12-1-2009

Joint model of a longitudinal process and informative time schedule data

Michael Bronsert

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

A JOINT MODEL OF A LONGITUDINAL PROCESS AND INFORMATIVE TIME SCHEDULE DATA

A Dissertation Submitted in Partial Fulfillment
Of the Requirements for the Degree of
Doctor of Philosophy

Michael Richard Bronsert

College of Education and Behavioral Sciences
School of Educational Research Leadership and Technology
Department of Applied Statistics and Research Methods

December, 2009

This Dissertation by: Michael Richard Bronsert

Entitled: *A Joint Model of a Longitudinal Process and Informative Time Schedule Data*

has been approved as meeting the requirement of the Degree of Doctor of Philosophy in College of Education and Behavioral Sciences in School of Educational Research Leadership and Technology, Program of Applied Statistics and Research Methods

Accepted by the Doctoral Committee

Khalil Shafie, Ph. D., Chair

Daniel J. Mundfrom, Ph. D., Committee Member

Jay R. Schaffer, Ph. D., Committee Member

Robert L. Heiny, Ph. D., Faculty Representative

Date of Dissertation Defense 7 May 2009

Accepted by the Graduate School

Robbyn R. Wacker, Ph. D.
Assistant Vice President for Research
Dean of the Graduate School & International Admissions

ABSTRACT

Bronsert, Michael Richard. *A Joint Model of a Longitudinal Process and Informative Time Schedule Data.* Published Doctor of Philosophy Dissertation, University of Northern Colorado, 2009.

Longitudinal studies are commonly encountered in a variety of research areas in which the scientific interest is in the pattern of change in a response variable over time. These observations are traditionally scheduled prospectively and therefore common fixed time interval models for repeated measurements are adequate. Conversely, in informative schedule studies in which subsequent observations are scheduled on the basis of prior response outcomes, time between observations now becomes informative in the longitudinal process. Traditional fixed time approaches, however, are unable to utilize the informative nature of the data lessening the inferences achieved by these approaches. Therefore, the purpose of this research was the development of a joint model of a longitudinal process and informative time schedule data. Maximum likelihood estimates (MLE) for two special cases of the proposed model were obtained from Monte Carlo simulated data by employing the Multivariate Newton-Raphson optimization routine implemented in a SAS/IML call statements. Parameter estimates were determined for a few select cases of subject and observation length and included parameter estimates for rectangular and nonrectangular observation matrices. Finally, estimates obtained from

PROC MIXED and from the proposed model were compared for accuracy and efficiency by examining their bias, variance, mean square error (MSE), and relative efficiency.

ACKNOWLEDGEMENTS

I would like to thank all the people who guided and supported me throughout my graduate studies, including this dissertation. First of all, I would like to thank Dr. Khalil Shafie for his guidance and support throughout the development and writing of my doctoral dissertation. Without his advice and deep understanding of mathematics, I would not have succeeded. I would also like to thank Dr. Daniel Mundfrom, Dr. Jay Schaffer, and Dr. Robert Heiny for their constructive criticism and support of the development of this dissertation.

I also wish to thank my parents, Neal and Patty Bronsert for their continual support and love. I especially would like to thank my wife, Bridget Bronsert for her support and understanding of my educational and career goals and the sacrifices she made so that I might obtain those goals. Most importantly, I would like to thank my son, Jonah Bronsert, and daughter, Makenna Bronsert for their gift of love and patience. I am indebted to them both.

Finally, I would like to thank and remember my grandmother, Gertrude 'Trudy' Long. She was indisputably the kindest and the most empathetic individual that I ever had the fortune to know and love. Her genuine interest in the lives of all who knew her and her ability to make even the worst day a little brighter will be sorely missed by all. I love you Grandma and I will meet you under the pine tree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Repeated measurement data arises when measurements of the same response variable are taken repeatedly on each of a number of experimental units or subjects which may be allocated to one of several treatment schemes. Repeated measurement data are in contrast to cross-sectional designs in which a single measurement of the response variable is taken on each subject which also may be allocated to one of several treatment schemes. The major advantage of repeated measurement designs over cross-sectional designs is their capacity to separate inherent between-subject from within-subject variability (Diggle, Heagerty, Liang, & Zeger, 2002). This separation of the two variability sources allows for the characterization in the change of the response variable across observations and the factors that influence that change (Fitzmaurice, Laird & Ware, 2004). However, repeated measurement designs, in general, require more complex computational approaches than cross-sectional designs since observations on each subject are considered correlated, i.e., subsequent response measurements are dependent on prior measurement values. For example, the amount of weight an individual is able to lose following the administration of a weight loss pill is dependent on his or her prior weight, i.e., heaver individuals may have more opportunity to lose more weight than lighter individuals. This correlation, if ignored, would potentially result in overestimation of the

sampling variability since the excess amount of variability shared between correlated observations would not be removed from the estimates of variability obtained for each observation separately. In essence, this failure to remove the overlapping variability would result in its inclusion in the overall estimate twice. Consequently, this overestimation of variability would in turn lead to an overly pessimistic estimate of precision which ultimately could result in misleading inferences obtained from the use of this variability estimate (Fitzmaurice et al., 2004).

The term ‘longitudinal data’ has also been applied to the study of repeated measurements in which the response variable is observed over a given time period (Davis, 2002). These studies are commonly encountered in epidemiology, clinical trials and social science studies where the scientific interest is in the pattern of change in a response variable over time (Hedeker & Gibbons, 2006). That is to say, that time of observation is considered a factor in the explanation of the change in the response variable along with other planned factors of interest. For example, this approach allows research practitioners to evaluate how a set of given factors or a single factor (e.g., preventive care protocols, novel drug treatments, skills training, etc.) effects changes in a response variable (e.g., disease progression, biomarker changes, results on a skill assessment test, etc.) across a given time period. Furthermore, this statistical method also allows practitioners to characterize changes in a response variable (e.g., aneurysm size, tumor growth, etc.) over a given time period in the absences of other explanatory factors other than time itself.

A cornucopia of longitudinal methods has been developed to accommodate several different study designs along with a variety of response and explanatory variable

types (c.f., Crowder & Hand, 1990: Laird, Donnelly & Ware, 1992: Lindsey, 1993: Everitt, 1995: Keselman, Alginas & Kowalchuk, 2001). These methods range from a simple univariate approach to the more complex mixed-effects models, but in general each method is often utilized more often for a specific discipline or developed to solve a particular research objective that other approaches fail to address adequately (Davis, 2002). That is to say, the choice of a particular method utilized by a researcher depends on the objective of the research project, the particular design of the study protocol, and the nature of the process that generates the responses observed during the study. For example, mixed-effects models were developed to address research objectives that traditional repeated measurement approaches were unable to achieve due to the overly restrictive assumption of constant variances and the inability to analyze datasets containing missing values in these models. However, these traditional approaches may be preferred for some designs in which these limitations are of less importance or absent altogether since they are, in general, less computationally complex. Thus, when choosing a statistical model one methodological approach's strength may be its weakness given a different set of research objectives and the underlying process that generates the observed response variable.

Statement of the Problem

Despite the variety of approaches to the analysis of longitudinal data, a common characteristic of each method is the assumption that time of observation is a fixed factor. This assumption limits the inferential scope or the explanatory ability of the model to the specific times observed within the given study protocol (Montgomery, 2005). Indeed, in

experimental longitudinal studies in which observation times are prospectively scheduled on the bases of theoretical, pharmacokinetic or convenience reasons, this assumption is valid by design. Here, each subject regardless of treatment group would be observed more or less at the same time periods resulting in relatively consistent time intervals across subjects. However, this approach is in contrast to a so called 'observational' longitudinal study in which a different stochastic structure is present in the data collection protocol. In this design, observation periods are not prospectively scheduled but are adaptively determined on the bases of prior response outcomes, i.e., subsequent observation periods are determined based on the outcome of the response variable of the previous observations. This adaptive scheduling approach based on prior response outcomes, therefore, assumes that time between subsequent observations has inherent information to contribute to the explanation of the changes in the response variable or assumes an 'informative schedule' design. The informative nature of this design can be appreciated in that shorter time frames between two given observations would most likely have smaller changes observed in the response variable while longer time frames would most likely have correspondingly larger changes in the response variable. It is also important to note that, potentially, each subject would have different informative lengths of observations suggesting that time in this model is no longer a fixed factor in the explanation of changes observed in the response variable. Therefore, applying existing longitudinal models with the assumption that time is a fixed factor would result in incorrect estimates of the sampling variability and could therefore result in misleading inferences when applied to data having this stochastic structure. This inability for existing longitudinal models to account for the informative nature of observed time

schedules suggests a need for a method that jointly models the distribution of informative intermittent times and corresponding measured responses, and not the usual conditional models, measured responses given the schedule times. The utilization of this joint model on informative schedule data would potentially result in more accurate estimates of sampling variability and therefore, would improve the overall generalizabiltiy of the given study.

Purpose and Research Questions

The purpose of this study, as mentioned above, was the development of a novel approach that jointly models a longitudinal process with the addition of an informative component for time of observation. The addition of informative time schedules, as opposed to fixed time schedules employed in traditional longitudinal methods, would potentially broaden the inferential scope obtained when applied to informative schedule data. This increased scope of inference allows for improved modeling of the change in a response variable over time by utilizing the additional information captured in the informative schedules. To achieve this goal of modeling an informative component for time along with repeatedly measured responses, this study investigated the following research questions:

1. Can a novel approach be developed that would jointly model a longitudinal response variable with a set of corresponding intermittent informative time intervals of observations?

2. Can an efficient numerical iterative method be developed to determine the maximum likelihood estimates for the proposed model?

3. In the presence of Monte Carlo simulated informative schedule data, how accurate and efficient is this proposed model in estimating the population parameters?

4. How are these maximum likelihood estimates influenced by a few select variations in subject sample size, total number of observations for each subject, and the degree of variation in observation lengths for each subject contained in the simulated sample?

5. Finally, how does the proposed model's parameter estimates compare on accuracy and efficiency with common parameter estimates obtained by the mixed-effects model implemented in SAS PROC MIXED when analyzing simulated informative schedule data?

Justification for This Study

Traditionally in longitudinal studies, observational times are prospectively scheduled based on some design protocol prior to the initiation of the study as mentioned above. Despite the underlying rationale for the chosen protocol, a prospective observational schedule may not be the best approach for all research questions. In these situations an informative schedule paradigm incorporating an adaptive observation schedule may be more beneficial in achieving the research objectives not to mention improving patient care over traditional approaches. This benefit can be seen in a study on the enlargement of Abdominal Aortic Aneurysm (AAA) in which patients' aneurysms were observed over a given time period to better characterize rate of growth and the accompanying risk of rupture without surgical repair. At some time, in this time interval, patients would enter the experiment and their aneurysm's sizes would be measured and depending on the observed size would be randomly assigned to either a surgical repair or a surveillance group. Patients belonging to the surveillance group would have their aneurysms measured by ultrasonography during each physician visit. Depending on the measurement observed the next observation time would be scheduled, where presumably

the larger the size of the aneurysm the closer the next appointment would be and hence the smaller the changes in size of the aneurysms would be observed. Observations for each patient would continue until a predetermined size was reached in which the patient would then enter the surgical repair group. Surgical repair was eventually performed on all patients to prevent the risk of rupture which could be fatal (Ingoldby, Wujanto & Mitchell, 1986).

In the above study protocol it is obvious that a fixed schedule paradigm may not be in the best interest of patient care given the risk of rupture in patients with larger aneurysms. The use of an adaptive schedule approach therefore, would allow shorter observation intervals for patients with larger aneurysm sizes reducing the risk of a rupture occurring between physician visits. This scheduling approach, as presented above, would subsequently result in different observation intervals for each patient that would be dependent at least on the last observed size of the patient's aneurysm. That is, even if each patient was observed a fixed number of times, the intervals between observations would not be the same and since the magnitude of each interval is dependent on the prior outcome, these interval measurements would contribute informatively to the process of change in aneurysm size in these individuals. Furthermore, the number of observations for each subject would most likely not be equal since some individual's aneurysms would take longer to reach the critical size for surgery requiring a longer observation period than others. These two conditions would therefore result in each subject's observation vectors being of different lengths and having different intervals between each observation resulting in a 'nonrectangular' schedule design. This nonrectangular characteristic of the sample matrix obtained from informative schedule designs prevents the use of traditional

analytic strategies for longitudinal studies, such as repeated measures analysis of variance (ANOVA) or multivariate analysis of variance (MANOVA) since they typically require all data to be available on all subjects at each measurement point (Diggle et al., 2002). The use of these methods would therefore require that the resulting data obtained under an informative schedule design be modified to accommodate their model assumptions. However, transformation methods such as deleting missing data or imputing missing observations can lead to substantial bias and undermine the validity of the results obtained (Lavori, 1992; Gibbons et al., 1993; Taylor & Amir, 1994). Furthermore, deleting or imputing data would essentially weaken the informative nature of the time intervals observed in this design by either removing them altogether or substituting misleading observations into the sample, respectively. Another approach would be the use of mixed-effects models which by the nature of their design allow for the analysis of nonrectangular sample matrices (Laird & Ware, 1982). However, these methods still require the assumption that time is a fixed factor in the explanation of the response process and therefore would result in the loss of the informative nature of the time intervals. While mixed-effects models would allow for estimations of the growth process to be obtained, these results would essentially restrict the generalizabiltiy of the given study since they treat the observed time schedules as a fixed factor.

It should also be noted that the total number of observations for each patient could also be different, given that not all patients would start at the beginning of the surveillance nor would each patient's initial aneurysm size be the same at the entrance of the study further contributing to the nonrectangular aspect of the sample designs. This latter issue of different initial aneurysm sizes would potentially result in truncation of the

aneurysm enlargement process. That is, patients with initially larger aneurysms and therefore more serious risk of rupture would more quickly reach the size requiring surgical repair causing them to leave the surveillance group sooner than patients with smaller initial sizes. The patients that experience the event of surgery earlier would therefore have their observations underrepresented in the sample which would result in the bias of the actual growth rate estimate. This condition is referred to as informative censoring as discussed by Wu & Carroll (1998), Hogan & Laird (1997a, b), and many other authors. Consequently, data containing informative censoring has been shown to give biased results when analyzed by mixed-effects models suggesting an analytical weakness in these model designs when faced with nonrectangular sample matrices. However, this issue of informative censoring would potentially be less of an analytical problem in an informative schedule design due to the assumed observation schedule protocol. In other words, individuals that have a more progressive or serious condition requiring early surgical intervention which would potentially result in informative censoring occurring would also have shorter observation schedules and subsequently more observations measured. This increased observational schedule would allow for individuals with rapid aneurysm growth to have more influence on the overall estimate of the growth process by the inclusion of more observations in the obtained sample matrix. Thus, the use of mixed-effects models to analyze informative schedule designs would not only result in restricted inferences but would potentially result in biased estimates in the growth process itself.

The purpose for studying the nature of aneurysm growth in these individuals was to better characterize the average and inter-patient variability in AAA expansion which

would allow for the development of more accurate protocols outlining when surgical intervention is necessary for patients with AAA. This better understanding of the growth process of AAA is necessary since it has been shown that a policy of early elective surgery for small aneurysms does not generally improve mortality rate (Lederle, Wilson, Johnson, Reinke, Littooy, Acher et al., 2002). Therefore, the utilization of an informative scheduling design would allow for improved estimates of the AAA growth process and would subsequently allow for more accurate determination of surveillance protocols and ultimately improved patient survivability. Furthermore, this model's utility is not necessarily limited to the above research project but can be beneficial in any study design in which an informative schedule model would be beneficial to the study participants by increasing the frequency of observation or when improved accurate estimates of the response variable are required especially in the presence of informative censoring. For example, the biological behavior of early gastric carcinoma, especially its growth rate, is not well documented and remains a significant cause of cancer deaths (Jemal et al., 2005). Furthermore, long term survival after surgery for gastric cancer is poor but prognosis improves with early detection, which suggests the need for accurate estimation of early development of gastric carcinoma (Heemsker, Lentze, Hulsewe & Hoofwijk, 2007). This area of research could potentially benefit from an informative scheduling design in light of the finding that some malignancies can grow rapidly (Haruma et al., 1991). The employment of an informative design in the evaluation of gastric carcinoma would allow for increased observation of patients with greater potential for the development of gastric carcinomas. This increased observation of patients with aggressive conditions would have the added benefit of improved estimates of gastric

carcinoma development since informative censoring would potentially be an issue in these patients who enter surgical intervention quickly. Other research areas that might benefit from this design would be in the field of psychiatry. For example, an increase in patient evaluation would better characterize the benefit of some novel psychopharmacological agent or psychotherapeutic method, especially in cases where failure to elicit improvements in patient conditions could result in adverse mental states or potentially result in patient suicide (Simon & Savarino, 2008). Once again, the use of an informative schedule design would allow for this desired increase in patient observation especially in patients that are responding poorly to prescribed therapeutic treatments. These poor responders are more likely to experience an event, such as suicide, which would result in informative censoring and consequently would result in biased estimates of the benefits of the prescribed psychotherapeutic intervention or psychopharmacological agent. On the other hand, improved estimation of therapeutic values of the prescribed treatment would aid practitioners in better understanding the mental disease process and hopefully improve quality of life for these individuals. These examples suggest that other research questions or fields of study would also benefit from an informative schedule design approach especially where patient care would be improved with increased physician or healthcare practitioner observation.

Terminology

The following terms that will be used frequently throughout the study will be formally defined here:

1. *Fixed time* is the assumption that levels associated with the time factor are the only levels of interest and therefore any analysis would be limited to drawing

conclusions on the specified levels included in the study protocol. These levels are typically prospectively determined and, as the name implies, are fixed across the study interval.

2. *Informative time* is the assumption that levels associated with the time factor will vary in length with their magnitudes dependent on prior observational outcomes suggesting that they contribute informatively on subsequent response observations. The magnitude of each informative time interval is adaptively determined for each subject and will vary across the study interval.

3. *Longitudinal data* are a set of observations of a response variable or variables that is measured repeatedly on each subject over a given time period. These measurements are scheduled on some prospective fixed time interval and limit the analysis to conclusions on the specific time intervals used in the study.

4. *Informative schedule data* are a set of observations of a response variable or variables that is measured repeatedly on each subject over a given time period. These measurements are scheduled on some adaptive time interval and their lengths are dependent on the prior observations suggesting that the magnitudes of the intervals are informative to the change realized in the response variable or variables.

Limitations

The limitations of this study that should be considered by researchers would be the following:

1. This study was limited to a single normally distributed response variable and therefore should not be applied to studies that might contain multivariate and/or non-normal response variables.

2. Furthermore, the present study made the assumption that time was exponentially distributed or that the log of time was normally distributed, these assumptions should be considered before applying the results to other studies which may have different time factor distribution assumptions.

3. As will be outlined in chapter three, a single set of model parameter coefficients will be utilized in simulating informative schedule data and a limited set of sample and observation sizes along with observation lengths will be simulated.

4. Furthermore, the evaluation of parameter estimates will be limited to three different criteria as outlined in chapter three.

5. Finally, common parameter estimates obtained from the analysis of the proposed model will be compared to a single traditional longitudinal approach and therefore may not be compared to parameter estimates obtained from other analysis approaches not included in this study.

Conclusion

Currently, traditional approaches to longitudinal analysis require the assumption that time be a fixed factor in the explanation of changes in the response variable. This analytical approach is generally adequate for most research designs in which subjects are observed on a prospective fixed observation schedule. However, this traditional approach does not hold in cases of adaptive schedule designs in which subsequent observation are determined following the observance of the response variable. Since times between observations are adaptively determined and informative in the response trajectory, models with fixed time assumptions are incapable of analyzing the informative nature of the data lessening the inference achievable. This inability for traditional approaches to capture the full informative nature of informative schedule data suggests a need for a novel approach. Consequently, this study proposes the development and evaluation of a novel model that jointly models an informative time component with a longitudinal measured response variable that can be utilized for the analysis of informative schedule study designs.

To better understand the issues presented in the introduction, chapter two provides a more comprehensive review of traditional approaches to longitudinal analysis and other joint model designs found in the literature along with other pertinent information necessary. Chapter three introduces the proposed model along with an outline of the

specific methods utilized to evaluate its efficacy in analyzing Monte Carlo simulated informative schedule data. Chapter four presents the results of the evaluation of the simulated data by the proposed informative schedule model while chapter five discusses the implications of the results and future research directions.

CHAPTER II

REVIEW OF LITERATURE

As discussed in chapter one, the purpose of this study was the development of a model that incorporates an informative time component along with a corresponding set of longitudinal measurements of a response variable. While there is plenty of literature that covers the development of longitudinal models with time as a nonrandom component, we are unaware of any research conducted on the joint modeling of longitudinal and informative schedule data at the time of this study. However, there is a growing presence in the literature of research investigating the joint modeling of survival time and longitudinal data which might be pertinent to the present study.

For the reader to achieve a contextual understanding of the relevant issues to this study a review of the literature is presented that introduces several tactics to longitudinal analysis. This review of the methodological approaches presented in the literature is divided into five sections. The first section presents simple methods for analyzing longitudinal data that consist of condensing the repeated observations into a single variable used in a subsequent analytical approach. The second section presents historical methods of analyzing longitudinal data that preserves the temporal nature of the data but has, in general, become obsolete due to unrealistic assumptions and requirements which are inherent to these models. The third section presents mixed-effects models that

incorporate random effects that are specific to each subject and are generally utilized in most longitudinal research studies. The fourth section presents methods that jointly model longitudinal and survival data which includes a single random event time or survival time associated with the occurrence of the event of interest. Finally, the fifth section presents a short introduction of Vector Autoregressive (VAR) models that are utilized in time series analysis and have the common objective to the proposed informative schedule model of modeling a set of repeatedly measured observations conditioned on prior responses outcomes.

Simple Longitudinal Models

In many research studies the objective is to evaluate changes in a response variable over time by observing repeated measures on a single subject. These repeated measurements result in observations that are correlated within-subject and therefore require more sophisticated statistical methods to account for this dependency of observations. One of the earliest methods for dealing with correlated data were presented by Student (1908) in his development of the *t*-test which avoids the issue of correlation by calculating a single summary variable for each subject used to analyze changes in a response variable from a pre-test to a post-test condition. Essentially, this approach constructs a single independent observation by obtaining the differences between the pre-test and the post-test for each subject which, subsequently, simplifies the analysis approach substantially. However, this method is of little use for any complex analysis involving more than two observation times and therefore, would be of little help with informative schedule designs which typically involves multiple observation occasions.

Other approaches that involve the conversion of a set of correlated observations measured on a single subject into a single response variable have also been developed. These approaches essentially convert the analysis from a longitudinal one with correlated data to a univariate problem void of dependency issues and have been termed in the literature as summary-statistic approach (Dawson & Lagakos, 1991, 1993: Frison & Pocock, 1992: Dawson, 1994), response feature analysis (Crowder & Hand, 1990), or derived variable analysis (Diggle et al, 2002). Matthew, Altman, Campbell and Royston, (1990) summarized several different approaches including (a) the use of the overall mean, (b) comparing the area under the curve, (c) the maximum or minimum value for each group, (d) time to maximum or minimum response and (e) regression coefficients to evaluate the rate of change between groups. Despite their ease of use, these methods have several drawbacks in that the analysis loses temporal aspects preventing the use of time-varying covariates. In addition, there is in general a substantial loss of statistical power and there is a level of uncertainty in the derived summary variable potentially violating the assumption of homoscedasticity (Hedeker & Gibbons, 2006). Furthermore, the removal of temporal aspects in the data clearly prevents the use of time as an informative component in the change of the response variable in these summary statistical methods and therefore would not be a candidate method for analysis of informative schedule data.

Historical Longitudinal Models

Traditional approaches to repeated measures designs in which temporal aspects of the data have been preserved have centered on two models: the univariate repeated measures analysis of variance (ANOVA; Winer, 1971) and the multivariate analysis of

variance (MANOVA) approach to repeated measurements (Cole & Grizzle, 1966). These methods persistent presence as an analytical tool in the study of repeated measurements can be attributed to their familiar methodology and ease of interpretation despite their inherent shortcomings. Here the primary focus of the analysis for both methods is on the comparisons of mean group responses for varying observations and neither model is informative about subject-specific changes across time. Furthermore, time points at which the response variables are observed are assumed to be fixed across subjects for both models and are treated as a classification variable (Hedeker & Gibbons, 2006). This fixed-time assumption, intrinsic in these methods, precludes the use of time as an informative component in the change of the response variable observed within-subjects and therefore, is of little use in achieving the present study's objectives. In addition, these models are of limited general use for most complex research situations because of the unrealistic assumption of equal variance-covariance structure and difficulties associated with missing data across time points (Everitt, 1995).

Mixed-effects Longitudinal Models

A more informative and practical approach to the analysis of longitudinal data are the use of mixed-effects models which includes the addition of random effects that are unique to a particular subject allowing for the evaluation of individual changes in the response variable along with fixed effects of the mean response for each group across time (Laird & Ware, 1982). More specifically, the mixed-effects model extends the general linear model (GLM) by modeling the combination of sample population characteristics that are assumed to be shared by all subjects, and subject-specific effects

that are unique to a particular individual (Fitzmaurice et. al., 2004). For this reason, mixed-effects models have become increasingly popular for modeling longitudinal data due to their more informative or subject-specific evaluation of the response variable of interest. Consequently, a variety of different approaches to mixed-effects models have been developed with varying assumptions underlying the random effect components and methods of obtaining model parameter estimates (Davis, 2002). These models are identified with a variety of descriptive names, e.g., variance component models (Dempster, Rubin, & Tsutakawa, 1981), random effects models (Laird & Ware, 1982), empirical Bayes models (Hui & Berger, 1983), random coefficient models (de Leeuw & Kreft, 1986), mixed models (Longford, 1987), two-stage models (Bock, 1989), multilevel models (Goldstein, 1995), and hierarchical linear models (Raudenbush & Bryk, 2002). Despite their differences in component assumptions and estimation methods, mixed-effects models, in general, allow for the analysis of unbalanced designs associated with missing data due to subject attrition (Hedeker & Gibbons, 2006), a common problem in many longitudinal studies. Although mixed-effects models allows for the analysis of non-rectangular designs, time of observation in these models are still considered fixed, limiting inferences to the time points present in the data vectors and ultimately preventing their use as an informative component in explaining changes in the response variable. Thus, once again, the nonrandomness assumption for time intrinsic in mixed-effects models prevents the use of these methodologies in an informative schedule design.

