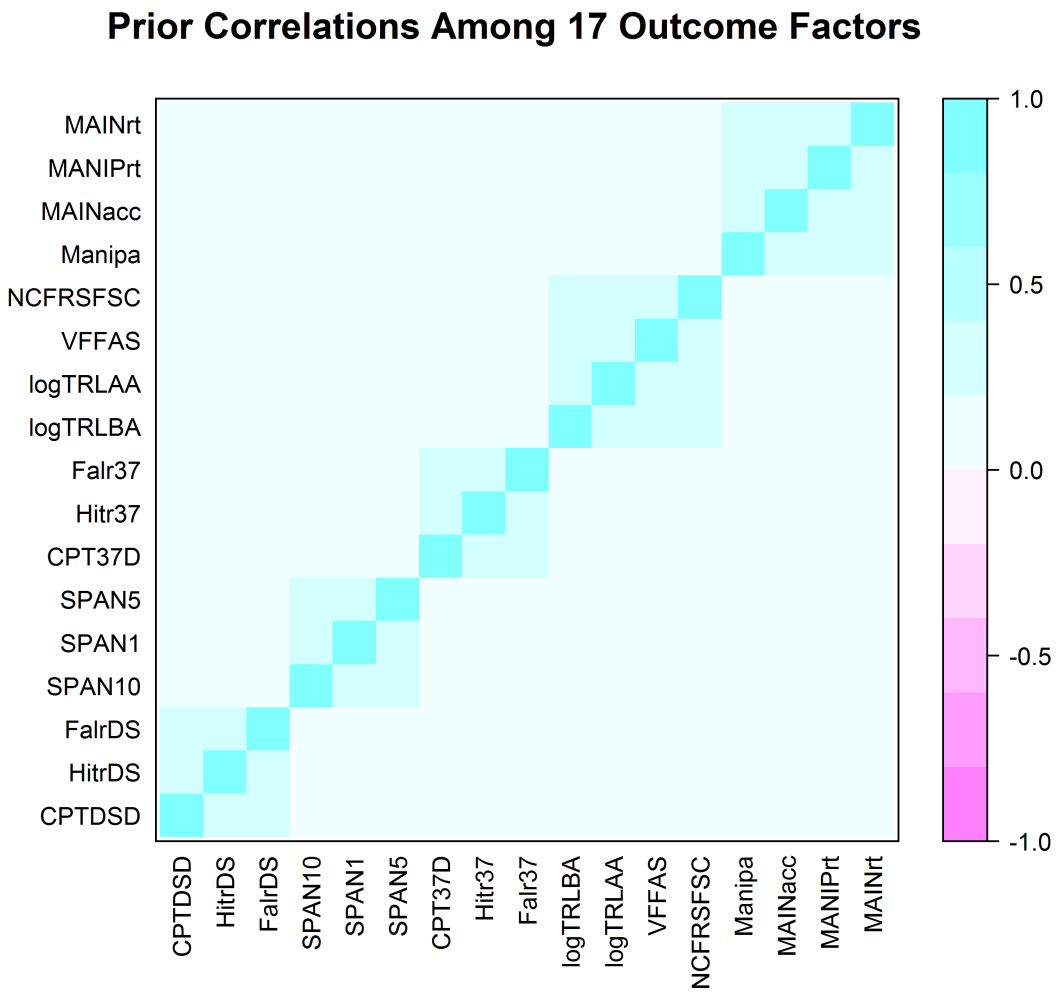


Figure 5.3: A heat map of prior correlations among the 17 outcome measures. Cyan and pink represent positive and negative values, respectively. All diagonal elements are 1, while prior correlations in the diagonal blocks equal to 0.4 and all other prior correlations are set to be 0.2.

5.3 Posterior Distribution Summary

This section provides the summary of fitting the BFFM to 17 outcomes from the UCLA NSF. It is desirable to compare parameters estimated by BFFM and FIML-QNO. However, when the CFA model with both family member factors and outcome factors were fit to the data, the FIML-QNO algorithm fail to converge, as the covariance matrix is not positive definite.

The BFFM is successfully fit to the familial data with 17 outcomes using $20,000$ primary iterations after an initial burn-in of 2000 iterations. Assuming $K = 17$ distinct error variances $\psi_1, \ldots, \psi_{17}$, the total number of free parameters is $(2JK + J^2/2 + K^2/2 - J/2 + K/2 + P) = 431$. As this number is large, I summarize the posterior distributions by visualizing the posterior means and posterior probability $p(\theta < 0|\boldsymbol{Y})$ for testing whether the parameter estimates are equal to zero.

5.3.1 Regression Coefficients

First, I look at the posterior densities of the regression coefficients, β_{pjk}, which are group means of the $K = 17$ outcomes measured on $J = 4$ types of family members: probands, siblings, fathers and mothers, where $p = 1, 2$ correspond to the control and SZ families. Figures F.1, F.2 and F.3 in Appendix F show the posterior density plots for these $P = 2JK = 136$ parameters, grouped by the K outcomes. The corresponding posterior means are listed in Table F.1. Overall, members of SZ families perform worse than members of control families. For all outcomes, the posterior means of β_{pjk} for the SZ probands are lower than those for the control probands. Fathers and mothers of SZ probands (green and blue dashed lines) also have smaller posterior means in some outcome measures than the control parents (green and blue solid lines). There are no obvious differences in posterior means between SZ siblings and control siblings for the group means. Furthermore, the CPT37D, hitr37, MAINacc and MANIPrt measures show a similar pattern in which the SZ probands (black dashed lines) have much smaller

means than all others.

To visualize the patterns across family members and between groups, these $P = 136$ posterior means of group means are standardized as follows: for each outcome, I compute the average of 8 group means and calculate the relative posterior group means

$$\text{Relative Group Means} = \frac{(\text{Group Mean} - \text{Average})}{|\text{Average}|}.$$

A heat map of the relative posterior group means is shown in the left part of Figure 5.4, in which pink indicates worse performance in cognitive tests, while cyan indicates better performance. This plot is consistent with the posterior density plots, showing smaller posterior group means for SZ families (the right half of the plot) in general.

The $JK = 68$ differences in group means, (control minus SZ, or $\beta_{2jk} - \beta_{1jk}$), are summarized in the right half of Table F.1 in Appendix F. I visualize the corresponding posterior probability, $p(\theta < 0|\boldsymbol{Y})$ for testing whether these posterior means are equal to 0, $p(\beta_{2jk} - \beta_{1jk} > 0)$, using a heat map in the right part of Figure 5.4. This heat map suggests that the SZ probands perform worse in all measures than the control probands. There are no obvious differences in group means between siblings in SZ and control families. The differences in group means show similar patterns within clusters of outcomes for parents: no significant differences in the Maintenance and Manipulation cluster; only mothers show significant differences in the Degraded Stimulus-CPT and the Span of Apprehension clusters; and both parents show significant differences for most of the Memory-Load CPT and Trail making test clusters. To further identify the pattern of significant differences in group means, Figure 5.5 plots the heat map of $p(\theta < 0|\boldsymbol{Y})$ with dendrograms added to the left side and to the top; the rows and columns are re-ordered according to row and column means. This plot suggests that among the relatives of probands, the SZ mothers have significantly smaller group means on most outcomes than the control mothers, while the differences in group means between siblings are the smallest. The dendrogram in Figure 5.2 based on the Bayesian

Family Factor Model retains some of the same clusters as the dendrogram for the clusters based on the raw data, from top to bottom, I can see the Degraded Stimulus-CPT cluster (HitrDS, CPTDSD and FalrDS), some of the Maintenance and Manipulation cluster (MAINrt, MANIPrt, MAnipa), some of the Memory-Load CPT cluster (Falr37 and CPT37), and the last two clusters mixed together. New information is obtained after accounting for family structure.

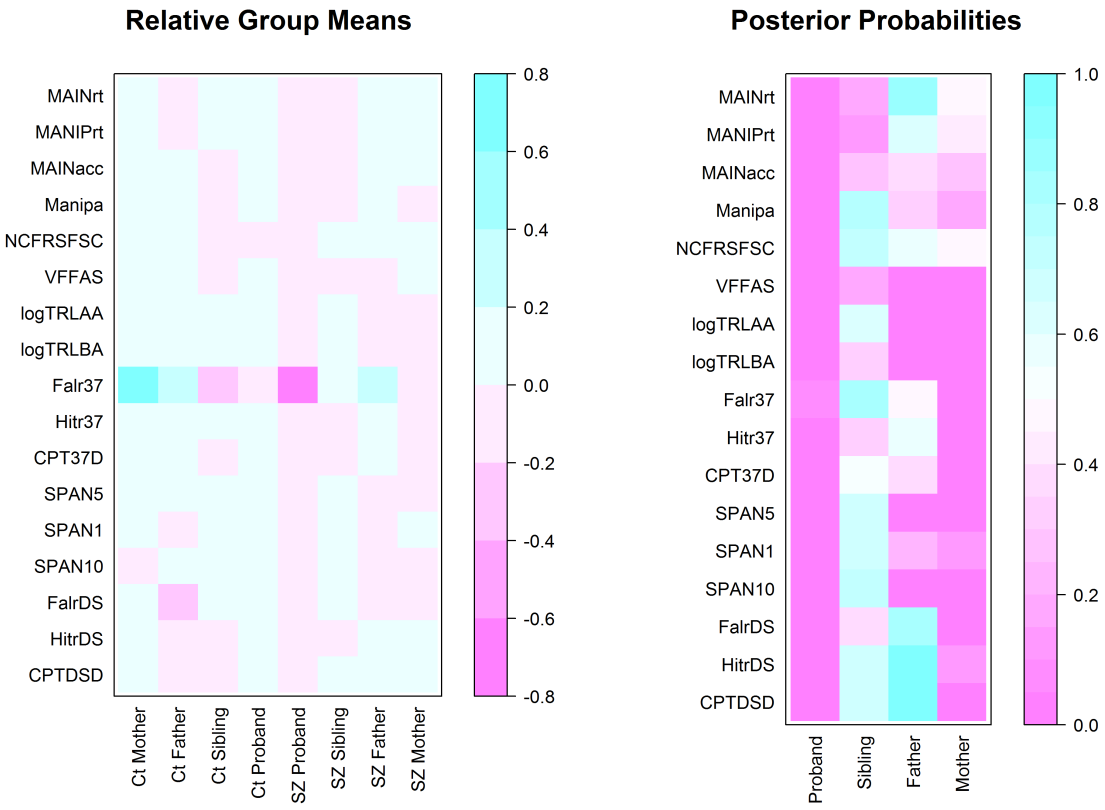


Figure 5.4: Heat maps of relative posterior means of group means by family member type and control or SZ families (left) and of the corresponding posterior probabilities of the parameters being smaller than 0, $p(\theta < 0|\boldsymbol{Y})$ (right).

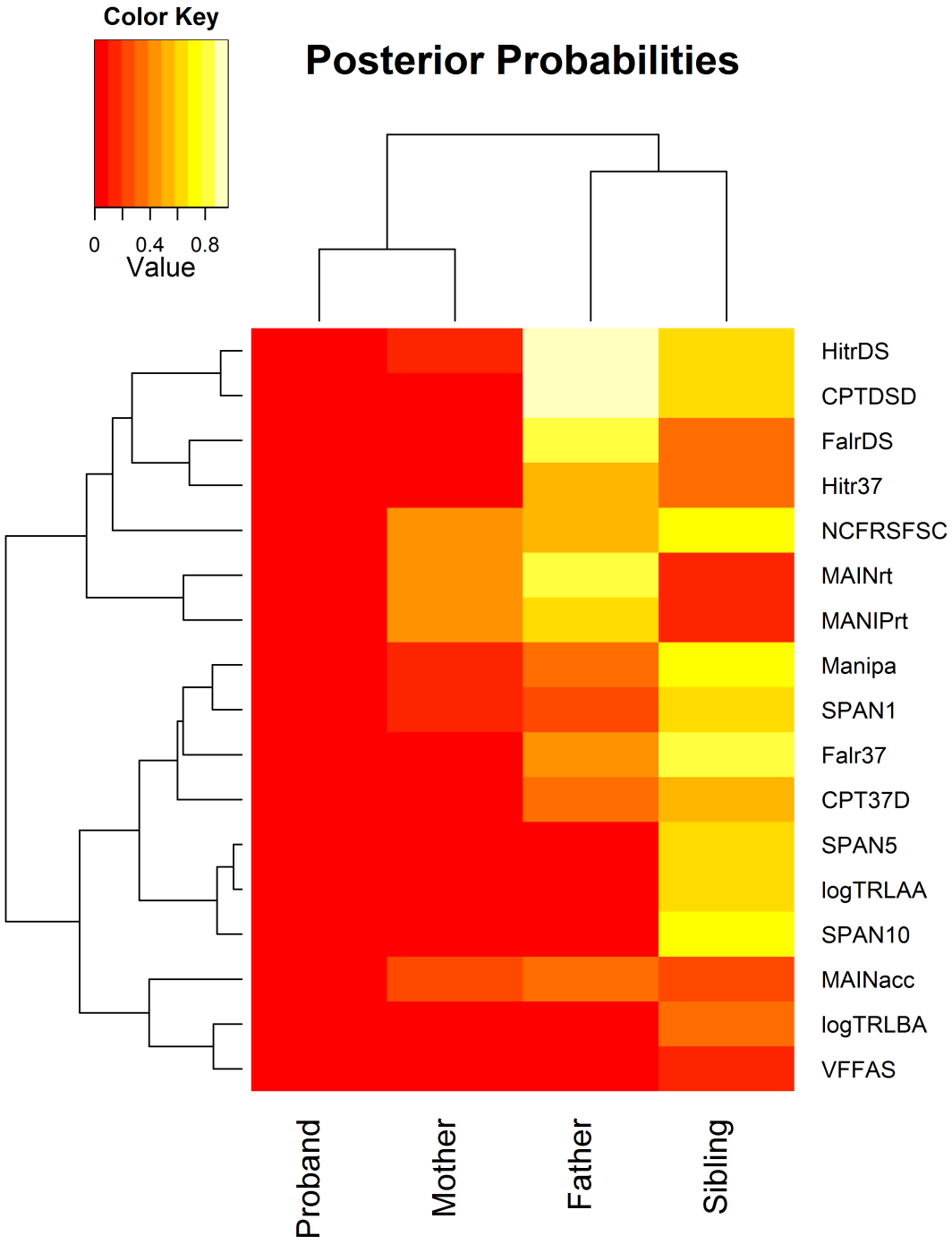


Figure 5.5: A heat map of posterior probabilities of the parameters being smaller than 0, $p(\theta < 0|\boldsymbol{Y})$, based on the BFFM. The rows and columns are clustered and re-ordered.

5.3.2 Factor Covariance and Correlation Matrices

The posterior means of family member factor variances and covariances, as well as the corresponding probabilities of the parameters being smaller than 0, $p(\theta < 0|\boldsymbol{Y})$, are listed in Table 5.4. The corresponding heat maps are shown in the top left and top right of Figure 5.6. All family member factor covariances are significantly greater than 0. Furthermore, all family member factor correlations are positive, but only the covariance between probands and siblings and that between probands and mothers are significantly greater than 0 with $p(\theta < 0|\boldsymbol{Y}) < 0.05$. The same heat maps for the analyses of the data with 5 outcomes are presented in the bottom left and bottom right of Figure 5.6 for comparison. Posterior means of family member factor correlations are consistent in the two analysis, except that the correlation between fathers and probands are slightly higher in the analysis of data with 17 outcomes. Compared with the analysis when $K = 5$, the posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$, are smaller when $K = 17$, suggesting that combining strength across more outcomes improves estimation of model parameters.

Similarly, posterior means of outcome factor correlations and variances are summarized in Tables F.2 and F.3, respectively, in Appendix F. Figure 5.7 shows a heat map visualizing these posterior means (left), as well as a heat map of the corresponding posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ (right). This figure suggests that the within-cluster posterior factor correlations are higher than the across-cluster posterior factor correlations. The heat map of posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$, indicate that almost all of these posterior correlations are significantly greater than zero. Although the posterior correlations between certain outcomes and MAINrt or MANIPrt are negative, none of these negative correlations are significantly different from 0.

	Factor Correlations				Factor
	Proband	Sibling	Father	Mother	Variance
Proband	1.000				55.89
Sibling	0.321	1.000			21.26
Father	0.194	0.082	1.000		18.90
Mother	0.183	0.141	0.174	1.000	20.93
	$p(\theta < 0 \mid \boldsymbol{Y})$				
	Proband	Sibling	Father	Mother	
Proband					
Sibling	<0.001				
Father	0.05	0.23			
Mother	0.02	0.10	0.09		

Table 5.4: Posterior means of the family member factor correlations and variances (top) and the corresponding posterior probabilities of the parameters being smaller than 0, $p(\theta < 0|\boldsymbol{Y})$ (bottom).

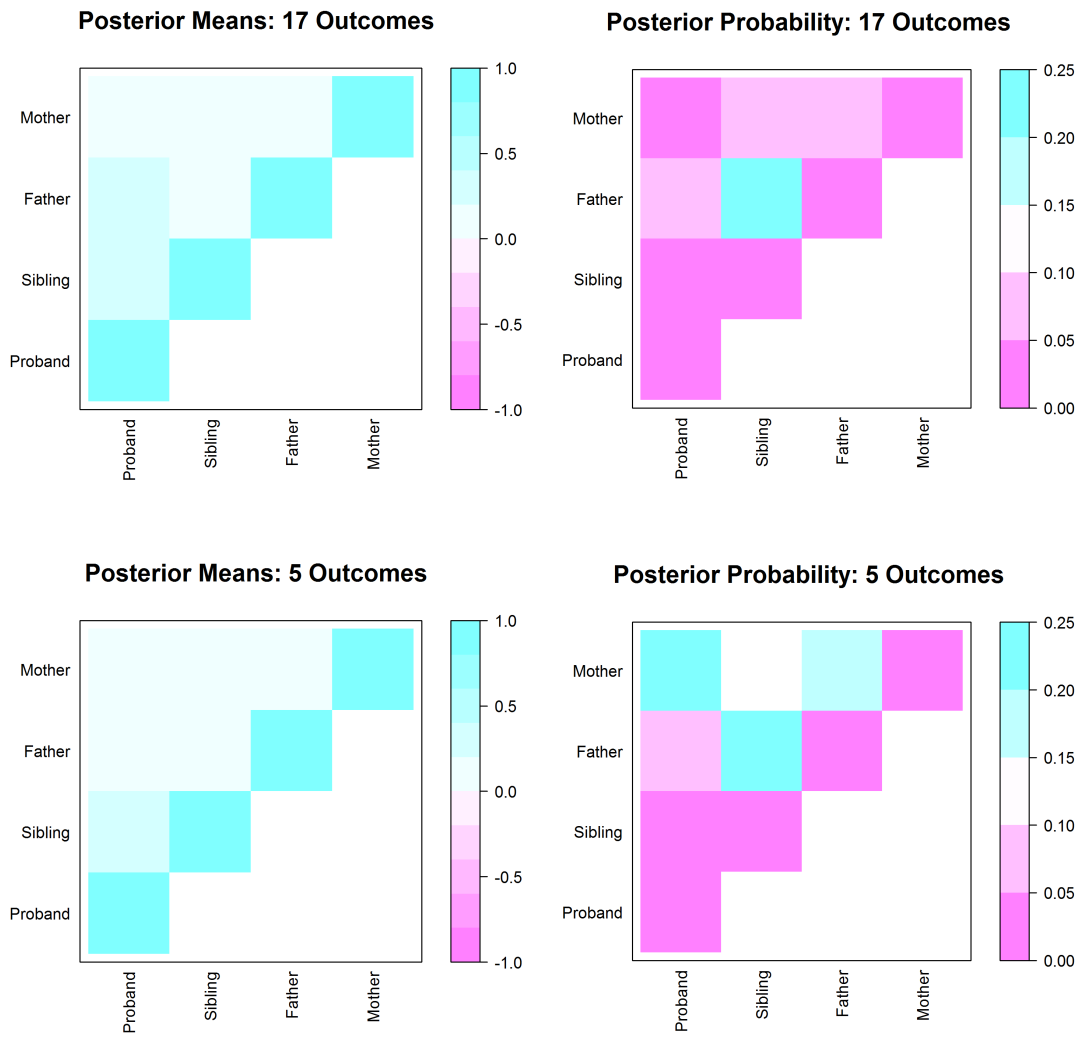


Figure 5.6: Heat maps of posterior means of outcome factor correlations estimated in BFFM with 17 outcomes (top left) and 5 outcomes (bottom left), and the corresponding posterior probabilities of parameters being smaller than 0, $p(\theta < 0|\boldsymbol{Y})$ (bottom left and bottom right), for the 17 and 5 outcome-models, respectively.

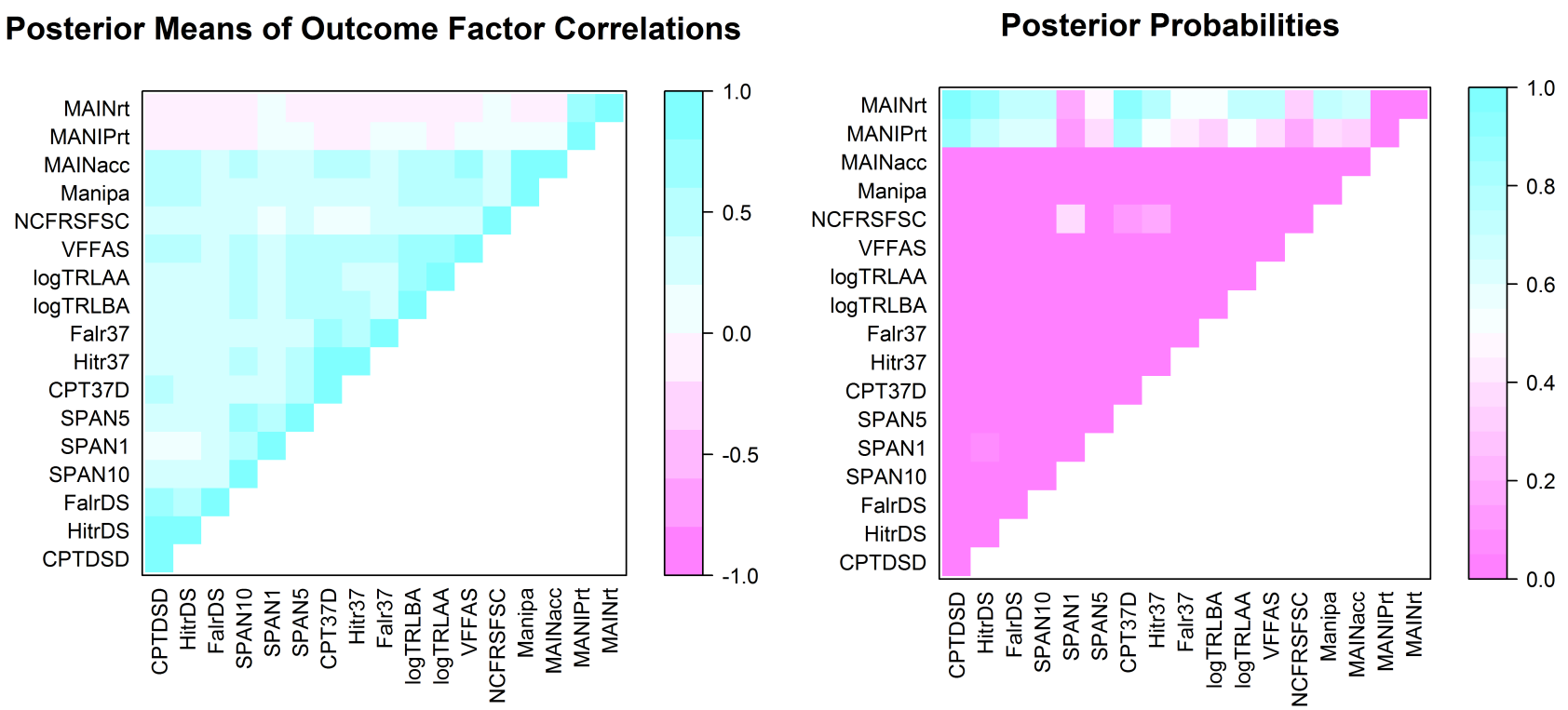


Figure 5.7: Heat maps of the posterior means of outcome factor correlations (left) and the corresponding posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$(right).

