Correlated GMM Logistic Regression Models with Time-Dependent Covariates and Valid Estimating Equations

by

Jianqiong Yin

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2012 by the
Graduate Supervisory Committee:

Jeffrey Wilson, Chair
Ming-Hung Kao
Mark Reiser

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

When analyzing longitudinal data it is essential to account both for the correlation inherent from the repeated measures of the responses as well as the correlation realized on account of the feedback created between the responses at a particular time and the predictors at other times. A generalized method of moments (GMM) for estimating the coefficients in longitudinal data is presented. The appropriate and valid estimating equations associated with the time-dependent covariates are identified, thus providing substantial gains in efficiency over generalized estimating equations (GEE) with the independent working correlation. Identifying the estimating equations for computation is of utmost importance. This paper provides a technique for identifying the relevant estimating equations through a general method of moments. I develop an approach that makes use of all the valid estimating equations necessary with each time-dependent and time-independent covariate. Moreover, my approach does not assume that feedback is always present over time, or present at the same degree. I fit the GMM correlated logistic regression model in SAS with PROC IML. I examine two datasets for illustrative purposes. I look at rehospitalization in a Medicare database. I revisit data regarding the relationship between the body mass index and future morbidity among children in the Philippines. These datasets allow us to compare my results with some earlier methods of analyses.

ACKNOWLEDGMENTS

I would like to thank my committee members, Dr. Mark Reiser and Dr. Ming-Hung (Jason) Kao, for their excellent instruction in my coursework, their contributions to my thesis, and their support throughout my education at Arizona State University.

In particular I would like to express my greatest gratitude to my advisor, Dr. Jeffrey Wilson, for his time, his patience, his guidance and his support. This thesis would never have been written if not for his guidance. He is more than an advisor to me. I cannot thank him enough for everything he has done for me.

TABLE OF CONTENTS

Page

APPENDIX Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

GENERALIZED METHOD OF MOMENTS

In the analysis of marginal models for longitudinal continuous data, Lai and Small (2007) use a generalized method of moments (GMM) approach, which makes optimal use of the information provided by time-dependent covariates, when obtaining estimates. In their approach, the time-dependent covariates are classified into one of three types: I, II, and III. The time-independent covariate is treated as type III. Each type of covariate requires a different set of estimating equations to be used in finding the corresponding coefficient's estimate. They compare their estimates to the generalized estimating equations (GEE) with the independent working correlation structure. They find that their GMM approach provides substantial gains in efficiency over the GEE if the covariates are type I or type II, and still remain consistent and comparable in efficiency when the covariates are type III. Thus, it is clear through omission or inclusion of estimating equations that the conclusions we make about covariates affecting our responses over time can vary.

However, I present a method somewhat different from Lai and Small (2007) and for binary data though I show how it is applicable to other types of data. I postulate that there are more than three types of covariates. I argue that there can be theoretically more than three types of covariates and as such I present an extended method that will best describe the covariates before I proceed to model the binary outcomes. Communications with Dr. Small has confirmed that there can be other classifications. I compare my results with existing methods for classification.

MAKING USE OF VALID ESTIMATING EQUATIONS

This paper concentrates on using valid and appropriate estimating equations for time-dependent covariates while determining their impact on the responses over time. In particular, I provide a method to choose valid equations in determining the effect of time-dependent covariates on binary responses. I provide some insights into fitting models without having to classify the covariates as belonging to a particular type. In particular, I fit logistic regression models with time-dependent covariates using GMM estimates in SAS with PROC IML. I also provide the necessary steps if one decides to fit a continuous response.

Lai and Small (2007) look at conditional normal models and compute the necessary results in R; I look at binary models and conduct my computations in SAS with PROC IML. In addition, the GMM models are compared under conditions where we allow certain covariates to have estimating equations not valid in method of moments procedures. Chapter 2 reviews longitudinal models in light of the class of generalized linear models (GLM), the class of GEE models with independent working correlation, and the class of GMM models for binary responses. A new method for choosing valid equations is presented and discussed in Chapter 3. For illustrative and comparative purposes, I analyze binary data pertaining to rehospitalization and revisit data pertaining to predicting morbidity among children in the Bukidnon region in the Philippines. I fit GMM models in SAS with PROC IML in Chapter 4. Some conclusions are made in Chapter 5.