Despite the underlying assumption inherent in mixed-effects models that prevent their usage in the analysis of informative schedule data, their prevalence as a statistical tool for analyzing longitudinal data and ability to analyze non-rectangular observation

vectors makes this approach the most likely comparative candidate for the analysis of informative schedule data and therefore warrants a more in-depth evaluation. Therefore, if we assume that samples of *m* individuals are measured repeatedly over time, the resulting observation for the *i*th individual on the *j*th occasion would be, y_{ij} which would be observed at time, t_{ij}. The complete set of observations realized for the *i*th individual would result in a vector of observations of the response variable, $\mathbf{y}_i = \left(y_{i1}, \mathrm{K}\ , y_{in_i}\right)$ and a corresponding vector of observed times, $\mathbf{t}_i = \left(t_{i1}, \mathrm{K}\ , t_{in_i}\right)$ collected over n_i repeated measurements. These vectors of observations and times allow for, but do not require, each individual to have a unique sequence of measurement occasions hinting to this methods ability to handle non-rectangular designs. Using vector and matrix notation, the mixed-effects model can be expressed as

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i$$

where $\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed effects, $\boldsymbol{\gamma}_i$ is a $q \times 1$ vector of random effects with a mean of zero and covariance matrix of G_i, $\mathbf{X}_i$ is a $n_i \times p$ matrix of covariates, $\mathbf{Z}_i$ is a $n_i \times q$ matrix of covariates, with $q \le p$, and $\boldsymbol{\varepsilon}_i$ is a $n_i \times 1$ vector of errors assumed to be independent of $\boldsymbol{\gamma}_i$, and also with a mean of zero and a covariance of R_i (Laird & Ware, 1982; Jennrich & Schluchter, 1986). Ordinarily, it is further assumed that R_i is the diagonal matrix, $\sigma^2 I_{n_i}$, where I_{n_i} denotes an $n_i \times n_i$ identity matrix (Fitzmaurice et. al., 2004). With these definitions, the matrix $\mathbf{X}_i$ is a known design matrix containing *p*

covariate vectors of fixed effects (e.g., time of observation, gender, age, treatment group, etc) associated with each repeated measure for the *i*th individual and contains the information that relates the unknown vector of regression coefficients, $\boldsymbol{\beta}$ to the mean of the vector of responses, $\mathbf{y}_i$. In essence, the mixed-effects model is a GLM where everything is the same and has the same general sample population interpretation except for the addition of the known design matrix, $\mathbf{Z}_i$ and the vector of unknown random effects, $\boldsymbol{\gamma}_i$ that are subject-specific. Here, $\mathbf{Z}_i$ is a design matrix that is a subset of the columns of $\mathbf{X}_i$ which links the vector of random effects, $\boldsymbol{\gamma}_i$ to the response vector, $\mathbf{y}_i$ for the *i*th individual. The addition of the vector of random effects associated with the *i*th individual describes a subset of regression parameters and how they deviate from the sample population fixed effects. Simply put, each individual has a set of subject-specific coefficients that describe how their mean responses deviates from the sample population mean. Furthermore, these subject-specific deviations obtained by the inclusion of the *i*th random effects vector results in two different mean response profiles. The conditional or subject-specific mean for $\mathbf{y}_i$, given by $\boldsymbol{\gamma}_i$, is $E(\mathbf{y}_i \mid \boldsymbol{\gamma}_i) = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\gamma}_i$ and the marginal or population-averaged mean is determined by $E(\mathbf{y}_i) = \mathbf{X}_i\boldsymbol{\beta} = \boldsymbol{\mu}$ since, $\boldsymbol{\gamma}_i$ is zero when averaged over the distribution of the random effects (Fitzmaurice et. al., 2004). This ability of the mixed-effects model to not only calculate the mean response of the sample, but to also determine the subject-specific mean responses, makes this model an attractive and more informative approach to longitudinal analysis. Consequently, the addition of the PROC MIXED procedure in the SAS system allows for the analysis of repeated measures or longitudinal designs by implementing the mixed-effects model and by

modeling the covariance structures directly through the use of likelihood based methods (Littell, Henry, & Ammerman, 1998).

Survival and Longitudinal models

As mentioned above, there is an increasing presence in the literature of research investigating the joint modeling of survival time and longitudinal data. This approach has some bearing on the present study because it jointly models a sequence of observations with a single random time event, which is similar to this study's objective of jointly modeling a sequence of observations with a set of corresponding informative time events or schedules. Therefore, the general approaches implemented in the literature of joint modeling of survival and longitudinal data may be of informative value in the development of the model in this study.

The primary goal of survival analysis is to estimate causal or predictive models in which the risk of an event depends on covariates or predictor variables (Kaplan & Meier, 1958). Cox (1972) introduced a model for the analysis of time to event data using proportional hazards regression methods in which the predictor variables can be either constant or vary across time. When the predictor variables vary across time and are observed multiple times during the experiment the resulting data set can be considered as repeated measurements. Consequently, methods investigating the joint modeling of longitudinal measurement and survival time data have been developed.

The usefulness of any survival analysis is dependent on the accuracy of the estimation of the regression parameters used in the expression of the hazard function which suggests that a complete knowledge of the predictor variable history is important.

Unfortunately, in most cases, time-dependent predictors are measured only periodically and with measurement error which can lead to biased estimation of regression parameters used in the survival analysis (Prentice, 1982). Furthermore, even when measurement error is unimportant, a complete knowledge of the predictor variables must be known to maximize the partial likelihood used in this analysis (Cox, 1975). To improve on the estimation of model parameters, Tsiatis, DeGruttola, and Wulfsohn (1995) used a two-stage approach in which the response variable trajectory is initially determined by using a mixed-effects longitudinal model and the second stage uses the estimates from the previous model to improve the covariate history that enters the hazard function of the Cox model. Essentially, the authors used a mixed-effects model to summarize the history or trajectory of the response variable up to some given time point where this obtained estimate is utilized in the subsequent proportional hazards model as a predictor variable or covariate in the estimation of the survival parameters. Once the obtained estimates from the mixed-effects model have entered the proportional hazards model, the survival parameters are estimated by maximizing the partial likelihood as usual. Faucett and Thomas (1996) used a similar approach of a repeated measures random effects model to estimate the response variable parameters and the survival process parameters simultaneously allowing for a more precise and accurate estimate of the relationship between the response variable and survival time event. They specified their model into two submodels where one describes the relationship of the observed covariate measurements as a function of the true, unobserved covariate values and the other describes the relationship between the risk of disease and the true, unobserved time-dependent covariate. The first model, the covariate tracking model, is essentially a

subject-specific linear model of the true, unobserved covariate or response variable at a given time, measured with some error, while the second model, the disease risk model, is the proportional hazards model that depends on the unobserved covariate from the first model at the same given time point. To estimate the unknown parameters for the overall model, Faucett and Thomas (1996) used Gibbs sampling which is a Monte Carlo method for generating samples from the joint posterior distribution of unknown parameters in a model, conditional only on the observed data. The use of this sampling approach allows for the estimation of the unknown parameters for both submodels simultaneously since the joint distribution of their proposed model is not conjugate. Wulfsohn and Tsiatis (1997) also modeled the response variable parameters and the survival process simultaneously to improve on parameter estimation due to measurement error. Their approach, once again, used a mixed-effects model to summarize the history or trajectory of the response variable or covariate and the Cox's proportional hazards model to determine the survival or event time parameters. However, the estimation of their model's unknown parameters was obtained by maximizing the joint likelihood for the covariate process and the failure time process of the observed data by using the expectation-maximization (EM) algorithm which they argued is a superior approach. Henderson, Diggle, and Dobson (2000) approached the modeling of event times and longitudinal analysis by conditioning on an unobserved or latent zero-mean bivariate Gaussian process that drives a pair of linked submodels. Here the two submodels, the measurement and intensity models, are in essence the mixed-effects model and the proportional hazards model that are conditionally independent given the latent association process which, subsequently, links the two models. Here, the association between the

two models is described through the cross-correlation between the latent processes and, when absent, suggests that the joint model does not improve on the estimation of the parameters over the two models separately. These latent coefficients enter into the proportional hazard model and measure the association induced by the mixed-effects model parameters on the estimation of the survival analysis. Ultimately, these parameter estimates, including the latent process coefficients, were obtained by the maximization of the joint model using the EM estimation algorithm. Wang and Taylor (2001) also jointly modeled longitudinal and survival processes through the use of the mixed-effects and proportional hazards models, but included an Ornstein-Uhlenbeck (IOU) stochastic process to better estimate the time-dependent parameters. The IOU stochastic process allows the response trajectory to vary around a straight line that is realized by each subject's path, since the slope of the response can vary over time. The inclusion of the IOU stochastic process allows for better estimation of the mixed-effects parameters that are used in the subsequently linked proportional hazards model. Parameters of their model were estimated by employing the Markov Chain Monte Carlo (MCMC) which is an iterative process that samples from the desired distribution and constructs a Markov chain that has the desired distribution as its equilibrium distribution. Lin, Turnbull, McCulloch and Slate, (2002) jointly modeled longitudinal time-dependent predictor variables with a latent class process modeled by a multinomial distribution, which describes the probability of an individual belonging to a specific latent class. Each subpopulation has its own model for the longitudinal process which is determined by the mixed-effects model with subpopulation differences entering the mean. This model captures common characteristics of the response trajectories within the subpopulation

through the latent classes resulting in improved estimations of covariates that enter into the proportional hazards model. Tseng, Hsieh and Wang, (2005) jointly modeled longitudinal data using linear mixed-effects models with accelerated failure time (AFT) analysis; an alternative method that allows a parametric approach that is considered more robust to unmeasured confounders when compared to Cox proportional hazard model. Here AFT is a linear model of the log of the predicted failure time related by the response variable and determined by the mixed-effects model which allows for the influence of the entire covariate history on subject-specific risk. The parameter estimates for the joint model of mixed-effects responses and the AFT process was determined by the use of the EM algorithm for the conditional distribution. Finally, Elashoff, Li, and Li (2007) developed a method to jointly model longitudinal measurements and competing risk failure time data which allows for the addition of more than one type of event included. However, this approach still models a single random event occurrence but allows for a variety of events to be considered in the model. The proposed model can be divided into three sub-models with the longitudinal response outcome being modeled by the mixed-effects approach, the second model assuming a multinomial distribution that models the probability that a specific risk has occurred for the given individual and the third model is the hazard function for the specific risk observed. Essentially, this model allows for a separate longitudinal and proportional hazards model for each of the specified risk components and incorporates the probability of the specified risk occurring in those models. The parameters associated with this model were also determined by maximum likelihood estimation via an EM algorithm.

As mentioned above, a common characteristic of each of the above approaches is the modeling of a single random event of interest by utilizing the information obtained through the measurement of a response variable across time. These approaches, despite including a single random time event, still include the assumption that response variable measurements are taken on a fixed time interval which prevents them from being utilized in an informative schedule design. Furthermore, a common problem that seems to be the impetus for most of these joint models is the need to improve on the evaluation of the response trajectory to prevent biased estimates obtained from the subsequent proportional hazard or accelerated failure time analysis. While improved estimation is always an objective in any study, this particular issue of accurately estimating the complete response trajectory or history of the response variable was not a direct concern for this study.

Vector Autoregressive

Time series analysis is concerned with modeling stochastic processes and for constructing predictions based on the developed models (Lutkepohl, 1991). This analytical ability to model time-dependent processes for the purpose of predicting or forecasting future observations is the reason that these models have become increasingly popular in the area of econometrics where the goal is to determine the future direction of economic indices. These models have also become popular in the area of meteorology where the prediction of future environmental conditions is a particular research goal, along with many other fields of study that contain stochastic data.

Time series data shares a remarkable similarity to longitudinal data in that the response variables are measured repeatedly over a given time interval and that these measured responses are correlated. Despite the similarities, time series data usually consist of a small number of long sequences of repeated measurements, whereas longitudinal data consist of a large number of relatively short sequences of repeated measures (Fitzmaurice et al., 2004). However, time series models also share a common assumption with longitudinal data in that repeated measures taken closer together in time are expected to be more highly correlated than repeated measures taken further apart in time. This assumption of decreasing correlation over time is a key component of Vector Autoregressive (VAR) models which describe the evolution and interdependencies of a set of variables over the same sample period as a linear function of only their past evolutions (Hipel, Mcleod, & Lennox, 1977). In essence, VAR models assume that past response outcomes are informative in the realization of current observations. For example, in a two variable case, we can let the time path for the response, $y_{1,t}$ be affected by current and past realizations of the response, $y_{2,t}$ and let the time path of, $y_{2,t}$ be affected by current and past realizations of the response, $y_{1,t}$ at time *t*. This would give a simple bivariate formula of the following:

$$y_{1,t} = C_1 + A_{1,1}y_{1,t-1} + A_{1,2}y_{2,t-1} + e_{1,t}$$
$$y_{2,t} = C_2 + A_{2,1}y_{1,t-1} + A_{2,2}y_{2,t-1} + e_{2,t}$$

Or, equivalently, in vector and matrix form:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}$$

Here, the vector $\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$ is a set of constants or intercepts, the matrix $\begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}$ is a set of regression coefficients, the vector $\begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}$ is Gaussian white noise with a mean of zero and covariance of Σ, and the vector $\begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix}$ is termed the lag which is essentially the prior observations for y_1 and y_2 at time $t-1$ (Lutkepohl, 1991). The addition of the vector of lags in the previous equation allows for current realizations of the response variables to be a linear function of prior responses. Furthermore, the inclusion of the lag vector suggests that each element or past realization of a single response affects the observed path of every variable included in the model, that is, each response variable is influenced by its own past realization along with the past realizations of the other response outcomes. The degree that past realizations affect the path of the current outcomes is not limited to first order lags as the above model demonstrates but can include any combination of *p* lags. Also, the amount of variables included in the model is not limited to a bivariate outcome but can be modeled for *k* variables. For example, in a $k \times 1$ vector of responses, y_t collected up to time *t* and including *p* lags, would have the following structure:

$$y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \Lambda + A_p y_{t-p} + e_t$$

Here the path for the vector of responses, y_t would be influenced by or a linear function of the *p* lags and a $k \times k$ matrix of regression coefficients, A_i where $i = 1,\mathrm{K}$, p along with the vector of errors, e_t for time *t*. Essentially, this model would be termed a VAR with *p* lags or VAR(*p*). It should also be noted that the *k* variables for time *t* would be a function of the *p* lags for the *k* variables, similar to the bivariate model presented above. More precisely, each variables path is not only affected by its own lags but is also affected by the lags for all other variables contained in the model.

Conclusion

In many different research areas, longitudinal studies play an important role in our understanding of the research objectives which cannot be obtained by other analytical approaches. Consequently, the literature is filled with a variety of different longitudinal approaches and model assumptions to accommodate the variety of response variable types and design issues faced by many researchers. Despite the multitude of different approaches, the underlying assumption of fixed time effects is common to all model approaches, which prevents their utilization in the analysis of informative schedule data. Furthermore, while there is a growing presence of joint models for longitudinal data and survival time analysis, these model's research objectives are not consistent with the objectives of this study and therefore are of limited use in this study. For the reader to achieve a better understanding, the proposed joint model and the methods employed in the evaluation of that model are presented in chapter three. Chapter four presents and discusses the results obtained from the evaluation of the proposed model and the conclusion of those results and future research directions is presented in chapter five.

CHAPTER III

METHODOLOGY

As discussed in chapters one and two, traditional approaches utilized in the analysis of longitudinal data have several shortcomings in the explanatory ability of these methods when applied to observations collected by an informative schedule design suggesting the need for a different approach that better explains their nature. To this end, the purpose of the present study was the development of a joint model for a longitudinal process and time of observation with improved explanatory ability when applied to informative schedule data.

To accomplish this study's purpose, chapter three begins with a discussion of the notation that was employed in the development of the proposed model. The second section presents the general structure of the informative schedule model and two special cases of that model that are considered further in this study. Also, this section includes the associated likelihood equation for one of the special cases and the SAS likelihood call statement for the other case that is subsequently used for model parameter estimation. The third section presents a discussion of the method of maximum likelihood estimation employed in obtaining the parameter coefficients for this model and the competing mixed-effects model design. The fourth section discusses the particulars of the optimization algorithm constructed to numerically determine model parameter estimates.

The fifth section describes the design issues associated with obtaining Monte Carlo simulated sample data used in the evaluation of these two models and the final section presents a discussion of the methods and criteria used to evaluate the effectiveness of the coefficient estimates and compared estimates obtained from the informative schedule models to the mixed-effects approach.

Notation

Suppose we have a set of *m* subjects or individuals followed over an interval from $[0, \tau)$. The *i*th individual provides a vector of quantitative observations, $\mathbf{y}_i = (y_{i1}, \mathrm{K}\ , y_{in_i})$ with a corresponding informative vector of time schedules, $\mathbf{t}_i = (t_{i1}, \mathrm{K}\ , t_{in_i})$ where the observations and time intervals range from $j = 1, \mathrm{K}\ , n_i$ and the individuals range from $i = 1, \mathrm{K}\ , m$. It should be noted that this notation allows for each individual to have a different observation schedule length. The resulting joint distribution of $\mathbf{y}_i$ and $\mathbf{t}_i$ is in general $f_{\Theta_i}(\mathbf{y}_i, \mathbf{t}_i)$, where $\mathbf{\Theta}_i$ is a matrix of unknown parameters needing to be estimated. The resulting function of $\mathbf{y}_i$ is conditioned on the vector of corresponding time schedules, namely:

$$f_{\Theta_i}(\mathbf{y}_i, \mathbf{t}_i) = f_{\Theta_i}(\mathbf{y}_i \mid \mathbf{t}_i) f_{\Theta_i}(\mathbf{t}_i). \tag{3.1}$$

If $\mathbf{t}_i$ has no information on $\mathbf{\Theta}_i$ then the joint distribution reduces to the following:

$$f_{\Theta_i}(\mathbf{y}_i, \mathbf{t}_i) = f_{\Theta_i}(\mathbf{y}_i \mid \mathbf{t}_i) f(\mathbf{t}_i) \tag{3.2}$$

and from the likelihood point of view, the model will be the same as a traditional analysis of longitudinal data in that time is no longer an informative component.

Proposed Model

The model we considered for the *i*th individual considers a one step dependency and has the following general form:

$$f_{\Theta_i}(\mathbf{y}_i, \mathbf{t}_i) = f_{\Theta_i}(y_{i1} \mid t_{i1}) f(t_{i1}) \prod_{j=2}^{n_i} f_{\Theta_i}(y_{ij} \mid t_{ij}, t_{ij-1}, y_{ij-1}) f_{\Theta_i}(t_{ij} \mid t_{ij-1}, y_{ij-1}). \qquad (3.3)$$

We assume $f(t_{i1})$ does not depend on $\boldsymbol{\Theta}_i$, so for the purpose of likelihood function we can ignore it. Furthermore, the resulting function of the initial observation, y_{i1} is conditioned on time of observation, t_{i1} which is the same approach found in traditional longitudinal models. However, subsequent observations of the response variable, y_{ij} are no longer exclusively conditioned on time of observation, t_{ij} alone but are now also conditioned on the most recent pervious observation, y_{ij-1} and time of observation.

The likelihood function for model (3.3) is the product of the terms for *m* individuals, namely:

$$L(\boldsymbol{\Theta}, y_1, \mathrm{K}\ , y_m, t_1, \mathrm{K}\ , t_m) = \prod_{i=1}^{m} f_{\Theta_i}(\mathbf{y}_i, \mathbf{t}_i)$$

$$= \prod_{i=1}^{m} f_{\Theta_i}(y_{i1} \mid t_{i1}) \prod_{j=2}^{n_i} f_{\Theta_i}(y_{ij} \mid t_{ij}, t_{ij-1}, y_{ij-1}) f_{\Theta_i}(t_{ij} \mid t_{ij-1}, y_{ij-1}) \qquad (3.4)$$

where $\boldsymbol{\Theta} = (\Theta_1, \mathrm{K}\ , \Theta_m)$. It should be noted from the above equation that the initial observation is a function of the unknown parameters and conditioned on time of observation alone, while subsequent observations are conditionally dependent on the most recent prior observation and time interval along with the unknown model parameters. This conditional dependence on the prior responses is what allows for the schedule times of observation to be informative in this proposed joint model (i.e., the present depends on the recent past). It should also be noted that since dependence is limited to the prior observed response or is of first-order, the model assumes that correlations between response observations decay as time separation increases, which is a common assumption found in many time series models. As a matter of fact, longitudinal data share remarkable similarities to time series data, despite differing analytical goals and general structure of data collection, in that measurements of a response variable are measured repeatedly over a given time period and are assumed to be correlated. Consequently, one special case of the model in (3.3) can be represented in a general time series structure. This special case, which is termed the Vector Autoregressive (VAR) model, can be represented as the following:

$$\begin{bmatrix} y_{ij} \\ \log(t_{ij}) \end{bmatrix} = \begin{bmatrix} \mu_{i,1} \\ \mu_{i,2} \end{bmatrix} + \begin{bmatrix} \phi_{i,11} & \phi_{i,12} \\ \phi_{i,21} & \phi_{i,22} \end{bmatrix} \left(\begin{bmatrix} y_{ij-1} \\ \log(t_{ij-1}) \end{bmatrix} - \begin{bmatrix} \mu_{i,1} \\ \mu_{i,2} \end{bmatrix} \right) + \begin{bmatrix} Z_{ij,1} \\ Z_{ij,2} \end{bmatrix}. \qquad (3.5)$$

Here, $\begin{bmatrix} Z_{ij,1} \\ Z_{ij,2} \end{bmatrix}$ is a vector of Gaussian white noise with zero mean and covariance Σ, while $\begin{bmatrix} \phi_{i,11} & \phi_{i,12} \\ \phi_{i,21} & \phi_{i,22} \end{bmatrix}$ is a matrix of autoregressive coefficients and $\begin{bmatrix} \mu_{i,1} \\ \mu_{i,2} \end{bmatrix}$ is vector of mean

constants for the *i*th individual and which are dependent on some explanatory variables by $\boldsymbol{\mu}_i = \beta X_i$, where β is a vector of coefficients associated with some explanatory variables and X_i is a design matrix for the *i*th individual. This mean constants vector $\boldsymbol{\mu}_i$ is composed of a mean, $\mu_{i,1}$ associated with the response variable, y_{ij} and a mean, $\mu_{i,2}$ associated with the log of time of observation, t_{ij}. Finally, $\left(\begin{bmatrix} y_{ij-1} \\ \log(t_{ij-1}) \end{bmatrix} - \begin{bmatrix} \mu_{i,1} \\ \mu_{i,2} \end{bmatrix}\right)$ is the mean adjusted effect of the prior response outcome and time interval for the *i*th individual.

In the VAR case the response variable is considered to be normally distributed while the log of time is consider to also be normally distributed or log normal. These normality assumptions for both the response variable and time of observation results, essentially, in a bivariate normal model. Furthermore, the inclusion of the mean adjusted prior response and time interval as regression coefficients contributes to this models informative schedule nature.

To simplify the notation for model (3.5), let $\boldsymbol{\mu}_i = \begin{bmatrix} \mu_{i,1} \\ \mu_{i,2} \end{bmatrix}$, $\mathbf{z}_{ij} = \begin{bmatrix} z_{ij,1} \\ z_{ij,2} \end{bmatrix}$, $\boldsymbol{\varphi}_i = \begin{bmatrix} \phi_{i,11} & \phi_{i,12} \\ \phi_{i,21} & \phi_{i,22} \end{bmatrix}$, and $\mathbf{W}_{ij} = \begin{bmatrix} y_{ij} \\ \log(t_{ij}) \end{bmatrix} - \begin{bmatrix} \mu_{i,1} \\ \mu_{i,2} \end{bmatrix}$.

With these notations the model has the reduced form of:

$$\mathbf{W}_{ij} = \boldsymbol{\varphi}_i \mathbf{W}_{ij-1} + \mathbf{Z}_{ij} . \qquad (3.6)$$

This model looks like a Vector Autoregressive of order 1 for each individual, which is a common model utilized in econometrics in modeling the dynamic behavior of economic and financial time series and in forecasting models (Lutkepohl, 1991). Consequently, SAS has incorporated a call statement, VARMALIK into SAS/IML (Interactive Matrix Language) procedure that will compute the log-likelihood function for a Vector Autoregressive Moving-Average model (SAS Institute, 2004). The approach implemented in the call statement utilizes the conditional approximation to the log-likelihood equation (Reinsel, 1997) and is computed as -0.5 x the sum of log determinant of the innovation variance and the weighted sum of squares of residuals (SAS Institute, 2004). However, an iterative numerical method, such as the multivariate Newton-Raphson, is still required to solve for estimates of model parameters and consequently the development of this iterative numerical approach is the primary purpose of the proposed study.