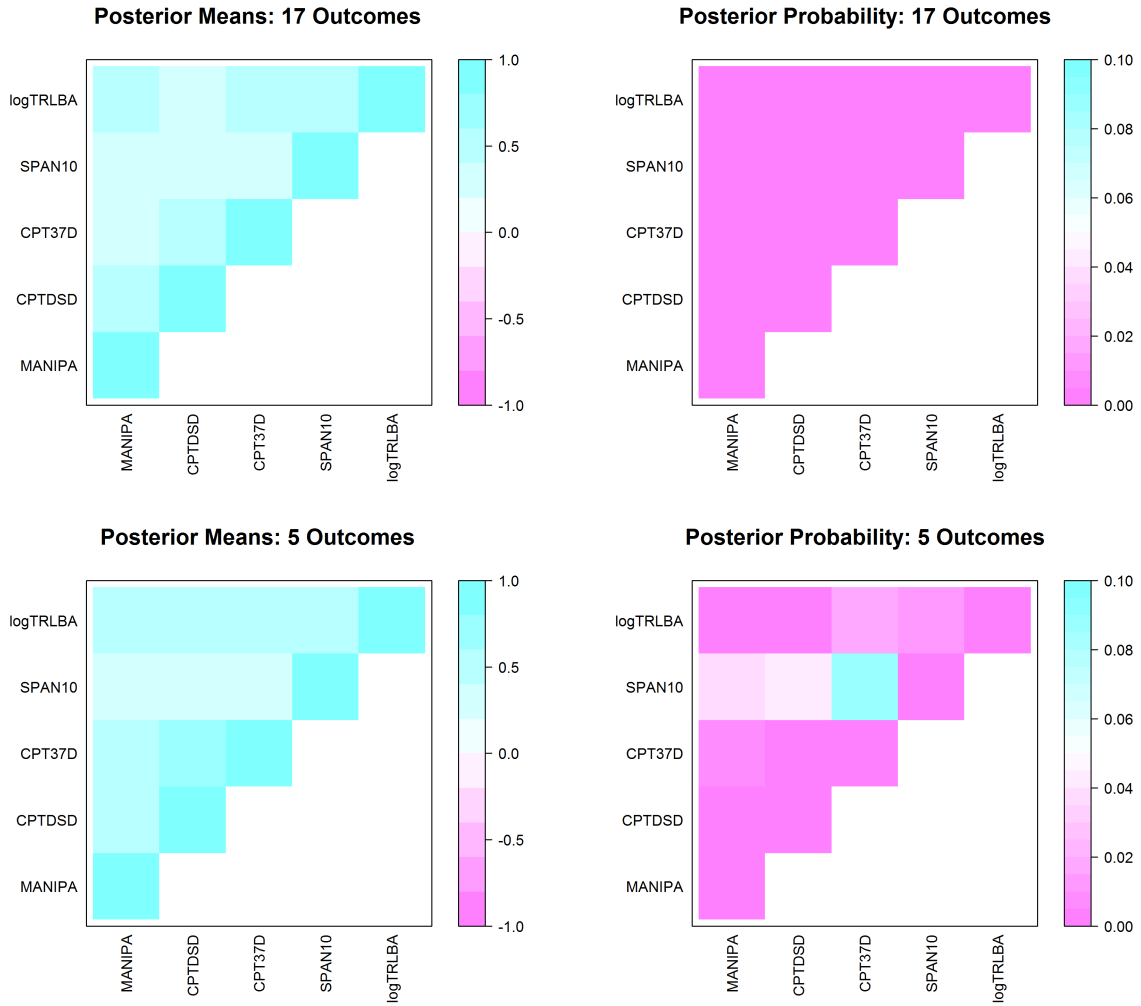


Figure 5.8: Heat map of the posterior means of outcome factor correlations (left) and heat map of the posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ (right).

5.3.3 Factor Loadings

This section presents a summary of the posterior distributions for the factor loadings. In general, factor loadings are interpreted as regression slopes for predicting the observed variables from the factors. Table F.3 in Appendix F presents the posterior means of both family member and outcome factor loadings. Note that the family member factor loadings for the first outcome (CPTDSD), a_{1k}, are fixed to 1, as this outcome

is chosen as a reference. A family member factor loading, a_{jk}, for $J = 1, \ldots, J$ and $k = 2, \ldots, K$, is interpreted as the amount of change in the k^{th} outcome on the j^{th} subject, y_{ijk} associated with 1 unit increase in f_{Aij} with all else fixed. In addition, as $a_{jk}/a_{j1} = a_{jk}/1 = a_{jk}$, this parameter is also the ratio of the effect of f_{Aij} on y_{ijk} to the effect of f_{Aij} on y_{ij1}. Figure 5.9 includes the heat map of posterior means of family member factor loadings (left) and the heat map of the corresponding posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ (right). For each outcome the posterior means of family member factor loadings have similar scales across family members. Almost all of these posterior means are greater than 0 with $p(\theta < 0|\boldsymbol{Y}) < 0.05$, except for the loadings of MAINrt and MANIPrt for probands and mothers.

An outcome factor loading, b_{jk}, for $J = 1, \ldots, J$ and $k = 1, \ldots, K$, can be interpreted as the amount of change in y_{jk} associated with 1 unit increase in f_{Bik}. Furthermore, as $b_{jk}/b_{1k} = b_{jk}/1 = b_{jk}$ for $j = 2, \ldots, J$, b_{jk} is also the ratio of the effect of f_{Bik} on y_{ijk} to that on y_{i1k}. Again, the outcome factor loadings for the first family member (proband), b_{1k}, are fixed to 1, as the proband is chosen as the reference. The posterior means of family member factor loadings and the corresponding posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ are visualized in Figure 5.10 using heat maps. All of these posterior means are positive, ranging from 0.03 to 1.35 and most are significantly greater than 0. Furthermore, factor loadings for siblings are in general larger than those for fathers and mothers.

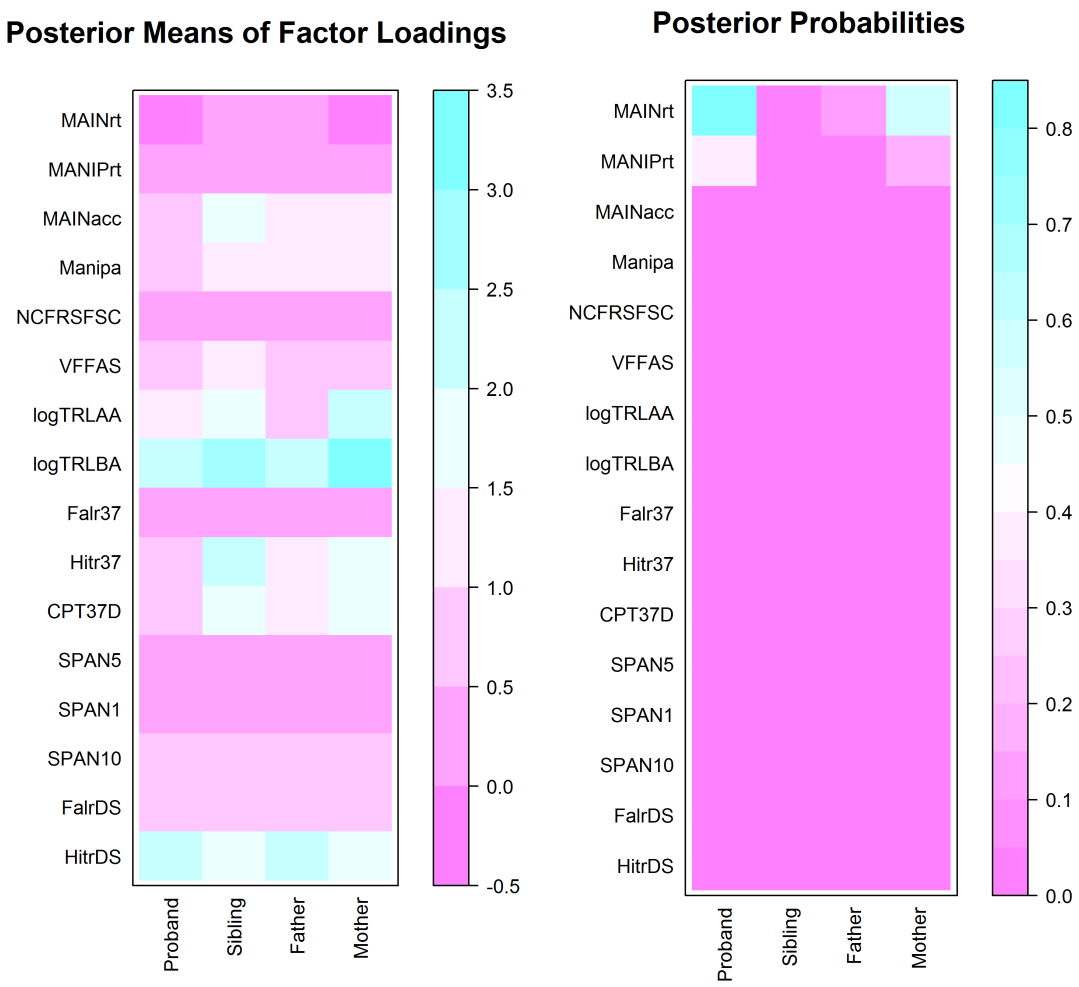


Figure 5.9: Heat map of the posterior means of family member factor loadings (left) and heat map of the posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ (right).

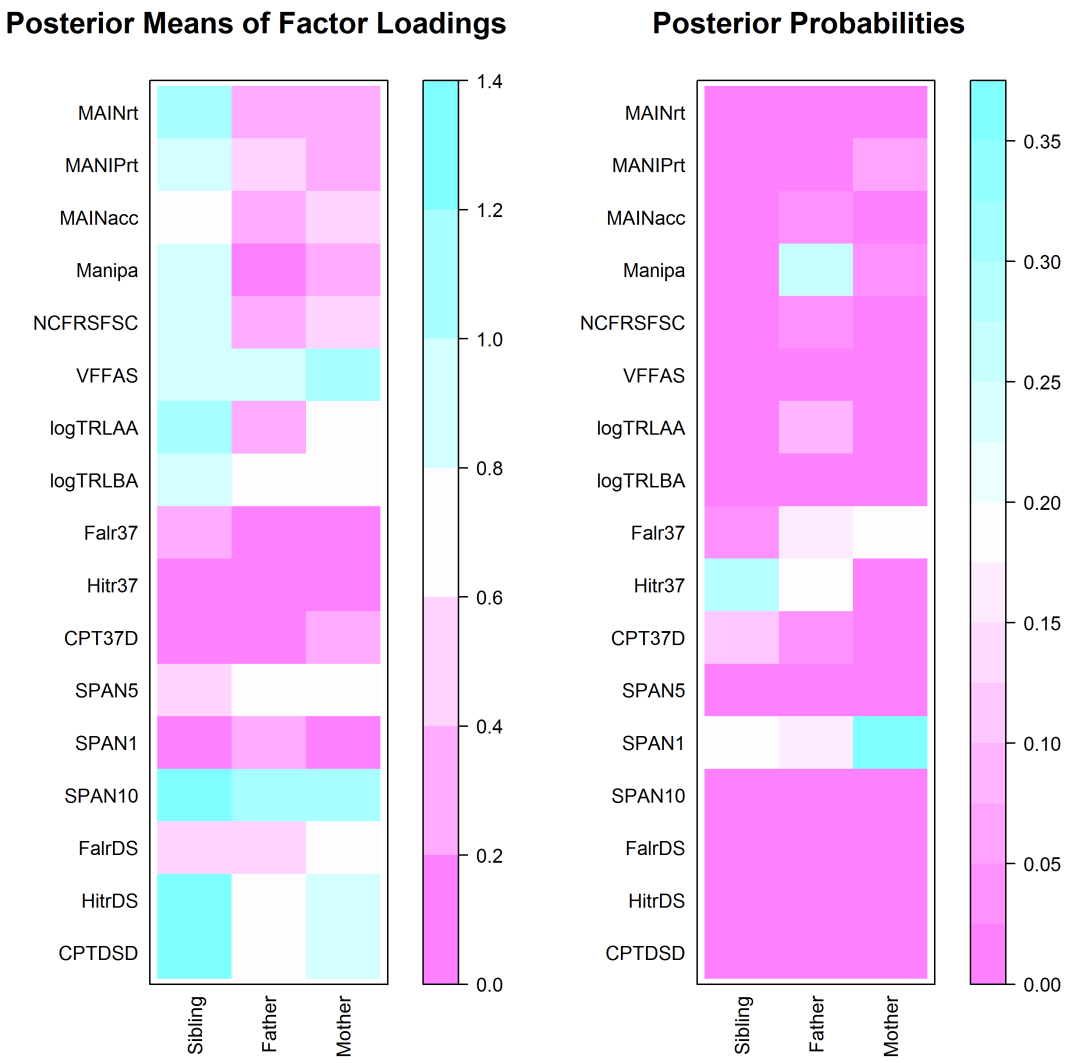


Figure 5.10: Heat map of the posterior means of outcome factor loadings (left) and heat map of the posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ (right).

CHAPTER 6

Hypothesis Testing using Bayes Factors

Besides estimating model parameters, it is of interest to test various hypotheses about these parameters. In Bayesian inference, testing a null hypothesis against an alternative can be regarded as comparing two corresponding models, $\mathcal{M}_0$ and $\mathcal{M}_1$. A Bayes factor (Berger, 1985; Kass and Raftery, 1995; Chib, 1995) is a summary of evidence provided by the data in favor of $\mathcal{M}_0$ as opposed to $\mathcal{M}_1$

$$B_{01} = \frac{p(\boldsymbol{Y}|\mathcal{M}_0)}{p(\boldsymbol{Y}|\mathcal{M}_1)} = \frac{p(\mathcal{M}_0|\boldsymbol{Y})}{p(|\mathcal{M}_1|\boldsymbol{Y})} \Big/ \frac{p(\mathcal{M}_0)}{p(\mathcal{M}_1)}, \tag{6.1}$$

where $p(\boldsymbol{Y}|\mathcal{M}_\ell) = \int p(\boldsymbol{y}|\boldsymbol{\Theta}_\ell)p(\boldsymbol{\Theta}_\ell)d\boldsymbol{\Theta}_\ell$ for $\ell = 0, 1$ is the marginal likelihood of the data $\boldsymbol{Y}$ given the model $\mathcal{M}_\ell$, $p(\mathcal{M}_0|\boldsymbol{Y})/p(|\mathcal{M}_1|\boldsymbol{Y})$ is the posterior odds of $\mathcal{M}_0$ to $\mathcal{M}_1$, and $p(\mathcal{M}_0)/p(\mathcal{M}_1)$ is the prior odds of $\mathcal{M}_0$ to $\mathcal{M}_1$. The Bayes factor is the ratio of two marginal likelihoods.A scale for interpretation of Bayes factors (BF) (Kass and Raftery, 1995) is given in Table 6.1. A value of BF > 1 means that $\mathcal{M}_0$ is more strongly supported by the data under consideration than $\mathcal{M}_1$.

Let $\mathcal{M}_1$ denote a general model indexed by $\boldsymbol{\Theta} = (\boldsymbol{\omega}^T, \boldsymbol{\Upsilon}^T)^T$, where $\boldsymbol{\omega}$ denotes the vector of parameters of interest, $\boldsymbol{\Upsilon}$ denotes the vector of all the remaining "nuisance parameters", $p(\boldsymbol{\Theta}|\mathcal{M}_1)$ denotes the prior density under $\mathcal{M}_1$ and $p(\boldsymbol{Y}|\boldsymbol{\Theta}, \mathcal{M}_1)$ denote the sampling density under $\mathcal{M}_1$. A nested model, denoted $\mathcal{M}_0$, is constructed by setting $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, while leaving $\boldsymbol{\Upsilon}$ unconstrained: $p(\boldsymbol{Y}|\mathcal{M}_0) = p(\boldsymbol{Y}|\boldsymbol{\omega} = \boldsymbol{\omega}_0, \mathcal{M}_1)$. The prior density under $\mathcal{M}_0$ satisfies $p(\boldsymbol{\Upsilon}|\mathcal{M}_0) = p(\boldsymbol{\Upsilon}|\boldsymbol{\omega} = \boldsymbol{\omega}_0, \mathcal{M}_1)$ and the sampling density

Bayes Factor	Strength of Evidence
$-\infty < BF \leq 0.01$	Very strong against $\mathcal{M}_0$
$0.01 < BF \leq 0.1$	Strong against $\mathcal{M}_0$
$0.1 < BF \leq \frac{1}{3}$	Moderate against $\mathcal{M}_0$
$\frac{1}{3} < BF \leq 1$	Barely worth mentioning against $\mathcal{M}_0$
$1 < BF \leq 3$	Barely worth mentioning for $\mathcal{M}_0$
$3 < BF \leq 10$	Moderate for $\mathcal{M}_0$
$10 < BF < 100$	Strong for $\mathcal{M}_0$
$100 < BF < \infty$	Very strong for $\mathcal{M}_0$

Table 6.1: A scale for interpretation of Bayes factors (Kass and Raftery, 1995).

under $\mathcal{M}_0$ is $p(\boldsymbol{Y}|\boldsymbol{\Upsilon}, \mathcal{M}_0)$. From Bayes Theorem,

$$p(\boldsymbol{Y}|\boldsymbol{\omega} = \boldsymbol{\omega}_0, \mathcal{M}_1) = \frac{p(\boldsymbol{\omega} = \boldsymbol{\omega}_0|\boldsymbol{Y}, \mathcal{M}_1)p(\boldsymbol{Y}|\mathcal{M}_1)}{p(\boldsymbol{\omega} = \boldsymbol{\omega}_0|\mathcal{M}_1)}, \tag{6.2}$$

so the Bayes factor can be expressed as the Savage-Dickey density ratio (Dickey and Lientz, 1970; Verdinelli and Wasserman, 1995; Morey et al., 2011)

$$B_{01} = \frac{p(\boldsymbol{Y}|\mathcal{M}_0)}{p(\boldsymbol{Y}|\mathcal{M}_1)} = \frac{p(\boldsymbol{Y}|\boldsymbol{\omega} = \boldsymbol{\omega}_0, \mathcal{M}_1)}{p(\boldsymbol{Y}|\mathcal{M}_1)} = \frac{p(\boldsymbol{\omega} = \boldsymbol{\omega}_0|\boldsymbol{Y}, \mathcal{M}_1)}{p(\boldsymbol{\omega} = \boldsymbol{\omega}_0|\mathcal{M}_1)}. \tag{6.3}$$

The marginal prior density $p(\boldsymbol{\omega} = \boldsymbol{\omega}_0|\mathcal{M}_1)$ can be easily calculated from the prior. The marginal posterior, $p(\boldsymbol{\omega} = \boldsymbol{\omega}_0|\boldsymbol{Y}, \mathcal{M}_1)$, can be estimated using MCMC outputs from the unrestricted model, which provide approximate samples from the marginal posterior $p(\boldsymbol{\omega}|\boldsymbol{Y}, \mathcal{M}_1)$. Different methods to calculate the marginal posteriors include

1. The Normal approximation:

   As the marginal posterior $p(\boldsymbol{\omega}|\boldsymbol{Y})$ is often approximately normal, $p(\boldsymbol{\omega}|\boldsymbol{Y})$ can be approximated using a multivariate normal distribution with mean and variance equal to the posterior mean and variance of $\boldsymbol{\omega}$ estimated from the MCMC output and thus an estimate of $p(\boldsymbol{\omega}|\boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ can be obtained. Deviations of the

posterior from normality can lead to problems in estimating the Savage-Dickey density ratio.

2. Conditional marginal density estimation (CMDE):
   When the full conditional posterior distribution $p(\boldsymbol{\omega}|\boldsymbol{\Upsilon}, \boldsymbol{Y})$ is known completely, the marginal posterior density at $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ can be approximated by an average of the full conditional posterior density of $\boldsymbol{\omega}$ at $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ over all iterations

$$\begin{aligned} p(\boldsymbol{\omega}|\boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \int_{\Upsilon} p(\boldsymbol{\omega}|\boldsymbol{\Upsilon}, \boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} p(\boldsymbol{\Upsilon}|\boldsymbol{Y}) d\boldsymbol{\Upsilon} \\ &= E_{\Upsilon|y}[p(\boldsymbol{\omega}|\boldsymbol{\Upsilon}, \boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}] \\ &\approx \frac{1}{T}\sum_{t=1}^{T} p(\boldsymbol{\omega}|\boldsymbol{\Upsilon}^{(t)}, \boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}, \end{aligned} \tag{6.4}$$

   where $\boldsymbol{\Upsilon}^{(t)}$ is the component of $\boldsymbol{\Upsilon}$ from the t^{th} MCMC sample, out of a total of T MCMC iterations (Morey et al., 2011).

3. Multivariate kernel density estimation (KDE):
   Using the MCMC outputs, one can obtain a kernel density estimate of $p(\boldsymbol{\omega}|\boldsymbol{Y})$ evaluating at $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. This can be implemented using nonparametric kernel smoothing package np in R. KDE always over-estimates the variance (Morey et al., 2011).

6.1 Testing Hypotheses for Familial Data with Multiple Outcomes

In the UCLA Family Study, hypotheses of interest include whether group means are equal (1) between schizophrenia and control families and (2) across family member types within schizophrenia families. Both scenarios are equivalent to testing whether particular linear combinations of the regression coefficients are simultaneously equal to zero

$$\boldsymbol{\omega}_{(l\times 1)} = \boldsymbol{L}_{(l\times P)}\boldsymbol{\beta}_{(P\times 1)} = \mathbf{0}_{(l\times 1)},$$

where $\boldsymbol{L}$ is a full rank matrix. Let $\boldsymbol{M}$ denote any $(P-l)\times P$ full rank matrix so that rank($[\boldsymbol{L}^T$,$\boldsymbol{M}^T]$) $= P$ and let $\boldsymbol{\omega}^{\perp} = \boldsymbol{M}\boldsymbol{\beta}$. Let $\mathcal{M}_1$ denote a general model where $\boldsymbol{\beta}$ is freely estimated, which has parameters $\boldsymbol{\Theta} = (\boldsymbol{\beta}, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Phi}_A, \boldsymbol{\Phi}_B, \boldsymbol{F}_A, \boldsymbol{F}_B, \boldsymbol{\Psi})$ and let $\boldsymbol{\Upsilon} = (\boldsymbol{\omega}^{\perp}, \boldsymbol{\Lambda}_A, \boldsymbol{\Lambda}_B, \boldsymbol{\Phi}_A, \boldsymbol{\Phi}_B, \boldsymbol{F}_A, \boldsymbol{F}_B, \boldsymbol{\Psi})$. Then $\boldsymbol{\Theta}$ can be reparametrized as $\boldsymbol{\Theta}^* = (\boldsymbol{\omega}, \boldsymbol{\Upsilon})$. A nested model, denoted $\mathcal{M}_0$, is constructed by setting $\boldsymbol{\omega} = \boldsymbol{L}\boldsymbol{\beta} = \boldsymbol{0}$. As

$$\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\mu}_{\beta 0}, \boldsymbol{\Sigma}_{\beta 0}),$$

the prior distribution of $\boldsymbol{\omega} = \boldsymbol{L}\boldsymbol{\beta}$ is

$$\boldsymbol{\omega}|\mathcal{M}_1 \sim \mathcal{N}(\boldsymbol{L}\boldsymbol{\mu}_{\beta 0}, \boldsymbol{L}\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{L}^T),$$

so we can obtain the denominator of the Savage-Dickey density ratio, $p(\boldsymbol{\omega}|\mathcal{M}_1)|_{\boldsymbol{\omega}=\boldsymbol{0}}$.