CHAPTER 2

LONGITUDINAL MODELS

LONGITUDINAL STUDIES

Longitudinal studies address among other things, how each unit changes over time; and what determines the differences among units in their change over time. Longitudinal data often contain repeated measurements of units at multiple time points. Such correlated observations are commonly encountered in studies in healthcare, polling, marketing and other types of behavioral research.

One major advantage of a longitudinal study is its capacity to separate change over time within unit and differences among units (Diggle, Heagerty, Liang and Zeger, 2002). However, when dealing with longitudinal data not only do the response variables change over time, but the predictors or covariates can also change over time. Thus the treatment of time-dependent covariates in the analysis of longitudinal data allows strong statistical inferences about dynamic relationships and provides more efficient estimators than can be obtained using cross-sectional data (Hedeker and Gibbons, 2006).

The generalized linear models (GLM) are inappropriate in analyzing longitudinal data due to the clustering, which results in non-independence thereby leading to overdispersion. The presence of such overdispersion or extravariation when fitting marginal regression models has shown to be best modeled through the use of GEE (Liang and Zeger, 1986; Zeger and Liang, 1986). However, when fitting GLM and GEE models it is assumed that the covariates are time-independent. Thus, neither the GLM nor GEE models takes the inherent correlation into account due to the fact that the covariates are time-dependent.

GENERALIZED LINEAR MODELS

Nelder and Wedderburn (1972), through the recognition of the “nice” properties of the normal distribution, present a wider class of distributions, the exponential family of distributions. For such cases they extend the numerical methods to estimate the vector of parameters $\boldsymbol{\beta}$ from the linear model to the situation where there is some non-linear function $\mathrm{g}(\mu) = \mathbf{X}\boldsymbol{\beta}$, where g is the link function, a monotone, twice-differentiable function, μ is the mean vector for the response vector Y and **X** are the data matrix of explanatory variables (McCullagh and Nelder, 1989). These models have now been further generalized to situations where the functions may be estimated numerically; and such is the case with generalized additive models (Hastie and Tibshirani, 1990). However, my interest is in correlated observations measured over time with or without feedback.

GENERALIZED ESTIMATING EQUATIONS

The analyses of longitudinal data with marginal models, and more generally, of correlated response data have received considerable attention in Zeger and Liang (1992) among others. Marginal models are appropriate when inferences about the population average are our primary interest (Diggle et al., 2002) or when we require the expectation of the response variable to be a function of current covariates in order to make future applications of the results (Pepe and Anderson, 1994).

For unit i, let $\mathbf{y}_{\mathrm{i}} = (\mathrm{y}_{\mathrm{i}1}, \dots, \mathrm{y}_{\mathrm{i}\mathrm{T}})'$ be a $\mathrm{T} \times 1$ vector of outcomes associated with matrix $\mathbf{X}_{\mathrm{i}} = \begin{bmatrix} \mathrm{x}_{\mathrm{i}11} & \cdots & \mathrm{x}_{\mathrm{i}1\mathrm{J}} \\ \vdots & \ddots & \vdots \\ \mathrm{x}_{\mathrm{i}\mathrm{T}1} & \cdots & \mathrm{x}_{\mathrm{i}\mathrm{T}\mathrm{J}} \end{bmatrix}$, where at time t the row vector, $\mathbf{x}_{\mathrm{it.}} = \left(\mathrm{x}_{\mathrm{it}1}, \dots, \mathrm{x}_{\mathrm{itJ}}\right)$ and for the j^{th} covariate the column vector $\mathbf{x}_{\mathrm{i.j}} = \left(\mathrm{x}_{\mathrm{i}1\mathrm{j}}, \dots, \mathrm{x}_{\mathrm{iTj}}\right)'$ such that $\mathrm{t} = 1, \dots, \mathrm{T}$; and $\mathrm{j} = 1, \dots, \mathrm{J}$. The observation times and correlation matrix may differ from subject to subject, but the structure for the form of the correlation matrix among the T observations, $\mathbf{R}_{\mathrm{i}}(\alpha)$