In many natural processes, random variation conforms to a particular probability distribution known as the normal distribution, which is the most commonly observed probability distribution. Therefore, a second special case for model (3.3) can be represented in this more familiar distributional form. This special case, which will be termed the Gaussian-Exponential model (GE), can be represented as the following:

$$f_{\Theta_i}(\mathbf{y}_i, \mathbf{t}_i) = \frac{1}{\sqrt{2\pi}\left(\sigma_i^2\right)^{\frac{1}{2}}} e^{-\frac{1}{2}\frac{(y_{i1}-\mathbf{X}'_{i1}\boldsymbol{\beta})^2}{\sigma_i^2}}$$
$$\times \prod_{j=2}^{n_i} \frac{1}{\sqrt{2\pi\left(\sigma_i^2\right)^{\frac{1}{2}}\sqrt{\left(1-\rho_i^2\right)}}} e^{-\frac{1}{2}\frac{\left(y_{ij}-t_{ij}\gamma_i-y_{ij-1}\phi_i-\mathbf{X}'_{ij}\boldsymbol{\beta}\right)^2}{\sigma_i^2\left(1-\rho_i^2\right)}} \times \exp\left(\alpha+\delta_i y_{ij-1}\right)\exp\left(-e^{\left(\alpha+\delta_i y_{ij-1}\right)}t_{ij}\right). \quad (3.7)$$

In the Gaussian-Exponential case the response variable is considered to be conditionally normal given time while time of observation is assumed to be distributed exponentially. Furthermore, the initial observation is assumed to be a function of the unknown regression parameters only, while the subsequent responses are conditioned on the unknown parameters along with the affects of the prior response outcome and time of observation. This conditional association on prior response outcomes contributes to this model's ability to analyze informative schedule data.

The above model would result in a log-likelihood for the *i*th individual of:

$$\begin{aligned}\ln\left(L_i\right) = C + \log\left[\frac{1}{\sqrt{2\pi}\left(\sigma_i^2\right)^{\frac{1}{2}}} e^{-\frac{1}{2}\frac{(y_{i1}-\mathbf{X}'_{i1}\boldsymbol{\beta})^2}{\sigma_i^2}}\right] \\ + \sum_{j=2}^{n_i} \log\left[\frac{1}{\sqrt{2\pi}\left(\sigma_i^2\right)\sqrt{\left(1-\rho_i^2\right)}} e^{-\frac{1}{2}\frac{\left(y_{ij}-t_{ij}\gamma_i-y_{ij-1}\phi_i-\mathbf{X}'_{ij}\boldsymbol{\beta}\right)^2}{\sigma_i^2\left(1-\rho_i^2\right)}}\right] \\ + \sum_{j=2}^{n_i} \log\left[\exp\left(\alpha + \delta_i y_{ij-1}\right)\exp\left(e^{-\left(\alpha+\delta_i y_{ij-1}\right)} t_{ij}\right)\right].\end{aligned} \tag{3.8}$$

The log-likelihood function for the GE model for all individuals would be the sum of the terms for *m* individuals, namely:

$$\begin{aligned}\ln(L) = mC + \sum_{i=1}^{m}\left[-\frac{1}{2}\log\left(\sigma_i^2\right) - \frac{1}{2}\frac{\left(y_{i1}-\mathbf{X}'_{i1}\boldsymbol{\beta}\right)^2}{\sigma_i^2}\right] \\ + \sum_{i=1}^{m}\sum_{j=2}^{n_1}\left[-\frac{1}{2}\log\left(\sigma_i^2\right) - \frac{1}{2}\log\left(1-\rho_i^2\right) - \frac{1}{2}\frac{\left(y_{ij}-t_{ij}\gamma_i-y_{ij-1}\phi_i-\mathbf{X}'_{ij}\boldsymbol{\beta}\right)^2}{\sigma_i^2\left(1-\rho_i^2\right)}\right] \\ + \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\alpha + \delta_i y_{ij-1} - \exp\left(\alpha+\delta_i y_{ij-1}\right)\cdot t_{ij}\right).\end{aligned} \tag{3.9}$$

Here, the log-likelihood function for the *m* individuals has a complicated form, forcing the use of some numerical iterative method to determine maximum likelihood estimates for the GE model. As mentioned above, the development of the procedure to determine the numerical method to estimate the parameters is the primary purpose of this study. Furthermore, the construction of the first-order derivatives was necessary to improve efficient estimation of model parameters and was calculated with the aid of Maple software (see Appendix C for Maple code).

Parameter Estimation

Given that distributional assumptions have been made about the vector of responses $\mathbf{y}_i$ for both special cases of the model, a very general approach to estimation of the model parameter $\boldsymbol{\Theta}$ can be obtained by assuming an iterative method to find the maximum likelihood estimates (MLE). In this method the estimates used in the model are iteratively obtained and are estimates for $\boldsymbol{\Theta}$ that would maximize the log of the likelihood functions of the proposed model, i.e., the estimated value of $\boldsymbol{\Theta}$ that best explains or models the observed data given the distributional assumptions. In general, ML estimators have the added benefit of having large sample consistency, that is there is a high probability that the derived estimate is close to the true population estimate, and are asymptotically unbiased in that as the sample size gets larger the parameters being estimated approach the true population values (Fitzmaurice et al., 2004).

Parameter estimations for both special cases of model (3.3) were accomplished by utilizing the nonlinear optimization call module available through SAS/IML (SAS Institute, 2004). This module offers a set of optimization subroutines for minimizing or

maximizing a user or subroutine supplied continuous function to determine estimate values. The log-likelihood function for the both models and the derivatives for each parameter of the GE model were constructed in SAS/IML as a user defined module and made available to the nonlinear optimization subroutine. In the case of the VAR model, the call subroutine VARMALIK which computes the conditional approximate log-likelihood values was utilized and made available to the nonlinear optimization subroutine. In both cases estimates were obtained by employing the numerical iterative method of the multivariate Newton-Raphson method. This numerical iterative method seeks to find an approximation of the MLE of $\boldsymbol{\Theta}$ or the vector of unknown model parameters by solving the following equations:

$$\hat{\boldsymbol{\Theta}}_m = \hat{\boldsymbol{\Theta}}_{m-1} - H\left(\hat{\boldsymbol{\Theta}}_m\right)^{-1} g\left(\hat{\boldsymbol{\Theta}}_m\right)$$

so that $\hat{\boldsymbol{\Theta}}_m \to \hat{\boldsymbol{\Theta}}$ as $m \to \infty$ and where $H\left(\hat{\boldsymbol{\Theta}}_m\right)$ is the Hessian matrix of the log-likelihood function, and $g\left(\hat{\boldsymbol{\Theta}}_m\right)$ is the derivative of the log-likelihood function or the gradient vector. In essence, this approach produces a series of parameter estimates that become closer and closer to the ML estimates. The use of this iterative method was chosen due to its preferred characteristics of a quick convergence of parameter estimates and the assurance of a positive-definite covariance matrix at each iteration step (Lindstrom & Bates, 1988). Furthermore, this method is also the preferred approach implemented by the PROC MIXED procedure utilized in the analysis of repeated measures data which allows for a more direct comparison between the proposed model estimates and the estimates obtained through the utilization of the mixed-effects procedure. Finally, it should be

noted that the second-order derivatives for both models and first-order derivatives for the VAR model were approximated by finite difference method (Gill, Murray, Saunders, & Wright, 1983) through log-likelihood function calls and therefore, will not be presented here.

Optimization Algorithm

The primary purpose of this study was the development of an efficient method for the estimation of model parameters for the two proposed special cases of the informative schedule model. To accomplish this goal, we took advantage of the extensive library of optimization routines callable from the matrix programming language of SAS available to solve nonstandard estimation problems (SAS Institute, 2004). The optimization subroutine used in this study relied on the calculated results of user-supplied callable modules for determining parameter estimates. In the case of the GE model two modules where constructed in which one returned the maximum likelihood value or objective function and the other which calculated the vector of gradient results (see Appendix B for SAS code). In the case of the gradient vector, first-order derivatives (see Appendix A for derivatives) were determined for each parameter and constructed in a call subroutine made available to the optimization algorithm. In the case of the VAR model, a module was developed that incorporated the conditional log-likelihood module VARMALIK (see Appendix B for SAS code) and was made available to the nonlinear optimization function to calculate the likelihood value of the simulated data.

The optimization algorithms utilized for this study was the double dogleg or NLPDD subroutine which combines the ideas of the quasi-Newton and trust-region

methods. The quasi-Newton optimization method was selected for this study since this subroutine allows for the approximation of the inverse Hessian matrix based on changes in the gradient vector between iterations. The primary advantage of this modified numerical method is that the Hessian matrix does not need to be approximated at each point, which may be computationally expensive (Jöreskog, 1967). This improved efficiency was especially important in the case of the VAR model in which both the gradient and the Hessian matrix needed to be estimated by finite difference method. The inclusion of the trust-region method was chosen since this method allows for the optimization of a restricted region of a quadratic approximation of the nonlinear objective function as opposed to the entire objective function, i.e., at each iteration the step size must remain within a specified trust-region (Dennis, Gay, & Welsch, 1981). Hence, this subroutine utilizes the dual quasi-Newton update method but does not require a line search to be performed. The specific update method employed in this study was the dual Broyden, Fletcher, Goldfarb, and Shanno (DBFGS) method of updating the Cholesky factor of an approximate Hessian matrix which is related by $H^* = R'R$, where H^* is the approximated Hessian matrix and R is the Cholesky decomposition factor (Davidon, 1959; Fletcher & Powell, 1963). Furthermore, the initial determination of the second-order derivatives or Hessian matrix for both models and the first-order or Gradient vector for the VAR model was computed by the numerically more expensive central difference formula (Gill et al, 1983) which allowed for improved accuracy in the approximation of the starting Hessian matrix for both methods and the gradient vector for the VAR model. Finally, the true parameter values (see Table 1 for values) were supplied as the initial starting values to both nonlinear optimization subroutines with the goal that these values

would improve the likelihood of obtaining an efficient and rapid convergence of the objective function.

Finally, Monte Carlo simulated data for both special cases was analyzed by implementing the PROC MIXED procedure and utilizing the maximum likelihood estimation option. The simulated data for both special cases was subsequently analyzed by the mixed-effects method where time of observation was assumed to be sequential and evenly distributed. Furthermore, the variance-covariance structures of the data were assumed to follow a compound symmetry structure.

Table 1.

Parameter values for both special cases of the proposed informative schedule model.

Fixed Model Parameter Values

Vector Autoregressive		Gaussian-Exponential	
Parameter	True value	Parameter	True value
β_1	4	β_0	0.2
β_2	2	β_1	0.5
β_3	3	σ^2	4
β_4	1	ρ	0.5
ϕ_{11}	0.8	ϕ	0.2
ϕ_{12}	0.3	γ	0.5
ϕ_{21}	0.2	α	2.0
ϕ_{22}	0.5	δ	0.04
σ_{11}	4		
σ_{22}	0.1		
σ_{12}	2		

Data Simulations

Monte Carlo simulations of known parameter conditions were generated in SAS/IML for both special cases of the proposed models. For parsimonious reasons, parameters were assumed to be constant across subject, i.e., the subscript *i* was not included in parameter estimations. The fixed population parameters for each special case is outlined in table 1 and were chosen for the purpose of illustrating the proposed model's utility only.

In the VAR case the observations for the response variable were assumed to follow a normal distribution for the measurement error while the observation for the time intervals were assumed to follow a log-normal distribution. Simulated data for the VAR model was accomplished by utilizing the SAS call subroutine VARMASIM which generates a random sequence of time series data in a user defined given structure (see Appendix B for SAS code). For the GE case, observations once again were assumed to follow the normal distribution conditioned for time of observation while the observations for the time intervals were assumed to follow an exponential distribution. Simulated data were accomplished for the GE model by generating random normal values adjusted by the appropriate mean and variance values in the case of the response variable and random exponential values adjusted by mean in the case of time of observation (see Appendix B for SAS® code). Since, the generated observations included the effects of prior outcomes, the resulting data matrices were considered to be correlated. In either special case the sample sizes and the lengths of the individual subject's observation vectors were varied following the patterns outlined in table 2.

Table 2.

Sample size, number of observations, observation scheme, and total number of observations utilized for each simulation study.

Monte Carlo Simulation Scheme

Sample Size	Number of Observations	Observation Design Scheme	Total Number of Observations	Scheme Number
20	5	Rectangular	100	1
	5 & 3	Nonrectangular	80	2
	10	Rectangular	200	3
	10 & 7	Nonrectangular	170	4
	20	Rectangular	400	5
	20 & 14	Nonrectangular	340	6
50	5	Rectangular	250	7
	5 & 3	Nonrectangular	200	8
	10	Rectangular	500	9
	10 & 7	Nonrectangular	425	10
	20	Rectangular	1000	11
	20 & 14	Nonrectangular	850	12
100	5	Rectangular	500	13
	5 & 3	Nonrectangular	400	14
	10	Rectangular	1000	15
	10 & 7	Nonrectangular	850	16
	20	Rectangular	2000	17
	20 & 14	Nonrectangular	1700	18

In essence, three different sample sizes were simulated with three levels of observations for each subject under two differing observation length protocols resulting in a total of 18 different sample schemes. The first observation length protocol would result in a rectangular design for all subjects, (i.e., each subject has the same number of observations), while the second protocol would result in half of the subjects obtaining a reduction in the lengths of their observation vectors resulting in a nonrectangular design. Furthermore, a two factor design matrix (e.g., gender, pre- and post-treatment, etc.) was included to demonstrate the models ability to include the possibility of multiple treatment factors. This design matrix included a random assignment to each subject the inclusion

of the estimation of the second β parameter(s), i.e., approximately half of the subjects would include both β parameters, $(X_i = [1 \quad 1])$ while the other half would have a single β parameter, $(X_i = [1 \quad 0])$ thus allowing for separate estimates based on different factors. Finally, 5,000 iterations of Monte Carlo simulated data were generated for both models. These simulated data were then subsequently analyzed by the appropriate proposed informative schedule model, i.e., GE and VAR model, and by the traditional longitudinal approach of mixed-effects model to obtain parameter estimates.

In the special case of the VAR model eleven parameters were utilized in the construction of the Monte Carlo simulated data. These parameters included a vector of explanatory variables or $\boldsymbol{\beta}$ parameters used to determine mean outcome for both the response variable and log of time of observation. Here, β_1 would be associated with the mean response for the observed data while β_3 would be the mean log time of observations for all subjects included in the data matrix. While, β_2 and β_4 are additive to the other two β parameters dependent on the inclusion of the explanatory variable supplied by the design matrix, respectively. The variance-covariance of the response variable and log of time of observation also need to be estimated. The parameters, σ_{11} and σ_{22} are the variance estimates for the response variable and log of time, respectively. While the parameter, σ_{12} is the covariance shared between the response variable and log of time. The VAR model also includes a matrix of regression coefficients, $\boldsymbol{\varphi}$ which maps the mean adjusted prior response outcomes onto the current observed response variable and log of time of observation.

For the GE model, eight parameters were utilized in the construction of simulate data. These included a vector of explanatory variables or $\boldsymbol{\beta}$ parameters where, β_0 is the intercept and, β_1 would be additive to the intercept coefficient dependent on the inclusion of the explanatory variable supplied by the design matrix. Included with the overall mean responses are the inclusions of the parameters that account for the prior response outcome and the current time of observation. Here the coefficient, ϕ accounts for the effect of the prior response outcome on the mean response while the coefficient, γ accounts for the effect of the current log of time of observation on the mean response. Parameters associated with modeling time of observation include a constant parameter, α and a coefficient that maps time of observation, δ. Finally, two parameters were included that estimated the amount of variance, σ^2 and correlation, ρ seen between the responses.

Model Evaluation

While there are, in theory, a multitude of parameter estimates that can model a given observed process, there are in general some characteristics of estimators that make them better than others. Parameter estimates obtained from the analysis of the proposed model and by the mixed-effects model were evaluated by examining their biases, variance and mean square errors of the simulated data.

Bias was defined as the difference between the estimator obtained and the true parameter being estimated, that is if *T* is an estimator of $\tau(\boldsymbol{\Theta})$, then the bias is given by:

$$bias(T) = E(T - \boldsymbol{\Theta})$$

With this definition an estimator that is closest on average to the true parameter being estimated will have the smallest bias. However, a slightly biased estimator that is highly centered on the parameter of interest and is less variable may be preferable to an unbiased estimator that is less concentrated (Bain & Engelhardt, 1992). The mean square error (MSE) is a reasonable criterion that considers both the variance and the bias of an estimator and is defined as the following:

$$MSE(T) = Var(T) + [bias(T)]^2$$

The use of MSE can be used to evaluate two or more estimators in how well they estimate the unknown parameters.

Finally, a direct comparison between the proposed model and the mixed-effects model was accomplished by comparing the relative efficiency of the common parameter estimates of the two models. Comparisons involving the variances of estimators can be used to determine which makes more efficient use of the data. This determination can be obtained by examining the relative efficiency of the estimator T of $\tau(\boldsymbol{\Theta})$ to another estimator T^* of $\tau(\boldsymbol{\Theta})$ and is given by:

$$rel(T, T^*) = \frac{MSE(T^*)}{MSE(T)}$$

This definition suggests that the estimator T^* is said to be efficient if $rel(T, T^*) \leq 1$ for another estimator of T. In each case of the proposed model, an estimate for β was

common with the traditional approach of a mixed-effects model (see Chapter two for mixed-effects model parameters). These parameter estimates obtained from the simulated data for both informative schedule models and by mixed-effects approach was compared by examining their biases, mean square errors, and relative efficiency.

Conclusion

The following study exploited the flexibility and versatility of the maximum likelihood approach of parameter estimation to evaluate the proposed model efficiency when compared to analysis by way of mixed-effects implemented in the SAS PROC MIXED subroutine. This evaluation was performed on Monte Carlo simulated informative schedule data with known parameters and data structure generated for each special case of the proposed model. Parameter estimations of the two special cases and the traditional approach were evaluated on the bases of bias, mean squared error, and the relative efficiency of the estimated parameters. These parameter estimate evaluation approaches were utilized to compare common parameters between the proposed model and the mixed-effects model. Finally, the results of this study are presented and discussed in chapter four while chapter five provides the conclusions of this research and future research directions.

CHAPTER IV

RESULTS AND DISCUSSION

The purpose of this study was the development of a novel approach that jointly models a longitudinal process and the informative component for time of observation. To achieve this goal of modeling an informative time component along with a repeatedly measured response variable, this study investigated the following research questions:

1. Can a novel approach be developed that would jointly model a longitudinal response variable with a set of corresponding intermittent informative time intervals of observation?

2. Can an efficient numerical iterative method be developed to determine the maximum likelihood estimates for the proposed model?

3. In the presence of simulated informative schedule data, how accurate and efficient is this proposed model in estimating known population parameters?

4. How are these maximum likelihood estimates influenced by a few select variations in subject sample size, number of observations, and the degree of variation in observation lengths for each subject?

5. Finally, how does the proposed model's parameter estimates compare on accuracy and efficiency with common parameter estimates obtained by the mixed-effects model when analyzing the same simulated informative schedule data?

Chapter four begins by evaluating the constructed nonlinear optimization algorithms used to estimate parameters for both special cases of the proposed informative schedule models. Secondly, this chapter summarizes the definitions that were utilized to

evaluate the obtained estimates and outlines the alterations in the data matrices that were evaluated. Thirdly, this chapter summarizes and discusses the average parameter estimates obtained from the VAR simulated data along with the average variance, bias, and MSE for each alteration in subject and number of observation along with alteration in sample matrices. This section also includes a comparison of estimates obtained from the mixed-effects model implemented by PROC MIXED when analyzing the same Monte Carlo simulated data. The fourth section includes a similar summarization and discussion of the GE model parameter estimates along with the bias, variance, and MSE evaluations and comparison of mixed-effects estimates. The fifth section discusses the resulting estimates and evaluations obtained from both models and all 18 different simulation schemes. And finally, the resulting estimates for all simulation schemes and both model approaches are presented in tables 7 through 42.

Joint Model of Informative Schedule Data

In the special case of the Vector Autoregressive (3.5) model eleven parameters were utilized in the construction of the Monte Carlo simulated data while in the special case of the Gaussian-Exponential (3.9) model nine parameters were utilized in the construction of the simulated data. In both cases, randomly generated data of know distributions was shaped accordingly to the established model parameters and simulated to known observation lengths and matrices designs before being analyzed by either of the two developed optimization subroutines and by the mixed-effects method. The resulting simulated data matrix for each model was presented to the constructed optimization algorithm which also had the appropriate log-likelihood equation made available in a call

function for both models and the gradient vector in the case of the GE model. In both cases, the developed optimization algorithms resulted in convergence in nearly every case and estimates for each model parameter were obtained (Tables 7 through 42) suggesting that an efficient numerical iterative method could be developed to jointly model a longitudinal process with informative time schedules. However, when sample sizes and the number of observations were at their smallest amounts both developed numerical iterative methods demonstrated a small proportion of cases (maximum of 1.94% for both models) where convergence was not achieved. This was not surprising since optimization algorithms are known to be less efficient when analyzing samples with small number of observations. In fact, as the number of observation increased the occurrence where convergence was not obtained decreased dramatically for both developed optimization algorithms and at the larger number of observations convergence occurred in every case.

Parameter Estimate Evaluation

One of the purposes of this study was to evaluate the proposed models accuracy and precision in estimating model parameters (see Chapter III for mathematical definitions). Here we defined accuracy in terms of the amount of bias or deviation the resulting estimates showed on average in relationship to the true parameter value while estimate precision would be defined in terms of the average amount of spread or variation in the obtained estimates. A third approach utilized in the evaluation of the informative schedule model parameters was the use of MSE which combines the contribution of both variance and bias of the parameter estimate into a single value. This latter approach

allows for the evaluation of the relative contribution that bias and variation have on the obtained estimate. A second purpose of this study was to evaluate the effect of a few select changes in the simulated data matrix has on the accuracy and precision of the obtained estimates. Here simulated data were generated at three different subject amounts along with three different levels of observations resulting in 18 different simulation schemes. Furthermore, these 18 simulation schemes were also generated for sample matrices in which half of the subjects had shorter observation lengths which were utilized to evaluate the effects that nonrectangular designs might have on parameter estimation. A final purpose of this study was to evaluate a single parameter estimate from the proposed model in comparison to the mixed-effects approach. Here the use of relative efficiency, which is a ratio of the MSEs for both models, was utilized for estimate evaluation along with bias and variation.

Vector Autoregressive Parameters

The VAR model includes a matrix of $\boldsymbol{\beta}$ parameters that along with the design matrix determines the mean outcome for both the response variable and the log of time of observation. Here, all four mean parameter estimates followed similar patterns of accuracy and precision as number of observations increased for both sample matrices designs. When numbers of observations were at their lowest amounts, the obtained estimates showed a substantial amount of variation, i.e., obtained parameter estimates were less precise at low sample numbers (Figure 1 through 4 and Tables 7 through 24). In addition, at lower number of observations a small amount of bias in obtained estimates was also seen. However, the non-directionality of the bias suggests that the observed

inaccuracy of estimates maybe due more to imprecision in the estimates than a systematic bias. For example, if all average estimates were negative in value this would suggest that the optimization method was systematically under estimating the population parameter. In addition, the amount of bias in estimation was relatively small compared to the amount of variation of obtained estimates. This fact is supported by the overwhelming influence that variation has on the calculation of the MSE values suggesting that the inaccuracy in the estimation is relatively small compared to the amount of imprecision in estimation. Furthermore, as number of observations increased the amount of variation and observed bias decreased substantially and, essentially, estimates become centered at 850 observations for all four parameters. Finally, estimates obtained from nonrectangular sample matrices showed a slightly larger amount of variation in obtained estimates when compared to estimates obtained from rectangular sample matrices at similar number of observations. In the two cases where rectangular and nonrectangular sample matrices have the same amount of number of observations the observed averaged variation for nonrectangular estimates was larger than the average variation for rectangular estimates. Thus, a rectangular sample matrix improves the precision in obtained estimates over nonrectangular designs. However, rectangular design matrices showed little effect on the amount of bias when compared to nonrectangular design matrices.

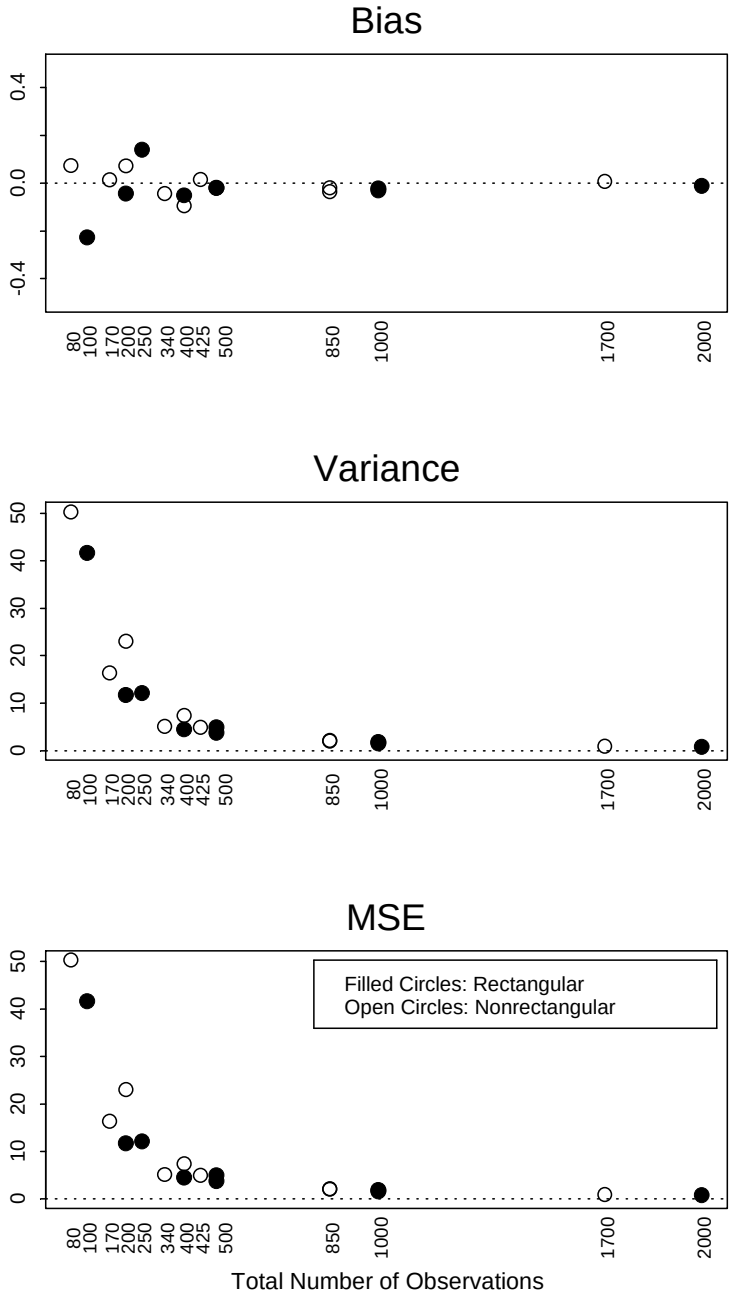

Figure 1. Bias, variance, and MSE for β_1 of VAR model with both rectangular and nonrectangular sample estimates.