Let $\widehat{\boldsymbol{\mu}}_{\beta}$ and $\widehat{\boldsymbol{\Sigma}}_{\beta}$ denote the posterior mean and variance of $\boldsymbol{\beta}$ estimated from the MCMC output. The marginal posterior distribution can be approximated as

$$\boldsymbol{\omega}|\boldsymbol{Y}, \mathcal{M}_1 \overset{\text{approx.}}{\sim} \mathcal{N}(\boldsymbol{L}\widehat{\boldsymbol{\mu}}_{\beta}, \boldsymbol{L}\widehat{\boldsymbol{\Sigma}}_{\beta}\boldsymbol{L}^T),$$

which gives an approximation to $p(\boldsymbol{\omega}|\boldsymbol{Y}, \mathcal{M}_1)|_{\boldsymbol{\omega}=\boldsymbol{0}}$.

The conditional marginal density estimator (CMDE) approximates $p(\boldsymbol{\omega}|\boldsymbol{Y}, \mathcal{M}_1)|_{\boldsymbol{\omega}=\boldsymbol{0}}$ using an average of the full conditional posterior density of $\boldsymbol{\omega}$ evaluated at $\boldsymbol{\omega} = \boldsymbol{0}$ over all T MCMC iterations

$$p(\boldsymbol{\omega} = \boldsymbol{0}|\boldsymbol{Y}, \mathcal{M}_1) \;\approx\; \tfrac{1}{T}\sum_{t=1}^{T} p(\boldsymbol{\omega}|\boldsymbol{\Upsilon}, \boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{0}, \boldsymbol{\Upsilon}=\boldsymbol{\Upsilon}^{(t)}}, \tag{6.5}$$

where $\boldsymbol{\Upsilon}^{(t)}$ is value of $\boldsymbol{\Upsilon}$ from the t^{th} MCMC sample. The full conditional posterior density of $\boldsymbol{\omega}$ is

$$p(\boldsymbol{\omega}|\boldsymbol{\Upsilon}, \boldsymbol{Y}) = p(\boldsymbol{\omega}|\boldsymbol{M}\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{Y}),$$

where $\boldsymbol{\Sigma} = \text{var}(\boldsymbol{y}_i|\boldsymbol{\beta}) = \boldsymbol{\Lambda}_A\boldsymbol{\Phi}_A\boldsymbol{\Lambda}_A^T + \boldsymbol{\Lambda}_B\boldsymbol{\Phi}_B\boldsymbol{\Lambda}_B^T + \boldsymbol{\Psi}$ is the variance of $\boldsymbol{y}_i$ given $\boldsymbol{\beta}$.

The full conditional posterior distribution of $\boldsymbol{\beta}$ given $\boldsymbol{\Sigma}$ is multivariate normal

$$\boldsymbol{\beta}|\boldsymbol{\Sigma}, \boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}_{\beta p}, \boldsymbol{\Sigma}_{\beta p}),$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_{\beta p} &= (\boldsymbol{\Sigma}_{\beta 0}^{-1} + \sum_{i=1}^{N} \boldsymbol{X}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{X}_i)^{-1}, \\ \text{and} \\ \boldsymbol{\mu}_{\beta p} &= \boldsymbol{\Sigma}_{\beta p}(\boldsymbol{\Sigma}_{\beta 0}^{-1}\boldsymbol{\mu}_{\beta 0} + \sum_{i=1}^{N} \boldsymbol{X}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y}_i), \end{aligned}$$

as described in Equation 2.4.2 in Section 2.4. The joint posterior distribution of $\boldsymbol{\omega} = \boldsymbol{L}\boldsymbol{\beta}$ and $\boldsymbol{\omega}^{\perp} = \boldsymbol{M}\boldsymbol{\beta}$ conditional on $\boldsymbol{\Sigma}$ and $\boldsymbol{Y}$ is multivariate normal

$$\begin{bmatrix} \boldsymbol{L}\boldsymbol{\beta} \\ \boldsymbol{M}\boldsymbol{\beta} \end{bmatrix} |\boldsymbol{\Sigma}, \boldsymbol{Y} \sim \mathcal{MVN} \left(\begin{bmatrix} \boldsymbol{L}\boldsymbol{\mu}_{\beta p} \\ \boldsymbol{M}\boldsymbol{\mu}_{\beta p} \end{bmatrix}, \begin{bmatrix} \boldsymbol{L}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{L}^T & \boldsymbol{L}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{M}^T \\ \boldsymbol{M}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{L}^T & \boldsymbol{M}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{L}^T \end{bmatrix} \right).$$

Therefore the full conditional posterior distribution of $\boldsymbol{\omega}$ given $\boldsymbol{\omega}^{\perp}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{Y}$ is

$$\boldsymbol{\omega}|\boldsymbol{\omega}^{\perp} = \boldsymbol{M}\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{Y} \sim \mathcal{N}\left(\boldsymbol{\mu}_{L\beta}, \boldsymbol{\Sigma}_{L\beta}\right)$$

where

$$\begin{aligned} \boldsymbol{\mu}_{L\beta} &= \boldsymbol{L}\boldsymbol{\mu}_{\beta p} + \boldsymbol{L}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{M}^T(\boldsymbol{M}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{M}^T)^{-1}\boldsymbol{M}(\boldsymbol{\beta} - \boldsymbol{\mu}_{\beta p}), \\ \boldsymbol{\Sigma}_{L\beta} &= \boldsymbol{L}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{L}^T - \boldsymbol{L}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{M}^T(\boldsymbol{M}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{M}^T)^{-1}\boldsymbol{M}\boldsymbol{\Sigma}_{\beta p}\boldsymbol{L}^T, \end{aligned} \tag{6.6}$$

.

Therefore, the conditional marginal density estimator of $\boldsymbol{\omega}$ is

$$p(\boldsymbol{\omega} = \boldsymbol{0}|\boldsymbol{Y}, \mathcal{M}_1) \approx \frac{1}{T}\sum_{t=1}^{T} p(\boldsymbol{\omega}|\boldsymbol{\omega}^{\perp}, \boldsymbol{\Sigma}, \boldsymbol{Y})|_{\boldsymbol{\omega}=\boldsymbol{0}, \boldsymbol{\omega}^{\perp}=\boldsymbol{M}\boldsymbol{\beta}^{(t)}, \boldsymbol{\Sigma}=\boldsymbol{\Sigma}^{(t)}}, \tag{6.7}$$

where $\boldsymbol{\beta}^{(t)}$ is the component of $\boldsymbol{\beta}$ from the t^{th} iteration of MCMC output, $\boldsymbol{\Sigma}^{(t)} = \boldsymbol{\Lambda}_A^{(t)}\boldsymbol{\Phi}_A^{(t)}\boldsymbol{\Lambda}_A^{(t)T} + \boldsymbol{\Lambda}_B^{(t)}\boldsymbol{\Phi}_B^{(t)}\boldsymbol{\Lambda}_B^{(t)T} + \boldsymbol{\Psi}^{(t)}$ is calculated from the t^{th} iteration of the MCMC

output.

6.2 Illustration of Hypothesis Testing

For the UCLA Neurocognitive Family Study, researchers are interested in differences in cognitive measurements both between the control and SZ families and among different members in the SZ family. I illustrate this with 3 prototypical examples

1. Testing whether the means of all outcomes for SZ and control probands are equal (the number of contraints, NC $= K = 5$). The linear combination of interest is

$$\boldsymbol{L}_{1(K\times 2JK)} = \left[\begin{array}{cccccccc} \boldsymbol{I}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} & -\boldsymbol{I}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} \end{array} \right],$$

where $\boldsymbol{0}_{K\times K}$ is a $K \times K$ matrix of 0's. A full rank matrix, $\boldsymbol{M}_{1((2J-1)K\times 2JK)}$, so that rank$([\boldsymbol{L}_1^T, \boldsymbol{M}_1^T]) = 2JK$ is

$$\boldsymbol{M}_1 = \left[\begin{array}{cccccccc} \boldsymbol{I}_{K\times K} & & & & \boldsymbol{I}_{K\times K} & & & \\ & \boldsymbol{I}_{K\times K} & & & & & & \\ & & \boldsymbol{I}_{K\times K} & & & & & \\ & & & \boldsymbol{I}_{K\times K} & & & & \\ & & & & & \boldsymbol{I}_{K\times K} & & \\ & & & & & & \boldsymbol{I}_{K\times K} & \\ & & & & & & & \boldsymbol{I}_{K\times K} \end{array} \right],$$

where all the rest of the elements in the matrix are 0.

2. Testing whether the means of SZ probands are equal to the means of the average of siblings, fathers and mothers of SZ probands (NC $= K = 5$). The linear combination of interest is

$$\boldsymbol{L}_{1(K\times 2JK)} = \left[\begin{array}{cccccccc} \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times K} & \boldsymbol{I}_K & -\frac{1}{3}\boldsymbol{I}_K & -\frac{1}{3}\boldsymbol{I}_K & -\frac{1}{3}\boldsymbol{I}_K \end{array} \right],$$

where $\mathbf{0}_{K\times K}$ is a $K \times K$ matrix of 0's;

A full rank matrix, $\boldsymbol{M}_{2((2J-1)K\times 2JK)}$, so that rank$([\boldsymbol{L}_2^T, \boldsymbol{M}_2^T]) = 2JK$ is

$$\boldsymbol{M}_2 = \begin{bmatrix} \boldsymbol{I}_{K\times K} & & & & & & & \\ & \boldsymbol{I}_{K\times K} & & & & & & \\ & & \boldsymbol{I}_{K\times K} & & & & & \\ & & & \boldsymbol{I}_{K\times K} & & & & \\ & & & & \boldsymbol{I}_{K\times K} & 3\boldsymbol{I}_{K\times K} & & \\ & & & & \boldsymbol{I}_{K\times K} & & 3\boldsymbol{I}_{K\times K} & \\ & & & & \boldsymbol{I}_{K\times K} & & & 3\boldsymbol{I}_{K\times K} \end{bmatrix},$$

3. Testing whether for a particular variable, say CPT37D ($k = 3$), the means of SZ and control families are equal across all members (NC $= J = 4$). Let $\boldsymbol{d} = (0, 0, 1, 0, 0)$ be a 1×5 row vector, then the linear combination of interest is

$$\boldsymbol{L}_3 = \begin{bmatrix} \boldsymbol{I}_J \otimes \boldsymbol{d} & -\boldsymbol{I}_J \otimes \boldsymbol{d} \end{bmatrix}.$$

Let

$$\boldsymbol{m}_{3((K-1)\times K)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

A full rank matrix, $\boldsymbol{M}_{3((2K-1)J\times 2JK)}$, so that rank$([\boldsymbol{L}_3^T, \boldsymbol{M}_3^T]) = 2JK$ is

$$\boldsymbol{M}_3 = \begin{bmatrix} \boldsymbol{I}_J \otimes \boldsymbol{d} & \boldsymbol{I}_J \otimes \boldsymbol{d} \\ \boldsymbol{I}_J \otimes \boldsymbol{m}_3 & \mathbf{0}_{J(K-1)\times JK} \\ \mathbf{0}_{J(K-1)\times JK} & \boldsymbol{I}_J \otimes \boldsymbol{m}_3 \end{bmatrix}.$$

6.2.1 Comparing Different Methods Using Simulated Data Sets

First, I analyze simulated data sets to evaluate the performance of the normal approximation estimator, the conditional marginal density estimator (CMDE) and the kernel density estimator (KDE) for Bayes factors (BF) corresponding to matrices $\boldsymbol{L}_1$, $\boldsymbol{L}_2$ and $\boldsymbol{L}_3$. As Bayes factors for a single data set can be affected greatly by random errors, I calculate the BFs for each method using the same 200 data sets with $K = 5$ outcomes and $J = 4$ family members simulated in the scenario where the true covariate matrix is close to non-positive definite, as described and analyzed in Section 3.2. I choose the scenario in which the true covariance matrix is close to non-positive definite, so that I can examine the potential differences in the estimates of BFs between the data sets which the quasi-Newton optimization (QNO) fails and those which the QNO does not fail to fit.

The true parameters values for the group means of the $K = 5$ outcomes on $J = 4$ family members are listed in Table 6.2.1, reflecting more severe neurocognitive deficits in SZ probands and SZ siblings compared with other family members in the control and SZ families. Therefore, I will expect the BFs corresponding to all 3 matrices to be strongly against the null hypotheses of no difference, $\boldsymbol{L}\boldsymbol{\beta} = \boldsymbol{0}$.

Figure 6.1 shows scatter plots of the BFs using the normal approximation method on the x-axis against those using the CDME (red) or the KDE (green) on the y-axis, for all 3 hypotheses. If the points are close to the diagonal line with slope 1, the CDME or KDE estimates are close to the normal approximation estimates. Circles represent data sets which the FIML-QNO was successful in fitting, while crosses represent data sets which the FIML-QNO failed to fit

In general, the BFs calculated using the normal approximation and the CDME are very similar, as the red points are mostly along the diagonal, while the KDE sometimes produces much smaller BFs (green points below the diagonal line). For hypotheses 1 and 2, all BF are $< 10^{-2}$, which is consistent with the setting of true parameters having

Group	Member	Outcome 1	Outcome 2	Outcome 3	Outcome 4	Outcome 5
Control	Proband	1	2	5	8	10
	Sibling	1	2	5	8	10
	Father	1	2	5	8	10
	Mother	1	2	5	8	10
SZ	Proband	0.5	1	2.5	4	5
	Sibling	0.75	1.5	4	6	8
	Father	1	2	5	8	10
	Mother	1	2	5	8	10

Table 6.2: True parameter values for group means of $K = 5$ outcomes on $J = 4$ family members in the control and SZ families for the 200 simulated data sets analyzed.

SZ probands different from their relatives and also different from control probands. For hypothesis 3, testing equality in means of the 3^{rd} outcome between the SZ and control families across all family members, BFs for some data sets are greater than 1, which is possible because the means of the 3^{rd} outcome for parents are set to be equal in both SZ and control families. Looking at Figure 6.1, there are no obvious differences between the data sets on which the FIML-QNO was successful or not.

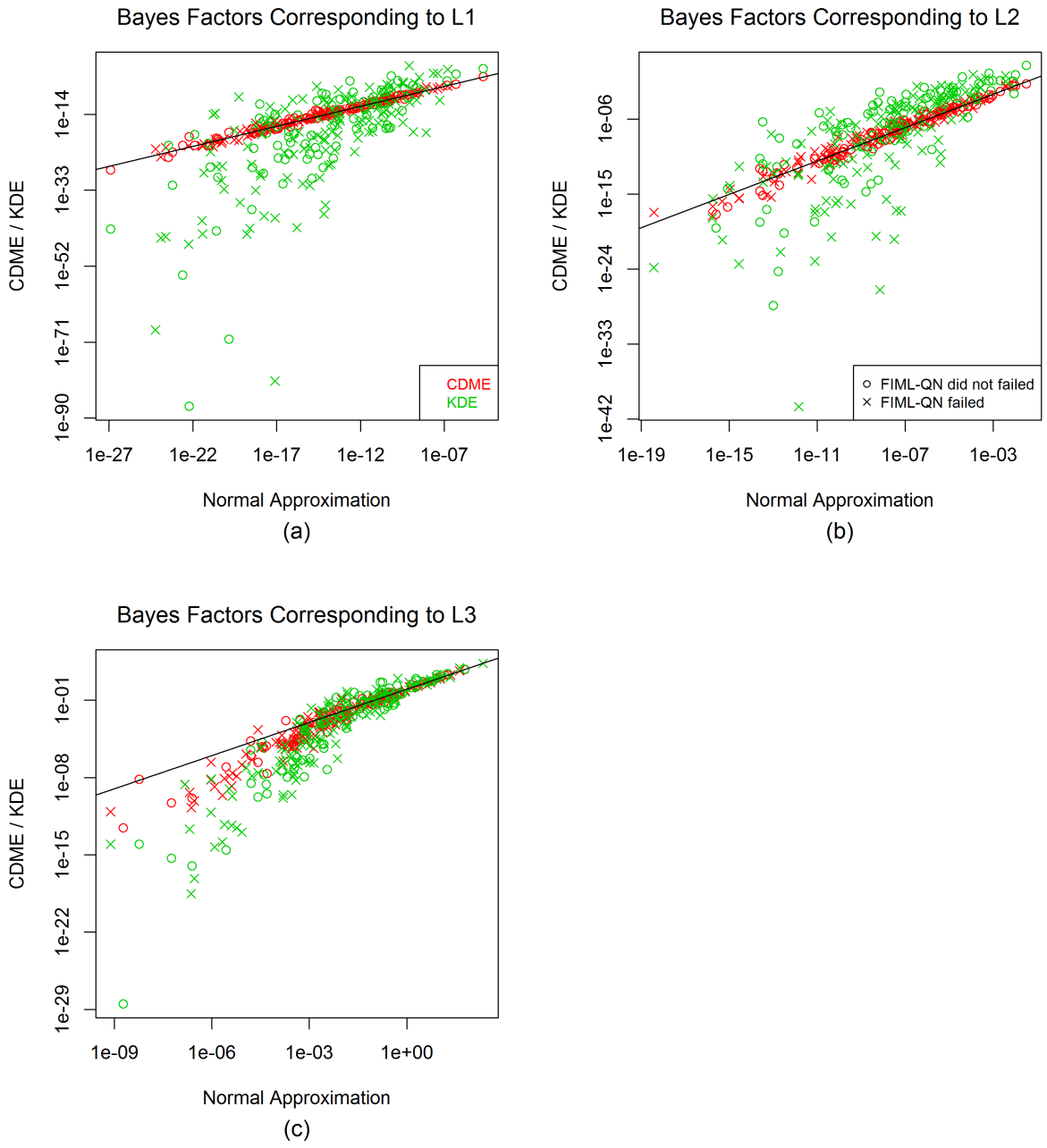


Figure 6.1: Scatter plots of BFs using normal approximation method on the x-axis against those using the CDME (red) and the KDE (green) on the y-axis, for all 3 hypotheses. If the points are close to the diagonal line, the CDME or KDE estimates are close to the normal approximation estimates. Circles represent data sets which the FIML-QNO was successful in fitting, while crosses represent data sets which the FIML-QNO failed to fit.

Hypothesis	df	Bayes Factor	Strength of Evidence
$\boldsymbol{L}_1$	5	8.97	Barely worth mentioning for
$\boldsymbol{L}_2$	5	0.00014	Strong against
$\boldsymbol{L}_3$	4	0.0010	Strong against

Table 6.3: Bayes factors for testing different hypotheses on the UCLA Family Study Data. The hypothesis associated with $\boldsymbol{L}_1$ tests whether for the probands, the means of all outcomes in the schizophrenia and control probands are equal (NC $= 5$). The hypothesis associated with $\boldsymbol{L}_2$ tests whether the means of schizophrenia probands are equal to the means of the average of siblings, fathers and mothers of schizophrenia probands (the number of constraints, NC $= 5$). The hypothesis associated with $\boldsymbol{L}_3$ tests whether for the memory load CPT, the means of the schizophrenia and control families are equal across all members (NC $= 4$).

6.2.2 Results for the UCLA Neurocognitive Family Study

Similar Bayes factors (BF) are obtained using the three methods, (the normal approximation, KDE and CMDE), for estimating $p(\boldsymbol{\omega}|\boldsymbol{Y}, \mathcal{M}_0)|_{\boldsymbol{\omega}=\mathbf{0}}$, so only the results using the normal approximation are presented in Table 6.2.2. The BF corresponding to $\boldsymbol{L}_1$ is 8.97, which suggests that the differences in the 5×1 vectors of means of the probands between the two groups are "barely worth mentioning" (Kass and Raftery, 1995), even if it points slightly towards the hypothesis of equal means. The BF corresponding to $\boldsymbol{L}_2$ (NC=5) is 0.0014, suggesting that the means of probands are quite different from those of their relatives for schizophrenia families. The BF corresponding to $\boldsymbol{L}_3$ (NC=4) is 0.0010, suggesting strong evidence against equality in means of CPT37D between SZ and control families, which is consistent with the 1-sided p-values, $p(\beta_{113} - \beta_{213} > 0|\boldsymbol{Y})$, for testing the difference in means between the two groups.

This hypothesis testing approach can be generalized to test equality constraints across family members on the following parameters: the family factor loadings $\boldsymbol{a}_j$, the outcome factor loadings $\boldsymbol{B}_j$, the error variances $\boldsymbol{\Psi}_j$ and the mean parameters $\boldsymbol{\mu}_j$.

CHAPTER 7

Discussion

I propose the Bayesian Family Factor Model (BFFM), which extends the classical confirmatory factor analysis (CFA) to jointly model multiple outcomes in familial data. This model explains the covariances among observed variables using a combination of family-member factors and outcome factors. Bayesian methods incorporating informative priors mitigate the problem of empirical under-identification, non-convergence and invalid solutions. The choice of conditionally conjugate priors enable the sampling from the posterior distributions using a Gibbs algorithm. The proposed framework has the advantage of being able to handle missing data, incorporate mean structure and test hypotheses easily.

I performed simulations to compare the BFFM to the full information likelihood (FIML) estimation using quasi-Newton algorithm, in settings that the true covariance matrix is not identified, close to not identified and identified. For these settings, the quasi-Newton algorithm fails to find a fit to the data in 85%, 57% and 13% of the cases, respectively, due to non-convergence or invalid estimates, while the BFFM provides stable estimates. Moreover, when both methods successfully fit the data, the BFFM estimates have smaller variances, as well as comparable mean squared errors and bias squared. The BFFM is used to analyze the UCLA Neurocognitive Family Study data to compare the degree of abnormality between schizophrenia families and control families and to determine correlations among measurements from relatives using hypothesis testing.

In the current analyses, factors are assumed positively correlated to the observed

variables. Choosing positive truncated normal priors can solve the problem of improper solutions where factor loading estimates are negative. In the current study, I use normal priors which are conditionally conjugate priors for the ease of computation. The Bayesian model can be modified to use positive truncated normal priors in the future.

An interesting extension to the current model would be to perform Bayesian model selection and Bayesian model averaging on different sub-models. For example, sub-models with different equality constraints across family members on regression coefficients, factor loadings, and unique errors variances. Other extensions include incorporating covariates in the factor loading parameters and the residual variances, to examine covariate effects on variance structure. This model can also be modified to fit other kinds of data with similar structure, like the multitrait-multimethod (MTMM) data used to examine construct validity in psychology. Finally, the current model is developed under the assumption that all outcomes are normally distributed. It will be useful to extend the current model to analyze mixed types of outcomes, for example normal, Poisson, exponential, gamma and binomial, which are all within the exponential family, by modeling the canonical parameters.

APPENDIX A

Values of True Parameters for Simulation Studies

		Group Means				
		Outcome 1	Outcome 2	Outcome 3	Outcome 4	Outcome 5
Ctrl	Member 1	1	2	5	8	10
	Member 2	1	2	5	8	10
	Member 3	1	2	5	8	10
	Member 4	1	2	5	8	10
SZ	Member 1	0.5	1	2.5	4	5
	Member 2	0.75	1.5	4	6	8
	Member 3	1	2	5	8	10
	Member 4	1	2	5	8	10
		Measurement Error Variance				
		Outcome 1	Outcome 2	Outcome 3	Outcome 4	Outcome 5
		0.1	0.4	2.5	6.4	10

Table A.1: Values of true parameters for simulation studies.