for the i^{th} subject, is fully specified by $\boldsymbol{\alpha}$. A valuable feature of modeling correlation with the GEE approach is that it accounts for the $\mathrm{s} \times 1$ parameter vector $\boldsymbol{\alpha}$, $\mathrm{s} \leq \mathrm{t}$. Liang and Zeger (1986) show that when $\mathbf{R}_{\mathrm{i}}(\boldsymbol{\alpha}) = \mathbf{I}$, the GEE estimating equations can be simplified to the score functions as from a likelihood analysis that assumes independence among repeated observations from a subject. The GEE estimates for $\boldsymbol{\beta}$ are consistent regardless of the choice of working correlation structure for time-independent covariates, although a correct specification of the working correlation structure does enhance efficiency. Further, the GEE method allows the user to specify any working correlation structure for a subject's outcomes $\mathbf{y}_{\mathrm{i}}$ such that its variance

$$\mathbf{V}_{\mathbf{i}}(\boldsymbol{\alpha}) = \mathbf{A}_{\mathbf{i}}^{\mathbf{1/2}}\mathbf{R}_{\mathbf{i}}(\boldsymbol{\alpha})\mathbf{A}_{\mathbf{i}}^{\mathbf{1/2}}\boldsymbol{\phi},$$

where $\mathbf{A}_{\mathrm{i}}$ is a diagonal matrix representing the variance under the assumption of independence. Thus the generalized estimating equations over N subjects

$$\mathbf{U}(\boldsymbol{\beta}) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left(\frac{\partial \boldsymbol{\mu}_{\mathrm{i}}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)^{\mathrm{T}} \mathbf{V}_{\mathrm{i}}^{-1}\{\mathbf{y}_{\mathrm{i}} - \boldsymbol{\mu}_{\mathrm{i}}(\boldsymbol{\beta})\} = \mathbf{0},$$

from which the parameter estimates are obtained. Liang and Zeger (1986) establish that the vector $\widehat{\boldsymbol{\beta}}$ that satisfies $\mathbf{U}(\widehat{\boldsymbol{\beta}}) = \mathbf{0}$ is asymptotically unbiased in the sense that $\lim_{\mathrm{N}\to\infty}\left[\mathbf{E}_{\boldsymbol{\beta}_{\mathbf{0}}}\{\mathbf{U}(\boldsymbol{\beta}_{\mathbf{0}})\}\right] = \mathbf{0}$, under suitable regularity conditions. Diggle et al. (2002) show that the GEE approach is usually satisfactory when the data consist of short, essentially complete, sequences of measurements observed at a common set of times on many experimental units, and a conservative selection in the choice of a working correlation matrix.

However, consistency may not hold for arbitrary working correlation structures if the covariates are time-dependent (Pepe and Anderson, 1994). Dobson (2002) argues that it is necessary to choose a correlation structure likely to reflect the relationships between the observations. In any case the correlation parameters are usually not of particular interest and are often seen as nuisance parameters, although they must be included in the model to obtain consistent estimates of the vector **β** of parameters and their standard errors. Nevertheless, it has been shown that when there are time-dependent covariates, GMM is an alternative and an even better choice.

GENERALIZED METHOD OF MOMENTS

When time-dependent covariates are present GMM provides more efficient estimators than using the GEE estimators based on the independent working correlation under certain conditions (Lai and Small, 2007). However, they show through a simulation study that when there are time-dependent covariates, some of the estimating equations combined by the GEE method with an arbitrary working correlation structure are not valid. They maintain that the GEE approach with time-independent covariates is an attractive approach as it provides consistent estimates under all correlation structures for subjects' repeated measurements. More so the GEE estimates with time-independent covariates produce efficient estimates if the working correlation structure is correctly specified and remain consistent as well as providing correct standard errors if the working correlation structure is incorrectly specified.

In particular, when there are time-dependent covariates, Hu (1993) and Pepe and Anderson (1994) have pointed out that the consistency of GEEs is not assured with arbitrary working correlation structures unless a key assumption is satisfied. However, the consistency is assured regardless of the validity of the key assumption if a subject's repeated measurements are independent (the independent working correlation) is employed. Pepe and Anderson (1994) suggest the use of the independent working correlation when using GEE with time-dependent covariates as a "safe" choice of analysis.