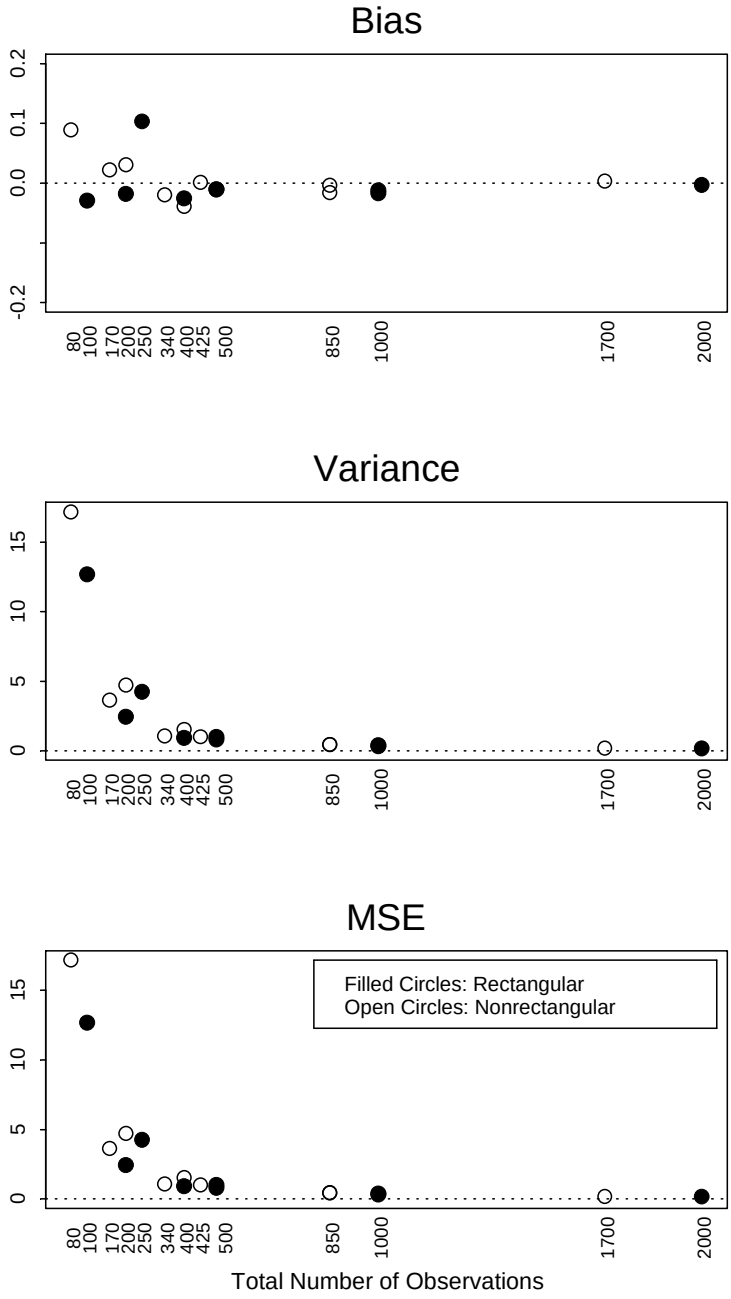

Figure 2. Bias, variance, and MSE for β_2 of VAR model with both rectangular and nonrectangular sample estimates.

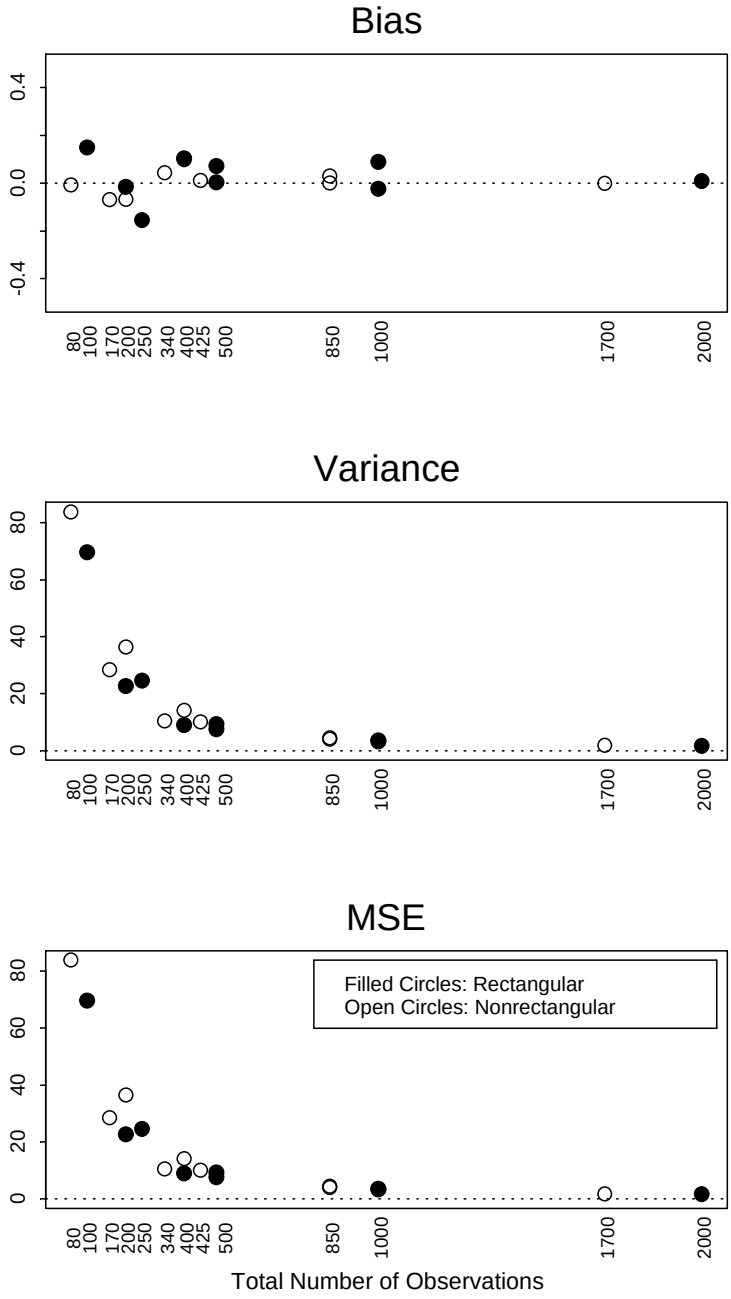

Figure 3. Bias, variance, and MSE for β_3 of VAR model with both rectangular and nonrectangular sample estimates.

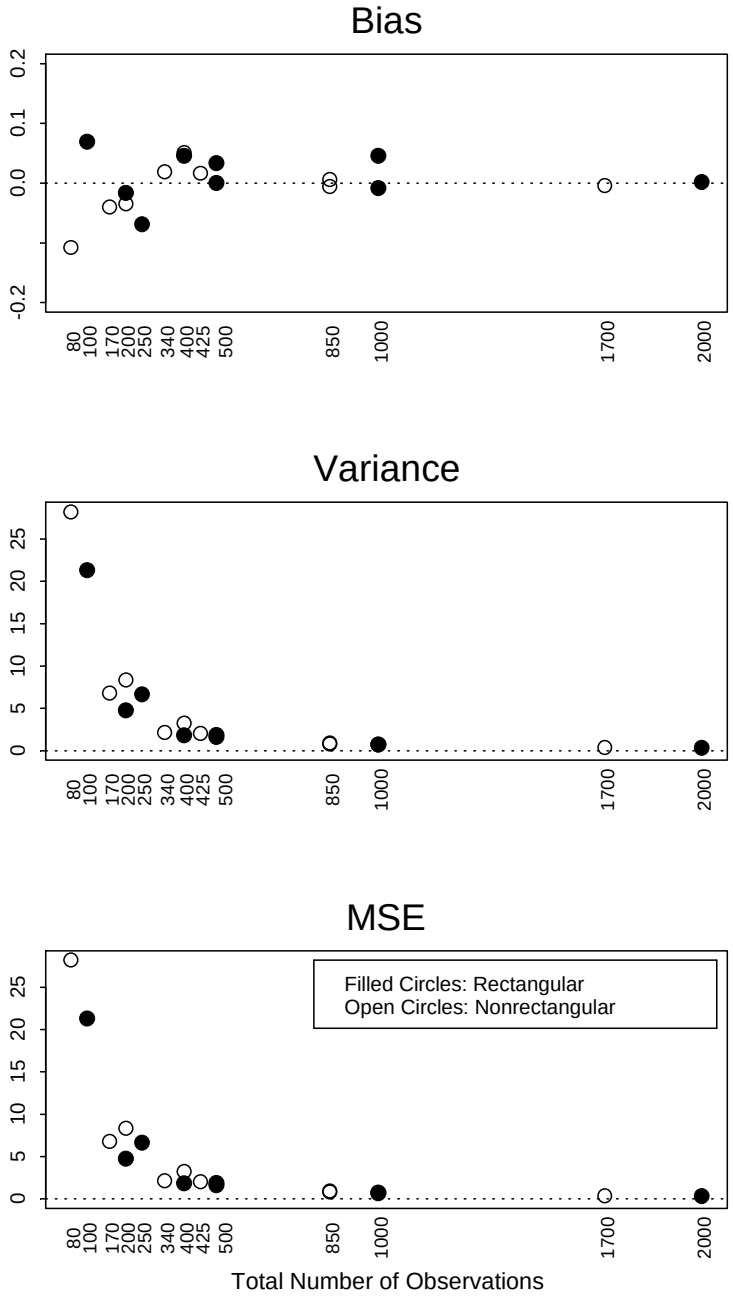

Figure 4. Bias, variance, and MSE for β_4 of VAR model with both rectangular and nonrectangular sample estimates.

The VAR model also contains a symmetrical matrix of variance-covariance parameters where σ_{11} is the variance associated with the response variable, σ_{22} is the variance associated with the log of time, and σ_{12} is the covariance between both response outcomes. In the case of the two variance parameters a similar pattern of precision and accuracy was observed while the covariance parameter showed a slightly different pattern. Both variance parameters demonstrated a systematic negative bias in estimation which was observation dependent, i.e., estimates became less negatively bias as the number of observations increased (Figure 5 and 7 and Tables 7 through 24). In other words, the estimates obtained for the variance parameters for the VAR model are asymptotically unbiased. On the other hand, the covariance parameter did not show any systematic pattern in bias estimates but at lower number of observations obtained estimates did show some small amount of non-directional bias which may be due more to imprecision of estimation (Figure 6). For all three variance-covariance parameters, estimates demonstrated a large amount of variation at smaller number of observations. However, as the number of observations increased the variation in obtained estimates decreased dramatically. Also, the amount of bias in estimation was relative small compared to the amount of variation of estimates for all three variance-covariance parameters which was supported by the MSE values. Finally, rectangular samples matrices demonstrated a small decrease in the average amount of variation in obtained estimates for variance-covariance parameters when compared to estimates obtained from nonrectangular sample matrices. Also, in the case of the variance parameters estimates obtained from rectangular sample designs showed less bias when compared to estimates obtained from nonrectangular designs.

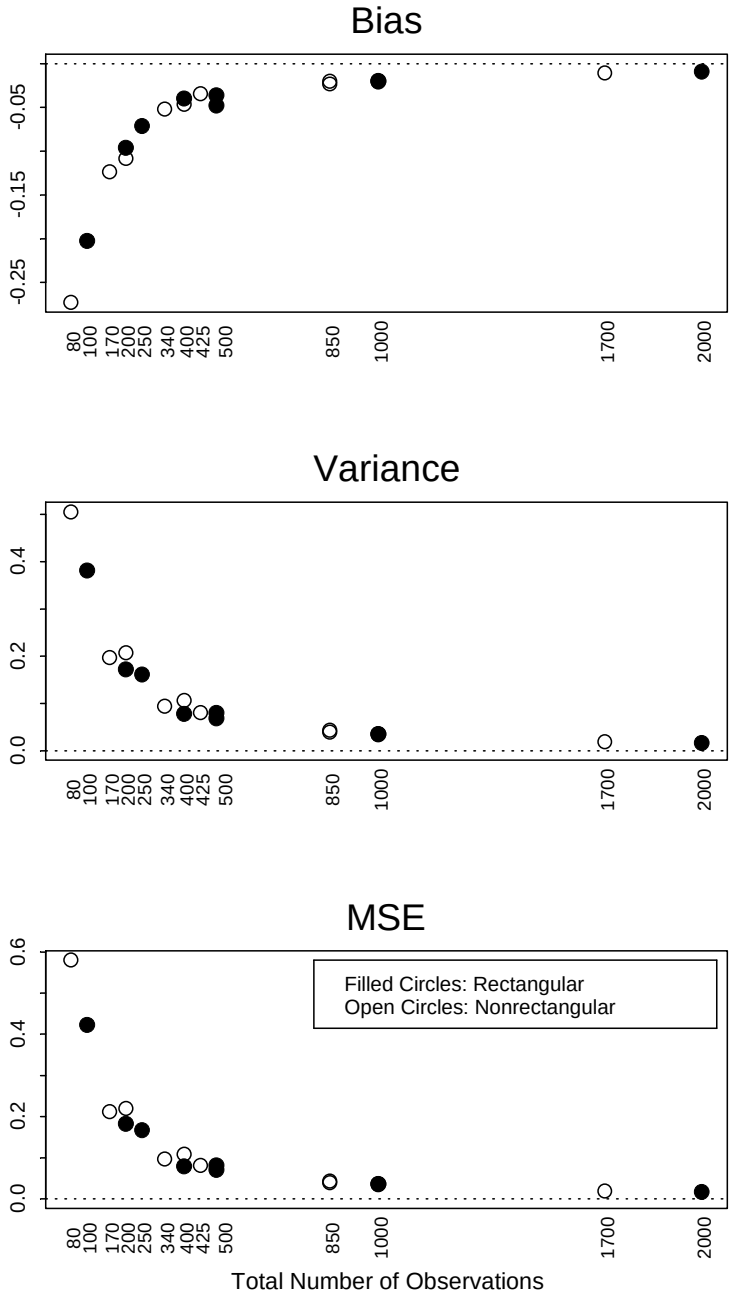

Figure 5. Bias, variance, and MSE for σ_{11} of VAR model with both rectangular and nonrectangular sample estimates.

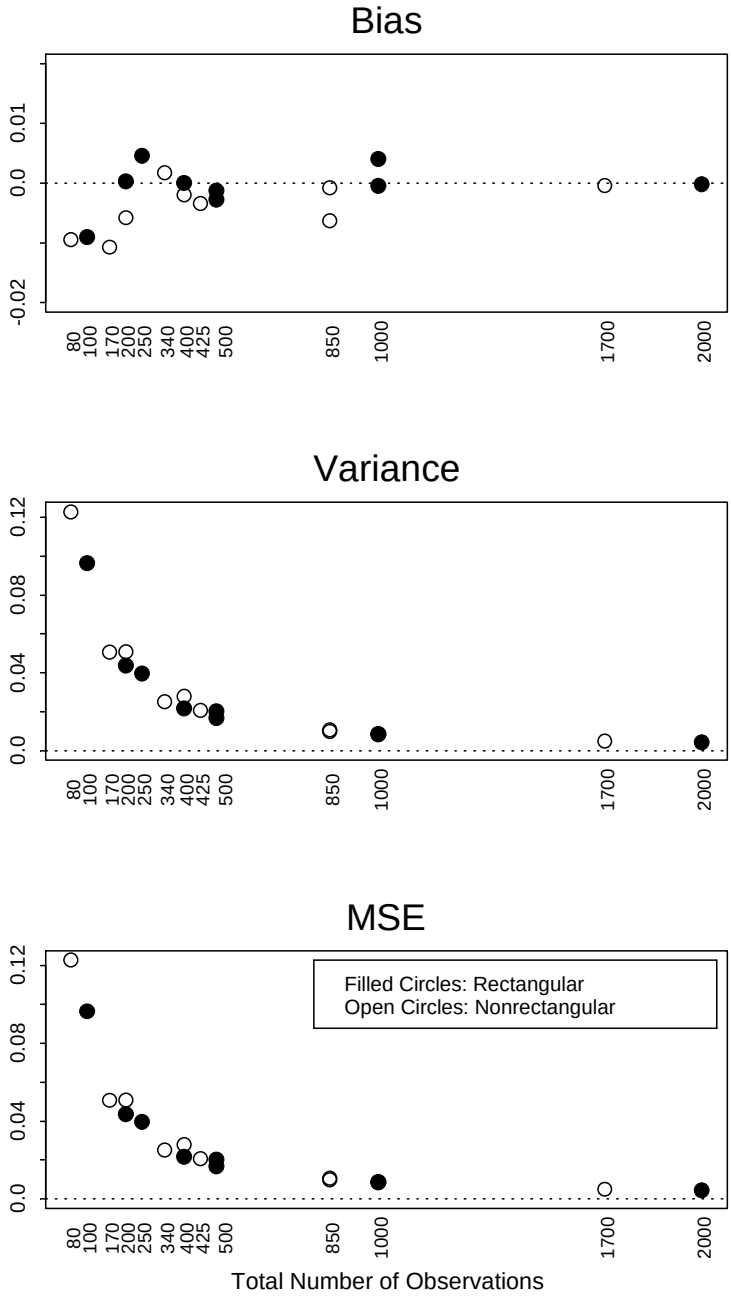

Figure 6. Bias, variance, and MSE for σ_{12} of VAR model with both rectangular and nonrectangular sample estimates.

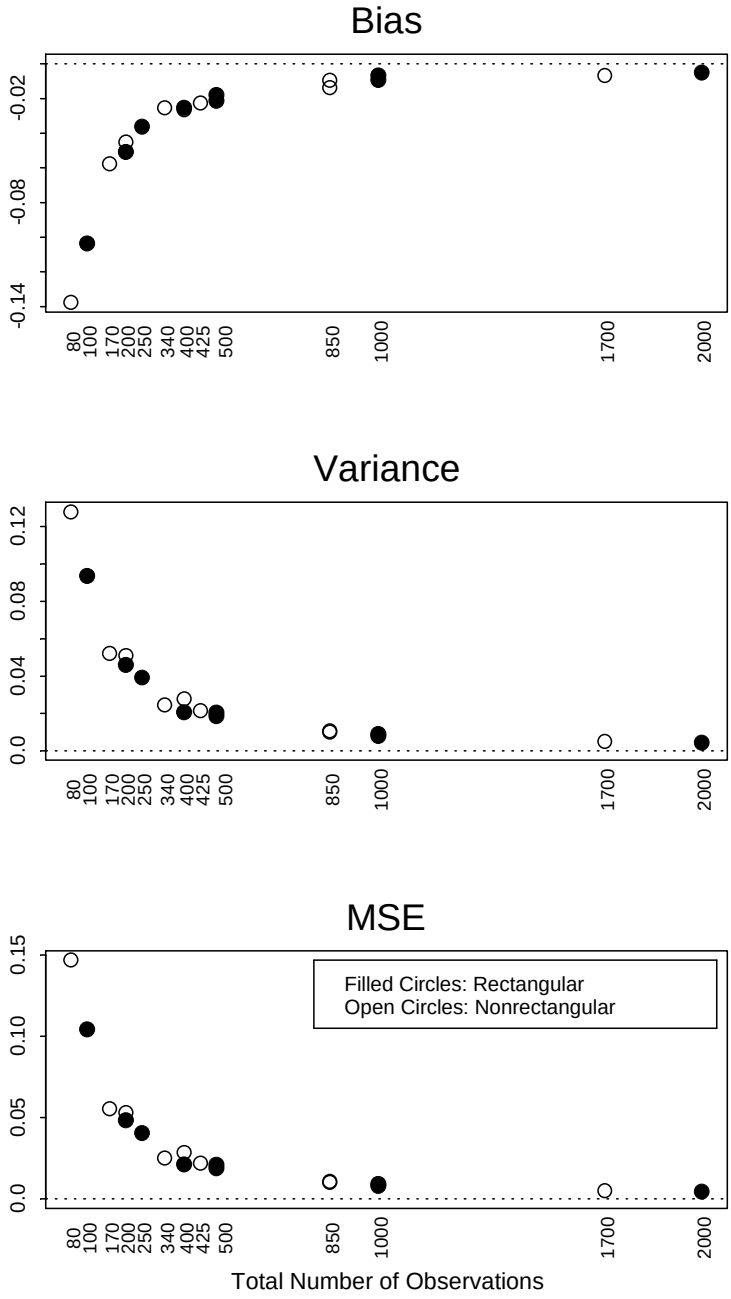

Figure 7. Bias, variance, and MSE for σ_{22} of VAR model with both rectangular and nonrectangular sample estimates.

Finally, the VAR model contains a matrix of regression coefficients that maps the mean adjusted prior response outcomes onto the current observed response variable and log of time of observation. The diagonal elements of the regression coefficients demonstrated similar patterns of accuracy and precision while the off-diagonal regression coefficients demonstrate similar patterns of precision and accuracy to each other. In the case of the diagonal elements there was a systematic negative bias in obtained estimates while for the off-diagonal elements there was a systematic positive bias in obtained estimates which in both cases where asymptomatically unbiased (Figure 8 through 11 and Tables 7 through 24). For all four regression coefficients, obtained estimates demonstrated a large amount of variation at smaller number of observations which progressively became more precise as the number of observations increased and essentially became centered by 850 observations. In addition, the relative contribution of the bias had little affect on the obtained MSE values, suggesting that estimate precision was more responsible for the observed results than the accuracy of the obtained estimates. Finally, the amount of variation in estimates for rectangular designs was less when compared to variations seen for nonrectangular designs at similar number of observations, suggesting that rectangular matrices improve estimate precision. But this trend was not clearly seen in the case of biasness which suggests that rectangular designs do not necessarily improve estimate accuracy. In fact, in a few cases the estimates obtained from nonrectangular sample matrices resulted in less bias estimates than for estimates obtained from rectangular sample matrices at similar number of observations but was not the case every time.

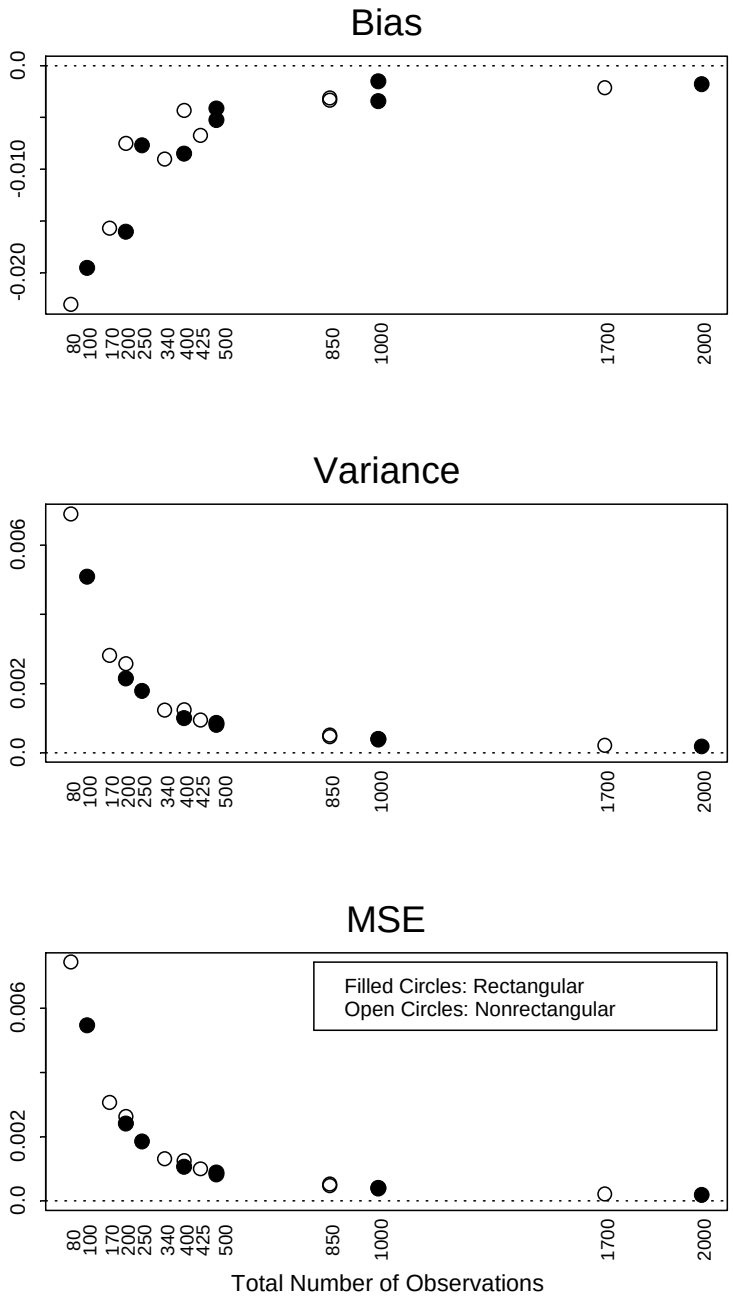

Figure 8. Bias, variance, and MSE for ϕ_{11} of VAR model with both rectangular and nonrectangular sample estimates.