Factor Varaince and Covariance	Identified	Close to Under-identified	Under-identified
ϕ_{A11}	1.959	1.271	1.076
ϕ_{A12}	0.488	0.563	0.549
ϕ_{A22}	0.651	1.284	1.002
ϕ_{A13}	0.585	0.483	0.583
ϕ_{A23}	0.484	0.550	0.430
ϕ_{A33}	1.422	1.275	1.238
ϕ_{A14}	0.520	0.601	0.510
ϕ_{A24}	0.670	0.631	0.634
ϕ_{A34}	0.553	0.581	0.389
ϕ_{A44}	0.832	0.880	1.144
ϕ_{B11}	1.523	1.221	1.172
ϕ_{B12}	0.206	0.316	0.373
ϕ_{B22}	0.961	1.008	1.174
ϕ_{B13}	0.129	0.227	0.306
ϕ_{B23}	0.327	0.243	0.218
ϕ_{B33}	0.615	1.544	0.504
ϕ_{B14}	0.254	0.471	0.310
ϕ_{B24}	0.339	0.139	0.341
ϕ_{B34}	0.211	0.219	0.167
ϕ_{B44}	1.422	1.002	0.961
ϕ_{B15}	0.309	0.258	0.319
ϕ_{B25}	0.098	0.240	0.283
ϕ_{B35}	0.138	0.298	0.251
ϕ_{B45}	0.298	0.342	0.361
ϕ_{B55}	1.960	0.795	0.883

Table A.2: Values of true parameters for simulation studies.

Factor Loading	Identified	Close to Under-identified	Under-identified
a_{12}	2.235	2.517	2
a_{13}	2.848	2.691	3
a_{14}	5.229	3.515	4
a_{15}	3.775	5.856	5
a_{22}	2.409	1.407	2
a_{23}	3.978	4.634	3
a_{24}	4.162	6.214	4
a_{25}	6.126	3.265	5
a_{32}	1.401	1.345	2
a_{33}	1.307	2.835	3
a_{34}	2.839	2.038	4
a_{35}	4.344	5.303	5
a_{42}	2.599	1.484	2
a_{43}	3.029	3.869	3
a_{44}	3.110	3.509	4
a_{45}	6.461	3.563	5
b_{21}	1.230	1.215	1.1
b_{22}	1.664	1.492	1.1
b_{23}	0.670	1.655	1.1
b_{24}	-0.060	0.604	1.1
b_{25}	1.384	0.985	1.1
b_{31}	1.790	1.361	1.2
b_{32}	0.793	1.673	1.2
b_{33}	1.682	0.876	1.2
b_{34}	1.832	1.077	1.2
b_{35}	1.589	1.418	1.2
b_{41}	1.436	1.796	1.3
b_{42}	1.545	2.199	1.3
b_{43}	1.250	1.977	1.3
b_{44}	1.661	1.054	1.3
b_{45}	2.356	2.057	1.3

Table A.3: Values of true parameters for simulation studies.

APPENDIX B

The Performance of BFFM and Quasi-Newton Optimization in Simulation Studies

The relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in 3 scenarios where the true covariance matrix is identified, close to not identified and not identified are summarized in tables below.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
ψ_1	1.20E+00	1.17E+00	3.58E-02	4.56E-02	1.42E-02	3.15E-02	3.12E-02	9.83E-03	2.18E-02
ψ_2	9.52E-02	8.85E-02	7.14E-03	1.63E-02	1.33E-02	3.13E-03	1.59E-02	1.41E-02	2.35E-03
ψ_3	1.18E-03	1.10E-03	8.86E-05	6.63E-03	6.08E-03	5.85E-04	6.15E-03	6.39E-03	2.30E-06
ψ_4	2.02E-04	1.44E-04	5.88E-05	1.01E-02	6.04E-03	4.14E-03	1.44E-02	5.17E-03	9.39E-03
ψ_5	7.52E-05	6.55E-05	1.01E-05	9.46E-03	7.10E-03	2.41E-03	9.54E-03	6.25E-03	3.53E-03
ϕ_{A11}	7.35E-03	7.34E-03	5.55E-05	6.11E-02	1.72E-02	4.39E-02	6.18E-02	1.60E-02	4.64E-02
ϕ_{A12}	1.47E-01	1.47E-01	1.08E-03	1.20E-01	1.21E-02	1.08E-01	1.11E-01	1.23E-02	9.93E-02
ϕ_{A22}	1.56E-02	1.56E-02	1.36E-04	3.91E-02	1.48E-03	3.76E-02	3.55E-02	1.73E-03	3.38E-02
ϕ_{A13}	2.75E-01	2.73E-01	3.72E-03	3.08E-01	3.36E-02	2.75E-01	2.19E-01	2.69E-02	1.93E-01
ϕ_{A23}	1.32E-01	1.33E-01	2.50E-04	2.17E-01	6.16E-03	2.11E-01	1.99E-01	5.16E-03	1.94E-01
ϕ_{A33}	1.07E-02	1.07E-02	1.01E-04	1.94E-01	1.09E-02	1.84E-01	1.59E-01	7.78E-03	1.51E-01
ϕ_{A14}	1.39E-01	1.40E-01	5.06E-04	1.26E-01	1.32E-02	1.12E-01	1.32E-01	1.28E-02	1.20E-01
ϕ_{A24}	6.17E-02	6.17E-02	4.23E-04	1.83E-01	6.82E-03	1.77E-01	1.80E-01	8.38E-03	1.72E-01
ϕ_{A34}	1.16E-01	1.16E-01	1.50E-05	2.39E-01	6.82E-03	2.33E-01	2.31E-01	7.43E-03	2.24E-01
ϕ_{A44}	2.98E-02	2.98E-02	2.06E-04	1.30E-01	5.46E-03	1.24E-01	1.36E-01	6.35E-03	1.30E-01

Table B.1: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is identified. The parameters compared are unique error variances, ψ_k and family factor vairances and covariances, ϕ_{Alm}, for $l = 1, \ldots, 4$ and $m = 1, \ldots, l$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
ϕ_{B11}	1.03E-02	1.03E-02	2.31E-05	5.89E-02	2.19E-02	3.71E-02	4.53E-02	2.29E-02	2.33E-02
ϕ_{B12}	3.21E+00	3.23E+00	8.48E-04	8.33E-01	1.75E-01	6.59E-01	7.00E-01	1.85E-01	5.22E-01
ϕ_{B22}	6.40E-02	6.42E-02	1.37E-04	4.72E-01	4.90E-02	4.23E-01	4.26E-01	3.53E-02	3.92E-01
ϕ_{B13}	2.14E+01	2.15E+01	1.25E-02	1.26E+00	4.05E-01	8.60E-01	7.67E-01	2.38E-01	5.38E-01
ϕ_{B23}	5.53E+00	5.56E+00	6.21E-03	4.60E+00	6.46E-01	3.96E+00	3.65E+00	4.16E-01	3.25E+00
ϕ_{B33}	1.57E-01	1.56E-01	2.21E-03	5.70E-02	1.52E-02	4.19E-02	4.48E-02	1.01E-02	3.51E-02
ϕ_{B14}	2.81E+00	2.81E+00	1.72E-02	5.73E-01	1.86E-01	3.89E-01	3.60E-01	1.34E-01	2.31E-01
ϕ_{B24}	3.07E+01	3.05E+01	3.00E-01	2.36E+01	3.43E+00	2.02E+01	2.04E+01	3.24E+00	1.73E+01
ϕ_{B34}	3.58E+01	3.58E+01	1.26E-01	4.65E+00	1.09E+00	3.57E+00	3.45E+00	6.32E-01	2.84E+00
ϕ_{B44}	4.88E-01	4.76E-01	1.43E-02	3.15E-01	1.35E-01	1.80E-01	1.11E-01	7.98E-02	3.44E-02
ϕ_{B15}	8.76E+00	8.65E+00	1.54E-01	3.02E+00	1.14E+00	1.89E+00	2.23E+00	9.58E-01	1.31E+00
ϕ_{B25}	1.09E+02	1.10E+02	6.75E-02	8.95E+00	2.40E+00	6.57E+00	8.58E+00	2.32E+00	6.35E+00
ϕ_{B35}	4.62E+01	4.64E+01	1.86E-02	2.00E+00	1.07E+00	9.31E-01	1.82E+00	1.03E+00	8.33E-01
ϕ_{B45}	1.42E+01	1.43E+01	1.05E-04	2.31E+00	1.54E+00	7.76E-01	2.22E+00	1.52E+00	7.64E-01
ϕ_{B55}	4.60E-01	4.62E-01	8.26E-04	1.91E+00	7.31E-01	1.19E+00	2.28E+00	1.02E+00	1.30E+00

Table B.2: Comparison of the relative mean square errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is identified. The parameters compared are outcome factor variances and covariances, ϕ_{Blm}, for $l = 1, \ldots, 5$ and $m = 1, \ldots, l$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
β_{111}	3.90E-02	3.89E-02	3.74E-04	2.05E-02	2.05E-02	1.91E-04	8.59E-03	7.94E-03	9.60E-04
β_{112}	5.98E-03	5.96E-03	6.06E-05	1.20E-02	1.20E-02	8.19E-05	1.74E-02	1.80E-02	4.77E-05
β_{113}	2.38E-04	2.36E-04	2.46E-06	3.16E-03	3.16E-03	2.32E-05	5.91E-03	6.11E-03	3.68E-05
β_{114}	1.35E-04	1.33E-04	3.40E-06	4.57E-03	4.45E-03	1.38E-04	6.38E-03	6.62E-03	8.36E-06
β_{115}	4.08E-05	4.04E-05	6.74E-07	2.49E-03	2.47E-03	3.34E-05	3.37E-03	3.50E-03	3.34E-06
β_{121}	3.81E-02	3.83E-02	3.18E-05	2.31E-02	2.32E-02	5.55E-05	1.29E-02	1.22E-02	1.18E-03
β_{122}	3.97E-03	3.97E-03	1.60E-05	1.02E-02	1.02E-02	6.28E-05	1.73E-02	1.80E-02	1.35E-05
β_{123}	2.08E-04	2.09E-04	1.19E-07	3.49E-03	3.51E-03	3.16E-07	2.89E-03	3.00E-03	6.16E-07
β_{124}	4.23E-05	4.23E-05	2.04E-07	1.91E-03	1.91E-03	6.45E-06	2.30E-03	2.36E-03	3.73E-05
β_{125}	3.55E-05	3.57E-05	2.78E-08	2.21E-03	2.22E-03	8.13E-07	2.70E-03	2.77E-03	3.60E-05
β_{131}	7.42E-02	7.39E-02	6.78E-04	4.54E-02	4.53E-02	4.16E-04	5.40E-02	5.31E-02	3.01E-03
β_{132}	2.28E-03	2.23E-03	5.84E-05	6.56E-03	6.45E-03	1.44E-04	7.82E-03	7.92E-03	2.00E-04
β_{133}	1.32E-04	1.32E-04	1.04E-06	2.73E-03	2.73E-03	1.75E-05	2.29E-03	2.28E-03	9.50E-05
β_{134}	4.68E-05	4.60E-05	1.04E-06	2.33E-03	2.30E-03	3.60E-05	2.31E-03	2.38E-03	1.79E-05
β_{135}	4.14E-05	4.07E-05	8.56E-07	2.98E-03	2.96E-03	3.96E-05	3.66E-03	3.79E-03	1.34E-05
β_{141}	4.95E-02	4.97E-02	8.57E-05	2.97E-02	2.97E-02	1.14E-04	2.34E-02	1.95E-02	4.64E-03
β_{142}	4.52E-03	4.52E-03	2.78E-05	1.12E-02	1.11E-02	1.50E-04	1.33E-02	1.36E-02	2.47E-04
β_{143}	1.31E-04	1.31E-04	1.63E-07	2.18E-03	2.19E-03	7.49E-06	3.15E-03	3.27E-03	6.79E-07
β_{144}	3.81E-05	3.84E-05	9.48E-12	1.73E-03	1.74E-03	3.48E-07	2.02E-03	2.00E-03	9.35E-05
β_{145}	5.42E-05	5.45E-05	1.58E-08	3.42E-03	3.44E-03	3.85E-06	3.43E-03	3.51E-03	5.95E-05

Table B.3: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is identified. The parameters compared are regression coefficients for the control families, β_{pjk}, for $j = 1, \ldots, 4$, $k = 1, \ldots, K$ and $p = 1$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
β_{211}	5.23E-01	5.18E-01	8.07E-03	2.99E-01	6.69E-02	2.32E-01	2.42E-01	6.87E-02	1.76E-01
β_{212}	1.15E-01	1.15E-01	1.10E-03	3.33E-01	5.76E-02	2.76E-01	2.51E-01	4.80E-02	2.05E-01
β_{213}	5.36E-03	5.34E-03	4.59E-05	9.73E-02	1.77E-02	7.97E-02	5.49E-02	1.21E-02	4.33E-02
β_{214}	2.43E-03	2.44E-03	3.22E-06	1.20E-01	2.02E-02	9.97E-02	9.32E-02	2.04E-02	7.36E-02
β_{215}	6.37E-04	6.41E-04	1.24E-10	4.82E-02	9.57E-03	3.87E-02	4.07E-02	1.22E-02	2.89E-02
β_{221}	1.14E-01	1.13E-01	1.62E-03	5.33E-02	3.75E-02	1.61E-02	3.21E-02	2.36E-02	9.42E-03
β_{222}	1.66E-02	1.64E-02	2.98E-04	3.98E-02	2.20E-02	1.79E-02	1.70E-02	6.64E-03	1.06E-02
β_{223}	5.93E-04	5.93E-04	2.77E-06	1.11E-02	5.40E-03	5.72E-03	1.02E-02	5.15E-03	5.20E-03
β_{224}	2.02E-04	2.03E-04	5.39E-09	7.85E-03	4.59E-03	3.28E-03	9.89E-03	4.02E-03	6.02E-03
β_{225}	1.03E-04	1.03E-04	1.46E-07	7.80E-03	3.76E-03	4.06E-03	6.91E-03	4.07E-03	3.00E-03
β_{231}	6.84E-02	6.72E-02	1.61E-03	5.14E-02	3.98E-02	1.19E-02	4.92E-02	4.52E-02	5.76E-03
β_{232}	3.02E-03	2.99E-03	5.41E-05	1.10E-02	7.66E-03	3.37E-03	7.67E-03	7.14E-03	8.02E-04
β_{233}	1.24E-04	1.25E-04	4.23E-07	2.88E-03	2.40E-03	5.01E-04	2.40E-03	2.44E-03	5.47E-05
β_{234}	8.18E-05	8.10E-05	1.23E-06	4.66E-03	3.64E-03	1.04E-03	4.09E-03	3.94E-03	3.08E-04
β_{235}	5.52E-05	5.54E-05	1.85E-07	4.72E-03	3.79E-03	9.57E-04	2.83E-03	2.19E-03	7.22E-04
β_{241}	4.47E-02	4.38E-02	1.16E-03	3.66E-02	2.50E-02	1.18E-02	2.05E-02	1.78E-02	3.41E-03
β_{242}	6.99E-03	6.92E-03	1.05E-04	2.66E-02	1.58E-02	1.10E-02	1.27E-02	7.00E-03	5.96E-03
β_{243}	2.17E-04	2.15E-04	3.84E-06	5.89E-03	3.20E-03	2.71E-03	3.27E-03	2.40E-03	9.62E-04
β_{244}	6.06E-05	6.06E-05	3.62E-07	3.78E-03	2.46E-03	1.33E-03	3.29E-03	3.13E-03	2.77E-04
β_{245}	6.69E-05	6.71E-05	1.31E-07	6.40E-03	3.81E-03	2.61E-03	5.30E-03	3.59E-03	1.85E-03

Table B.4: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is identified. The parameters compared are regression coefficients for the SZ families, β_{pjk}, for $j = 1, \ldots, 4$, $k = 1, \ldots, K$ and $p = 2$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
a_{12}	1.20E-04	1.20E-04	1.66E-07	1.38E-03	5.34E-04	8.50E-04	1.70E-03	1.04E-03	7.04E-04
a_{13}	2.50E-04	2.49E-04	2.26E-06	1.26E-02	2.13E-03	1.05E-02	1.31E-02	1.28E-03	1.19E-02
a_{14}	1.18E-04	1.19E-04	1.06E-07	2.35E-02	3.40E-03	2.02E-02	2.39E-02	2.73E-03	2.13E-02
a_{15}	9.07E-05	9.06E-05	5.46E-07	9.23E-03	1.58E-03	7.65E-03	1.07E-02	1.52E-03	9.22E-03
a_{22}	2.13E-03	2.14E-03	1.12E-06	5.85E-03	4.64E-03	1.24E-03	6.70E-03	6.49E-03	4.58E-04
a_{23}	2.71E-04	2.72E-04	8.01E-07	4.51E-02	2.58E-03	4.25E-02	4.49E-02	3.96E-03	4.10E-02
a_{24}	2.20E-04	2.21E-04	2.93E-09	5.27E-02	4.13E-03	4.86E-02	4.21E-02	2.66E-03	3.96E-02
a_{25}	6.76E-04	6.77E-04	2.69E-06	4.58E-01	1.83E-02	4.40E-01	4.06E-01	2.78E-02	3.79E-01
a_{32}	1.54E-03	1.55E-03	1.42E-06	5.61E-02	5.47E-03	5.07E-02	4.32E-02	2.36E-03	4.09E-02
a_{33}	1.75E-03	1.76E-03	2.90E-07	5.93E-02	6.24E-03	5.31E-02	5.25E-02	2.44E-03	5.01E-02
a_{34}	1.79E-03	1.80E-03	2.86E-06	4.61E-01	3.37E-02	4.27E-01	4.25E-01	2.06E-02	4.05E-01
a_{35}	1.90E-04	1.90E-04	3.64E-07	1.14E-01	8.30E-03	1.06E-01	9.27E-02	5.41E-03	8.75E-02
a_{42}	1.17E-03	1.17E-03	3.67E-07	3.16E-03	2.84E-03	3.33E-04	3.43E-03	3.44E-03	1.20E-04
a_{43}	3.37E-04	3.32E-04	6.82E-06	5.01E-02	2.57E-03	4.75E-02	5.51E-02	4.17E-03	5.11E-02
a_{44}	7.21E-04	7.18E-04	6.73E-06	1.51E-01	7.77E-03	1.43E-01	1.72E-01	1.22E-02	1.60E-01
a_{45}	5.94E-04	5.93E-04	4.91E-06	4.66E-01	1.89E-02	4.47E-01	5.00E-01	2.48E-02	4.76E-01

Table B.5: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is identified. The parameters compared are family member factor loadings, a_{jk}, for $j = 1, \ldots, 4$ and $k = 2, \ldots, 5$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
b_{21}	7.08E-04	7.12E-04	1.60E-07	2.81E-03	5.81E-04	2.23E-03	3.37E-03	8.39E-04	2.56E-03
b_{22}	5.20E-03	5.14E-03	8.45E-05	2.73E-02	9.59E-04	2.64E-02	3.12E-02	1.21E-03	3.00E-02
b_{23}	3.57E-01	3.53E-01	5.49E-03	6.42E-02	5.37E-04	6.37E-02	6.75E-02	6.07E-04	6.70E-02
b_{24}	2.10E+02	2.11E+02	3.83E-01	2.83E+00	6.78E-03	2.82E+00	2.99E+00	5.11E-03	2.99E+00
b_{25}	8.78E-01	8.37E-01	4.59E-02	9.18E-02	1.05E-03	9.08E-02	9.42E-02	1.10E-03	9.31E-02
b_{31}	5.64E-04	5.66E-04	1.48E-06	6.23E-03	9.74E-04	5.26E-03	5.73E-03	1.08E-03	4.69E-03
b_{32}	5.42E-03	5.44E-03	8.64E-06	1.65E-03	7.07E-04	9.47E-04	1.50E-03	6.35E-04	8.84E-04
b_{33}	2.27E+00	2.18E+00	9.78E-02	2.64E-01	3.15E-03	2.61E-01	2.71E-01	4.33E-03	2.67E-01
b_{34}	3.32E-01	3.12E-01	2.20E-02	2.62E-01	1.82E-03	2.60E-01	2.90E-01	1.48E-03	2.89E-01
b_{35}	4.72E-01	4.50E-01	2.44E-02	6.97E-02	8.94E-04	6.89E-02	5.96E-02	6.22E-04	5.90E-02
b_{41}	2.86E-04	2.87E-04	1.10E-06	1.86E-03	3.24E-04	1.54E-03	1.40E-03	3.98E-04	1.02E-03
b_{42}	2.34E-03	2.35E-03	9.32E-06	2.68E-03	6.20E-04	2.07E-03	1.56E-03	5.02E-04	1.08E-03
b_{43}	2.68E-01	2.64E-01	5.31E-03	1.27E-03	5.58E-04	7.12E-04	1.13E-03	5.09E-04	6.38E-04
b_{44}	4.00E-01	3.77E-01	2.58E-02	1.05E-01	1.86E-03	1.03E-01	1.21E-01	1.59E-03	1.19E-01
b_{45}	1.66E-01	1.56E-01	1.08E-02	2.22E-01	5.50E-04	2.21E-01	2.25E-01	4.18E-04	2.25E-01

Table B.6: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is identified. The parameters compared are outcome factor loadings, b_{jk}, for $j = 1, \ldots, 4$ and $k = 2, \ldots, 5$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
ψ_1	8.90E-01	7.49E-01	1.50E-01	1.09E-02	8.56E-03	2.46E-03	1.69E-02	1.31E-02	3.90E-03
ψ_2	5.12E-02	4.52E-02	6.58E-03	1.11E-02	7.55E-03	3.65E-03	1.30E-02	9.19E-03	3.88E-03
ψ_3	1.24E-03	1.12E-03	1.33E-04	9.94E-03	7.12E-03	2.91E-03	1.26E-02	8.18E-03	4.48E-03
ψ_4	2.09E-04	1.72E-04	3.88E-05	5.44E-03	5.12E-03	3.77E-04	7.77E-03	5.69E-03	2.12E-03
ψ_5	7.26E-05	5.87E-05	1.46E-05	5.04E-03	4.90E-03	1.90E-04	4.85E-03	4.61E-03	2.83E-04
ϕ_{A11}	6.86E-03	6.74E-03	1.99E-04	1.41E-02	7.70E-03	6.49E-03	1.48E-02	8.85E-03	6.04E-03
ϕ_{A12}	1.20E-01	1.14E-01	7.40E-03	4.17E-02	2.24E-02	1.95E-02	4.24E-02	2.83E-02	1.44E-02
ϕ_{A22}	9.79E-03	8.81E-03	1.09E-03	2.53E-02	1.04E-02	1.51E-02	1.95E-02	7.97E-03	1.16E-02
ϕ_{A13}	1.99E-01	1.92E-01	8.40E-03	5.25E-02	2.68E-02	2.61E-02	5.26E-02	2.57E-02	2.71E-02
ϕ_{A23}	1.65E-01	1.58E-01	8.88E-03	5.56E-02	2.73E-02	2.86E-02	4.93E-02	2.32E-02	2.63E-02
ϕ_{A33}	8.81E-03	8.70E-03	2.19E-04	2.33E-02	9.17E-03	1.43E-02	2.59E-02	1.10E-02	1.50E-02
ϕ_{A14}	8.42E-02	8.29E-02	2.32E-03	4.63E-02	1.49E-02	3.16E-02	5.30E-02	1.74E-02	3.57E-02
ϕ_{A24}	9.76E-02	9.24E-02	6.32E-03	5.59E-02	1.80E-02	3.81E-02	5.28E-02	1.52E-02	3.78E-02
ϕ_{A34}	1.34E-01	1.27E-01	9.39E-03	6.56E-02	1.94E-02	4.65E-02	6.42E-02	1.93E-02	4.51E-02
ϕ_{A44}	3.98E-02	3.70E-02	3.20E-03	6.37E-02	1.31E-02	5.07E-02	5.99E-02	1.02E-02	4.98E-02