In this paper first I fit binary models with GMM estimates using PROC IML in SAS. I also consider a continuous response situation. I present a procedure of first identifying the estimating equations associated with each time-dependent covariate. Further I define a GMM estimate $\boldsymbol{\beta}_{\mathrm{GMM}}$ that uses all valid estimating equations. I look at bivariate correlations to determine the equations to use regarding each covariate. I show that incorrectly specifying the type of covariate may result in significant changes in the standard errors and thereby lead to erroneous conclusions. Therefore, I opt for entering the valid estimating equations rather than designating a covariate of a certain type (i.e. a set of estimating equations). I fit logistic regression models and normal regression models with different types of time-dependent covariates.

CHAPTER 3

IDENTIFYING VALID EQUATIONS

In this chapter, I consider GMM estimators that take advantage of the appropriate and valid estimating equations to produce consistent and more efficient estimators as opposed to the class of GEE estimators. In order to obtain such GMM estimates, we need to know which of these estimating equations, also called moment conditions, are valid. I recall the procedures of Lai and Small (2007) and then present an extended approach. In so doing my procedures are based on an examination of each of the estimating equations for $s \neq t$.

THREE-TYPE CLASSIFICATION

Suppose that we have repeated observations taken over T times on N subjects with J covariates such that $(y_{it}, \mathbf{x}_{it.})$ for subjects $i = 1, \dots, N$; for covariates $j = 1, \dots, J$; and times $t = 1, \dots, T$; where y_{it} denotes the observation for subject i at time t, whose marginal distribution given the time-varying vector $\mathbf{x}_{it.}$ of covariates follows a generalized linear model. We assume that observations y_{is} and y_{kt} are independent whenever $i \neq k$ but not necessarily when $i = k$ and $s \neq t$. To obtain GMM estimates, we need to make use of the estimating equations (Fitzmaurice, 1995; Zeger, Liang and Albert, 1988)

$$E\left[\frac{\partial \mu_{is}(\boldsymbol{\beta})}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta})\}\right] = 0, \tag{1}$$

for appropriately chosen s, t, and j, where $\mu_{it}(\boldsymbol{\beta})$ denotes expectation of y_{it} based on the vector of covariate values, $\mathbf{x}_{it.}$ where $\boldsymbol{\beta}$ is the vector of parameters in the systematic component that describes the marginal distribution of y_{it}. However, for certain types of time-dependent covariates, there are valid estimating equations available that are not exploited or considered by the usual GEE estimators. The assumption that the marginal distribution of y_{it} given $\mathbf{x}_{it.}$ ensures that equation (1) holds for $t = s$, and obviously holds for all t and s when x_{isj} does not vary with time. In such a case we would use the GEE method. However, when $t \neq s$ and x_{isj} varies with time, this is not the case. A sufficient condition for equation (1) to hold is $E(y_{it}|\mathbf{x}_{is.}, \mathbf{x}_{it.}) = E(y_{it}|\mathbf{x}_{it.})$ (Lai and Small, 2007). When this equation fails, it is typically for one or both of the reasons; there is a time-series effect which causes early covariate vectors $\mathbf{x}_{is.}$ to affect the expectations of later observations y_{it}, or early responses y_{it} have an effect on later covariate vectors $\mathbf{x}_{is.}$, which means that knowing the value of $\mathbf{x}_{is.}$ gives us some information about the value of y_{it}.

To identify valid moment conditions, Lai and Small (2007) introduce the notion of classification of time-dependent covariates into types I, II, or III. The j^{th} covariate is said to be type I if equation (1) holds for all s and t. An obvious situation in which this occurs (for all covariates) is when the y_{it} are all independent. Another relatively straightforward case is when the differences between individuals' observations can be modeled via the introduction of random effects covariates into the generalized linear model. Type I covariates plausibly satisfy a condition that their outcomes are independent of past and future outcomes of the response. A sufficient condition for covariate $\mathbf{x}_{i.j}$ in a linear model to be type I is that

$$f(x_{i1j}, \dots, x_{iTj}|y_{it}, \mathbf{x}_{it.}) = f(x_{i1j}, \dots, x_{iTj}|\mathbf{x}_{it.}),$$

so that $\mathbf{x}_{i.j}$ satisfies

$$E_{\boldsymbol{\beta}_0}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta}_0)\}\right] = 0,$$

for all s, $t = 1, \dots, T$. In my analysis there will be T^2 estimating equations valid for each type I covariate.