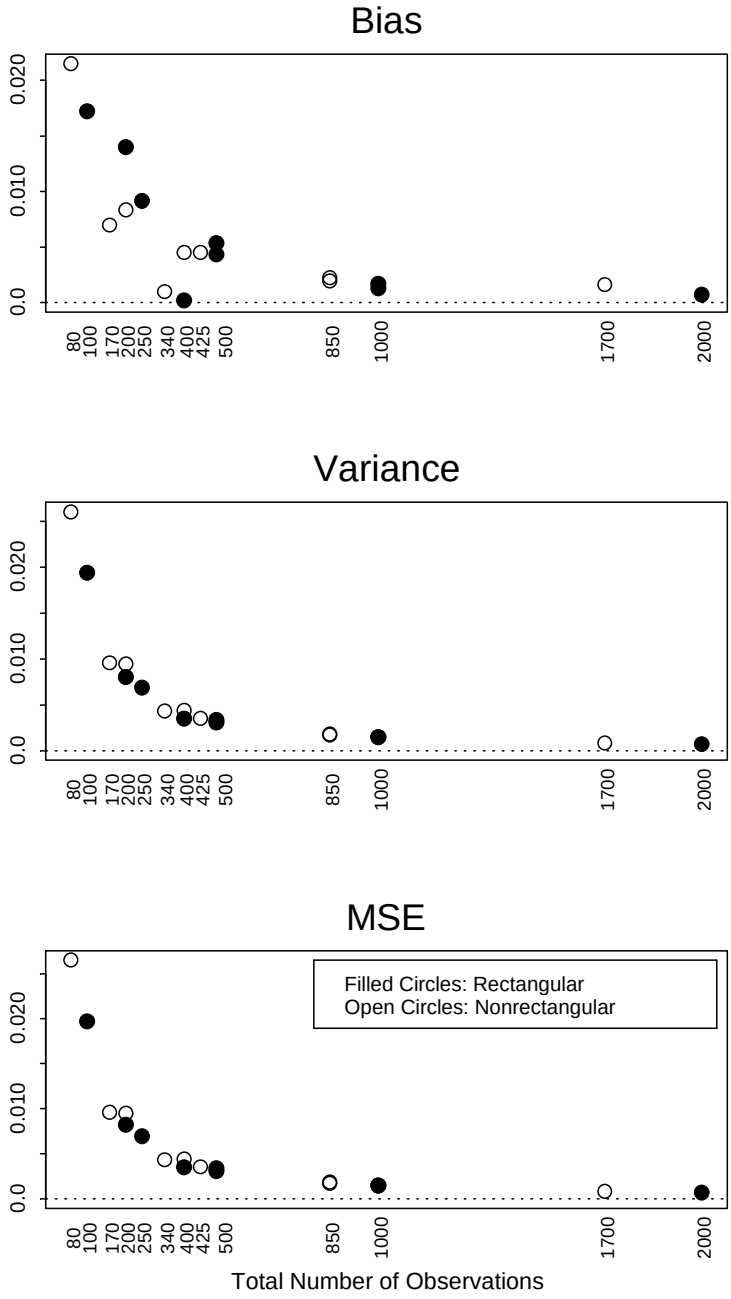

Figure 9. Bias, variance, and MSE for ϕ_{12} of VAR model with both rectangular and nonrectangular sample estimates.

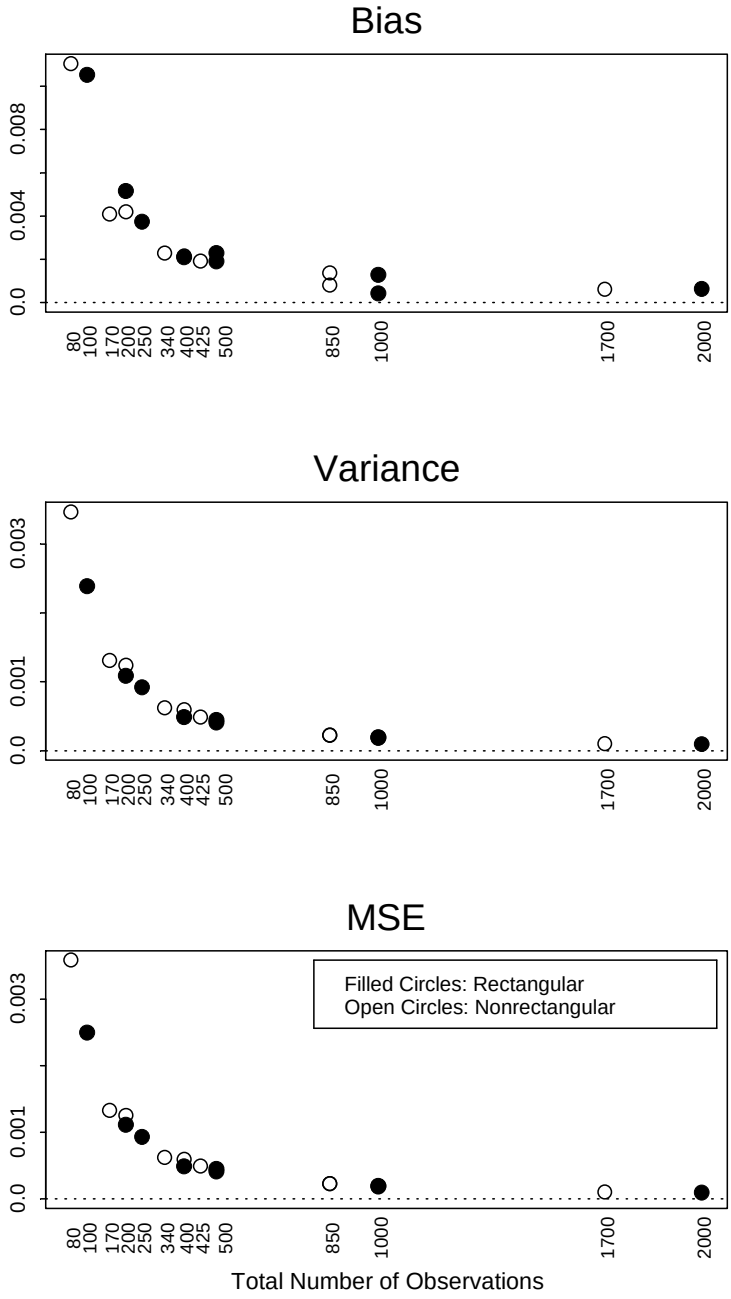

Figure 10. Bias, variance, and MSE for ϕ_{21} of VAR model with both rectangular and nonrectangular sample estimates.

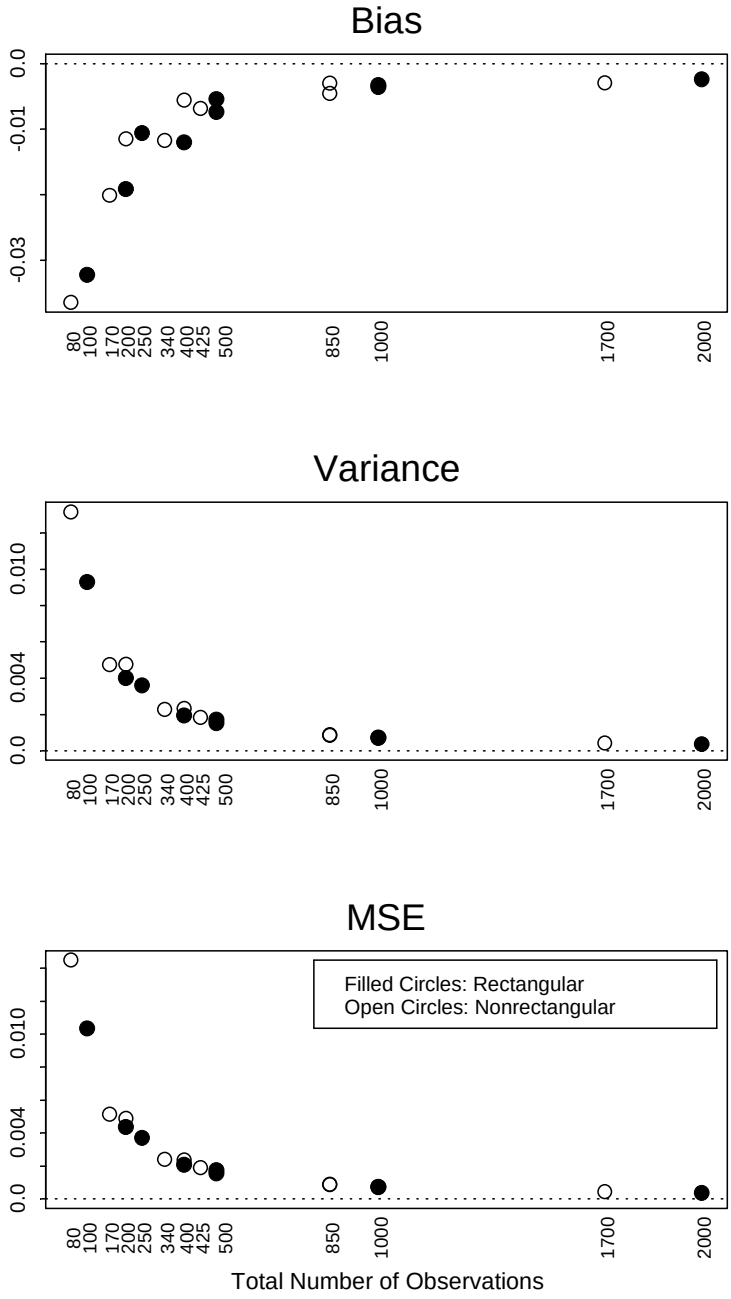

Figure 11. Bias, variance, and MSE for ϕ_{22} of VAR model with both rectangular and nonrectangular sample estimates.

VAR: Mixed-Effects Comparison

The same Monte Carlo simulated data used to estimate VAR parameters were also analyzed by the mixed-effects model. The analysis approach of the mixed-effects data assumed that the variance-covariance matrix followed a compound symmetrical structure and that observations of the response variable were correlated with each other across time. Furthermore, for the mixed effects approached observation lengths were assumed to be evenly spaced, i.e., the time between observations was no longer considered to be informative.

For all number of observations and for both rectangular and nonrectangular designs, the mixed-effects approach showed a substantial negative bias in estimates as compared to the response parameter β_1 (Tables 3 and 4). This observed substantial bias suggests that the analysis of informative schedule data by traditional longitudinal methods could substantially underestimate model parameters. Furthermore, the amount of variation seen in estimates obtained by the mixed-effects approach was much less than the amount of variation seen in estimates obtained by the VAR model. In addition, both approaches demonstrated decreasing variation as number of observations increased, but this trend was much more pronounced in the VAR model. Consequently, with a larger decrease in the amount of variation of estimates and a substantially less bias, the relative efficiency of the VAR model parameter was larger when compared to the mixed-effects model.

Table 3.

Mixed-effects parameter estimates for Vector Autoregressive with rectangular design.

Vector Autoregressive				
Observations	Bias	Variance	MSE	Relative Efficiency
20 Subjects				
100	*-0.2278* -7.0329	*41.5787* 5.2637	*41.6306* 54.7255	1.3145
200	*-0.0438* -7.0124	*11.7070* 4.8266	*11.7089* 54.0010	4.6120
400	*-0.0519* -7.0516	*4.4727* 3.6995	*4.4754* 53.4248	11.9374
50 Subjects				
250	*0.1398* -7.0397	*12.0888* 1.9783	*12.1084* 51.5362	4.2562
500	*-0.0212* -7.0017	*3.7542* 1.8650	*3.7546* 50.8888	13.5535
1000	*-0.0315* -6.9854	*1.5550* 1.3943	*1.5559* 50.1907	32.2574
100 Subjects				
500	*-0.0194* -6.9709	*4.9053* 0.9965	*4.9056* 49.5896	10.1087
1000	*-0.0239* -7.0047	*1.7999* 0.8780	*1.8005* 49.9436	27.7393
2000	*-0.0123* -6.9921	*0.8133* 0.7435	*0.8135* 49.6332	61.0120

Note: Italicized results are for Vector Autoregressive model.

Table 4.

Mixed-effects parameter estimates for Vector Autoregressive with nonrectangular design.

Vector Autoregressive				
Observations	Bias	Variance	MSE	Relative Efficiency
20 Subjects				
80	*0.0734* -6.9895	*50.2752* 5.4942	*50.2806* 54.3476	1.0809
170	*0.0144* -6.9169	*16.3259* 4.7735	*16.3261* 52.6170	3.2229
340	*-0.0436* -7.0568	*5.0974* 3.8492	*5.0993* 53.6477	10.5205
50 Subjects				
200	*0.0721* -6.9619	*23.0159* 2.1203	*23.0211* 50.5890	2.1975
425	*0.0147* -7.0162	*4.9416* 1.8451	*4.9418* 51.0718	10.3346
850	*-0.0357* -7.0235	*2.0163* 1.5293	*2.0176* 50.8589	25.2079
100 Subjects				
400	*-0.0955* -7.0231	*7.3969* 1.0671	*7.4061* 50.3911	6.8040
850	*-0.0197* -6.9921	*2.1595* 0.9794	*2.1599* 49.8687	23.0884
1700	*0.0063* -6.9899	*0.9196* 0.7913	*0.9196* 49.6494	53.9891

Note: Italicized results are for Vector Autoregressive model.

Gaussian-Exponential Parameters

The GE model includes a vector of $\boldsymbol{\beta}$ parameters that along with the design matrix determines the mean outcome for the response variable. Here, both mean parameter estimates followed similar patterns of precision as number of observations increased for both sample matrices designs but a slight difference in accuracy patterns was observed. When the number of observations were at their lowest amounts, the obtained estimates for both parameters showed a substantial amount of variation that became less pronounced as the number of observations increased suggesting that estimates become more centered as the number of observations increase (Figure 12 and 13 and Tables 25 through 42). Although, it should be noted that in both cases of the mean parameters a small amount of variation in the obtained estimates was still present even at the largest number of observations. In the case of the β_0 parameter the average estimates showed a systematic positive bias in obtained estimates that became asymptomatically unbiased as number of observations also increased. However, this trend was not as consistent in the obtained estimates for β_1 and in a few cases the average estimate demonstrated a negative bias. Once again, the relative contribution of the bias had little effect on the obtained MSE values, suggesting that estimate precision was more responsible for the observed results than the accuracy of the obtained estimates. Finally, the estimates obtained from rectangular sample matrices, once again seemed to result in less variation in the obtained estimates when compared to nonrectangular estimates but this trend was not necessarily observed in the case of bias.

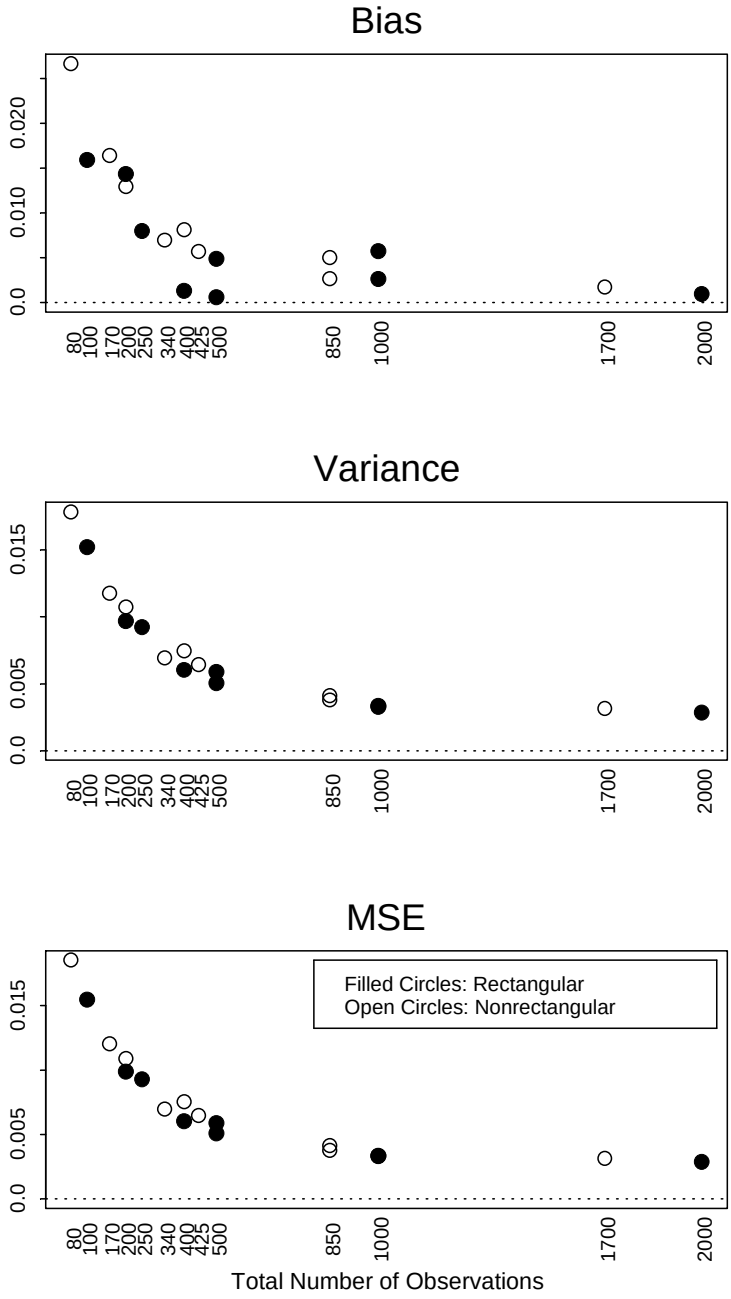

Figure 12. Bias, variance, and MSE for β_0 of GE model with both rectangular and nonrectangular sample estimates.

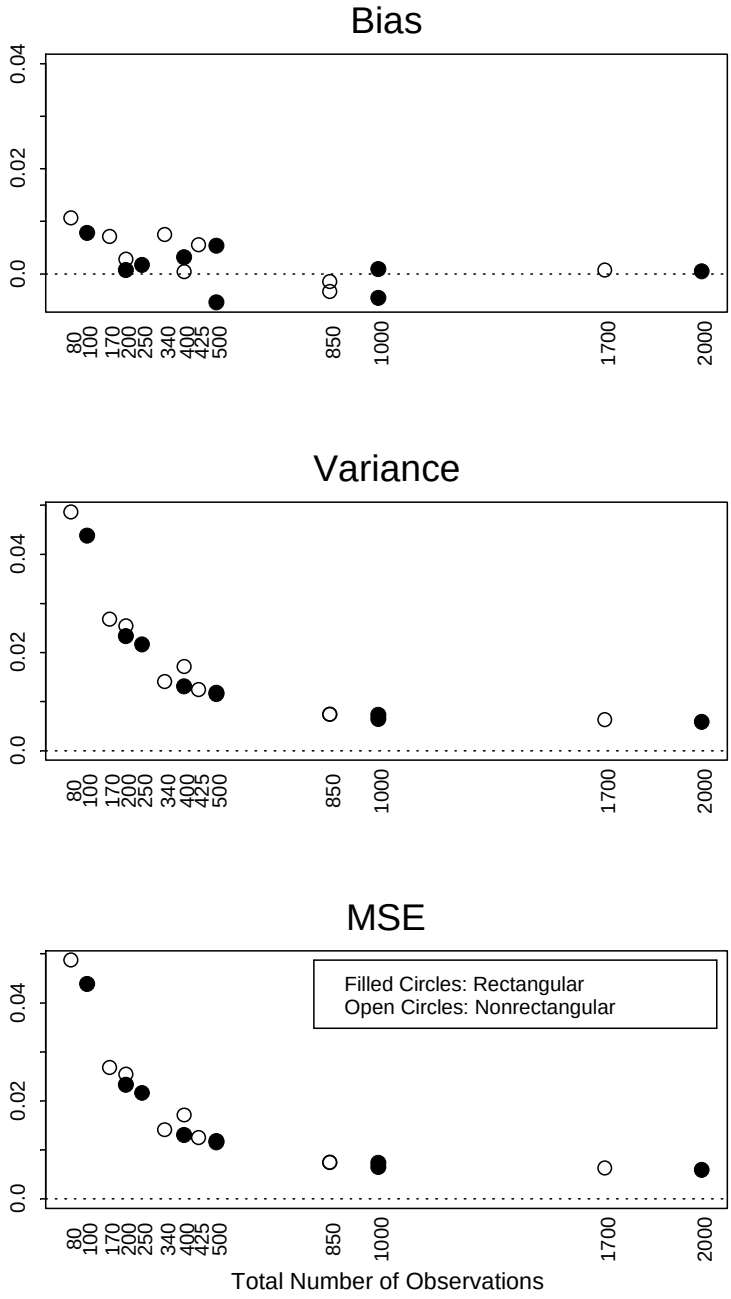

Figure 13. Bias, variance, and MSE for β_1 of GE model with both rectangular and nonrectangular sample estimates.

The GE model also contained parameters for the variance of the responses, σ^2 and a parameter for the correlation between subsequent responses, ρ. For both of these parameters the trends for precision and accuracy appeared to be similar as number of observations increased. Both parameters demonstrated a systematic negative bias in estimates at low number of observations which became less pronounced as the number of observations increased (Figures 14 and 15 and Tables 25 through 42). These observations suggest that the GE model tends to underestimate the variance and correlation parameters but that they are asymptotically unbiased as the number of observations in the sample matrix increases. Both parameters also demonstrated an observation dependent decease in the amount of variation in the obtained estimates and the obtained estimates essentially became centered by 850 observations, albeit more pronounced for the variance parameter. Furthermore, both parameter estimates appear to be more influence by estimate precision than by the accuracy of estimates in that bias values had little effect on the calculated MSE values. Also, estimates from nonrectangular sample matrices demonstrated a slight increase in variation of estimates when compared to estimates obtained from rectangular designs. However, this was not observed in every case for the variance parameter which might suggest that there might be the effect of number of subjects. Finally, estimates obtained from nonrectangular matrices did not appear to decease or increase the amount of bias seen for either parameter.

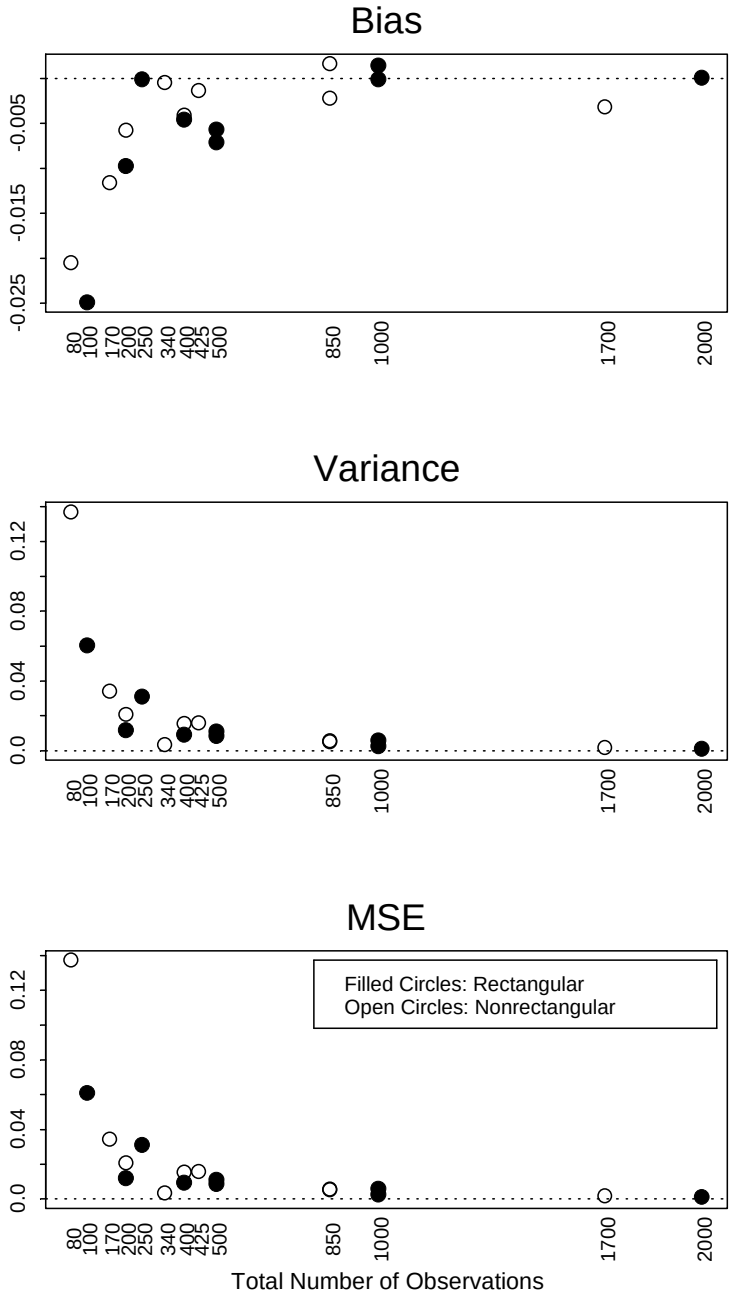

Figure 14. Bias, variance, and MSE for σ^2 of GE model with both rectangular and nonrectangular sample estimates.