Table B.7: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is close to not identified. The parameters compared are unique error variances, ψ_k and family factor vairances and covariances, ϕ_{Alm}, for $l = 1, \ldots, 4$ and $m = 1, \ldots, l$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
ϕ_{B11}	1.19E-02	1.21E-02	1.16E-08	1.45E-02	1.29E-02	1.78E-03	1.23E-02	9.13E-03	3.23E-03
ϕ_{B12}	1.74E+00	1.73E+00	2.84E-02	1.44E-01	9.14E-02	5.35E-02	1.73E-01	8.61E-02	8.76E-02
ϕ_{B22}	8.66E-02	8.45E-02	3.10E-03	1.12E-01	1.79E-02	9.42E-02	1.17E-01	1.98E-02	9.75E-02
ϕ_{B13}	1.36E+01	1.28E+01	9.77E-01	5.13E-01	3.73E-01	1.44E-01	6.42E-01	4.37E-01	2.09E-01
ϕ_{B23}	1.50E+01	1.42E+01	1.04E+00	5.40E-01	3.49E-01	1.95E-01	6.08E-01	3.41E-01	2.71E-01
ϕ_{B33}	5.58E-02	5.41E-02	2.38E-03	8.36E-02	3.96E-02	4.45E-02	8.72E-02	4.03E-02	4.73E-02
ϕ_{B14}	1.61E+00	1.61E+00	1.98E-02	6.59E-02	6.50E-02	1.70E-03	5.59E-02	5.51E-02	1.22E-03
ϕ_{B24}	2.98E+02	2.76E+02	2.47E+01	1.49E+00	1.09E+00	4.10E-01	1.35E+00	7.81E-01	5.74E-01
ϕ_{B34}	1.07E+02	8.82E+01	2.02E+01	5.87E-01	5.41E-01	5.25E-02	6.20E-01	5.85E-01	4.05E-02
ϕ_{B44}	1.04E+00	8.49E-01	2.01E-01	2.23E-02	2.01E-02	2.44E-03	2.14E-02	1.11E-02	1.04E-02
ϕ_{B15}	3.29E+01	3.04E+01	2.92E+00	4.79E-01	3.02E-01	1.81E-01	4.09E-01	2.19E-01	1.92E-01
ϕ_{B25}	6.97E+01	6.41E+01	6.32E+00	7.94E-01	3.81E-01	4.17E-01	9.07E-01	3.43E-01	5.68E-01
ϕ_{B35}	4.66E+01	3.90E+01	8.10E+00	3.56E-01	3.53E-01	7.51E-03	3.69E-01	3.42E-01	2.95E-02
ϕ_{B45}	5.39E+01	4.40E+01	1.04E+01	2.39E-01	1.96E-01	4.48E-02	1.36E-01	1.20E-01	1.75E-02
ϕ_{B55}	6.78E+00	5.32E+00	1.52E+00	6.08E-02	3.94E-02	2.18E-02	5.16E-02	4.04E-02	1.15E-02

Table B.8: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is close to not identified. The parameters compared are outcome factor variances and covariances, ϕ_{Blm}, for $l = 1, \ldots, 5$ and $m = 1, \ldots, l$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
β_{111}	2.58E-02	2.55E-02	6.32E-04	1.87E-02	1.86E-02	3.44E-04	1.88E-02	1.89E-02	1.03E-04
β_{112}	6.42E-03	6.43E-03	6.41E-05	1.87E-02	1.88E-02	1.07E-04	1.78E-02	1.79E-02	1.17E-04
β_{113}	2.14E-04	2.16E-04	9.53E-07	4.12E-03	4.17E-03	4.31E-06	3.76E-03	3.78E-03	1.53E-05
β_{114}	6.28E-05	6.07E-05	2.84E-06	3.10E-03	3.01E-03	1.28E-04	3.37E-03	3.39E-03	1.69E-05
β_{115}	5.65E-05	5.64E-05	7.99E-07	4.19E-03	4.21E-03	3.61E-05	4.38E-03	4.42E-03	1.58E-06
β_{121}	2.93E-02	2.85E-02	1.11E-03	2.13E-02	2.08E-02	6.96E-04	2.56E-02	2.53E-02	4.72E-04
β_{122}	3.38E-03	3.36E-03	6.24E-05	1.02E-02	1.02E-02	1.33E-04	1.11E-02	1.11E-02	1.12E-05
β_{123}	5.68E-04	5.54E-04	2.06E-05	1.02E-02	9.97E-03	3.05E-04	9.97E-03	9.95E-03	1.05E-04
β_{124}	1.11E-04	1.04E-04	8.78E-06	4.96E-03	4.64E-03	3.76E-04	6.49E-03	6.19E-03	3.55E-04
β_{125}	2.37E-05	2.32E-05	7.37E-07	1.86E-03	1.83E-03	4.93E-05	2.12E-03	2.11E-03	3.31E-05
β_{131}	3.33E-02	3.30E-02	7.08E-04	2.49E-02	2.48E-02	4.29E-04	2.73E-02	2.71E-02	3.68E-04
β_{132}	3.99E-03	4.03E-03	8.37E-06	1.27E-02	1.29E-02	1.17E-05	1.38E-02	1.39E-02	3.16E-06
β_{133}	2.40E-04	2.42E-04	3.41E-07	4.68E-03	4.73E-03	1.92E-06	3.61E-03	3.64E-03	2.16E-07
β_{134}	3.06E-05	3.06E-05	3.43E-07	1.67E-03	1.68E-03	1.47E-05	1.89E-03	1.82E-03	7.85E-05
β_{135}	5.04E-05	5.10E-05	1.32E-08	3.95E-03	3.99E-03	4.11E-07	2.54E-03	2.57E-03	4.69E-09
β_{141}	4.57E-02	4.51E-02	1.10E-03	3.50E-02	3.47E-02	7.27E-04	4.20E-02	4.22E-02	2.06E-04
β_{142}	5.21E-03	5.27E-03	8.63E-06	1.70E-02	1.72E-02	7.38E-06	1.67E-02	1.68E-02	3.37E-05
β_{143}	2.76E-04	2.76E-04	3.53E-06	5.25E-03	5.27E-03	4.05E-05	6.33E-03	6.36E-03	2.23E-05
β_{144}	3.88E-05	3.77E-05	1.56E-06	2.06E-03	2.02E-03	6.46E-05	2.24E-03	2.21E-03	4.32E-05
β_{145}	3.23E-05	3.24E-05	3.28E-07	2.72E-03	2.74E-03	1.71E-05	2.21E-03	2.23E-03	8.69E-07

Table B.9: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is close to not identified. The parameters compared are regression coefficients for the control families, β_{pjk}, for $j = 1, \ldots, 4$, $k = 1, \ldots, K$ and $p = 1$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
β_{211}	4.02E-01	3.97E-01	9.98E-03	9.84E-02	7.04E-02	2.88E-02	1.11E-01	7.70E-02	3.50E-02
β_{212}	1.06E-01	1.05E-01	1.81E-03	1.21E-01	7.53E-02	4.65E-02	1.37E-01	6.68E-02	7.05E-02
β_{213}	3.65E-03	3.67E-03	1.81E-05	2.84E-02	1.73E-02	1.13E-02	3.20E-02	1.80E-02	1.42E-02
β_{214}	1.01E-03	1.02E-03	5.20E-06	2.05E-02	1.28E-02	7.91E-03	2.15E-02	1.29E-02	8.71E-03
β_{215}	7.66E-04	7.39E-04	3.63E-05	2.28E-02	1.36E-02	9.44E-03	3.21E-02	1.59E-02	1.64E-02
β_{221}	1.22E-01	1.23E-01	4.44E-04	5.80E-02	4.97E-02	8.96E-03	3.93E-02	3.14E-02	8.12E-03
β_{222}	1.35E-02	1.37E-02	2.08E-05	2.89E-02	2.26E-02	6.55E-03	2.52E-02	2.02E-02	5.12E-03
β_{223}	1.95E-03	1.95E-03	2.72E-05	3.30E-02	2.21E-02	1.11E-02	2.11E-02	1.36E-02	7.67E-03
β_{224}	5.65E-04	5.59E-04	1.29E-05	2.23E-02	1.34E-02	9.08E-03	1.71E-02	1.08E-02	6.38E-03
β_{225}	7.46E-05	7.51E-05	3.93E-07	5.03E-03	3.65E-03	1.42E-03	3.58E-03	2.73E-03	8.79E-04
β_{231}	4.26E-02	4.19E-02	1.20E-03	3.15E-02	3.13E-02	6.66E-04	2.79E-02	2.61E-02	2.11E-03
β_{232}	3.67E-03	3.60E-03	1.15E-04	1.08E-02	1.07E-02	2.35E-04	1.19E-02	1.04E-02	1.65E-03
β_{233}	3.35E-04	3.39E-04	4.39E-07	7.27E-03	6.87E-03	4.83E-04	5.32E-03	4.83E-03	5.35E-04
β_{234}	4.76E-05	4.79E-05	2.25E-07	2.90E-03	2.68E-03	2.53E-04	2.13E-03	1.79E-03	3.56E-04
β_{235}	6.95E-05	7.00E-05	2.54E-07	5.77E-03	5.52E-03	3.08E-04	4.76E-03	3.89E-03	9.06E-04
β_{241}	5.06E-02	5.07E-02	4.86E-04	4.02E-02	3.79E-02	2.70E-03	3.57E-02	3.38E-02	2.21E-03
β_{242}	4.78E-03	4.77E-03	6.88E-05	1.53E-02	1.44E-02	1.14E-03	1.95E-02	1.71E-02	2.47E-03
β_{243}	4.18E-04	4.21E-04	1.18E-06	1.03E-02	7.85E-03	2.53E-03	1.02E-02	7.75E-03	2.54E-03
β_{244}	5.59E-05	5.64E-05	1.33E-07	3.54E-03	2.82E-03	7.62E-04	3.86E-03	3.05E-03	8.43E-04
β_{245}	2.68E-05	2.71E-05	7.70E-08	2.41E-03	2.09E-03	3.46E-04	2.78E-03	2.24E-03	5.62E-04

Table B.10: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is close to not identified. The parameters compared are regression coefficients for the SZ families, β_{pjk}, for $j = 1, \ldots, 4$, $k = 1, \ldots, K$ and $p = 2$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
a_{12}	1.97E-04	1.99E-04	8.31E-10	2.51E-03	1.12E-03	1.40E-03	2.69E-03	1.38E-03	1.32E-03
a_{13}	3.52E-04	3.51E-04	4.90E-06	4.86E-03	2.66E-03	2.24E-03	4.44E-03	2.73E-03	1.74E-03
a_{14}	3.20E-04	3.20E-04	4.09E-06	7.54E-03	4.00E-03	3.59E-03	7.81E-03	3.86E-03	3.99E-03
a_{15}	6.60E-05	6.64E-05	4.20E-07	5.28E-03	2.27E-03	3.04E-03	5.37E-03	2.32E-03	3.06E-03
a_{22}	1.39E-03	1.41E-03	2.24E-07	6.11E-03	2.85E-03	3.29E-03	4.77E-03	2.09E-03	2.70E-03
a_{23}	1.11E-04	1.11E-04	1.11E-06	8.02E-03	2.23E-03	5.82E-03	7.54E-03	2.23E-03	5.33E-03
a_{24}	6.86E-05	6.94E-05	4.54E-08	6.73E-03	2.45E-03	4.31E-03	6.09E-03	2.41E-03	3.70E-03
a_{25}	6.89E-04	6.96E-04	1.40E-06	1.66E-02	7.13E-03	9.51E-03	1.70E-02	8.12E-03	8.96E-03
a_{32}	1.96E-03	1.92E-03	6.70E-05	7.49E-03	3.70E-03	3.83E-03	7.85E-03	3.28E-03	4.60E-03
a_{33}	3.83E-04	3.86E-04	1.88E-06	9.70E-03	3.07E-03	6.67E-03	1.03E-02	3.73E-03	6.57E-03
a_{34}	3.15E-03	3.17E-03	2.10E-05	2.55E-02	1.33E-02	1.24E-02	2.36E-02	9.94E-03	1.37E-02
a_{35}	1.23E-04	1.24E-04	1.97E-11	1.37E-02	4.04E-03	9.75E-03	1.19E-02	4.28E-03	7.61E-03
a_{42}	4.16E-03	4.14E-03	7.54E-05	2.26E-02	7.94E-03	1.47E-02	2.26E-02	5.83E-03	1.68E-02
a_{43}	5.17E-04	5.23E-04	6.68E-07	3.49E-02	6.37E-03	2.86E-02	4.13E-02	7.36E-03	3.40E-02
a_{44}	1.46E-03	1.47E-03	1.52E-05	4.24E-02	1.37E-02	2.89E-02	3.57E-02	7.98E-03	2.78E-02
a_{45}	1.43E-03	1.44E-03	2.49E-06	7.55E-02	1.53E-02	6.04E-02	8.25E-02	1.76E-02	6.50E-02

Table B.11: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is close to not identified. The parameters compared are family member factor loadings, a_{jk}, for $j = 1, \ldots, 4$ and $k = 2, \ldots, 5$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
b_{21}	1.23E-03	1.13E-03	1.17E-04	1.12E-03	1.13E-03	4.53E-07	7.93E-04	7.73E-04	2.66E-05
b_{22}	6.15E-03	6.23E-03	2.40E-07	1.61E-02	1.96E-03	1.41E-02	1.53E-02	1.54E-03	1.38E-02
b_{23}	1.83E-02	1.67E-02	1.85E-03	2.74E-02	4.11E-03	2.34E-02	3.29E-02	3.63E-03	2.93E-02
b_{24}	2.16E+01	1.96E+01	2.17E+00	6.24E-01	2.28E-02	6.02E-01	6.20E-01	1.86E-02	6.02E-01
b_{25}	3.57E+00	3.29E+00	3.22E-01	1.51E-02	5.46E-03	9.75E-03	1.39E-02	5.78E-03	8.16E-03
b_{31}	9.16E-04	8.70E-04	5.64E-05	1.19E-03	1.20E-03	1.62E-06	1.16E-03	1.06E-03	1.17E-04
b_{32}	4.82E-03	4.81E-03	6.08E-05	1.86E-02	1.65E-03	1.70E-02	1.60E-02	2.23E-03	1.38E-02
b_{33}	1.12E-01	1.09E-01	3.84E-03	1.96E-02	1.96E-02	2.21E-04	1.87E-02	1.88E-02	2.11E-05
b_{34}	1.51E+00	1.46E+00	6.75E-02	1.65E-02	6.72E-03	9.90E-03	1.61E-02	8.16E-03	8.01E-03
b_{35}	4.96E+00	4.99E+00	2.83E-02	2.18E-02	3.19E-03	1.86E-02	2.31E-02	4.11E-03	1.91E-02
b_{41}	4.23E-04	4.24E-04	4.41E-06	1.05E-03	9.50E-04	1.10E-04	8.96E-04	7.25E-04	1.77E-04
b_{42}	2.63E-03	2.66E-03	5.68E-06	1.80E-02	1.57E-03	1.65E-02	1.70E-02	1.40E-03	1.56E-02
b_{43}	1.07E-02	1.04E-02	4.62E-04	3.22E-02	4.24E-03	2.80E-02	3.35E-02	3.57E-03	3.00E-02
b_{44}	3.44E+00	3.20E+00	2.84E-01	4.77E-02	7.45E-03	4.04E-02	4.74E-02	1.05E-02	3.70E-02
b_{45}	1.03E+01	1.04E+01	1.39E-04	1.11E-01	1.83E-03	1.09E-01	1.13E-01	1.61E-03	1.12E-01

Table B.12: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is close to not identified. The parameters compared are outcome factor loadings, b_{jk}, for $j = 1, \ldots, 4$ and $k = 2, \ldots, 5$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
ψ_1	1.34E+00	1.13E+00	2.52E-01	6.97E-02	1.03E-02	5.98E-02	7.05E-02	1.50E-02	5.56E-02
ψ_2	5.22E-02	5.31E-02	9.77E-04	1.17E-02	8.83E-03	3.19E-03	2.17E-02	1.53E-02	6.48E-03
ψ_3	2.19E-03	1.46E-03	7.78E-04	9.71E-03	6.54E-03	3.39E-03	1.04E-02	5.83E-03	4.63E-03
ψ_4	3.90E-04	2.26E-04	1.72E-04	1.05E-02	6.57E-03	4.11E-03	8.06E-03	4.46E-03	3.63E-03
ψ_5	1.15E-04	5.68E-05	6.04E-05	5.56E-03	4.32E-03	1.38E-03	7.74E-03	5.30E-03	2.47E-03
ϕ_{A11}	1.82E-02	1.65E-02	2.26E-03	6.88E-02	6.48E-03	6.25E-02	6.09E-02	4.84E-03	5.61E-02
ϕ_{A12}	1.18E-01	1.00E-01	2.12E-02	1.34E-01	6.93E-03	1.28E-01	1.44E-01	9.50E-03	1.34E-01
ϕ_{A22}	1.03E-02	9.02E-03	1.63E-03	6.50E-02	3.08E-03	6.20E-02	6.73E-02	4.39E-03	6.30E-02
ϕ_{A13}	2.22E-01	2.07E-01	2.19E-02	1.90E-01	1.40E-02	1.77E-01	2.06E-01	2.08E-02	1.85E-01
ϕ_{A23}	2.32E-01	1.96E-01	4.25E-02	1.13E-01	9.95E-03	1.04E-01	1.23E-01	1.19E-02	1.11E-01
ϕ_{A33}	8.73E-03	7.36E-03	1.63E-03	7.89E-02	6.39E-03	7.27E-02	8.09E-02	8.07E-03	7.28E-02
ϕ_{A14}	2.00E-01	1.69E-01	3.73E-02	1.29E-01	6.42E-03	1.22E-01	1.29E-01	8.79E-03	1.20E-01
ϕ_{A24}	9.60E-02	8.07E-02	1.81E-02	1.50E-01	6.66E-03	1.44E-01	1.65E-01	7.99E-03	1.57E-01
ϕ_{A34}	4.38E-01	3.86E-01	6.50E-02	1.06E-01	9.32E-03	9.69E-02	1.12E-01	1.14E-02	1.01E-01
ϕ_{A44}	3.84E-02	3.54E-02	4.16E-03	1.85E-01	9.99E-03	1.75E-01	1.99E-01	1.08E-02	1.88E-01

Table B.13: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is not identified. The parameters compared are unique error variances, ψ_k and family factor vairances and covariances, ϕ_{Alm}, for $l = 1, \ldots, 4$ and $m = 1, \ldots, l$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
ϕ_{B11}	1.65E-02	1.67E-02	3.34E-04	1.51E-02	1.38E-02	1.86E-03	1.51E-02	1.05E-02	4.65E-03
ϕ_{B12}	2.63E+00	2.49E+00	2.26E-01	3.24E-01	1.69E-01	1.61E-01	2.88E-01	1.27E-01	1.62E-01
ϕ_{B22}	1.12E-01	1.05E-01	1.08E-02	5.32E-02	3.26E-02	2.17E-02	5.47E-02	2.85E-02	2.64E-02
ϕ_{B13}	1.83E+01	1.77E+01	1.26E+00	6.78E-01	2.98E-01	3.90E-01	6.63E-01	2.30E-01	4.34E-01
ϕ_{B23}	8.07E+01	6.15E+01	2.12E+01	8.10E-01	2.85E-01	5.34E-01	7.08E-01	2.47E-01	4.62E-01
ϕ_{B33}	9.56E-01	6.78E-01	3.01E-01	1.70E-02	7.33E-03	9.89E-03	1.31E-02	6.48E-03	6.63E-03
ϕ_{B14}	9.20E+00	7.92E+00	1.55E+00	1.86E-01	6.46E-02	1.23E-01	1.85E-01	6.70E-02	1.19E-01
ϕ_{B24}	2.75E+02	2.09E+02	7.32E+01	2.03E+00	8.48E-01	1.21E+00	2.70E+00	9.99E-01	1.70E+00
ϕ_{B34}	8.47E+02	5.60E+02	3.06E+02	5.58E-01	2.71E-01	2.96E-01	4.98E-01	2.23E-01	2.76E-01
ϕ_{B44}	5.15E+00	3.49E+00	1.78E+00	2.02E-02	1.48E-02	5.92E-03	3.13E-02	2.58E-02	5.62E-03
ϕ_{B15}	3.98E+01	3.88E+01	2.36E+00	6.09E-01	2.61E-01	3.58E-01	8.37E-01	2.65E-01	5.74E-01
ϕ_{B25}	2.03E+02	1.65E+02	4.42E+01	9.35E-01	2.83E-01	6.62E-01	1.23E+00	4.23E-01	8.13E-01
ϕ_{B35}	2.89E+02	2.19E+02	7.72E+01	3.07E-01	1.04E-01	2.07E-01	3.94E-01	1.49E-01	2.45E-01
ϕ_{B45}	3.15E+02	2.28E+02	9.50E+01	1.34E-01	5.58E-02	8.05E-02	2.27E-01	9.93E-02	1.28E-01
ϕ_{B55}	2.26E+01	1.48E+01	8.33E+00	5.14E-02	3.27E-02	1.98E-02	7.19E-02	4.10E-02	3.11E-02

Table B.14: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is not identified. The parameters compared are outcome factor variances and covariances, ϕ_{Blm}, for $l = 1, \ldots, 5$ and $m = 1, \ldots, l$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
β_{111}	2.97E-02	3.08E-02	3.38E-06	1.93E-02	2.00E-02	6.51E-07	1.34E-02	1.33E-02	1.91E-04
β_{112}	4.54E-03	4.70E-03	1.53E-07	1.08E-02	1.12E-02	1.25E-06	8.36E-03	8.27E-03	1.41E-04
β_{113}	2.37E-04	2.45E-04	3.22E-08	3.93E-03	4.07E-03	1.97E-06	4.05E-03	4.07E-03	4.50E-06
β_{114}	6.33E-05	6.16E-05	3.78E-06	2.69E-03	2.58E-03	2.00E-04	2.60E-03	2.60E-03	2.36E-05
β_{115}	3.83E-05	3.95E-05	1.23E-07	2.31E-03	2.39E-03	1.59E-06	2.70E-03	2.68E-03	3.39E-05
β_{121}	2.70E-02	2.79E-02	9.79E-05	1.79E-02	1.85E-02	4.20E-05	1.32E-02	1.32E-02	1.90E-05
β_{122}	3.87E-03	3.97E-03	3.62E-05	1.03E-02	1.06E-02	5.93E-05	8.57E-03	8.63E-03	2.64E-08
β_{123}	1.77E-04	1.75E-04	7.54E-06	2.97E-03	2.92E-03	1.57E-04	3.29E-03	3.31E-03	3.43E-06
β_{124}	5.99E-05	5.91E-05	2.76E-06	2.68E-03	2.66E-03	1.08E-04	2.75E-03	2.76E-03	2.18E-06
β_{125}	3.25E-05	3.21E-05	1.50E-06	2.06E-03	2.03E-03	9.73E-05	2.07E-03	2.08E-03	3.51E-06
β_{131}	4.86E-02	5.03E-02	8.05E-05	3.11E-02	3.21E-02	1.59E-04	1.42E-02	1.43E-02	2.76E-07
β_{132}	6.23E-03	6.42E-03	3.09E-05	1.49E-02	1.53E-02	1.72E-04	9.67E-03	9.68E-03	5.02E-05
β_{133}	3.64E-04	3.74E-04	3.05E-06	5.50E-03	5.58E-03	1.14E-04	3.42E-03	3.44E-03	2.46E-06
β_{134}	1.05E-04	1.04E-04	3.85E-06	4.09E-03	4.02E-03	2.10E-04	2.28E-03	2.28E-03	1.43E-05
β_{135}	5.94E-05	6.06E-05	9.30E-07	3.50E-03	3.56E-03	5.53E-05	2.49E-03	2.47E-03	3.16E-05
β_{141}	3.96E-02	4.10E-02	1.42E-05	2.82E-02	2.92E-02	3.09E-05	1.90E-02	1.91E-02	3.99E-05
β_{142}	3.48E-03	3.59E-03	8.19E-06	9.40E-03	9.68E-03	5.65E-05	1.18E-02	1.18E-02	2.50E-05
β_{143}	1.34E-04	1.32E-04	7.07E-06	2.15E-03	2.07E-03	1.46E-04	3.69E-03	3.70E-03	9.93E-06
β_{144}	4.57E-05	4.72E-05	1.31E-07	2.26E-03	2.33E-03	1.11E-05	3.04E-03	3.06E-03	2.71E-06
β_{145}	1.98E-05	2.00E-05	4.49E-07	1.40E-03	1.40E-03	5.20E-05	2.98E-03	2.98E-03	1.06E-05