The j^{th} covariate is said to be type II if equation (1) holds whenever $s \geq t$, but fails to hold for some $s < t$. This is the case in many, although not all, time-series models. If the expectation of y_{it} depends directly on previous values of y_{is}, which causes the underlying previous factors $\mathbf{x}_{is.}$ at time s to affect the expected value of y_{it} at time at time t, $t > s$, then a covariate will be type II. This occurs, for instance, in a linear model with autoregressive responses. Although not all time-series models result in type II covariates, it is the time-series nature of the data that would lead us to suspect that equation (1) might fail for $s < t$. Type II covariates plausibly satisfy a condition that their future outcomes are independent of previous outcomes of the response, i.e. there is no feed-back from the response process to the covariate process. A sufficient condition for a covariate $\mathbf{x}_{i.j}$ in a linear model to be type II is that

$$f\left(x_{i(t+1)j}, \dots, x_{iTj} \middle| y_{it}, \mathbf{x}_{it.}\right) = f\left(x_{i(t+1)j}, \dots, x_{iTj} \middle| \mathbf{x}_{it.}\right),$$

so that $\mathbf{x}_{i.j}$ satisfies

$$E_{\boldsymbol{\beta_0}}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right] = 0,$$

for all $s \geq t$, $t = 1, \dots, T$. In my analysis there will be $\frac{T(T+1)}{2}$ estimating equations valid for each type II covariate.

It is not straightforward to distinguish between types I and II covariates. Obviously, if the observations are independent or if the dependence between observations is due to a random-effects-type error term, as in the linear model $y_{it} = \boldsymbol{\beta}\mathbf{x}_{it.} + u_i + \varepsilon_{it}$ where ε_{it} and u_i are independent with zero mean and constant variance, the covariates will be type I. A random effects model will not generally apply when t indexes time. But the GMM approach could be used to analyze clustered outcomes with covariates differing within clusters whether or not t indexes time, and if not, a random effects model and all-type I covariates might make sense.

The j^{th} covariate is said to be type III if equation (1) fails to hold for any $s > t$. This will not occur if the x_{isj} are deterministic or are determined by a random process exogenous to the y_{it}. However, it can occur if the x_{isj} have a random distribution that is not independent of previous values of y_{it}; that is, if there is some feedback loop or common response to an omitted variable. A covariate $\mathbf{x}_{i,j}$ is said to be type III if it is not type II, i.e.

$$E_{\boldsymbol{\beta_0}}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right] \neq 0,$$

for some $s > t$. In my analysis there will be T estimating equations valid for each type III covariate.

The easiest decision to make on the basis of an expert's prior knowledge of the field is whether we should suspect that a covariate is type III. A deterministic or exogenous covariate cannot be type III. The only time we should be concerned that a covariate might be type III is when it changes randomly and we suspect that its distribution may depend on past values of the response. There are a number of real-world situations in which we would expect some covariates to be type III. In finance, for instance, we would expect a firm's stock performance to depend on its bond rating, so a marginal regression model of stock price would probably include the bond rating as a covariate. But the firm's bond rating will also depend on the past performance of its stock, creating feedback. Similar situations occur in health data, where many measurable variables are interrelated. An individual's likelihood of developing heart problems, for instance, depends heavily on the amount of exercise the individual gets. But individuals with poor heart health are less likely to exercise adequately, which, in turn, is likely to further worsen their heart health.

Pan and Connett (2002) develop a predictive mean-squared error approach for choosing among the class of usual GEE estimators when there are time-dependent covariates. Lai and Small (2007) provide a test which is useful when a researcher has a strong prior belief that the moment conditions are all valid and is using the test to see whether there is any evidence in the data against this belief. When a researcher has more uncertainty about the validity of some of the moment conditions the predictive mean-square error approach is useful. They argue that unless there are substantive reasons to think that a time-dependent covariate is type I or type II, we assume that it is type III. If there are substantive reasons to think that a covariate is type I (or type II), then test the null hypothesis that it is type I (or type II) versus the alternative that it is type III and, if the test is not rejected, we use the moment conditions in our GMM estimator. The GMM moment selection estimator gains efficiency for type I and type II covariates compared with GEEs with the independent working correlation when our hypothesis that a covariate is type I or type II is correct.