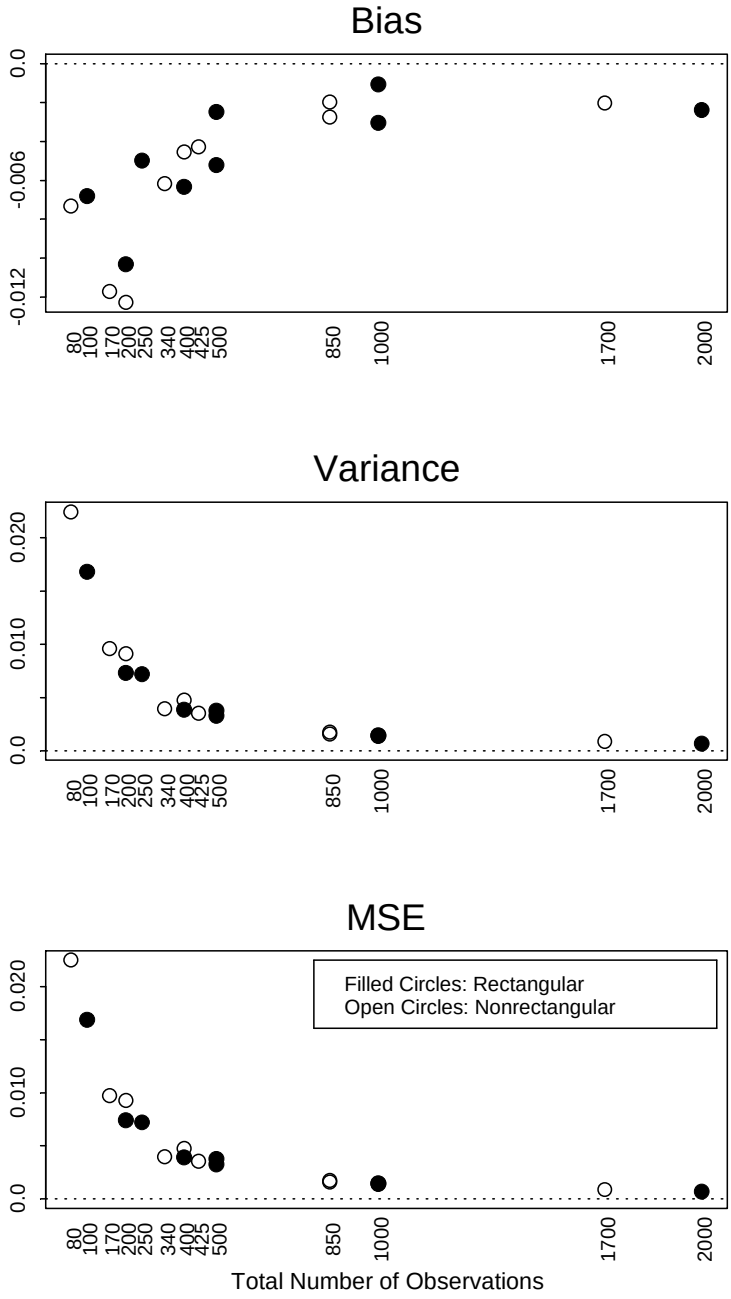

Figure 15. Bias, variance, and MSE for ρ of GE model with both rectangular and nonrectangular sample estimates.

Included with the overall mean response are the inclusion of the parameters that account for the prior response outcome and the current time of observation. Here the parameter, ϕ accounts for the effect of the prior response outcome on the mean response while the coefficient γ accounts for the effect of the current time of observation on the mean response. In the case of the prior response parameter, the obtained estimates demonstrated a slight negative bias at low number of observations that was weakly dependent on the number of observations (Figure 16 and Tables 25 through 42). This weak dependency on changes in number of observations might suggest that the bias in estimates might be due more to imprecision of the estimates. On the other hand, variation in the estimates did demonstrate a strong dependency on number of observations, in that as the number of observations increased the amount of observed variation in the estimates decreased. In the case of the current time of observation the obtained estimates demonstrated a systematic positive bias in obtained estimates that was clearly dependent on the number of observations (Figure 17 and Tables 25 through 42). The time parameter estimates also demonstrated a clear dependency on the number of observations with increased amount of estimate variation being seen at low number of observations. In both case of prior response parameters, the amount of bias seemed to have marginal influence on the calculated MSE values, once again suggesting that variation or precision is more influential in the obtained estimates. Finally, the effects on estimates obtained from nonrectangular sample matrices seemed to be limited to the variation in the estimates for both prior response parameters.

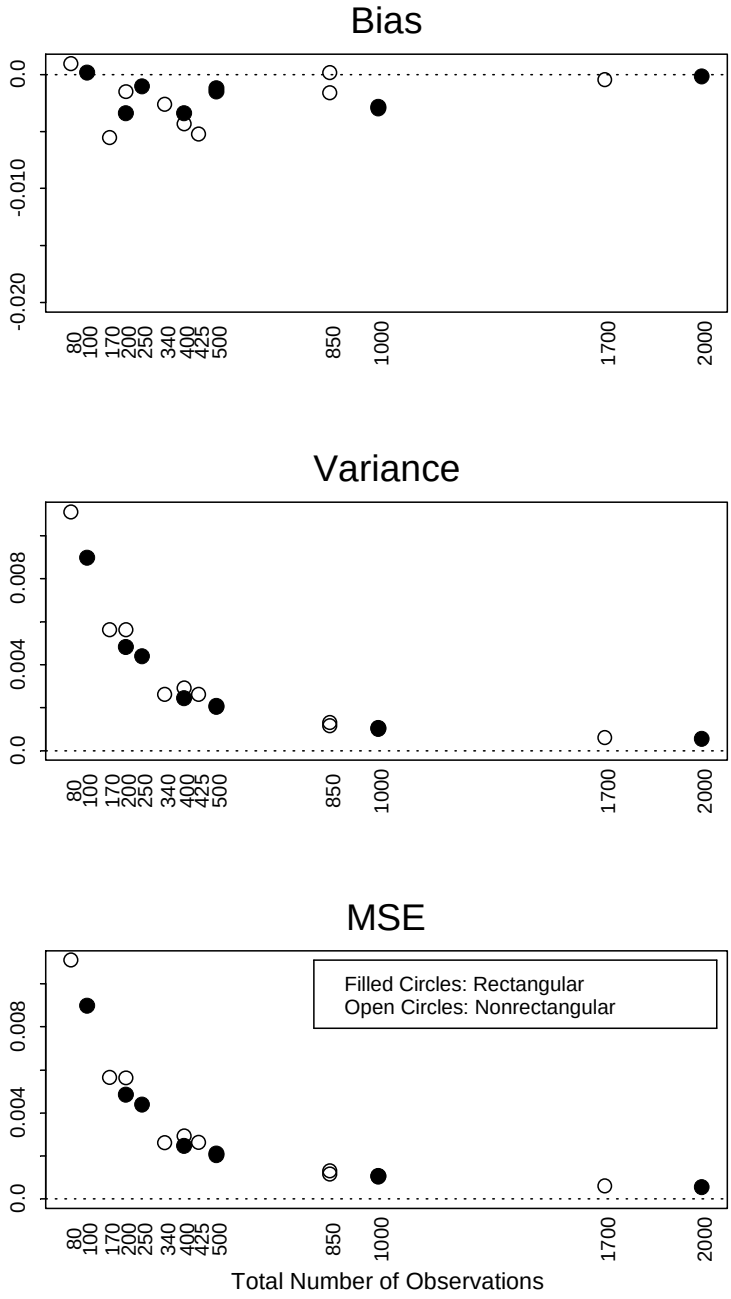

Figure 16. Bias, variance, and MSE for ϕ of GE model with both rectangular and nonrectangular sample estimates.

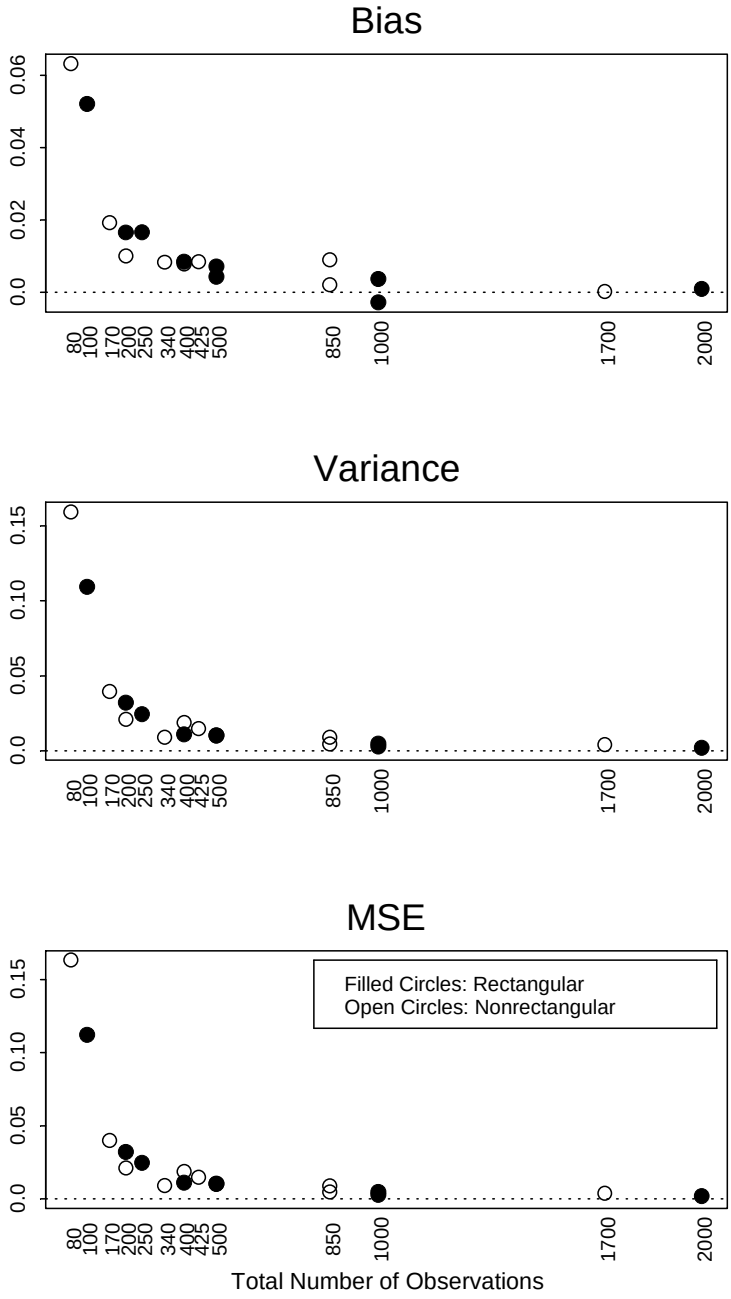

Figure 17. Bias, variance, and MSE for γ of GE model with both rectangular and nonrectangular sample estimates.

Finally, the GE model also contained parameters associated with modeling time of observation which included a constant parameter, α and a coefficient that maps time of observation, δ. For both of these parameters the variation in the estimates obtained showed a clear dependency on number of observations (Figure 18 and 19 and Tables 25 through 42). However, the estimates obtained for the constant parameter demonstrated a slight non-directional bias while the mapping coefficient demonstrated a clear systematic positive bias in estimates. Once again, precision of the estimates appeared to be more influential on the estimation of both parameters in that MSE values were essentially the same as the variance values. Finally, for both parameter estimates obtained from rectangular sample matrices seemed to have small amount of variation and bias when compared to nonrectangular obtained estimates.

GE: Mixed-Effects Comparison

For all number of observations and for both rectangular and nonrectangular designs, the mixed-effects approach showed a slight negative bias in estimates as compared to the response parameter β_0 (Tables 5 and 6). Furthermore, both approaches demonstrated a decrease in the variation of obtained estimates as the number of observations increased but this effect was more pronounced in the case of the GE model. Consequently, with a more pronounced decrease in the amount of variation of estimates and with a slightly less bias, the relative efficiency of the GE model parameter was larger when compared to the mixed-effects model suggesting improved estimation efficiencies.

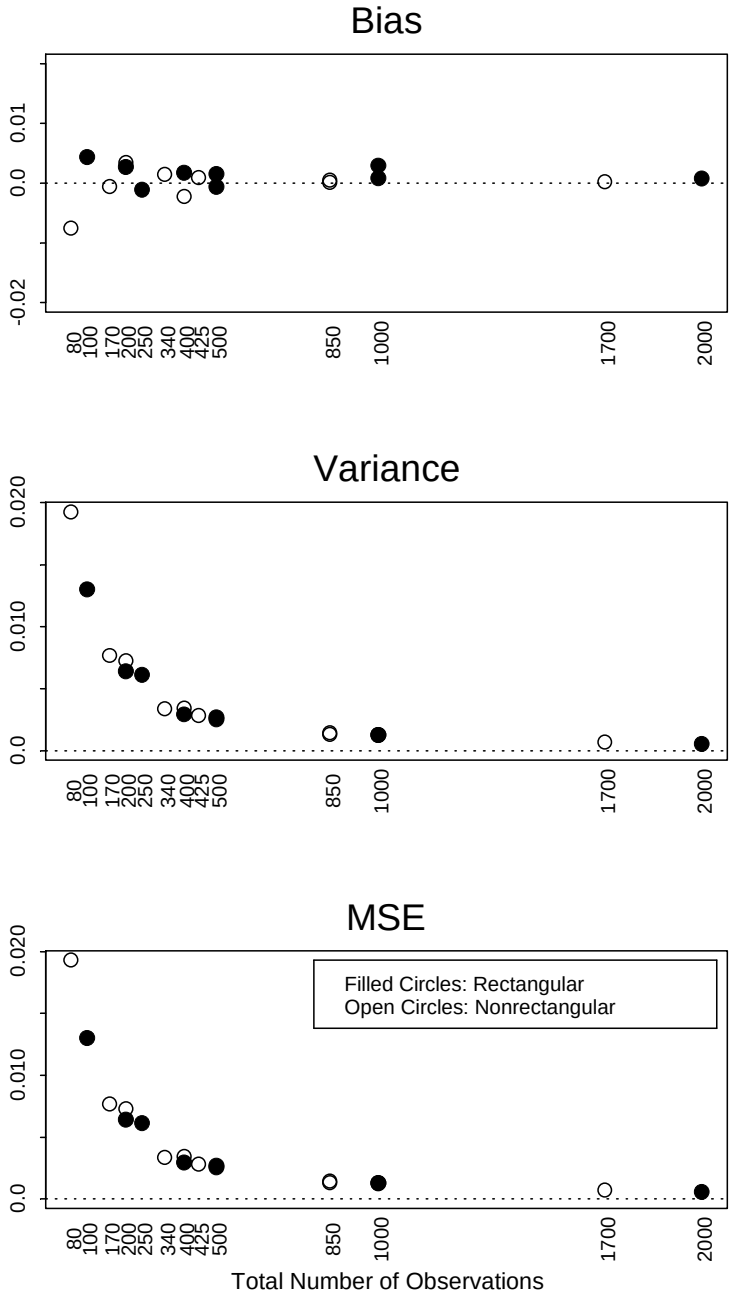

Figure 18. Bias, variance, and MSE for α of GE model with both rectangular and nonrectangular sample estimates.

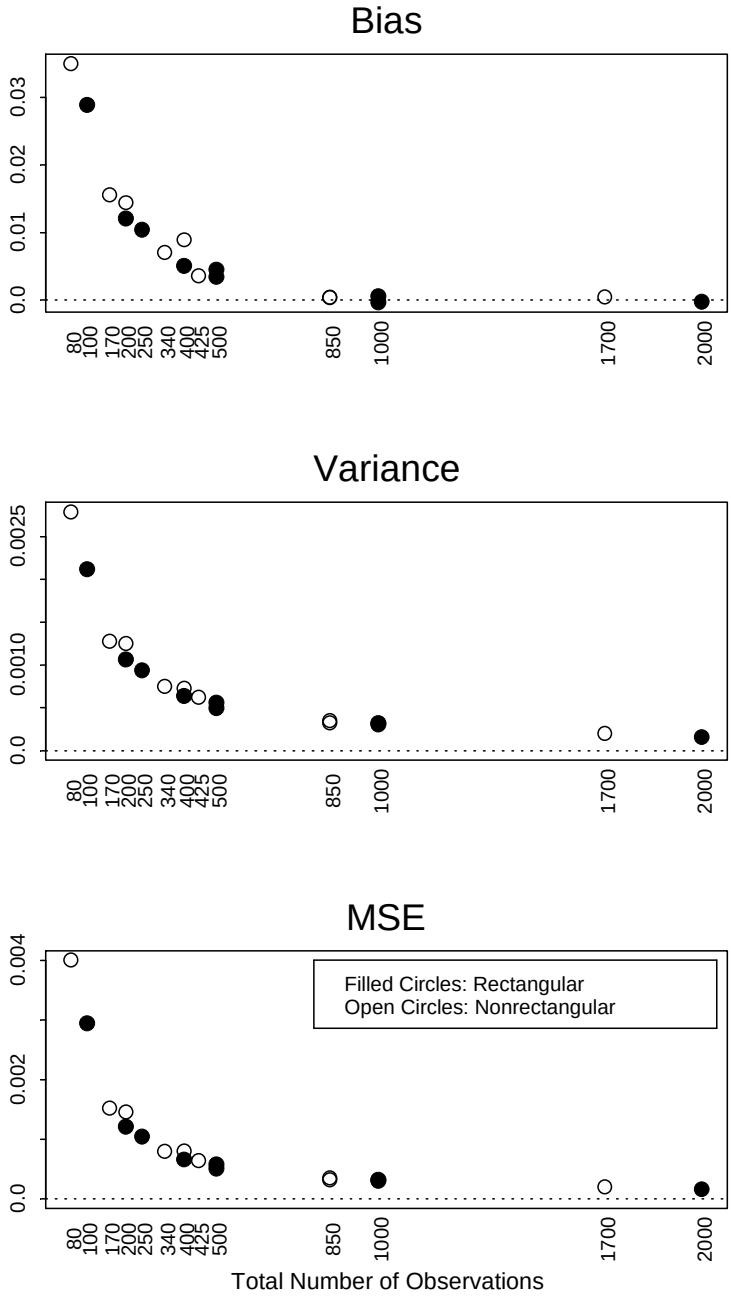

Figure 19. Bias, variance, and MSE for δ of GE model with both rectangular and nonrectangular sample estimates.

Table 5.

Mixed effect parameter estimates for Gaussian Exponential with rectangular data

Gaussian Exponential				
Observations	Bias	Variance	MSE	Relative Efficiency
20 Subjects				
100	*0.0159* -0.1997	*0.0152* 0.3177	*0.0154* 0.3576	23.1508
200	*0.0143* -0.2221	*0.0097* 0.1630	*0.0099* 0.2124	21.5208
400	*0.0013* -0.2233	*0.0060* 0.0921	*0.0060* 0.1420	23.5716
50 Subjects				
250	*0.0080* -0.2103	*0.0092* 0.1128	*0.0093* 0.1570	16.9391
500	*0.0049* -0.2052	*0.0051* 0.0617	*0.0051* 0.1038	20.4493
1000	*0.0057* -0.2073	*0.0033* 0.0364	*0.0033* 0.0794	23.9598
100 Subjects				
500	*0.0006* -0.2023	*0.0059* 0.0525	*0.0059* 0.0934	15.9413
1000	*0.0026* -0.1997	*0.0033* 0.0278	*0.0033* 0.0677	20.2368
2000	*0.0009* -0.1989	*0.0029* 0.0179	*0.0029* 0.0575	20.1704

Note: Italicized results are for Gaussian Exponential model.

Table 6.

Mixed effect parameter estimates for Gaussian Exponential with nonrectangular data

	Gaussian Exponential			
Observations	Bias	Variance	MSE	Relative Efficiency
20 Subjects				
80	*0.0267* -0.1864	*0.0178* 0.3882	*0.0185* 0.4230	22.8186
170	*0.0164* -0.2249	*0.0118* 0.2035	*0.0120* 0.2540	21.1180
340	*0.0069* -0.2408	*0.0069* 0.1116	*0.0070* 0.1695	24.3489
50 Subjects				
200	*0.0129* -0.2088	*0.0107* 0.1377	*0.0109* 0.1813	16.6695
425	*0.0057* -0.1999	*0.0064* 0.0720	*0.0065* 0.1120	17.3542
850	*0.0027* -0.2054	*0.0038* 0.0384	*0.0038* 0.0806	21.3207
100 Subjects				
400	*0.0081* -0.2121	*0.0075* 0.0654	*0.0075* 0.1104	14.6750
850	*0.0050* -0.2018	*0.0041* 0.0367	*0.0041* 0.0774	18.7685
1700	*0.0017* -0.2042	*0.0031* 0.0175	*0.0031* 0.0592	18.8589

Note: Italicized results are for Gaussian Exponential model.

Discussion

The results obtained from the analysis of both the VAR and GE simulated data indicate that model parameters can be estimated using the maximum likelihood method. These obtained estimates generally showed low bias especially for larger number of observations and in most cases approached the true parameter value as the number of observations increased. In a few cases the amount of bias observed, especially for low number of observations, demonstrated a systematic trend. Namely, the estimates for the two variance components of the VAR model showed evidence that the proposed model underestimates these parameters. However, this was not the case for the estimates of the covariance parameter in this model. Evidence of underestimation was also seen in the GE model for the variance and correlation parameters. Underestimation of variance is a common issue in maximum likelihood estimation especially when sample sizes are relatively small (Fitzmaurice et al. 2004) and this may be the issue seen in our models. Furthermore, in a few cases for both VAR and GE model parameter estimates demonstrated a systematic overestimation. However, for both models and for all parameters the amount of bias observed decreased as the number of observations increased and at larger number of observations was essentially equal to the true population parameter.

For both models and all parameters the amount of variation in estimates was substantially large at low number of observations but as the number of observations increased the amount of variation in estimates decrease. Also, for the most part evaluation of the estimate's MSE revealed the same patterns and approximately the same values as those observed for variation in estimates. This significant dependency on

variation in the calculation of the MSE values suggest that parameter estimate performance was largely influenced by the variation of the estimates and less by the amount of bias. More precisely, estimates obtained from each model show very little inaccuracy but have large amount of imprecision at low number of observations.

When estimates were obtained from nonrectangular designs, the overall patterns of bias and variation in estimates seen in rectangular designs held. However, in many cases the amount of variation of estimates obtained from rectangular sample matrices was slightly decrease when compared to nonrectangular samples matrices. And in a few cases this improved performance of estimates obtained from rectangular sample matrices was also seen in bias of obtained estimates. These results suggest that rectangular or complete sample matrices result in more accurate and precise estimates.

Finally, the estimates obtained from both proposed models showed improved performance when compared to estimates obtained from the mixed-effects model. This improved performance was most obvious in the VAR model in that the bias of the mixed-effects model was substantially larger. However, a 'fair' comparison between the informative schedule model and the mixed-effect was not strictly possible since there were very little overlap in common parameters. Although, it should be noted that the addition of parameters that allow for the estimation of prior response and time of observation effect on the observed response outcome can only contribute to better understanding of the process that generated the data.

Table 7.

Parameter estimates for 20 subjects with 100 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.7722	-0.2278	41.5787	41.6306
β_2	2	1.9708	-0.0292	12.6666	12.6674
β_3	3	3.1489	0.1489	69.5616	69.5838
β_4	1	1.0691	0.0691	21.2965	21.3013
σ_{11}	4	3.7972	-0.2028	0.3813	0.4224
σ_{12}	0.1	0.0909	-0.0091	0.0963	0.0964
σ_{22}	2	1.8961	-0.1039	0.0934	0.1042
ϕ_{11}	0.8	0.7805	-0.0195	0.0051	0.0055
ϕ_{12}	0.3	0.3172	0.0172	0.0194	0.0197
ϕ_{21}	0.2	0.2105	0.0105	0.0024	0.0025
ϕ_{22}	0.5	0.4677	-0.0323	0.0093	0.0103

Table 8.

Parameter estimates for 20 subjects with 200 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9562	-0.0438	11.7070	11.7089
β_2	2	1.9818	-0.0182	2.4234	2.4237
β_3	3	2.9842	-0.0158	22.6441	22.6444
β_4	1	0.9833	-0.0167	4.7488	4.7491
σ_{11}	4	3.9037	-0.0963	0.1725	0.1817
σ_{12}	0.1	0.1002	0.0002	0.0436	0.0436
σ_{22}	2	1.9490	-0.0510	0.0458	0.0484
ϕ_{11}	0.8	0.7839	-0.0161	0.0021	0.0024
ϕ_{12}	0.3	0.3140	0.0140	0.0080	0.0082
ϕ_{21}	0.2	0.2052	0.0052	0.0011	0.0011
ϕ_{22}	0.5	0.4808	-0.0192	0.0040	0.0044

Table 9.

Parameter estimates for 20 subjects with 400 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9481	-0.0519	4.4727	4.4754
β_2	2	1.9745	-0.0255	0.9161	0.9168
β_3	3	3.1034	0.1034	8.9457	8.9564
β_4	1	1.0455	0.0455	1.8189	1.8209
σ_{11}	4	3.9603	-0.0397	0.0779	0.0795
σ_{12}	0.1	0.1000	0.0000	0.0216	0.0216
σ_{22}	2	1.9745	-0.0255	0.0205	0.0212
ϕ_{11}	0.8	0.7915	-0.0085	0.0010	0.0011
ϕ_{12}	0.3	0.3002	0.0002	0.0035	0.0035
ϕ_{21}	0.2	0.2021	0.0021	0.0005	0.0005
ϕ_{22}	0.5	0.4880	-0.0120	0.0019	0.0021

Table 10.