Table B.15: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is not identified. The parameters compared are regression coefficients for the control families, β_{pjk}, for $j = 1, \ldots, 4$, $k = 1, \ldots, K$ and $p = 1$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
β_{211}	2.72E-01	2.62E-01	1.92E-02	1.56E-01	4.33E-02	1.14E-01	1.32E-01	6.09E-02	7.11E-02
β_{212}	5.55E-02	5.45E-02	2.86E-03	1.39E-01	3.26E-02	1.08E-01	1.22E-01	3.62E-02	8.62E-02
β_{213}	3.97E-03	3.63E-03	4.70E-04	6.94E-02	1.57E-02	5.43E-02	4.62E-02	1.51E-02	3.12E-02
β_{214}	6.64E-04	6.66E-04	2.12E-05	3.90E-02	6.84E-03	3.24E-02	3.85E-02	1.26E-02	2.59E-02
β_{215}	6.34E-04	6.48E-04	8.42E-06	4.08E-02	9.83E-03	3.13E-02	3.59E-02	1.05E-02	2.55E-02
β_{221}	6.27E-02	6.49E-02	1.26E-05	4.19E-02	2.53E-02	1.74E-02	5.19E-02	3.34E-02	1.87E-02
β_{222}	1.22E-02	1.25E-02	6.63E-05	3.78E-02	1.85E-02	2.00E-02	3.53E-02	1.68E-02	1.87E-02
β_{223}	4.97E-04	5.00E-04	1.47E-05	1.39E-02	5.45E-03	8.60E-03	1.09E-02	5.59E-03	5.38E-03
β_{224}	2.18E-04	2.26E-04	1.26E-09	1.05E-02	5.65E-03	5.04E-03	9.56E-03	4.78E-03	4.81E-03
β_{225}	6.19E-05	6.40E-05	2.32E-08	6.95E-03	2.68E-03	4.36E-03	9.03E-03	4.15E-03	4.90E-03
β_{231}	2.66E-02	2.67E-02	7.91E-04	2.69E-02	1.85E-02	9.02E-03	2.74E-02	2.21E-02	5.51E-03
β_{232}	3.32E-03	2.74E-03	6.72E-04	1.98E-02	7.18E-03	1.29E-02	1.77E-02	1.29E-02	4.83E-03
β_{233}	1.49E-04	1.40E-04	1.32E-05	5.21E-03	2.43E-03	2.86E-03	5.96E-03	4.77E-03	1.22E-03
β_{234}	6.13E-05	5.58E-05	7.36E-06	4.97E-03	2.33E-03	2.72E-03	3.72E-03	2.82E-03	9.23E-04
β_{235}	3.72E-05	3.53E-05	3.06E-06	4.64E-03	2.59E-03	2.14E-03	4.59E-03	3.65E-03	9.60E-04
β_{241}	2.78E-02	2.74E-02	1.30E-03	3.03E-02	1.92E-02	1.17E-02	3.40E-02	2.76E-02	6.57E-03
β_{242}	4.59E-03	4.58E-03	1.74E-04	1.97E-02	1.20E-02	8.11E-03	1.91E-02	1.27E-02	6.47E-03
β_{243}	1.74E-04	1.64E-04	1.54E-05	6.13E-03	3.01E-03	3.23E-03	4.99E-03	3.47E-03	1.55E-03
β_{244}	6.92E-05	6.88E-05	2.72E-06	4.62E-03	2.83E-03	1.89E-03	4.34E-03	3.09E-03	1.27E-03
β_{245}	3.37E-05	3.49E-05	4.44E-08	3.19E-03	2.19E-03	1.08E-03	4.54E-03	3.12E-03	1.44E-03

Table B.16: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is not identified. The parameters compared are regression coefficients for the SZ families, β_{pjk}, for $j = 1, \ldots, 4$, $k = 1, \ldots, K$ and $p = 2$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
a_{12}	4.73E-04	4.90E-04	2.50E-07	1.13E-02	1.35E-03	9.96E-03	1.07E-02	9.27E-04	9.75E-03
a_{13}	9.77E-04	1.01E-03	5.53E-06	6.84E-02	5.18E-03	6.34E-02	6.68E-02	4.72E-03	6.21E-02
a_{14}	5.34E-04	5.09E-04	4.26E-05	9.27E-02	8.14E-03	8.48E-02	9.43E-02	7.33E-03	8.70E-02
a_{15}	2.05E-04	2.10E-04	2.42E-06	5.64E-02	6.90E-03	4.97E-02	5.27E-02	4.46E-03	4.83E-02
a_{22}	2.41E-03	2.44E-03	4.81E-05	4.23E-02	4.75E-03	3.77E-02	4.18E-02	3.07E-03	3.88E-02
a_{23}	3.08E-04	2.94E-04	2.42E-05	2.86E-02	1.78E-03	2.69E-02	3.07E-02	1.88E-03	2.88E-02
a_{24}	1.48E-04	1.53E-04	4.61E-07	3.63E-02	1.61E-03	3.48E-02	3.81E-02	1.77E-03	3.64E-02
a_{25}	7.00E-04	6.95E-04	2.81E-05	2.18E-01	1.21E-02	2.06E-01	2.22E-01	1.21E-02	2.10E-01
a_{32}	9.58E-04	9.90E-04	2.03E-06	3.75E-02	3.42E-03	3.42E-02	3.62E-02	3.87E-03	3.24E-02
a_{33}	6.39E-04	6.17E-04	4.38E-05	6.27E-02	3.77E-03	5.91E-02	5.63E-02	4.42E-03	5.19E-02
a_{34}	1.35E-03	1.38E-03	2.32E-05	2.75E-01	2.19E-02	2.54E-01	2.34E-01	1.59E-02	2.18E-01
a_{35}	9.77E-05	9.54E-05	5.58E-06	6.77E-02	3.51E-03	6.43E-02	5.67E-02	4.40E-03	5.23E-02
a_{42}	1.47E-03	1.47E-03	5.67E-05	3.49E-02	3.79E-03	3.12E-02	3.85E-02	3.31E-03	3.52E-02
a_{43}	2.26E-04	2.24E-04	9.88E-06	4.60E-02	2.12E-03	4.39E-02	4.21E-02	2.36E-03	3.98E-02
a_{44}	6.59E-04	6.74E-04	8.63E-06	1.15E-01	5.24E-03	1.10E-01	1.25E-01	8.88E-03	1.16E-01
a_{45}	7.59E-04	7.79E-04	7.09E-06	1.88E-01	1.28E-02	1.76E-01	1.86E-01	9.66E-03	1.76E-01

Table B.17: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is not identified. The parameters compared are family member factor loadings, a_{jk}, for $j = 1, \ldots, 4$ and $k = 2, \ldots, 5$.

	QNO			BFFM (QNO Converges)			BFFM (QNO Fails)		
Parameter	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias	RMSE	Var	Sq Bias
b_{21}	1.90E-03	1.96E-03	1.03E-06	1.17E-03	1.21E-03	9.74E-09	9.10E-04	9.14E-04	1.37E-06
b_{22}	5.46E-03	5.40E-03	2.43E-04	1.52E-03	1.51E-03	6.71E-05	1.21E-03	1.11E-03	1.05E-04
b_{23}	1.60E-01	1.65E-01	1.28E-03	6.04E-04	6.25E-04	9.14E-07	4.31E-04	4.14E-04	1.95E-05
b_{24}	9.99E-01	1.03E+00	1.27E-03	2.56E-03	2.50E-03	1.45E-04	2.34E-03	2.34E-03	1.17E-05
b_{25}	4.93E+00	5.10E+00	4.36E-03	4.25E-04	4.21E-04	1.83E-05	3.92E-04	3.91E-04	3.41E-06
b_{31}	1.46E-03	1.51E-03	2.11E-06	1.13E-03	1.17E-03	2.25E-06	7.02E-04	7.06E-04	1.45E-09
b_{32}	2.84E-03	2.62E-03	3.12E-04	9.06E-04	7.81E-04	1.52E-04	9.59E-04	9.46E-04	1.82E-05
b_{33}	6.34E-01	6.56E-01	3.80E-04	2.97E-03	2.42E-03	6.39E-04	2.63E-03	2.54E-03	1.05E-04
b_{34}	2.12E-01	2.19E-01	4.71E-05	9.17E-04	9.09E-04	4.00E-05	8.48E-04	8.36E-04	1.63E-05
b_{35}	1.88E+00	1.95E+00	3.97E-05	5.35E-04	5.53E-04	4.31E-07	2.23E-04	2.15E-04	9.32E-06
b_{41}	1.05E-03	1.06E-03	2.85E-05	9.61E-04	9.63E-04	3.08E-05	4.89E-04	4.89E-04	2.94E-06
b_{42}	2.71E-03	2.56E-03	2.33E-04	1.00E-03	8.88E-04	1.45E-04	7.34E-04	7.08E-04	2.99E-05
b_{43}	7.14E-02	7.31E-02	8.21E-04	5.58E-04	4.46E-04	1.27E-04	4.83E-04	4.66E-04	1.99E-05
b_{44}	3.55E-01	3.52E-01	1.50E-02	1.36E-03	1.39E-03	9.72E-06	1.08E-03	1.06E-03	1.97E-05
b_{45}	9.62E-01	9.86E-01	1.04E-02	1.85E-04	1.89E-04	3.03E-06	1.87E-04	1.85E-04	3.19E-06

Table B.18: Comparison of the relative mean squared errors (MSEs), relative variances and relative squared biases from FIML-QNO when it is successful, from the BFFM when FIML-QNO is successful, and from the BFFM when FIML-QNO fails, in the scenario where the true covariance matrix is not identified. The parameters compared are outcome factor loadings, b_{jk}, for $j = 1, \ldots, 4$ and $k = 2, \ldots, 5$.

APPENDIX C

Summary Statistics for the UCLA Family Study Data

		Control				Schizophrenia			
	Variable	Prob	Sib	Fa	Mo	Prob	Sib	Fa	Mo
Mean	MANIPA	0.76	0.70	0.76	0.74	0.68	0.71	0.76	0.72
	CPTDSD	2.54	2.31	2.07	2.62	2.17	2.28	2.42	2.33
	CPT37D	4.47	4.35	4.49	4.86	3.87	4.3	4.48	4.29
	SPAN10	56.1	54.9	53.8	53.8	53.3	54.8	51.4	51.9
	logTRLBA	−1.39	−1.43	−1.4	−1.41	−1.51	−1.47	−1.48	−1.52
Std Dev	MANIPA	0.13	0.15	0.11	0.13	0.17	0.18	0.16	0.15
	CPTDSD	1.00	1.03	0.88	0.92	1.22	1.11	0.99	1.04
	CPT37D	0.82	0.94	0.74	0.63	1.09	0.95	0.59	0.90
	SPAN10	4.9	6.5	5.4	4.7	7.2	5.3	5.4	5.9
	logTRLBA	0.24	0.22	0.21	0.2	0.24	0.23	0.2	0.23

Table C.1: Raw group means and standard deviations of the 5 outcomes measured on probands, siblings, fathers and mothers in the schizophrenia and control families. Please refer to Table 4.2 for description of variables.

	Pr1	Pr2	Pr3	Pr4	Pr5	Sib1	Sib2	Sib3	Sib4	Sib5	Fa1	Fa2	Fa3	Fa4	Fa5	Mo1	Mo2	Mo3	Mo4
Pr1	1.00																		
Pr2	.35	1.00																	
Pr3	.38	.50	1.00																
Pr4	.35	.39	.54	1.00															
Pr5	.33	.42	.54	.50	1.00														
Sib1	.20	.13	.16	.12	.26	1.00													
Sib2	.23	.21	.28	.07	.30	.40	1.00												
Sib3	.09	.23	.28	.19	.30	.41	.39	1.00											
Sib4	.10	.17	.18	.22	.24	.25	.29	.44	1.00										
Sib5	.09	.18	.28	.20	.31	.47	.28	.53	.46	1.00									
Fa1	.08	.02	−.02	.04	−.03	.03	.01	.09	.00	0.03	1.00								
Fa2	−.01	.08	−.02	.04	.09	.08	.14	−.01	.10	-0.03	.12	1.00							
Fa3	.12	.19	.10	.04	.20	.11	.13	.04	−.01	0.05	.26	.57	1.00						
Fa4	.20	.20	.16	.18	.21	.09	.16	.15	.18	.04	.27	.28	.27	1.00					
Fa5	.21	.14	.10	.08	.05	.01	.06	−.06	−.04	.09	.06	.18	.31	.36	1.00				
Mo1	.17	.17	.09	.12	.21	.04	.03	−.01	−.01	-.13	−.13	.01	−.06	.02	.09	1.00			
Mo2	.21	.08	.12	.12	.15	.09	.14	.13	.16	.09	−.02	−.03	−.07	.00	.13	.22	1.00		
Mo3	.23	.16	.22	.13	.20	.00	−.04	.05	.01	.04	.01	−0.07	.03	.17	.24	.32	.32	1.00	
Mo4	.15	−.03	.04	.06	.09	.03	−.02	−.05	.21	-.01	−.05	.02	−0.02	.16	.13	.21	.39	.47	1.00
Mo5	.19	.12	.08	.11	.21	.12	−0.04	.05	.09	.12	−.07	.07	.05	.11	.22	.30	.30	.61	.52

Table C.2: Correlation matrix of 4 of measurements of 5 outcomes on 4 family members. Pr, Sib, Fa and Mo refer to proband, sibling, father and mother, respectively. Numbers 1, 2, 3, 4, 5 on variable names refers to the five outcomes MANIPA, CPTDSD, CPT37D, SPAN10 and logTRLBA, respectively. The block diagonal matrices are correlations between the five outcomes for the same family member.

		Proband	Sibling	Father	Mother
MANIPA	Proband	1.00			
	Sibling	0.20	1.00		
	Father	0.08	0.03	1.00	
	Mother	0.17	0.04	−0.13	1.00
CPTDSD	Proband	1.00			
	Sibling	0.31	1.00		
	Father	0.05	0.09	1.00	
	Mother	0.21	0.12	0.22	1.00
CPT37D	Proband	1.00			
	Sibling	0.22	1.00		
	Father	0.18	0.18	1.00	
	Mother	0.06	0.21	0.16	1.00
SPAN10	Proband	1.00			
	Sibling	0.28	1.00		
	Father	0.10	0.04	1.00	
	Mother	0.22	0.05	0.03	1.00
logTRLBA	Proband	1.00			
	Sibling	0.21	1.00		
	Father	0.08	0.14	1.00	
	Mother	0.08	0.14	−0.03	1.00

Table C.3: Sample correlations between measurements on different family members for the same outcome. For all five outcomes, the correlations between proband and sibling are the highest (about 0.2). Please refer to Table 4.2 for description of variables.

APPENDIX D

Summary of Posterior Distributions for the UCLA Family Study Data with Five Outcomes

Tables D.1, D.2 and D.3 give summaries of the posterior distributions including mean, SD and posterior probabilities, $p(\theta < 0|\boldsymbol{Y})$ from fitting the BFFM to the UCLA NSF data.

Figure D.1 plots the posterior distribution group means for all five outcomes measured on probands, fathers, mothers and siblings in schizophrenia and control families. Figures D.2 and D.3 plot the posterior distribution of non-zero family factor loadings and outcome factor loadings, grouped by family member and by outcome, respectively.

Param	Mean	SD	$p(\theta < 0 \mid \boldsymbol{Y})$	Param	Mean	SD	$p(\theta < 0 \mid \boldsymbol{Y})$
Unique Error Variances				Family Member Factor Loadings			
ψ_1	150.17	12.15		a_{12}	0.88	0.16	$< .0001$
ψ_2	65.41	5.28		a_{13}	1.01	0.15	$< .0001$
ψ_3	28.57	3.19		a_{14}	0.61	0.09	$< .0001$
ψ_4	15.77	1.65		a_{15}	2.13	0.33	$< .0001$
ψ_5	229.25	21.24		a_{22}	0.61	0.17	0.0008
Famliy Member Factor Vari -Covar				a_{23}	0.99	0.17	$< .0001$
ϕ_{A11}	50.57	13.64		a_{24}	0.46	0.09	$< .0001$
ϕ_{A12}	19.79	8.16	0.002	a_{25}	2.05	0.34	$< .0001$
ϕ_{A22}	50.81	14.69		a_{32}	0.92	0.23	$< .0001$
ϕ_{A13}	7.10	7.48	0.16	a_{33}	0.73	0.18	$< .0001$
ϕ_{A23}	2.87	7.20	0.34	a_{34}	0.47	0.12	$< .0001$
ϕ_{A33}	33.31	10.85		a_{35}	1.77	0.51	0.0006
ϕ_{A14}	3.49	5.70	0.27	a_{42}	0.77	0.18	$< .0001$
ϕ_{A24}	1.47	6.58	0.41	a_{43}	0.90	0.14	$< .0001$
ϕ_{A34}	3.67	6.49	0.27	a_{44}	0.52	0.09	$< .0001$
ϕ_{A44}	37.14	9.84		a_{45}	2.36	0.36	$< .0001$

Table D.1: Posterior means and SD of unique error vairances, ψ_k, family member factor covriance matrix, $\boldsymbol{\Phi}_A$ and family member factor loadings, a_{jk}, for $j = 1, \ldots 4$ corresponding to prbands, siblings, fathers and mothers, and $k = 1, \ldots, 5$ corresponding to MANIPA, CPTDSD, CPT37D, SPAN10 and logTRLBA. The posterior probabilities, $p(\theta < 0 \mid \boldsymbol{Y})$ are also listed.

Param	Mean	SD	$p(\theta < 0 \mid \boldsymbol{Y})$	Param	Mean	SD	$p(\theta < 0 \mid \boldsymbol{Y})$
Unique Error Variances				Family Member Factor Loadings			
Outcome Factor Var-Covar				Outcome Factor Loadings			
ϕ_{B11}	50.20	18.02		b_{21}	1.18	0.28	$< .0001$
ϕ_{B12}	20.66	8.51	0.0005	b_{22}	1.23	0.28	$< .0001$
ϕ_{B22}	24.06	7.63		b_{23}	0.84	0.37	0.01
ϕ_{B13}	14.26	7.05	0.008	b_{24}	1.52	0.30	$< .0001$
ϕ_{B23}	12.21	5.11	0.0005	b_{25}	0.81	0.28	0.002
ϕ_{B33}	16.70	5.92		b_{31}	0.18	0.33	0.29
ϕ_{B14}	5.35	3.45	0.03	b_{32}	0.66	0.29	0.01
ϕ_{B24}	3.76	2.23	0.02	b_{33}	0.60	0.29	0.01
ϕ_{B34}	2.57	1.93	0.06	b_{34}	1.08	0.36	0.0006
ϕ_{B44}	4.67	1.52		b_{35}	0.84	0.36	0.007
ϕ_{B15}	41.26	19.52	0.002	b_{41}	0.54	0.27	0.02
ϕ_{B25}	25.50	12.12	0.002	b_{42}	0.65	0.27	0.006
ϕ_{B35}	21.38	10.87	0.004	b_{43}	0.68	0.23	0.001
ϕ_{B45}	10.04	5.26	0.004	b_{44}	0.98	0.27	0.0002
ϕ_{B55}	97.69	37.86		b_{45}	0.76	0.26	0.001

Table D.2: Posterior means and SD of outcome factor variance-covariances, $\boldsymbol{\Phi}_B$ and outcome factor loadings, b_{jk}, for $j = 1, \ldots 4$ corresponding to prbands, siblings, fathers and mothers, and $k = 1, \ldots, 5$ corresponding to MANIPA, CPTDSD, CPT37D, SPAN10 and logTRLBA. The posterior probabilities, $p(\theta < 0 \mid \boldsymbol{Y})$ are also listed.

	Means of Control			Means of SZ			Diff in Means, Control-SZ			
	Param	Mean	SD	Param	Mean	SD	Param	Mean	SD	$p(\theta < 0)$
	β_{111}	76.01	1.89	β_{211}	66.56	1.60	$\beta_{111} - \beta_{211}$	9.45	2.47	0.0002
	β_{112}	25.27	1.25	β_{212}	21.38	1.10	$\beta_{112} - \beta_{212}$	3.89	1.67	0.01
Prob	β_{113}	44.63	1.08	β_{213}	38.44	0.93	$\beta_{113} - \beta_{213}$	6.19	1.42	<.0001
	β_{114}	56.05	0.68	β_{214}	53.32	0.60	$\beta_{114} - \beta_{214}$	2.73	0.91	0.00
	β_{115}	-139.35	2.60	β_{215}	-151.11	2.26	$\beta_{115} - \beta_{215}$	11.77	3.44	0.0004
	β_{121}	70.25	1.97	β_{221}	72.37	2.26	$\beta_{121} - \beta_{221}$	-2.12	2.99	0.77
	β_{122}	22.85	1.24	β_{222}	23.44	1.48	$\beta_{122} - \beta_{222}$	-0.59	1.93	0.61
Sib	β_{123}	43.36	1.06	β_{223}	43.77	1.24	$\beta_{123} - \beta_{223}$	-0.42	1.64	0.59
	β_{124}	54.79	0.68	β_{224}	55.33	0.80	$\beta_{124} - \beta_{224}$	-0.54	1.06	0.70
	β_{125}	-143.00	2.56	β_{225}	-145.10	2.98	$\beta_{125} - \beta_{225}$	2.10	3.95	0.30
	β_{131}	77.00	2.18	β_{231}	75.59	2.40	$\beta_{131} - \beta_{231}$	1.41	3.22	0.33
	β_{132}	20.83	1.51	β_{232}	24.02	1.65	$\beta_{132} - \beta_{232}$	-3.19	2.24	0.92
Fath	β_{133}	45.01	1.06	β_{233}	44.36	1.20	$\beta_{133} - \beta_{233}$	0.65	1.62	0.34
	β_{134}	53.88	0.78	β_{234}	51.13	0.84	$\beta_{134} - \beta_{234}$	2.75	1.15	0.01
	β_{135}	-139.48	2.92	β_{235}	-150.00	3.32	$\beta_{135} - \beta_{235}$	10.52	4.43	0.01
	β_{141}	74.39	1.75	β_{241}	71.84	1.63	$\beta_{141} - \beta_{241}$	2.55	2.38	0.14
	β_{142}	26.07	1.16	β_{242}	23.25	1.10	$\beta_{142} - \beta_{242}$	2.82	1.61	0.04
Moth	β_{143}	48.55	0.94	β_{243}	42.90	0.90	$\beta_{143} - \beta_{243}$	5.65	1.29	<.0001
	β_{144}	53.75	0.63	β_{244}	51.68	0.61	$\beta_{144} - \beta_{244}$	2.06	0.87	0.01
	β_{145}	-140.78	2.53	β_{245}	-151.76	2.46	$\beta_{145} - \beta_{245}$	10.98	3.52	<.0001

Table D.3: The left and middle panels present posterior means and SD of regression coefficients, β_{pjk}, which are the means of the k^{th} outcomes for the j^{th} family member in the control families ($p = 1$) and the SZ families ($p = 2$), for $j = 1, \ldots 4$ corresponding to probands, siblings, fathers and mothers, and $k = 1, \ldots, 5$ corresponding to MANIPA, CPTDSD, CPT37D, SPAN10 and logTRLBA. The right panel presents the posterior means, SD of difference in group means, control minus SZ, $\beta_{1jk} - \beta_{2jk}$. The posterior probabilities, $p(\beta_{1jk} - \beta_{2jk} < 0|\boldsymbol{Y})$ are also listed.