UNGROUPED ESTIMATING EQUATION MODELS

Lalonde and Wilson (2010) suggest that the challenge with the test for classification is that it requires us to test all of the moment conditions at once, and with J time-varying covariates, we are faced with 3^J possible choices of moment conditions. It is unrealistic to compare all of them, both because this requires a lot of computation and because of the usual problem with multiple comparison testing. They propose that since we do not expect that type I covariates will, in practice, coexist with other covariate types, then first test all of the covariates being type I against some other default choice. Alternatively we could ignore the possibility of type I at first and, as Lai and Small (2007) suggest, use expert advice and hypothesis tests to find the best model in terms of just type II and type III, and then test the all type I model against this. This has the advantage of being conservative, but it has the disadvantage that, if everything is type I, a lot of time was invested unnecessarily. Further testing all-type I against an expert-advice-based default choice of type II for everything that we think will not display feedback and type III for anything that we think might. This might be the best option in practice.

Further, Lalonde and Wilson (2010) suggest testing all type I against all type II. In practice this may be questionable, because if some are type II, then we are doing a goodness-of-fit test for two wrong models, and we are in doubt to be certain that we can expect the test to reject the "least desirable" model. We could also test all type I versus all type III. This test should provide useful results since the type III estimating equations will be appropriate for all covariates. However, one may wonder what will be the detrimental effect on the power of our test. If such is the case we would wind up rejecting the type I hypothesis too often. In this paper, I forgo these approaches and present a method based on correlation to determine the valid estimating equations.

I posit that in the classification of covariates based on Lai and Small (2007) there are other possible cases or types. For example we may refer to a type IV covariate. This would be in direct contrast to type II but completes the possible groupings. Thus, I classify a time-dependent covariate $\mathbf{x}_{\mathrm{i.j}}$ as being type IV if the future responses are not affected by the previous covariate process. There is no feed-back from the covariate process to the response process. A sufficient condition for a covariate $\mathbf{x}_{\mathbf{i.j}}$ in a linear model to be type IV is that

$$\mathrm{f}\left(\mathrm{x}_{\mathrm{i1j}}, \ldots, \mathrm{x}_{\mathrm{i(t-1)j}} \middle| \mathrm{y}_{\mathrm{it}}, \mathbf{x}_{\mathrm{it.}}\right) = \mathrm{f}\left(\mathrm{x}_{\mathrm{i1j}}, \ldots, \mathrm{x}_{\mathrm{i(t-1)j}} \middle| \mathbf{x}_{\mathrm{it.}}\right),$$

so that $\mathbf{x}_{\mathrm{i.j}}$ satisfies

$$\mathrm{E}_{\boldsymbol{\beta}_{\mathbf{0}}}\left[\frac{\partial \mu_{\mathrm{is}}(\boldsymbol{\beta}_{\mathbf{0}})}{\partial \beta_{\mathrm{j}}}\{\mathrm{y}_{\mathrm{it}} - \mu_{\mathrm{it}}(\boldsymbol{\beta}_{\mathbf{0}})\}\right] = 0,$$

for all $\mathrm{s} \leq \mathrm{t}$, $\mathrm{t} = 1, \ldots, \mathrm{T}$. In my analysis there will be $\frac{\mathrm{T(T+1)}}{2}$ estimating equations valid for each type IV covariate. These approaches (type I, type II, type III, and type IV) consider the grouping of the estimating equations in an effort to determine valid estimating equations by group. These do not consider cases when the feedback may be immediate but later ineffective. Thus I take a different approach.

I consider the $\mathrm{T}(\mathrm{T}-1)$ estimating equations to determine cases where

$$\mathrm{E}_{\boldsymbol{\beta}_0}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\{y_{it}-\mu_{it}(\boldsymbol{\beta}_0)\}\right]=0.$$