Parameter estimates for 50 subjects with 250 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	4.1398	0.1398	12.0888	12.1084
β_2	2	2.1030	0.1030	4.2328	4.2434
β_3	3	2.8447	-0.1553	24.5113	24.5354
β_4	1	0.9310	-0.0690	6.6238	6.6286
σ_{11}	4	3.9285	-0.0715	0.1609	0.1660
σ_{12}	0.1	0.1046	0.0046	0.0397	0.0397
σ_{22}	2	1.9636	-0.0364	0.0391	0.0404
ϕ_{11}	0.8	0.7923	-0.0077	0.0018	0.0018
ϕ_{12}	0.3	0.3091	0.0091	0.0069	0.0070
ϕ_{21}	0.2	0.2037	0.0037	0.0009	0.0009
ϕ_{22}	0.5	0.4894	-0.0106	0.0036	0.0037

Table 11.

Parameter estimates for 50 subjects with 500 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9788	-0.0212	3.7542	3.7546
β_2	2	1.9891	-0.0109	0.7946	0.7947
β_3	3	3.0710	0.0710	7.5976	7.6026
β_4	1	1.0334	0.0334	1.5924	1.5935
σ_{11}	4	3.9522	-0.0478	0.0682	0.0705
σ_{12}	0.1	0.0987	-0.0013	0.0167	0.0167
σ_{22}	2	1.9818	-0.0182	0.0184	0.0187
ϕ_{11}	0.8	0.7947	-0.0053	0.0008	0.0008
ϕ_{12}	0.3	0.3043	0.0043	0.0030	0.0031
ϕ_{21}	0.2	0.2023	0.0023	0.0004	0.0004
ϕ_{22}	0.5	0.4926	-0.0074	0.0015	0.0016

Table 12.

Parameter estimates for 50 subjects with 1000 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9685	-0.0315	1.5550	1.5559
β_2	2	1.9829	-0.0171	0.3234	0.3237
β_3	3	2.9754	-0.0246	3.2904	3.2910
β_4	1	0.9912	-0.0088	0.6788	0.6789
σ_{11}	4	3.9796	-0.0204	0.0352	0.0356
σ_{12}	0.1	0.0996	-0.0004	0.0086	0.0086
σ_{22}	2	1.9933	-0.0067	0.0079	0.0080
ϕ_{11}	0.8	0.7965	-0.0035	0.0004	0.0004
ϕ_{12}	0.3	0.3017	0.0017	0.0015	0.0015
ϕ_{21}	0.2	0.2004	0.0004	0.0002	0.0002
ϕ_{22}	0.5	0.4967	-0.0033	0.0007	0.0007

Table 13.

Parameter estimates for 100 subjects with 500 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9806	-0.0194	4.9053	4.9056
β_2	2	1.9889	-0.0111	0.9978	0.9979
β_3	3	3.0026	0.0026	9.2970	9.2970
β_4	1	1.0000	0.0000	1.8722	1.8722
σ_{11}	4	3.9638	-0.0362	0.0801	0.0815
σ_{12}	0.1	0.0972	-0.0028	0.0201	0.0202
σ_{22}	2	1.9785	-0.0215	0.0204	0.0209
ϕ_{11}	0.8	0.7959	-0.0041	0.0009	0.0009
ϕ_{12}	0.3	0.3053	0.0053	0.0034	0.0034
ϕ_{21}	0.2	0.2019	0.0019	0.0004	0.0004
ϕ_{22}	0.5	0.4945	-0.0055	0.0017	0.0017

Table 14.

Parameter estimates for 100 subjects with 1000 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9761	-0.0239	1.7999	1.8005
β_2	2	1.9876	-0.0124	0.3706	0.3707
β_3	3	3.0889	0.0889	3.5258	3.5337
β_4	1	1.0455	0.0455	0.7260	0.7281
σ_{11}	4	3.9801	-0.0199	0.0344	0.0348
σ_{12}	0.1	0.1040	0.0040	0.0084	0.0084
σ_{22}	2	1.9905	-0.0095	0.0090	0.0090
ϕ_{11}	0.8	0.7985	-0.0015	0.0004	0.0004
ϕ_{12}	0.3	0.3013	0.0013	0.0015	0.0015
ϕ_{21}	0.2	0.2013	0.0013	0.0002	0.0002
ϕ_{22}	0.5	0.4964	-0.0036	0.0007	0.0007

Table 15.

Parameter estimates for 100 subjects with 2000 observations in a rectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9877	-0.0123	0.8133	0.8135
β_2	2	1.9971	-0.0029	0.1676	0.1676
β_3	3	3.0078	0.0078	1.6703	1.6703
β_4	1	1.0014	0.0014	0.3410	0.3410
σ_{11}	4	3.9908	-0.0092	0.0165	0.0166
σ_{12}	0.1	0.0998	-0.0002	0.0043	0.0043
σ_{22}	2	1.9948	-0.0052	0.0043	0.0044
ϕ_{11}	0.8	0.7982	-0.0018	0.0002	0.0002
ϕ_{12}	0.3	0.3007	0.0007	0.0007	0.0007
ϕ_{21}	0.2	0.2006	0.0006	0.0001	0.0001
ϕ_{22}	0.5	0.4976	-0.0024	0.0004	0.0004

Table 16.

Parameter estimates for 20 subjects with 80 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	4.0734	0.0734	50.2752	50.2806
β_2	2	2.0890	0.0890	17.1728	17.1808
β_3	3	2.9918	-0.0082	83.7806	83.7807
β_4	1	0.8921	-0.1079	28.2052	28.2168
σ_{11}	4	3.7270	-0.2730	0.5056	0.5801
σ_{12}	0.1	0.0906	-0.0094	0.1227	0.1228
σ_{22}	2	1.8622	-0.1378	0.1279	0.1468
ϕ_{11}	0.8	0.7769	-0.0231	0.0069	0.0074
ϕ_{12}	0.3	0.3215	0.0215	0.0260	0.0265
ϕ_{21}	0.2	0.2110	0.0110	0.0035	0.0036
ϕ_{22}	0.5	0.4635	-0.0365	0.0132	0.0145

Table 17.

Parameter estimates for 20 subjects with 170 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	4.0144	0.0144	16.3259	16.3261
β_2	2	2.0223	0.0223	3.6316	3.6321
β_3	3	2.9298	-0.0702	28.4676	28.4726
β_4	1	0.9600	-0.0400	6.7809	6.7825
σ_{11}	4	3.8764	-0.1236	0.1966	0.2119
σ_{12}	0.1	0.0893	-0.0107	0.0506	0.0507
σ_{22}	2	1.9422	-0.0578	0.0520	0.0554
ϕ_{11}	0.8	0.7843	-0.0157	0.0028	0.0031
ϕ_{12}	0.3	0.3070	0.0070	0.0096	0.0096
ϕ_{21}	0.2	0.2041	0.0041	0.0013	0.0013
ϕ_{22}	0.5	0.4799	-0.0201	0.0047	0.0051

Table 18.

Parameter estimates for 20 subjects with 340 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9564	-0.0436	5.0974	5.0993
β_2	2	1.9803	-0.0197	1.0594	1.0598
β_3	3	3.0439	0.0439	10.4686	10.4705
β_4	1	1.0190	0.0190	2.1559	2.1563
σ_{11}	4	3.9480	-0.0520	0.0943	0.0970
σ_{12}	0.1	0.1017	0.0017	0.0251	0.0251
σ_{22}	2	1.9744	-0.0256	0.0245	0.0251
ϕ_{11}	0.8	0.7910	-0.0090	0.0012	0.0013
ϕ_{12}	0.3	0.3010	0.0010	0.0043	0.0043
ϕ_{21}	0.2	0.2023	0.0023	0.0006	0.0006
ϕ_{22}	0.5	0.4883	-0.0117	0.0023	0.0024

Table 19.

Parameter estimates for 50 subjects with 200 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	4.0721	0.0721	23.0159	23.0211
β_2	2	2.0310	0.0310	4.7024	4.7034
β_3	3	2.9322	-0.0678	36.3688	36.3734
β_4	1	0.9649	-0.0351	8.3279	8.3292
σ_{11}	4	3.8916	-0.1084	0.2075	0.2193
σ_{12}	0.1	0.0942	-0.0058	0.0508	0.0508
σ_{22}	2	1.9547	-0.0453	0.0510	0.0530
ϕ_{11}	0.8	0.7925	-0.0075	0.0026	0.0026
ϕ_{12}	0.3	0.3083	0.0083	0.0094	0.0095
ϕ_{21}	0.2	0.2042	0.0042	0.0012	0.0013
ϕ_{22}	0.5	0.4885	-0.0115	0.0048	0.0049

Table 20.

Parameter estimates for 50 subjects with 425 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	4.0147	0.0147	4.9416	4.9418
β_2	2	2.0010	0.0010	1.0043	1.0043
β_3	3	3.0113	0.0113	10.0859	10.0861
β_4	1	1.0164	0.0164	2.0348	2.0351
σ_{11}	4	3.9655	-0.0345	0.0805	0.0817
σ_{12}	0.1	0.0966	-0.0034	0.0208	0.0208
σ_{22}	2	1.9773	-0.0227	0.0214	0.0219
ϕ_{11}	0.8	0.7933	-0.0067	0.0009	0.0010
ϕ_{12}	0.3	0.3045	0.0045	0.0035	0.0035
ϕ_{21}	0.2	0.2019	0.0019	0.0005	0.0005
ϕ_{22}	0.5	0.4931	-0.0069	0.0018	0.0019

Table 21.

Parameter estimates for 50 subjects with 850 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9643	-0.0357	2.0163	2.0176
β_2	2	1.9838	-0.0162	0.4176	0.4179
β_3	3	3.0000	0.0000	3.9642	3.9642
β_4	1	0.9944	-0.0056	0.8118	0.8118
σ_{11}	4	3.9797	-0.0203	0.0388	0.0393
σ_{12}	0.1	0.0992	-0.0008	0.0098	0.0098
σ_{22}	2	1.9905	-0.0095	0.0099	0.0100
ϕ_{11}	0.8	0.7967	-0.0033	0.0005	0.0005
ϕ_{12}	0.3	0.3019	0.0019	0.0017	0.0017
ϕ_{21}	0.2	0.2008	0.0008	0.0002	0.0002
ϕ_{22}	0.5	0.4955	-0.0045	0.0008	0.0009

Table 22.

Parameter estimates for 100 subjects with 400 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9045	-0.0955	7.3969	7.4061
β_2	2	1.9608	-0.0392	1.5213	1.5228
β_3	3	3.0978	0.0978	14.1507	14.1602
β_4	1	1.0507	0.0507	3.2248	3.2274
σ_{11}	4	3.9536	-0.0464	0.1058	0.1080
σ_{12}	0.1	0.0981	-0.0019	0.0279	0.0279
σ_{22}	2	1.9734	-0.0266	0.0279	0.0286
ϕ_{11}	0.8	0.7957	-0.0043	0.0012	0.0013
ϕ_{12}	0.3	0.3045	0.0045	0.0044	0.0044
ϕ_{21}	0.2	0.2021	0.0021	0.0006	0.0006
ϕ_{22}	0.5	0.4944	-0.0056	0.0023	0.0024

Table 23.

Parameter estimates for 100 subjects with 850 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	3.9803	-0.0197	2.1595	2.1599
β_2	2	1.9963	-0.0037	0.4426	0.4427
β_3	3	3.0293	0.0293	4.4227	4.4236
β_4	1	1.0062	0.0062	0.9028	0.9028
σ_{11}	4	3.9769	-0.0231	0.0428	0.0433
σ_{12}	0.1	0.0937	-0.0063	0.0106	0.0106
σ_{22}	2	1.9860	-0.0140	0.0105	0.0107
ϕ_{11}	0.8	0.7969	-0.0031	0.0005	0.0005
ϕ_{12}	0.3	0.3022	0.0022	0.0018	0.0018
ϕ_{21}	0.2	0.2014	0.0014	0.0002	0.0002
ϕ_{22}	0.5	0.4970	-0.0030	0.0009	0.0009

Table 24.

Parameter estimates for 100 subjects with 1700 observations in a nonrectangular design.

Vector Autoregressive

Parameter	True Value	Estimate	Bias	Variance	MSE
β_1	4	4.0063	0.0063	0.9196	0.9196
β_2	2	2.0032	0.0032	0.1918	0.1918
β_3	3	2.9985	-0.0015	1.8004	1.8004
β_4	1	0.9958	-0.0042	0.3754	0.3754
σ_{11}	4	3.9893	-0.0107	0.0191	0.0192
σ_{12}	0.1	0.0996	-0.0004	0.0049	0.0049
σ_{22}	2	1.9933	-0.0067	0.0050	0.0050
ϕ_{11}	0.8	0.7979	-0.0021	0.0002	0.0002
ϕ_{12}	0.3	0.3016	0.0016	0.0008	0.0008
ϕ_{21}	0.2	0.2006	0.0006	0.0001	0.0001
ϕ_{22}	0.5	0.4971	-0.0029	0.0004	0.0004

Table 25.

Parameter estimates for 20 subjects with 100 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2159	0.0159	0.0152	0.0154
β_1	0.5	0.5077	0.0077	0.0438	0.0438
σ^2	4	3.9751	-0.0249	0.0603	0.0609
ρ	0.5	0.4932	-0.0068	0.0168	0.0169
ϕ	0.2	0.2001	0.0001	0.0090	0.0090
γ	0.3	0.5520	0.0520	0.1093	0.1120
α	2	2.0044	0.0044	0.0130	0.0130
δ	0.04	0.0688	0.0288	0.0021	0.0029

Table 26.

Parameter estimates for 20 subjects with 200 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2143	0.0143	0.0097	0.0099
β_1	0.5	0.5007	0.0007	0.0233	0.0233
σ^2	4	3.9902	-0.0098	0.0117	0.0118
ρ	0.5	0.4897	-0.0103	0.0073	0.0074
ϕ	0.2	0.1966	-0.0034	0.0048	0.0048
γ	0.3	0.5165	0.0165	0.0319	0.0322
α	2	2.0027	0.0027	0.0064	0.0064
δ	0.04	0.0520	0.0120	0.0011	0.0012

Table 27.

Parameter estimates for 20 subjects with 400 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2013	0.0013	0.0060	0.0060
β_1	0.5	0.5031	0.0031	0.0130	0.0131
σ^2	4	3.9954	-0.0046	0.0092	0.0092
ρ	0.5	0.4937	-0.0063	0.0039	0.0039
ϕ	0.2	0.1966	-0.0034	0.0024	0.0025
γ	0.3	0.5085	0.0085	0.0108	0.0109
α	2	2.0017	0.0017	0.0029	0.0029
δ	0.04	0.0450	0.0050	0.0006	0.0007

Table 28.

Parameter estimates for 50 subjects with 250 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2080	0.0080	0.0092	0.0093
β_1	0.5	0.5017	0.0017	0.0216	0.0216
σ^2	4	3.9999	-0.0001	0.0309	0.0309
ρ	0.5	0.4950	-0.0050	0.0072	0.0072
ϕ	0.2	0.1990	-0.0010	0.0044	0.0044
γ	0.3	0.5167	0.0167	0.0244	0.0247
α	2	1.9989	-0.0011	0.0061	0.0061
δ	0.04	0.0504	0.0104	0.0009	0.0010

Table 29.

Parameter estimates for 50 subjects with 500 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2049	0.0049	0.0051	0.0051
β_1	0.5	0.4945	-0.0055	0.0117	0.0118
σ^2	4	3.9943	-0.0057	0.0086	0.0086
ρ	0.5	0.4975	-0.0025	0.0032	0.0033
ϕ	0.2	0.1988	-0.0012	0.0020	0.0020
γ	0.3	0.5043	0.0043	0.0103	0.0103
α	2	2.0015	0.0015	0.0027	0.0027
δ	0.04	0.0434	0.0034	0.0005	0.0005

Table 30.

Parameter estimates for 50 subjects with 1000 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2057	0.0057	0.0033	0.0033
β_1	0.5	0.4954	-0.0046	0.0065	0.0065
σ^2	4	3.9999	-0.0001	0.0025	0.0025
ρ	0.5	0.4989	-0.0011	0.0014	0.0014
ϕ	0.2	0.1971	-0.0029	0.0010	0.0010
γ	0.3	0.5037	0.0037	0.0047	0.0047
α	2	2.0008	0.0008	0.0013	0.0013
δ	0.04	0.0405	0.0005	0.0003	0.0003

Table 31.

Parameter estimates for 100 subjects with 500 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2006	0.0006	0.0059	0.0059
β_1	0.5	0.5053	0.0053	0.0115	0.0115
σ^2	4	3.9929	-0.0071	0.0109	0.0109
ρ	0.5	0.4948	-0.0052	0.0037	0.0038
ϕ	0.2	0.1985	-0.0015	0.0021	0.0021
γ	0.3	0.5071	0.0071	0.0102	0.0102
α	2	1.9994	-0.0006	0.0025	0.0025
δ	0.04	0.0445	0.0045	0.0006	0.0006

Table 32.

Parameter estimates for 100 subjects with 1000 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2026	0.0026	0.0033	0.0033
β_1	0.5	0.5009	0.0009	0.0073	0.0073
σ^2	4	4.0014	0.0014	0.0059	0.0059
ρ	0.5	0.4969	-0.0031	0.0015	0.0015
ϕ	0.2	0.1970	-0.0030	0.0011	0.0011
γ	0.3	0.4972	-0.0028	0.0027	0.0027
α	2	2.0029	0.0029	0.0013	0.0013
δ	0.04	0.0396	-0.0004	0.0003	0.0003

Table 33.

Parameter estimates for 100 subjects with 2000 observations in a rectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2009	0.0009	0.0029	0.0029
β_1	0.5	0.5005	0.0005	0.0059	0.0059
σ^2	4	4.0001	0.0001	0.0011	0.0011
ρ	0.5	0.4976	-0.0024	0.0007	0.0007
ϕ	0.2	0.1998	-0.0002	0.0005	0.0005
γ	0.3	0.5010	0.0010	0.0018	0.0018
α	2	2.0008	0.0008	0.0006	0.0006
δ	0.04	0.0397	-0.0003	0.0002	0.0002

Table 34.

Parameter estimates for 20 subjects with 80 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2267	0.0267	0.0178	0.0185
β_1	0.5	0.5106	0.0106	0.0486	0.0487
σ^2	4	3.9795	-0.0205	0.1369	0.1374
ρ	0.5	0.4927	-0.0073	0.0224	0.0225
ϕ	0.2	0.2009	0.0009	0.0111	0.0111
γ	0.3	0.5632	0.0632	0.1592	0.1632
α	2	1.9925	-0.0075	0.0193	0.0193
δ	0.04	0.0749	0.0349	0.0028	0.0040

Table 35.

Parameter estimates for 20 subjects with 170 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2164	0.0164	0.0118	0.0120
β_1	0.5	0.5071	0.0071	0.0267	0.0268
σ^2	4	3.9884	-0.0116	0.0342	0.0343
ρ	0.5	0.4883	-0.0117	0.0096	0.0097
ϕ	0.2	0.1944	-0.0056	0.0056	0.0056
γ	0.3	0.5193	0.0193	0.0396	0.0400
α	2	1.9994	-0.0006	0.0077	0.0077
δ	0.04	0.0555	0.0155	0.0013	0.0015

Table 36.

Parameter estimates for 20 subjects with 340 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2069	0.0069	0.0069	0.0070
β_1	0.5	0.5074	0.0074	0.0140	0.0141
σ^2	4	3.9995	-0.0005	0.0034	0.0034
ρ	0.5	0.4938	-0.0062	0.0039	0.0040
ϕ	0.2	0.1974	-0.0026	0.0026	0.0026
γ	0.3	0.5083	0.0083	0.0089	0.0090
α	2	2.0014	0.0014	0.0034	0.0034
δ	0.04	0.0471	0.0071	0.0007	0.0008

Table 37.

Parameter estimates for 50 subjects with 200 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2129	0.0129	0.0107	0.0109
β_1	0.5	0.5028	0.0028	0.0254	0.0254
σ^2	4	3.9942	-0.0058	0.0207	0.0208
ρ	0.5	0.4877	-0.0123	0.0091	0.0093
ϕ	0.2	0.1985	-0.0015	0.0056	0.0056
γ	0.3	0.5100	0.0100	0.0210	0.0211
α	2	2.0034	0.0034	0.0073	0.0073
δ	0.04	0.0544	0.0144	0.0013	0.0015

Table 38.

Parameter estimates for 50 subjects with 425 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2057	0.0057	0.0064	0.0065
β_1	0.5	0.5055	0.0055	0.0125	0.0125
σ^2	4	3.9987	-0.0013	0.0158	0.0158
ρ	0.5	0.4957	-0.0043	0.0035	0.0035
ϕ	0.2	0.1948	-0.0052	0.0026	0.0026
γ	0.3	0.5085	0.0085	0.0147	0.0148
α	2	2.0009	0.0009	0.0028	0.0028
δ	0.04	0.0436	0.0036	0.0006	0.0006

Table 39.

Parameter estimates for 50 subjects with 850 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2027	0.0027	0.0038	0.0038
β_1	0.5	0.4985	-0.0015	0.0074	0.0074
σ^2	4	3.9978	-0.0022	0.0057	0.0057
ρ	0.5	0.4980	-0.0020	0.0017	0.0017
ϕ	0.2	0.2002	0.0002	0.0013	0.0013
γ	0.3	0.5021	0.0021	0.0046	0.0046
α	2	2.0001	0.0001	0.0013	0.0013
δ	0.04	0.0404	0.0004	0.0004	0.0004

Table 40.

Parameter estimates for 100 subjects with 400 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2081	0.0081	0.0075	0.0075
β_1	0.5	0.5004	0.0004	0.0171	0.0171
σ^2	4	3.9959	-0.0041	0.0154	0.0155
ρ	0.5	0.4954	-0.0046	0.0047	0.0048
ϕ	0.2	0.1957	-0.0043	0.0029	0.0029
γ	0.3	0.5078	0.0078	0.0187	0.0188
α	2	1.9977	-0.0023	0.0034	0.0034
δ	0.04	0.0489	0.0089	0.0007	0.0008

Table 41.

Parameter estimates for 100 subjects with 850 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2050	0.0050	0.0041	0.0041
β_1	0.5	0.4966	-0.0034	0.0074	0.0074
σ^2	4	4.0017	0.0017	0.0052	0.0052
ρ	0.5	0.4972	-0.0028	0.0016	0.0016
ϕ	0.2	0.1984	-0.0016	0.0012	0.0012
γ	0.3	0.5090	0.0090	0.0088	0.0089
α	2	2.0005	0.0005	0.0014	0.0014
δ	0.04	0.0404	0.0004	0.0003	0.0003

Table 42.

Parameter estimates for 100 subjects with 1700 observations in a nonrectangular design.

Gaussian-Exponential

Parameter	True Value	Estimate	Bias	Variance	MSE
β_0	0.2	0.2017	0.0017	0.0031	0.0031
β_1	0.5	0.5007	0.0007	0.0063	0.0063
σ^2	4	3.9968	-0.0032	0.0018	0.0018
ρ	0.5	0.4980	-0.0020	0.0009	0.0009
ϕ	0.2	0.1995	-0.0005	0.0006	0.0006
γ	0.3	0.5003	0.0003	0.0039	0.0039
α	2	2.0002	0.0002	0.0007	0.0007
δ	0.04	0.0405	0.0005	0.0002	0.0002

CHAPTER V

CONCLUSION AND RECOMMENDATION

Conclusion

The primary impetus for this study was the development of an approach that could jointly model a longitudinal process with informative schedule data. In this study two proposed models were developed that demonstrated that parameter estimates could be obtained from simulated data exhibiting an informative schedule structure. For both the Vector Autoregressive and Gaussian-Exponential models, parameter estimates showed much more bias and variability when observation numbers were at the lowest levels which was not surprising. However, in almost all cases the amount of bias and variability in the estimates decreased substantially when observation numbers increased. In fact, when observation numbers were at their highest levels the amount of bias and variation in estimates for all model parameters were relatively small compared to the value of the parameter being estimated. In essences, both proposed models demonstrated large sample consistency and were asymptotically unbiased which are two desirable characteristics of any estimator (Fitzmaurice et. al., 2004).

At small observation numbers, one would expect that there would be a certain amount of non-directional variation in obtained estimates due to inefficiency in the optimization algorithm. In fact, for both models several parameters did demonstrate a

non-directional bias in obtained estimates at small observation numbers. However, this was not necessarily the case for bias in estimates for the variance parameters for both proposed models and the correlation parameter in the GE model. In these cases, the estimates obtained demonstrated that the proposed models underestimated the true parameter values slightly which suggests a common issue in estimation between both approaches. This underestimation of variance components when estimates are obtained by maximum likelihood estimation is a common problem and arises because the error in estimating the other model parameters are not being accounted for in the estimation of the variance components (Fitzmaurice et. al., 2004; Wu, Gumpertz, & Boos, 2001). To account for this bias associated with the estimation of multiple parameters many different techniques have been developed with restricted maximum likelihood (REML) estimation being one of the more common approaches. Here estimates for the variance components are determined from the relevant part of the data separate from the part that is used to estimate the other parameters and can be achieved in a number of ways (Wu et. al., 2001). One possible way to obtain the REML would require that data be transformed into a linear combination that does not depend on the other parameters and then maximize a slightly modified log-likelihood equation to obtain estimates for the variance components (Wu et. al., 2001). Since both proposed informative schedule models included several parameters that need to be simultaneously estimated, an alternative approach to estimating the variance components with the goal of reducing the amount of observed bias would be a logical future approach. However, these variance components were not the only parameters that showed some level of underestimation. In fact, there were a few parameters from both models that also demonstrated a systematic

overestimation in obtained estimates. However, it should be noted that in every case the amount of bias observed for these parameters decreased substantially as observation numbers increased, suggesting that at least part of the observed systematic bias may be due to inefficiency in estimation of the developed algorithms at small observation numbers.