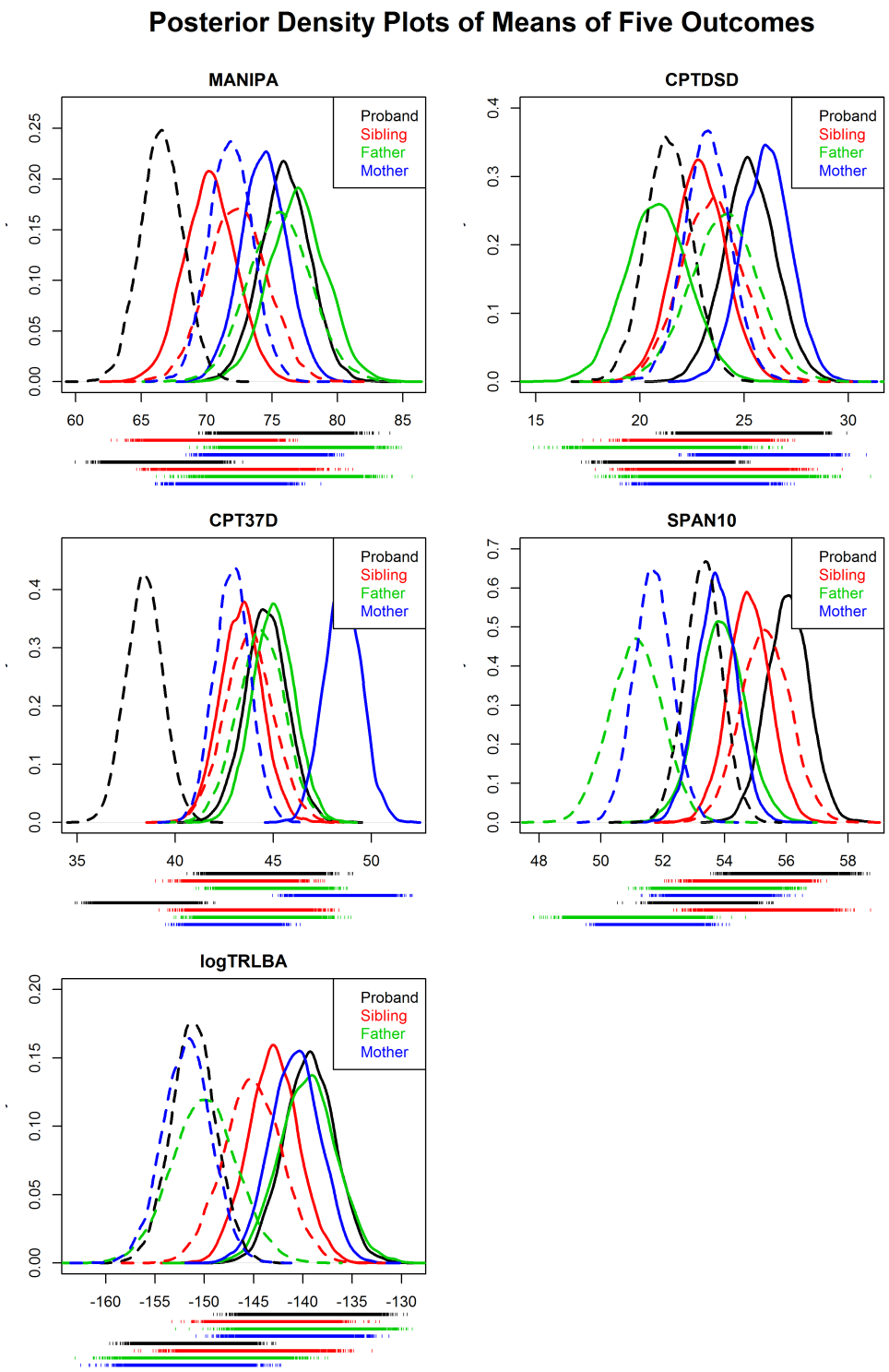


Figure D.1: Posterior density of group means for probands, siblings, fathers and mothers in the control (solid lines) and SZ (dashed lines) families.

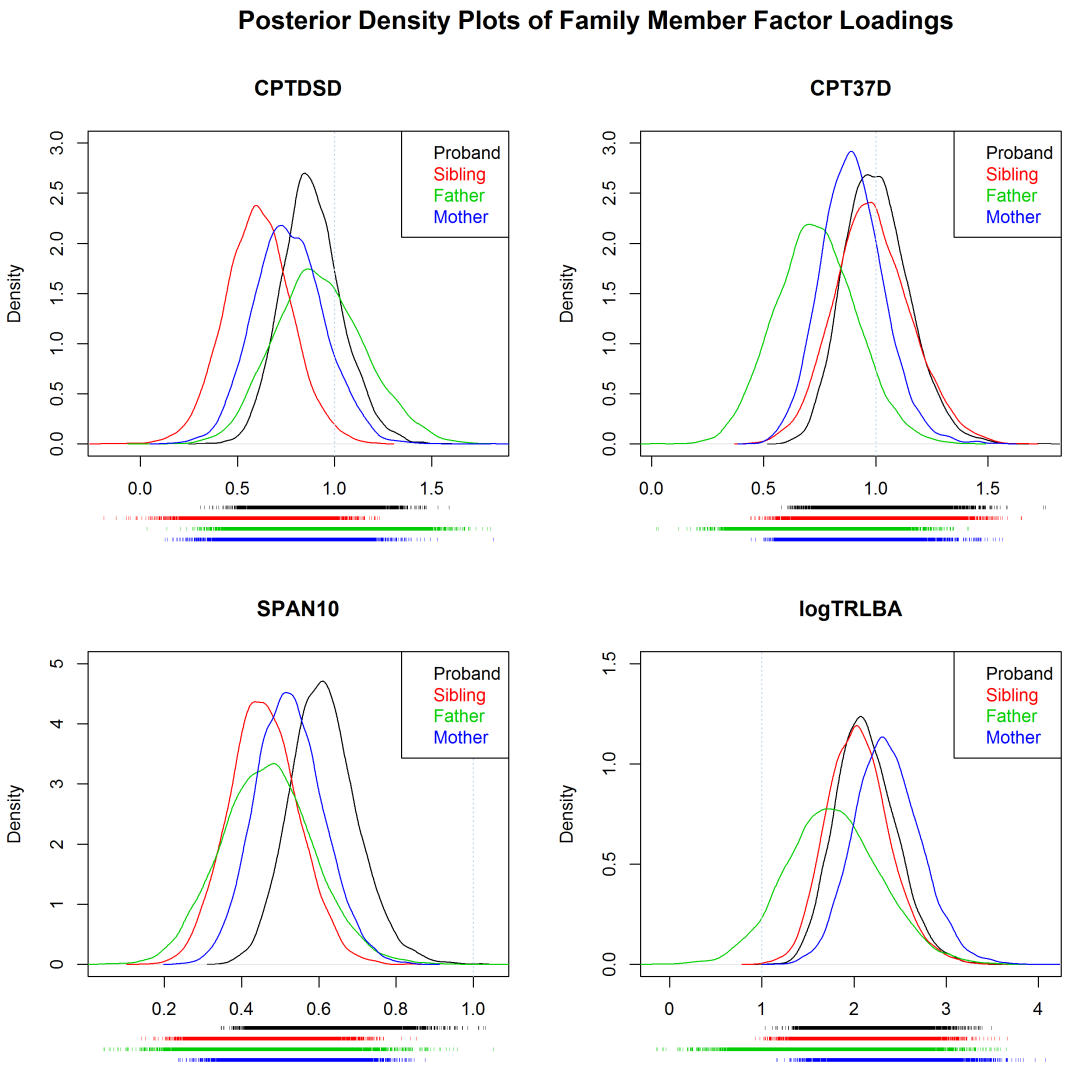


Figure D.2: Posterior density plots of family member factor loadings, a_{jk}, for $j = 1, \ldots, 4$ corresponding to probands, siblings, fathers and mothers and $k = 2, \ldots, 5$ corresponding to CPTDSD, CPT37D, SPAN10 and logTRLBA.

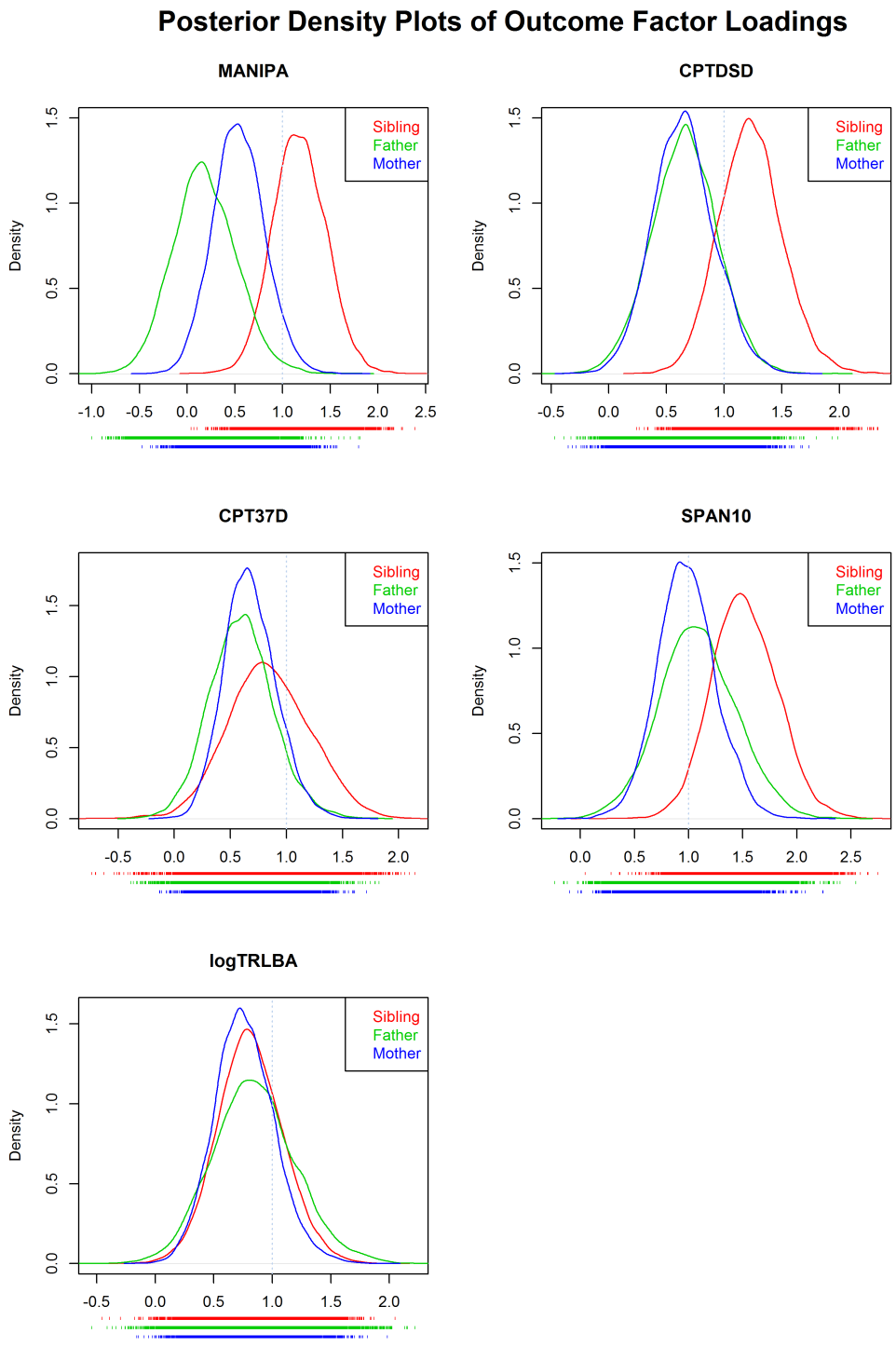


Figure D.3: Posterior density plots of outcome factor loadings, b_{jk}, for $j = 2, \ldots, 4$ corresponding to siblings, fathers and mothers and $k = 1, \ldots, 5$ corresponding to MANIPA, CPTDSD, CPT37D, SPAN10 and logTRLBA.

APPENDIX E

Descriptive Statistics and Prior Specification for the UCLA Family Study Data with Seventeen Outcomes

Table E.1 lists the sample correlations among observations of 17 outcomes from the UCLA NFS data, ignoring the family structure.

	CPTDSD	HitrDS	FalrDS	SPAN10	SPAN1	SPAN5	CPT37D	Hitr37	Falr37	logTRLBA	logTRLAA	VFFAS	NCFRSFSC	Manipa	MAINacc	MANIPrt
CPTDSDR	1.00															
HitrDS	.86	1.00														
FalrDS	.66	.35	1.00													
SPAN10	.34	.30	.32	1.00												
SPAN1	.25	.23	.28	.38	1.00											
SPAN5	.31	.29	.35	.74	.52	1.00										
CPT37D	.43	.40	.35	.43	.39	.44	1.00									
Hitr37	.35	.37	.27	.40	.36	.44	.88	1.00								
Falr37	.27	.24	.32	.31	.37	.40	.68	.52	1.00							
logTRLBA	.33	.30	.31	.47	.33	.47	.52	.45	.42	1.00						
logTRLAA	.27	.23	.20	.38	.21	.37	.44	.37	.38	.64	1.00					
VFFAS	.33	.30	.25	.32	.21	.30	.42	.37	.31	.51	.43	1.00				
NCFRSFSC	.27	.29	.18	.31	.19	.26	.34	.27	.34	.31	.32	.30	1.00			
Manipa	.29	.29	.19	.25	.22	.22	.37	.36	.24	.33	.32	.25	.18	1.00		
MAINacc	.32	.33	.25	.26	.26	.27	.48	.44	.34	.42	.34	.32	.26	.66	1.00	
MANIPrt	.00	.02	.02	-.02	.13	.04	.10	.13	.06	.05	.01	.11	.06	.10	.12	1.00
MAINrt	-.07	-.05	-.04	-.08	.09	-.02	.02	.03	.03	-.05	-.05	-.02	-.02	-.03	.05	.76

Table E.1: Sample correlations among observations of 17 outcomes from the UCLA NFS data, ignoring the family structure.

	β_{jk}		ψ_k		a_{jk}		b_{jk}		ϕ_{Bkk}
	Mean	Var	β	α	Mean	Var	Mean	Var	Mean
CPTDSD	28	121	194	10	1.000	1.000	1	1	48.4
HitrDS	63	529	846	10	2.091	4.372	1	1	211.6
FalrDS	7	225	360	10	1.364	1.860	1	1	90
SPAN10	50	25	40	10	0.455	0.207	1	1	10
SPAN1	60	20.25	32	10	0.409	0.167	1	1	8.1
SPAN5	60	59.29	95	10	0.700	0.490	1	1	23.7
CPT37D	41	81	130	10	0.818	0.669	1	1	32.4
Hitr37	95	361	578	10	1.727	2.983	1	1	144.4
Falr37	0.6	16	26	10	0.364	0.132	1	1	6.4
logTRLBA	-140	400	640	10	1.818	3.306	1	1	160
logTRLAA	-140	289	462	10	1.545	2.388	1	1	115.6
VFFAS	38	100	160	10	0.909	0.826	1	1	40
NCFRSFSC	23	17.64	28	10	0.382	0.146	1	1	7.056
MANIPA	70	169	270	10	1.182	1.397	1	1	67.6
MAINacc	70	169	270	10	1.182	1.397	1	1	67.6
MANIPrt	12	9	14	10	0.273	0.074	1	1	3.6
MAINrt	12	9	14	10	0.273	0.074	1	1	3.6

Table E.2: Means and variances of hyper-parameters for priors of regression coefficients/group means, μ, unique error variances, ψ_k, family member factor loadings, a_{jk}, outcome factor loadings, b_{jk} and family member factor variances, ϕ_{Bkk}, for $j = 1, \ldots, 4$.

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1																
2	.4	1															
3	.4	.4	1														
4	.2	.2	.2	1													
5	.2	.2	.2	.4	1												
6	.2	.2	.2	.4	.4	1											
7	.2	.2	.2	.2	.2	.2	1										
8	.2	.2	.2	.2	.2	.2	.4	1									
9	.2	.2	.2	.2	.2	.2	.4	.4	1								
10	.2	.2	.2	.2	.2	.2	.2	.2	.2	1							
11	.2	.2	.2	.2	.2	.2	.2	.2	.2	.4	1						
12	.2	.2	.2	.2	.2	.2	.2	.2	.2	.4	.4	1					
13	.2	.2	.2	.2	.2	.2	.2	.2	.2	.4	.4	.4	1				
14	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	1			
15	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.4	1		
16	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.4	.4	1	
17	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.2	.4	.4	.4	1

Table E.3: Prior correlation matrix of outcome factors, where outcomes $1, \ldots, 17$ are CPTDSD, DShitr, DSfalr, SPAN10, SPAN1, SPAN5, CPT37D, hitr37, falr37, logTRLBA, logTRLAA, VFFAS, NCFRSFSC, Manipa, MAINacc, MANIPrt and MAINrt.

APPENDIX F

Summary of Posterior Distributions for the UCLA Family Study Data with 17 Outcomes

Posterior densities of group means for schizophrenia and control families for 17 outcomes are presented in Figures F.1, F.2 and F.3. Table F.1 lists the posterior means of group means for control (Ctrl) and schizophrenia (SZ) families (left) and the posterior means of difference between group means (right). Table F.2 includes the posterior means of outcome factor correlations, while Table F.3 presents the posterior means of outcome factor variances, family member factor loadings and outcome factor loadings.

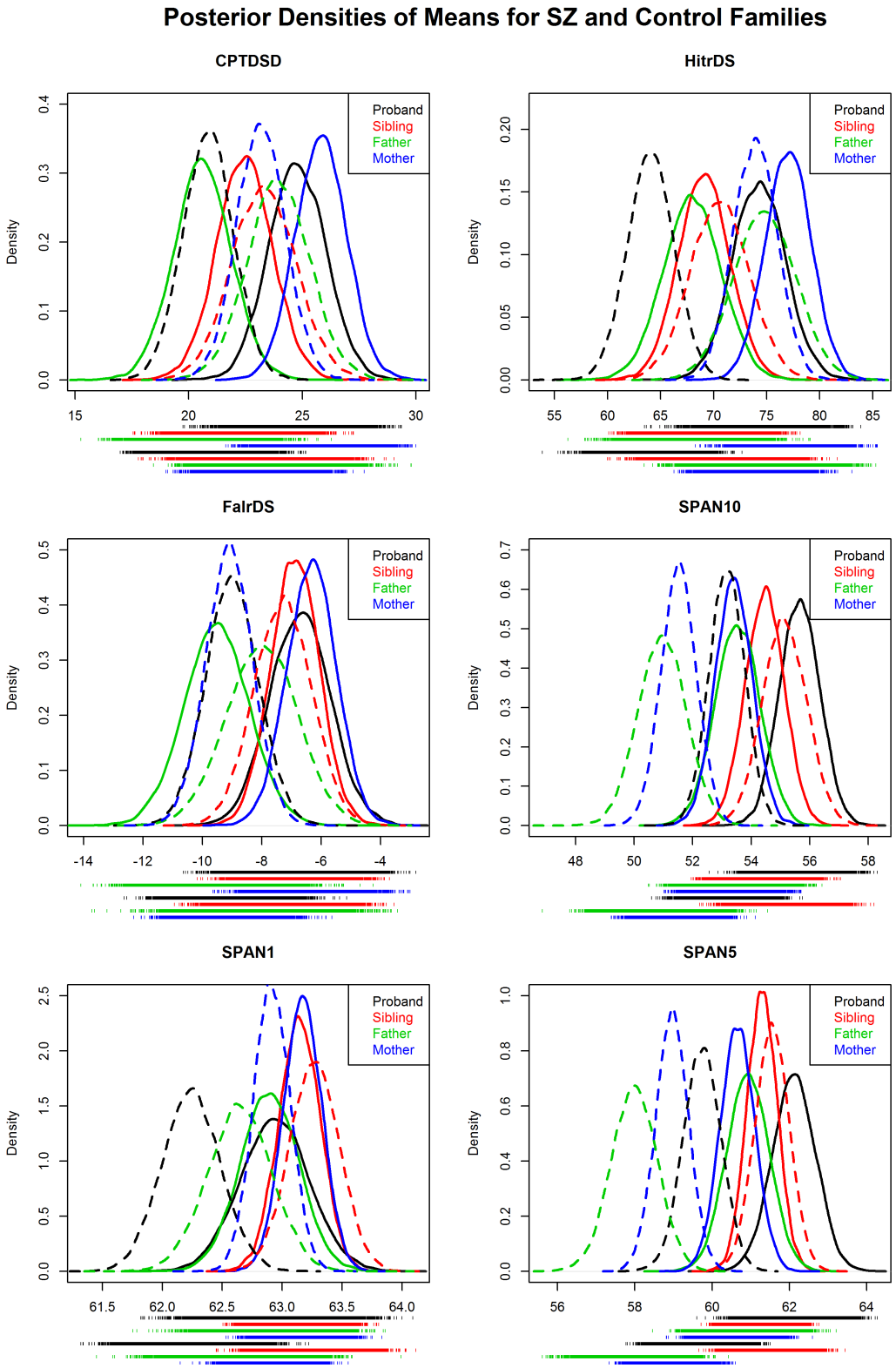


Figure F.1: Posterior densities of means for probands, fathers, mothers and siblings in schizophrenia family (dashed lines) and control family (solid lines).

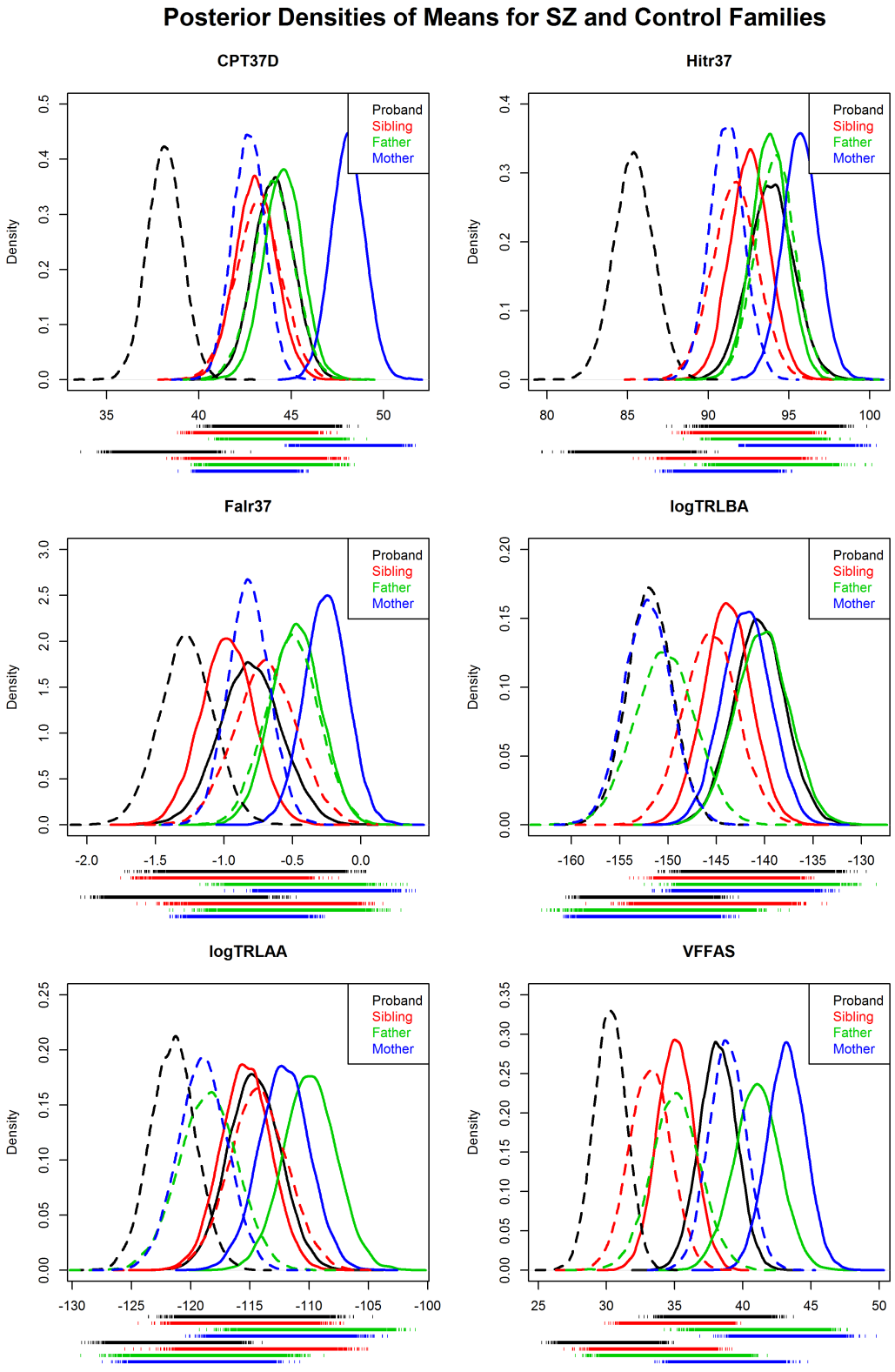


Figure F.2: Posterior densities of means for probands, fathers, mothers and siblings in schizophrenia family (dashed lines) and control family (solid lines), continued.

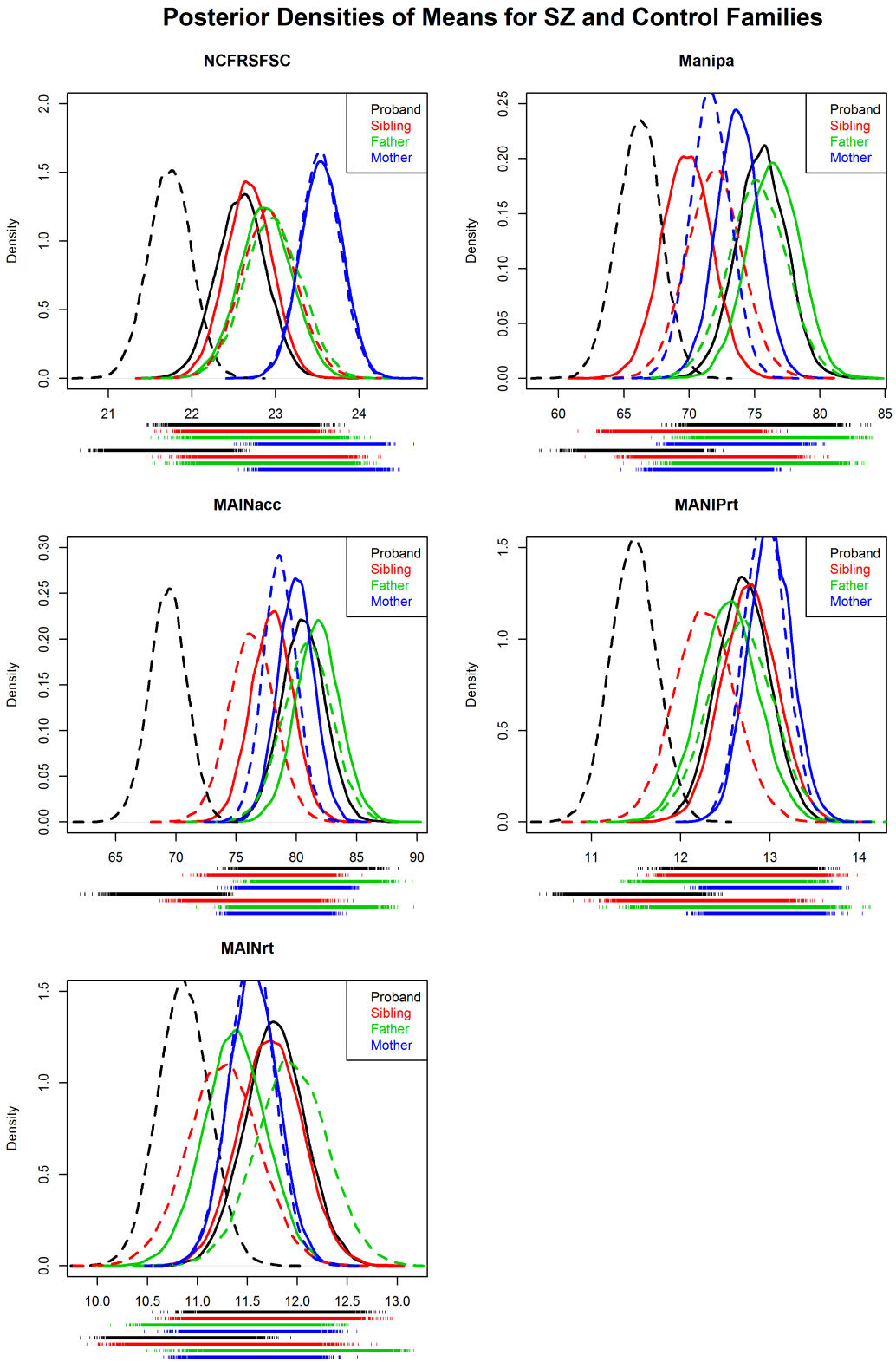


Figure F.3: Posterior densities of means for probands, fathers, mothers and siblings in schizophrenia family (dashed lines) and control family (solid lines), continued.

	Group Means								Diff between Group Means			
	Proband		Sibling		Father		Mother		Proband	Sibling	Father	Mother
Outcome	Ctrl	SZ	Ctrl	SZ	Ctrl	SZ	Ctrl	SZ	Control minus SZ			
CPTDSD	24.78	20.86	22.47	23.32	20.61	23.96	25.84	23.18	3.92	-0.84	-3.35	2.66
HitrDS	74.26	64.09	69.13	70.63	68.05	74.74	77.05	73.94	10.18	-1.50	-6.69	3.11
FalrDS	-6.72	-9.00	-6.87	-7.27	-9.54	-8.03	-6.32	-9.09	2.29	0.40	-1.51	2.77
SPAN10	55.65	53.17	54.47	55.14	53.54	50.96	53.41	51.51	2.48	-0.66	2.57	1.90
SPAN1	62.94	62.24	63.15	63.27	62.90	62.65	63.17	62.91	0.70	-0.12	0.25	0.26
SPAN5	62.12	59.74	61.30	61.54	60.92	57.98	60.66	58.96	2.38	-0.24	2.94	1.70
CPT37D	44.05	38.16	43.09	43.25	44.56	44.04	48.08	42.70	5.89	-0.16	0.52	5.38
Hitr37	93.90	85.38	92.58	91.66	93.78	94.15	95.74	91.15	8.52	0.92	-0.36	4.59
Falr37	-0.80	-1.27	-0.97	-0.70	-0.47	-0.50	-0.25	-0.82	0.47	-0.28	0.03	0.57
logTRLBA	-140.6	-151.9	-143.8	-145.6	-140.3	-150.4	-141.9	-152.1	11.3	1.8	10.1	10.2
logTRLAA	-114.7	-121.5	-115.3	-114.3	-110.0	-118.5	-112.1	-119.0	6.8	-1.0	8.5	6.9
VFFAS	38.16	30.28	35.05	33.22	40.97	35.12	43.20	38.87	7.88	1.82	5.85	4.33
NCFRSFSC	22.59	21.74	22.68	22.92	22.88	22.97	23.55	23.53	0.85	-0.24	-0.09	0.02
Manipa	75.52	66.17	69.86	71.93	76.46	75.20	73.69	71.60	9.35	-2.07	1.26	2.09
MAINacc	80.57	69.35	78.01	76.30	81.76	80.91	79.95	78.64	11.22	1.72	0.86	1.30
MANIPrt	12.70	11.47	12.77	12.27	12.53	12.67	13.00	12.94	1.23	0.51	-0.14	0.06
MAINrt	11.78	10.86	11.73	11.26	11.38	11.94	11.56	11.54	0.91	0.47	-0.56	0.02

Table F.1: Posterior means of group means for control (Ctrl) and schizophrenia (SZ) families (left) and the posterior means of difference between group means (right).

	Outcome Factor Correlations															
	CPTDSD	DShitr	DSfalr	SPAN10	SPAN1	SPAN5	CPT37D	hitr37	falr37	logTRLBA	logTRLAA	VFFAS	NCFRSFSC	Manipa	MAINacc	MANIPrt
CPTDSD	1.00															
DShitr	0.82	1.00														
DSfalr	0.67	0.58	1.00													
SPAN10	0.35	0.35	0.29	1.00												
SPAN1	0.20	0.18	0.28	0.40	1.00											
SPAN5	0.32	0.31	0.33	0.77	0.57	1.00										
CPT37D	0.44	0.36	0.34	0.38	0.30	0.40	1.00									
hitr37	0.37	0.33	0.28	0.44	0.35	0.49	0.86	1.00								
falr37	0.30	0.24	0.34	0.31	0.31	0.38	0.63	0.55	1.00							
logTRLBA	0.38	0.36	0.32	0.58	0.34	0.52	0.45	0.43	0.37	1.00						
logTRLAA	0.30	0.26	0.26	0.53	0.28	0.46	0.43	0.38	0.37	0.76	1.00					
VFFAS	0.45	0.42	0.33	0.56	0.28	0.49	0.53	0.49	0.41	0.79	0.71	1.00				
NCFRSFSC	0.25	0.26	0.22	0.30	0.03	0.21	0.12	0.10	0.21	0.36	0.39	0.38	1.00			
Manipa	0.47	0.42	0.36	0.37	0.23	0.32	0.39	0.36	0.23	0.54	0.51	0.56	0.26	1.00		
MAINacc	0.52	0.47	0.40	0.41	0.28	0.37	0.50	0.47	0.31	0.57	0.51	0.60	0.25	0.86	1.00	
MANIPrt	-0.13	-0.08	-0.04	-0.03	0.14	0.04	-0.11	-0.01	0.02	0.06	-0.02	0.03	0.12	0.03	0.05	1.00
MAINrt	-0.19	-0.13	-0.07	-0.07	0.12	0.00	-0.17	-0.08	-0.01	-0.01	-0.07	-0.06	0.06	-0.07	-0.05	0.79

Table F.2: Posterior means of outcome factor correlations.

	Outcome	Family Member Factor Loading				Outcome Factor Loading			
Outcome	Fac Var	Prob	Sib	Fa	Mo	Pro	Sib	Fa	Mo
CPTDSD	35.07	1	1	1	1	1	1.30	0.60	0.93
DShitr	116.87	2.02	1.63	2.31	1.75	1	1.34	0.67	0.89
DSfalr	27.72	0.70	0.53	0.89	0.50	1	0.58	0.49	0.64
SPAN10	9.27	0.58	0.60	0.55	0.63	1	1.27	1.02	1.01
SPAN1	3.01	0.19	0.16	0.18	0.04	1	0.07	0.21	0.03
SPAN5	10.44	0.43	0.32	0.40	0.48	1	0.54	0.75	0.63
CPT37D	52.73	0.85	2.00	1.48	1.62	1	0.15	0.18	0.22
hitr37	97.86	0.93	2.10	1.32	1.84	1	0.06	0.08	0.15
falr37	1.92	0.15	0.28	0.10	0.18	1	0.24	0.10	0.10
logTRLBA	151.98	2.10	2.59	2.17	3.18	1	0.87	0.78	0.66
logTRLAA	95.26	1.47	1.65	0.97	2.33	1	1.01	0.32	0.64
VFFAS	42.98	0.77	1.00	0.84	0.92	1	0.87	0.96	1.14
NCFRSFSC	2.43	0.17	0.18	0.20	0.20	1	0.92	0.35	0.52
Manipa	100.61	0.89	1.45	1.23	1.45	1	0.95	0.14	0.27
MAINacc	92.92	0.98	1.71	1.31	1.28	1	0.78	0.33	0.44
MANIPrt	2.36	0.01	0.14	0.12	0.04	1	0.99	0.43	0.24
MAINrt	2.46	-0.03	0.13	0.07	-0.01	1	1.10	0.39	0.31

Table F.3: Posterior means of outcome factor variances, family member factor loadings and outcome factor loadings.

BIBLIOGRAPHY

Aguilar, O. and West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics* **18**, 338–357.

Albright, J. J. and Park, H. M. (2009). Confirmatory factor analysis using Amos, LISREL, Mplus, and SAS/STAT CALIS. *The Trustees of Indiana University* **1**, 1–85.

Arbuckle, J. L. (2011). *IBM® SPSS® AMOS 20 Users Guide*. IBM Corporation.

Asarnow, R. F., Nuechterlein, K. H., Fogelson, D., Subotnik, K. L., Payne, D. A., Russell, A. T., Asamen, J., Kuppinger, H., and Kendler, K. S. (2001). Schizophrenia and schizophrenia-spectrum personality disorders in the first-degree relatives of children with schizophrenia: The UCLA Family Study. *Archives of General Psychiatry* **58**, 581–588.

Asarnow, R. F., Nuechterlein, K. H., Subotnik, K. L., Fogelson, D. L., Torquato, R. D., Payne, D. L., Asamen, J., Mintz, J., and Guthrie, D. (2002). Neurocognitive impairments in non-psychotic parents of children with schizophrenia and attention-deficit/hyperactivity disorder: The University of California, Los Angeles Family Study. *Archives of General Psychiatry* **59**, 1053–1060.

Bartholomew, D. J., Knott, M., and Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*. Wiley.

Basilevsky, A. T. (2009). *Statistical Factor Analysis and Related Methods: Theory and Applications*. Wiley.

Benton, A. L. (1994). *Contributions to Neuropsychological Assessment: A Clinical Manual*. Oxford University Press.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer, 2nd edition.

Blunch, N. (2012). *Introduction to Structural Equation Modeling Using IBM SPSS Statistics and AMOS*. Sage.

Bollen, K. A. (1998). *Structural Equation Models*. Wiley.

Brannick, M. T. and Spector, P. E. (1990). Estimation problems in the block-diagonal model of the multitrait-multimethod matrix. *Applied Psychological Measurement* **14**, 325–339.

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. The Guilford Press.

Browne, M. W. (1984). The decomposition of multitrait-multimethod matrices. *British Journal of Mathematical and Statistical Psychology* **37**, 1–21.

Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International Journal of Testing* **1**, 55–86.

Byrne, B. M. (2009). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. CRC Press.

Byrne, B. M. (2013). *Structural Equation Modeling with EQS: Basic Concepts, Applications, and Programming*. Routledge.

Campbell, D. T. and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* **56**, 81–105.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association* **90**, 1313–1321.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin* **105**, 317–327.

Dickey, J. M. and Lientz, B. P. (1970). The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov chain. *The Annals of Mathematical Statistics* **41**, 214–226.

Donner, A. and Koval, J. J. (1980). The estimation of intraclass correlation in the analysis of family data. *Biometrics* **36**, 19–25.

Eid, M., Lischetzke, T., and Nussbeck, F. W. (2006). *Structural Equation Models for Multitrait-multimethod Data.*, volume xiv, pages 283–299. American Psychological Association.

Fox, J. (2006). Teacher's corner: Structural equation modeling with the sem package in R. *Structural Equation Modeling* **13**, 465–486.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.

Geweke, J. and Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies* **9**, 557–587.

Ghosh, J. and Dunson, D. B. (2009). Default prior distributions and efficient posterior computation in Bayesian factor analysis. *Journal of Computational and Graphical Statistics* **18**, 306–320.

Grayson, D. and Marsh, H. (1994). Identification with deficient rank loading matrices in confirmatory factor analysis: Multitrait-multimethod models. *Psychometrika* **59**, 121–134.

Hamsten, A. and de Faire, U. (1987). Risk factors for coronary artery disease in families of young men with myocardial infarction. *The American Journal of Cardiology* **59**, 14–19.

Harman, H. H. (1960). *Modern Factor Analysis*. University of Chicago Press.

Harrap, S. B., Stebbing, M., Hopper, J. L., Hoang, H. N., and Giles, G. G. (2000). Familial patterns of covariation for cardiovascular risk factors in adults: The Victorian Family Heart Study. *American Journal of Epidemiology* **152**, 704–715.

Hunter, J. L. (2005). *Structural Equation Modeling with SAS-CALIS*. Ph.D. thesis, Kansas State University.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* **34**, 183–202.

Jöreskog, K. G. and Sörbom, D. (2012). Lisrel 9.1 [computer software]. *Lincolnwood, IL: Scientific Software International* .

Karlin, S., Cameron, E. C., and Williams, P. T. (1981). Sibling and parent–offspring correlation estimation with variable family size. *Proceedings of the National Academy of Sciences* **78**, 2664–2668.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**, 773–795.

Kenny, D. and Kashy, D. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin* **112**, 165–172.

Kim, J., Glahn, D. C., Nuechterlein, K. H., and Cannon, T. D. (2004). Maintenance and manipulation of information in schizophrenia: Further evidence for impairment in the central executive component of working memory. *Schizophrenia Research* **68**, 173–187.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. Guilford press.

Koide, T., Aleksic, B., Kikuchi, T., Banno, M., Kohmura, K., Adachi, Y., Kawano, N., Iidaka, T., and Ozaki, N. (2012). Evaluation of factors affecting continuous performance test identical pairs version score of schizophrenic patients in a Japanese clinical sample. *Schizophrenia Research and Treatment* **2012**, 1–5.

Kopelowicz, A., Liberman, R. P., Ventura, J., Zarate, R., and Mintz, J. (2005). Neurocognitive correlates of recovery from schizophrenia. *Psychological Medicine* **35**, 1165–1174.

Lance, C. E., Noble, C. L., and Scullen, S. E. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. *Psychological Methods* **7**, 228–244.

Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. Wiley, 2nd edition.

Loehlin, J. C. (2004). *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis*. Lawrence Erlbaum Associates, 4th edition.

Lopes, H. F. and West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica* **14**, 41–68.

Madans, J., Miller, K., Maitland, A., and Willis, G. (2011). *Question Evaluation Methods: Contributing to the Science of Data Quality*. Wiley.

Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement* **13**, 335–361.

Marsh, H. W. and Hocevar, D. (1988). A new, more powerful approach to multitrait-multimethod analyses: Application of second-order confirmatory factor analysis. *Journal of Applied Psychology* **73**, 107–117.

McCulloch, C. E. (2006). *Generalized Linear Mixed Models*. Wiley.

Morey, R. D., Rouder, J. N., Pratte, M. S., and Speckman, P. L. (2011). Using MCMC chain outputs to efficiently estimate Bayes factors. *Journal of Mathematical Psychology* **55**, 368–378.

Morris, N. J. (2009). *Multivariate and Structural Equation Models for Family Data*. Ph.D. thesis, Case Western Reserve University.

Muthén, L. K. and Muthén, B. O. (1998–2012). *Mplus*. Muthén & Muthén, 7th edition.

Naik, D. N. and Rao, S. S. (2001). Analysis of multivariate repeated measures data with a Kronecker product structured covariance matrix. *Journal of Applied Statistics* **28**, 91–105.

Nelson, B. D. (2001). Variable reduction for modeling using PROC VARCLUS. In *SAS Users Group International Proceedings*, pages 261–263.

Nuechterlein, K. H., Asarnow, R. F., Subotnik, K. L., Fogelson, D. L., Payne, D. L., Kendler, K. S., Neale, M. C., Jacobson, K. C., and Mintz, J. (2002). The structure of schizotypy: Relationships between neurocognitive and personality disorder features in relatives of schizophrenic patients in the UCLA Family Study. *Schizophrenia Research* **54**, 121–130.

Nuechterlein, K. H., Subotnik, K. L., Green, M. F., Ventura, J., Asarnow, R. F., Gitlin, M. J., Yee, C. M., Gretchen-Doorly, D., and Mintz, J. (2011). Neurocognitive predictors of work outcome in recent-onset schizophrenia. *Schizophrenia Bulletin* **37**, S33–S40.

Press, S. J. (2012). *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*. Courier Dover Publications.

Press, S. J. and Shigemasu, K. (1989). Bayesian inference in factor analysis. In L. J. Gleser, M. D. Perlman, S. J. Press, and A. R. Sampson, editors, *Contributions to Probability and Statistics*, chapter 15, pages 271–287. Springer.

Press, S. J. and Shigemasu, K. (1997). Bayesian inference in factor analysis (revised). Technical Report 243, University of California, Riverside.

Quinn, K. M. (2004). Bayesian factor analysis for mixed ordinal and continuous responses. *Political Analysis* **12**, 338–353.

Rao, C. R. (1955). Estimation and tests of significance in factor analysis. *Psychometrika* **20**, 93–111.

Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer, 2nd edition.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software* **48**, 1–36.

Rowe, D. B. (1998). *Correlated Bayesian Factor Analysis*. Ph.D. thesis, Univsersity of California, Riverside.

Rowe, D. B. (2003). *Multivariate Bayesian Statistics: Models for Source Separation and Signal Unmixing*. CRC Press.

Rubin, D. B. and Thayer, D. T. (1982). EM algorithms for ML factor analysis. *Psychometrika* **47**, 69–76.

Sammel, M., Lin, X., and Ryan, L. (1999). Multivariate linear mixed models for multiple outcomes. *Statistics in Medicine* **18**, 2479–2492.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. CRC press.

Schultz, S. H., North, S. W., and Shields, C. G. (2007). Schizophrenia: A review. *American Family Physician* **75**, 1821–1829.

Springate, B. and Fein, D. (2013). Halstead-Reitan neuropsychological test battery. In *Encyclopedia of Autism Spectrum Disorders*, pages 1479–1481. Springer.

Srivastava, M. S., von Rosen, T., and von Rosen, D. (2008). Models with a Kronecker product covariance structure: Estimation and testing. *Mathematical Methods of Statistics* **17**, 357–370.

Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. American Psychological Association.

Verdinelli, I. and Wasserman, L. (1995). Computing Bayes factors using a generalization of the Savage-Dickey density ratio. *Journal of the American Statistical Association* **90**, 614–618.

Wang, X., Guo, X., He, M., and Zhang, H. (2011). Statistical inference in mixed models and analysis of twin and family data. *Biometrics* **67**, 987–995.

West, M. (2003). Bayesian factor regression models in the "large p, small n" paradigm. *Bayesian Statistics* **7**, 723–732.

Wothke, W. (1984). *The Estimation of Trait and Method Components in Multitrait-Multimethod Measurement*. Ph.D. thesis, University of Chicago.

Wothke, W. and Browne, M. W. (1990). The direct product model for the MTMM matrix parameterized as a second order factor analysis model. *Psychometrika* **55**, 255–262.