When parameter efficiency was evaluated for nonrectangular samples, bias and variation in estimates showed similar trends and nearly similar values as seen in rectangular samples. In addition, estimates obtained from nonrectangular sample matrices also demonstrated large sample consistency and were asymptotically unbiased. However, in several cases the amount of variation in estimates and in a few cases the amount of bias observed was marginally larger in estimates obtained from nonrectangular sample matrices when compared to estimates obtained from rectangular sample matrices at similar observation numbers. In general, when there are missing data, there will be a level of loss of information and a reduction in the precision of obtained estimates which could account for this reduced efficiency for nonrectangular estimates (Lin & Stivers, 1975; Fitzmaurice et. al., 2004).

Finally, the comparison of common parameters with the mixed-effect approach demonstrated that both models were, in general, more efficient at estimating the true parameter values. For both models, the amount of bias observed for the mixed-effects model estimates was larger than for estimates from either informative schedule model. This underestimation of the mixed-effects model compared to the informative schedule model was more obvious in the Vector Autoregressive model. Also, estimates for both informative schedule and mixed-effects models resulted in a reduction in the observed

variation of estimates as observation numbers increased. However, this reduction in the amount of variation was more pronounced in the informative schedule models when compared to the mixed-effects model. Thus, both informative schedule models demonstrated increased relative efficiency when compared to estimates obtained from the mixed-effects model. However, it should be reminded to the reader that a single parameter was compared between the informative schedule models and the mixed-effects model. The efficiency of a particular model over another approach can not be ascertained in the evaluation of a single parameter in most cases. Although, a direct comparison of all informative schedule model parameters can not be performed since the mixed-effects model does not include estimates for many of these parameters and, in fact, is where the potential benefit of the informative schedule model resides. More precisely, the informative schedule model not only allows for the estimation of mean changes of the response variable but would also allow for the estimation of the effect that time intervals has on the obtained response outcome which would not be possible in analysis by traditional approaches.

This study, in conclusion, demonstrates that the two proposed informative schedule models were able to estimate parameters when data were simulated having informative schedule stochastic structure. The estimates obtained from informative schedule models also demonstrated that they could be estimated efficiently, especially when subject or observation numbers where large. Furthermore, this study demonstrated that efficient estimation can still be achieved even when sample matrices are unbalanced or nonrectangular. Finally, estimates for the informative model were as efficient or more efficient than estimates obtained by traditional longitudinal methods.

Recommendations for Future Researchers

The optimization algorithms developed for this study did well in estimating the Monte Carlo simulated informative schedule data constructed for each model approach. Although, there are other optimization algorithms available that may give different patterns in estimates than what was observed in this study. Therefore, one should exercise caution when applying other algorithms to the informative schedule models. Furthermore, optimization routines are susceptible to starting values and in many cases parameter estimates can be drastically different when other values are supplied (Cam, 1990; SAS Institute, 2004). Since, a single set of starting values was supplied to the subroutines other initial values might result in entirely different estimates and should therefore be considered when choosing starting values. It should also be noted that the starting values supplied in this study were close to the root of the supplied function to increase the likelihood of convergence. Other values, especially ones further from the root of the function or values not within the feasible range may result in different estimates not to mention changes in bias and variation seen in those estimates. Also, the utilized algorithm developed depends on the approximation of the Hessian matrix for both models and the approximation of the gradient vector in the GE model which would potentially result in less efficient estimations. Therefore, the determination of the Hessian matrix and gradient vector of the likelihood equations would potentially improve overall estimation efficiency. Finally, a model's utility is best demonstrated by the analysis of 'real' data which was not performed in the present study. The analysis of data that exhibits informative schedule stochastic structure and the subsequent interpretation of the obtained results would be a logical future approach.

REFERENCE

Bain L. J. & Engelhard, M. (1992). *Introduction to probability and mathematical statistics, 2nd edition.* Duxbury, California.

Bock, R. D. (1989). *Multilevel Analysis of Educational Data.* Academic Press, New York.

Cam L. L. (1990). Maximum likelihood: an introduction. *International Statistical Review*, 58: 153-171.

Cole, J. W. L. & Grizzle, J. E. (1966). Applications of multivariate analysis of variance to repeated measurements experiments. *Biometrics,* 22: 810-828.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society,* Series B, 34: 187-220.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62: 269-276.

Crowder, M. J. & Hand, D. J. (1990). *Analysis of Repeated Measures.* Chapman and Hall, London.

Davidon, W.C. (1959). Variable metric methods for minimisation. A.E.C. *Research and Development Report* ANL-5990.

Davis, C. S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. New York, NY: Springer.

Dawson, J. D. (1994). Comparing treatment groups on the basis of slopes, areas-under-the-curve, and other summary measures. *Drug Information Journal*, 28: 723-732.

Dawson, J. D. & Lagakos, S. W. (1991). Analyzing laboratory marker changes in AIDS clinical trials. *Journal of Acquired Immune Deficiency Syndromes*, 4: 667-676.

Dawson, J. D. & Lagakos, S. W. (1993). Size and power of two-sample tests of repeated measures data. *Biometrics*, 49: 1022-1032.

de Leeuw, J. & Kreft, I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics,* 11: 57-85.

Dempster, A. P., Rubin, D. B. & Tsutakawa, R. K. (1981). Estimation in covariance component models. *Journal of the American Statistical Society,* 76: 341-353.

Dennis, J. E., Gay, D. M., & Welsch, R. E. (1981). An adaptive nonlinear least-squares algorithm, *ACM Transactions on Mathematical Software,* 7: 348-368.

Diggle, P. J., Heagerty, P., Liang, K. Y. & Zeger, S. L. (2002). *Analysis of Longitudinal data, 2nd edition.* Oxford University Press, Oxford.

Elashoff, M., Li, G. & Li, N. (2007). An approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine,* 26: 2813-2835.

Everitt, B. S. (1995). The analysis of repeated measures: a practical review with examples. *The Statistician*, 44, 113-135.

Faucett, C. L. & Thomas, D. C. (1996). Simultaneously modeling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Statistics in Medicine,* 15: 1663-1685.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004) *Applied Longitudinal Analysis*. New York, NY: John Wiley & Sons, Inc.

Fletcher, R. & Powell, M.J.D. (1963). A rapidly convergent descent method for minimization. *Computer Journal,* 7: 149-154.

Frison, L. & Pocock, S. J. (1992). Repeated measures in clinical trials: Analysis using mean summary statistics and its implications for design. *Statistics in Medicine.* 11: 1685-1704.

Gibbons, R. D., Hedeker, D., Elkin, I., Waternaux, C., Kraemer, H. C., Greenhouse, J. B., Shea, M. T., Imber, S. D., Sotsky, S. M. & Watkins, J. T. (1993). Some conceptual and statistical issues in analysis of longitudinal psychiatric data. *Archives of General Psychiatry,* 50: 739-750.

Gill, E. P., Murray, W., Saunders, M. A., & Wright, M. H. (1983), Computing forward-difference intervals for numerical optimization. *SIAM J. Sci. Stat. Comput.*, 4: 310 - 321.

Goldstein, H. (1995). *Multilevel Statistical Models, 2nd edition.* Halstead Press, New York.

Haruma, K., Suzuki, T., Tsuda, T., Yoshihara, M., Sumii, K., & Kajiyama G. (1991). Evaluation of tumor growth rate in patients with early gastric carcinoma of the elevated type. *Gastrointestinal Radiology,* 16: 289-292.

Hedeker, D., & Gibbons, R. D. (2006) *Longitudinal Data Analysis*. New York, NY: John Wiley & Sons, Inc.

Heemskerk, V. H., Lentze, F., Hulsewe, K. W. E., & Hoofwijk, A. G. M. (2007). Gastric carcinoma: Review of the results of treatment in a community teaching hospital. *World Journal of Surgical Oncology,* 5: 81-87.

Henderson, R., Diggle, P. J. & Dobson, A. (2000). Joint modeling of longitudinal measurements and event time data. *Biostatistics*, 1: 465-480.

Hipel, K. W., McLeod, A. I., & Lennox, W. C. (1977) Advances in Box-Jenkins modeling 1. Model construction. *Water Resources Research,* 13: 567-575.

Hogan, J. W. & Laird, N. M., (1997a). Mixture models for the joint distributions of repeated measures and event times. *Statistics in Medicine*, 16: 239-257.

Hogan, J. W. & Laird, N. M., (1997b).Model-based approaches to analyzing incomplete longitudinal and failure time data. *Statistics in Medicine*, 16: 259-272.

Hui, S. L. & Berger, J. O. (1983). Empirical Bayes estimation of rates in longitudinal studies. *Journal of the American Statistical Association,* 78: 753-759.

Ingoldby, C.J.H., Wujanto, R. & Mitchell, J.E. (1986). Impact of vascular surgery on community mortality from ruptured aortic aneurysms. *British Journal of Surgery,* 73: 551-553.

Jemal A., Murray, T., Ward, E., Samuels, A., Tiwari, R. C., Ghafoor, A., Feuer, E. J., & Thun, M. J. (2005). Cancer statistics, 2005. *CA A Cancer Journal for Clinicians,* 55: 10-30.

Jennrich, R. I. & Schluchter, M. D. (1986). Unbalanced repeated-measures with structured covariance matrices. *Biometrics,* 42: 805-820.

Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika,* 32: 443-482.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association,* 53: 457-481.

Keselman, H. J., Alginas, J. & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematics and Statistical Psychology,* 54(1): 1-20.

Laird, N. M. & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics,* 38: 963-974.

Laird, N. M., Donnelly, C. & Ware, J. H. (1992). Longitudinal studies with continuous responses. *Statistical Methods in Medical Research,* 1: 225-247.

Lavori, P. W. (1992). Clinical trials in psychiatry: Should protocol deviation censor patient data? *Neuropsychopharmacology,* 6: 39-48.

Lederle, F. A., Wilson, S. E., Johnson G. R., Reinke, D. B., Littooy, F. N., Acher, C. W., et al. Aneurysm detection and Management Veterans Affairs Cooperative Study Group. (2002). Immediate repair compared with surveillance of small abdominal aortic aneurysms. *New England Journal of Medicine,* 346: 1437-1444.

Lin, H., Turnbull, B. W., McCulloch, C. E. & Slate, E. H. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association,* 97: 53-65.

Lin, P. E. & Stivers, L. E. (1975). Testing for equality of means with incomplete data on one variable: a Monte Carlo study. *Journal of American Statistical Association,* 70: 190-193.

Lindsey, J. (1993). *Models for Repeated Measurements.* Oxford: Oxford University Press.

Lindstrom, M. J. & Bates D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association,* 83: 1014-1022.

Littell, R. C., Henry, P. R., & Ammerman, C. B. (1998) Statistical analysis of repeated measures data using SAS procedures. *Journal of Animal Science,* 76: 1216-1231.

Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika,* 74: 817-827.

Lutkepohl, H. (1991). *Introduction to multiple time series analysis.* Springer-Verlag, Berlin.

Matthews, J. N. S., Altman, D. G., Campbell, M. J. & Royston, P. (1990) Analysis of serial measurements in medical research. *British Medical Journal,* 300: 230-235.

McCall, R. B. & Appelbaum, M. I. (1973). Bias in the analysis of repeated measures designs: some alternative approaches. *Child Development,* 44, 401-415.

Montgomery, D. C. (2005). *Design and Analysis of Experiments, 6th edition.* New York, NY: John Wiley & Sons, Inc.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models, 2nd edition.* Sage, Thousand Oaks, California.

Reinsel, G. C. (1997). *Elements of Multivariate Time Series Analysis, 2nd edition.* Springer-Verlag. New York, New York.

Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika,* 69: 331-342.

SAS Institute (2004). *SAS/IML user's guide, Version 9.1*. Cary, NC: Author.

Simon, G. E. & Savarino, J. (2008). Suicide attempts among patients starting depression treatment with medications or psychotherapy. *American Journal of Psychiatry,* 164: 1029-1034.

Student (1908). The probable error of a mean. *Biometrika,* 6: 1-25.

Taylor, M. A. & Amir, N. (1994). The problem of missing clinical data for research in psychopathology. *Journal of Nervous and Mental Disease,* 182: 222-229.

Tseng, Y. K., Hsieh, F. & Wang, J. L. (2005). Joint modeling of accelerated failure time and longitudinal data. *Biometrika,* 92(3): 587-603.

Tsiatis, A. A., DeGruttola, V. & Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Association,* 90: 27-37.

Wang, Y. & Taylor, J. M. G. (2001). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association,* 96: 895-905.

Winer, B. J. (1971). *Statistical Principles in Experimental Design, 2nd edition*. McGraw-Hill, New York.

Wu, C. T., Gumpertz, M. L. & Boos, D. D. (2001) Comparison of GEE, MINQUE, ML, and REML estimating equations for normally distributed data. *The American Statistician,* 55: 125-130.

Wu, M. C. & Carroll, R. J. (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics,* 44: 175-188.

Wulfsohn, M. S. & Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics,* 53:330-339.

APPENDIX A

GRADIENT DERIVATIVES FOR GAUSSIAN-EXPONENTIAL INFORMATIVE MODELS

For the second special case of the informative schedule model, the Gaussian-Exponential derivatives are given below. The derivative for the eight parameters can be summarized as:

$$\frac{dLog(L)}{d\beta_0} = \sum_{i=1}^{m}\left(\left(y_{i1} - \mathbf{X}_{i1}\boldsymbol{\beta}\right)\cdot\frac{\mathbf{X}_{i2}}{\sigma^2}\right) + \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\left(y_{ij} - t_{ij}\gamma - y_{ij-1}\phi - \mathbf{X}_{ij}\boldsymbol{\beta}\right)\cdot\frac{\mathbf{X}_{i2}}{\sigma^2\left(1-\rho^2\right)}\right)$$

$$\frac{dLog(L)}{d\beta_1} = \sum_{i=1}^{m}\left(\left(y_{i1} - \mathbf{X}_{i1}\boldsymbol{\beta}\right)\cdot\frac{\mathbf{X}_{i1}}{\sigma^2}\right) + \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\left(y_{ij} - t_{ij}\gamma - y_{ij-1}\phi - \mathbf{X}_{ij}\boldsymbol{\beta}\right)\cdot\frac{\mathbf{X}_{i1}}{\sigma^2\left(1-\rho^2\right)}\right)$$

$$\frac{dLog(L)}{d\sigma^2} = \frac{1}{2}\cdot\sum_{i=1}^{m}\left(\frac{\left(y_{i1} - \mathbf{X}_{i1}\boldsymbol{\beta}\right)^2}{\left(\sigma^2\right)^2}\right) + \frac{1}{2}\cdot\sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\frac{\left(y_{ij} - t_{ij}\gamma - y_{ij-1}\phi - \mathbf{X}_{ij}\boldsymbol{\beta}\right)^2}{\left(\sigma^2\right)^2\left(1-\rho^2\right)}\right) - \frac{\sum_{i=1}^{m} n_i}{2\sigma^2}$$

$$\frac{dLog(L)}{d\rho} = \frac{\sum_{i=1}^{m}\left(n_i - 1\right)\rho}{\left(1-\rho^2\right)} - \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\left(y_{ij} - t_{ij}\gamma - y_{ij-1}\phi - \mathbf{X}_{ij}\boldsymbol{\beta}\right)^2\cdot\frac{\rho}{\sigma^2(1-\rho^2)^2}\right)$$

$$\frac{dLog(L)}{d\gamma} = \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\left(y_{ij} - t_{ij}\gamma - y_{ij-1}\phi - \mathbf{X}_{ij}\boldsymbol{\beta}\right)\cdot\frac{t_{ij}}{\sigma^2\left(1-\rho^2\right)}\right)$$

$$\frac{dLog(L)}{d\phi} = \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\left(y_{ij} - t_{ij}\gamma - y_{ij-1}\phi - \mathbf{X}_{ij}\boldsymbol{\beta}\right)\cdot\frac{y_{ij-1}}{\sigma^2\left(1-\rho^2\right)}\right)$$

$$\frac{dLog(L)}{d\alpha} = \sum_{i=1}^{m}\left(n_i - 1\right) - \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(\exp\left(\alpha + y_{ij-1}\delta\right)\cdot t_{ij}\right)$$

$$\frac{dLog(L)}{d\delta} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij} - \sum_{i=1}^{m}\sum_{j=2}^{n_i}\left(y_{ij-1}\exp\left(\alpha + y_{ij-1}\delta\right)\cdot t_{ij}\right)$$

APPENDIX B

SAS CODE FOR VECTOR AUTOREGRESSIVE AND GAUSSIAN-EXPONENTIAL INFORMATIVE MODELS

Vector Autoregressive Function Call.

```
start  logl(x)  global(y, xmatrix, nobs, m);
nn=cusum(nobs);
opt={1};
beta=J(2,2,.);
beta[1,1]=x[1];
beta[1,2]=x[2];
beta[2,1]=x[3];
beta[2,2]=x[4];

sigma=J(2,2,.);
sigma[1,1]=x[5];
sigma[1,2]=x[6];
sigma[2,1]=x[6];
sigma[2,2]=x[7];
ss=det(sigma);

Phi=J(2,2,.);
Phi[1,1]=x[8];
Phi[1,2]=x[9];
Phi[2,1]=x[10];
Phi[2,2]=x[11];

index=1:nn[1];
mu=xmatrix[1,]*beta;
w1=y[index,1]-mu[ ,1];
w2=y[index,2]-mu[ ,2];
w= w1||w2;
call varmalik(lnl,w,phi, ,sigma,,,opt);
fun=lnl[1];

do k= 2 to m;

lb=nn[k-1]+1;
index=lb:nn[k];
mu=xmatrix[k,]*beta;

w1=y[index,1]-mu[ ,1];
   w2=y[index,2]-mu[ ,2];
w= w1||w2;

call varmalik(lnl,w,phi, ,sigma,,,opt);
      fun=fun+lnl[1];
end;

return(fun);
finish logl;

x0={4, 2, 3, 1, 4, 0.1, 2, 0.8, 0.3, 0.2, 0.5};

optn = {1 0 . 1 . . . 11};


call nlpdd (rc, xres, "logl", x0, optn,,,);
```

Gaussian-Exponential Function Call.

```
start maxlike(x) global(y, xmatrix, obsvec);
   m=ncol(y)/2;
   observ=m/2;
   j=1;
   fun=0;fun1=0;fun2=0;fun3=0;


   do k = 1 to m;
         if k <= observ then half=obsvec[1,1];
               else half=obsvec[2,1];

   f1 = -0.5*log(x[1])-0.5*(((y[1,j]-(xmatrix[1,k]*x[2])-
          (xmatrix[1,k+1]*x[3]))**2)/x[1]);

   fun1=fun1+f1;

   do i = 2 to half;

   f2=-0.5*log(x[1])-0.5*log(1-(x[4]**2))-0.5*(((y[i,j]-
         (y[i,j+1]*x[5])-(y[i-1,j]*x[6])-(xmatrix[1,k]*x[2])-
         (xmatrix[1,k+2]*x[3]))**2)/(x[1]*(1-(x[4]**2))));

   f3 =(x[7]+(y[i-1,j]*x[8])-exp(x[7]+(y[i-1,j]*x[8]))*y[i,j+1]);

               fun2=fun2+f2;
               fun3=fun3+f3;
         end;
   j=j+2;
   end;

   fun = m+fun1+fun2+fun3;

return(fun);
finish maxlike;
```

Gaussian-Exponential Gradient Call.

```
start maxlike(x) global(y, xmatrix, nvector, maxsub);

   i=1;tic=0;
   fun=0;fun1=0;fun2=0;fun3=0;

   do k = 1 to maxsub;

         mu=(xmatrix[1,i]*x[2])+(xmatrix[1,i+1]*x[3]);
         f1=-0.5*log(x[1])-0.5*(((y[tic+1,1]-mu)**2)/x[1]);

         fun1=fun1+f1;

         do j = 2 to nvector[1,k];
```

```
   f2=-0.5*log(x[1])-0.5*log(1-(x[4]**2))-0.5*(((y[tic+j,1]-
(y[tic+j,2]*x[6])-(y[tic+j-1,1]*x[5])-mu)**2)/(x[1]*(1-(x[4]**2))));
   f3=(x[7]+(y[tic+j-1,1]*x[8])-exp(x[7]+y[tic+j-
1,1]*x[8])*y[tic+j,2]);

                fun2=fun2+f2;
                fun3=fun3+f3;
          end;
          i=i+2;tic=tic+nvector[1,k];
   end;

   fun=maxsub+fun1+fun2+fun3;

return(fun);
finish maxlike;

start gradient(x) global(y, xmatrix, nvector, maxsub);

   i=1; tic=0;
   sigma1=0;sigma2=0;beta1a=0;beta2a=0;beta1a=0;beta1b=0;beta2b=0;rho=0
;gamma=0;phi=0;alpha=0;delta=0;sumy1=0;sumy2=0;sumni=0;summinus=0;
   g=j(1,8,.);

   do k =1 to maxsub;

                sumni=sumni+nvector[1,k];
                summinus=summinus+(nvector[1,k]-1);
                sumy1=sumy1+y[tic+1,1];

          mu=(xmatrix[1,i]*x[2])+(xmatrix[1,i+1]*x[3]);
          yi1=y[tic+1,1];

          sig1=((yi1-mu)**2)/(x[1]**2);
          bet1=(yi1-mu)*(xmatrix[1,i+1]/x[1]);
          bet2=(yi1-mu)*(xmatrix[1,i]/x[1]);

          sigma1=sigma1+sig1;
          beta1a=beta1a+bet1;
          beta2a=beta2a+bet2;

          do j = 2 to nvector[1,k];

                yij=y[tic+j,1];
                tij=y[tic+j,2];
                yijm=y[tic+j-1,1];

                rij=(yij-(tij*x[6])-(yijm*x[5])-mu);

                sig2=(rij**2)/((x[1]**2)*(1-(x[4]**2)));
                bet1b=rij*(xmatrix[1,i+1]/(x[1]*(1-(x[4]**2))));
                bet2b=rij*(xmatrix[1,i]/(x[1]*(1-(x[4]**2))));
                rho2=(rij**2)*(x[4]/(x[1]*((1-(x[4]**2))**2)));
                gam2=rij*(tij/(x[1]*(1-(x[4]**2))));
                phi2=rij*(yijm/(x[1]*(1-(x[4]**2))));
                alp2=(exp(x[7]+(yijm*x[8]))*tij);
                del2=(yijm*exp(x[7]+(yijm*x[8]))*tij);
```

```
                sumy2=sumy2+y[tic+j,1];

                sigma2=sigma2+sig2;
                beta1b=beta1b+bet1b;
                beta2b=beta2b+bet2b;

                rho=rho+rho2;
                gamma=gamma+gam2;
                phi=phi+phi2;

                alpha=alpha+alp2;
                delta=delta+del2;

          end;
          i=i+2;tic=tic+nvector[1,k];

          sumy2=sumy2-y[tic,1];

   end;

   sumy=sumy1+sumy2;

   g[1] = (0.5*sigma1)+(0.5*sigma2)-(sumni/(2*x[1]));
   g[2] = beta1a+beta1b;
   g[3] = beta2a+beta2b;
   g[4] =((summinus*x[4])/(1-(x[4]**2)))-rho;
   g[5] = phi;
   g[6] = gamma;
   g[7] = summinus-alpha;
   g[8] = sumy-delta;

return (g);
finish gradient;

x0 = {4, 0.2, 0.5, 0.5, 0.2, 0.5, 1, 0.04};

optn = {1 0 . 1 . . . 11};

con=j(1,8,.0000001)//j(1,8,.);

call nlpdd (rc, xres, "maxlike", x0, optn, con) grd="gradient";
```

APPENDIX C

MAPLE CODE FOR GAUSSIAN-EXPONENTIAL DERIVATIVES

```
> with(linalg):
> with(codegen, makeproc):
> p := 2 :
> m := 3 :
> n := 3 :
> beta:=vector[col](p);
```

$$\beta := array(1..2, [\,])$$

```
> X:=matrix(m,n);
```

$$X := array(1..3, 1..3, [\,])$$

```
> y := matrix(m, n);
```

$$y := array(1..3, 1..3, [\,])$$

```
> t := matrix(m, n);
```

$$t := array(1..3, 1..3, [\,])$$

```
>
```

$$L1 := \mathbf{proc}(\sigma, \beta, \rho, \kappa, \phi, \alpha, \delta) \quad -.5 \cdot sum\left(\ln(\sigma) + \frac{(y[i,1] - X[i,1]\cdot\beta[1] - X[i,2]\cdot\beta[2])^2}{\sigma}, i = 1..m\right) - .5 \cdot sum\left(sum\left(\ln(\sigma) + \ln(1-\rho^2) + \frac{1}{\sigma\cdot(1-\rho^2)}\left((y[i, j] - t[i,j]\cdot\kappa - y[i,j-1]\cdot\phi - X[i,1]\cdot\beta[1] - X[i,2]\cdot\beta[2])^2\right), j = 2..n\right), i = 1..m\right) + sum(sum(\alpha + \delta\cdot y[i,j-1] - \exp(\alpha + \delta \cdot y[i,j-1])\cdot t[i,j], j = 2..n), i = 1..m); \mathbf{end};$$

```
> l1sigma := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), σ), [σ, β, ρ, κ, ϕ,
      α, δ]);
> l1beta1 := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), β[1]), [σ, β, ρ, κ,
      ϕ, α, δ]);
> l1beta2 := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), β[2]), [σ, β, ρ, κ,
      ϕ, α, δ]);
> l1rho := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), ρ), [σ, β, ρ, κ, ϕ, α,
      δ]);
> l1other := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), κ), [σ, β, ρ, κ, ϕ,
      α, δ]);
> l1phi := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), ϕ), [σ, β, ρ, κ, ϕ, α,
      δ]);
> l1alpha := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), α), [σ, β, ρ, κ, ϕ,
      α, δ]);
> l1delta := makeproc (diff (L1(σ, β, ρ, κ, ϕ, α, δ), δ), [σ, β, ρ, κ, ϕ, α,
      δ]);
```