I take all cases of $s=t$ as the base set of T estimating equations that must be considered. I then examine simultaneously the $\mathrm{T}(\mathrm{T}-1)$ estimating equations associated with $s \neq t$ to determine which are valid. Consider for each time t the example the model:

$$\mathcal{L}_t = \operatorname{logit}(p_t) = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t}, \tag{2}$$

where p_t is the probability that $y_{it}=1$. Let e_t denote the residual at time t, estimates of the estimator ϵ_t. Let $\hat{\rho}_{e_t x_s}$ be the estimate for the correlation between the errors at t and the covariate at s, $\rho_{e_t x_s}$. We know by design $\rho_{e\,x}=0$ when $s=t$ but not necessarily when $s \neq t$. I posit that when $\rho_{e_t x_s}=0$ for $s \neq t$ then the corresponding estimating equation is valid. Thus I conduct a test for the correlation and ignore the equation when the correlations were significant. I justify my approach and assume

$$\mathrm{E}(y_{it}|\mathbf{x}_{it\cdot}) = \mu_{it} \equiv \mu_{it}(\boldsymbol{\beta}),$$

so that

$$\mathrm{E}(y_{it}-\mu_{it}) = \mathrm{E}[\mathrm{E}(y_{it}-\mu_{it}|\mathbf{x}_{it\cdot})] = \mathrm{E}[\mathrm{E}(y_{it}|\mathbf{x}_{it\cdot})-\mu_{it}] = \mathrm{E}(\mu_{it}-\mu_{it}) = 0,$$

and

$$\begin{aligned}
&\mathrm{Cov}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}, y_{it}-\mu_{it}(\boldsymbol{\beta}_0)\right] \\
&= \mathrm{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\{y_{it}-\mu_{it}(\boldsymbol{\beta}_0)\}\right] - \mathrm{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\right]\mathrm{E}[y_{it}-\mu_{it}(\boldsymbol{\beta}_0)] \\
&= \mathrm{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\{y_{it}-\mu_{it}(\boldsymbol{\beta}_0)\}\right].
\end{aligned}$$

By definition,

$$\operatorname{Cov}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}, y_{it} - \mu_{it}(\boldsymbol{\beta_0})\right]$$

$$= \operatorname{Corr}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}, y_{it} - \mu_{it}(\boldsymbol{\beta_0})\right]\sqrt{\operatorname{Var}\left(\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}\right)\operatorname{Var}\left(y_{it} - \mu_{it}(\boldsymbol{\beta_0})\right)}\,,$$

So

$$\operatorname{Corr}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}, y_{it} - \mu_{it}(\boldsymbol{\beta_0})\right] = 0 \iff \operatorname{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right] = 0.$$

Since in the logistic regression case

$$\operatorname{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta}_0)\}\right] = \operatorname{E}\left[x_{isj}\mu_{is}(\boldsymbol{\beta_0})\{1 - \mu_{is}(\boldsymbol{\beta_0})\}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right],$$

we need to examine $\operatorname{Corr}\left(x_{isj}\ \mu_{is}(\boldsymbol{\beta_0})[1 - \mu_{is}(\boldsymbol{\beta_0})],\ y_{it} - \mu_{it}(\boldsymbol{\beta_0})\right)$ to check the validity of $\operatorname{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right] = 0$. Thus, our investigation includes the correlation between the residuals from the logistic regression based on all the covariates at time t with the weighted particular covariate at time s. I postulate that testing for this correlation is sufficient to determine the valid estimating equations.

Since in the normal regression case

$$\operatorname{E}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right] = \operatorname{E}\left[x_{isj}\{y_{it} - \mu_{it}(\boldsymbol{\beta_0})\}\right],$$

we need to examine $\operatorname{Corr}\left(x_{isj},\ y_{it} - \mu_{it}(\boldsymbol{\beta_0})\right)$ to check the validity of each estimating equation. Thus, our investigation includes the correlation between the residuals from the normal regression based on all the covariates at time t with the particular covariate at time s.

DISTRIBUTION OF BIVARIATE CORRELATION

To determine whether or not the correlation is significant we need to obtain the asymptotic distribution of residuals at time t and past or future covariates. Suppose $[e_i, X_i]$ for $i = 1, \dots, N$; represent independent, identically distributed bivariate random values with mean $(0,0)$ and variance such that $\Omega = \begin{pmatrix} 1 & \rho_{ex} \\ \rho_{ex} & 1 \end{pmatrix}$. Denote the sample correlation between e_i and X_i by

$$\hat{\rho}_{ex} = \frac{\sum_{i=1}^{N}(e_i - \bar{e})(X_i - \bar{X})}{\sqrt{\sum_{i=1}^{N}(e_i - \bar{e})^2 \sum_{i=1}^{N}(X_i - \bar{X})^2}} = \frac{\sum_{i=1}^{N} e_i X_i}{\sqrt{\sum_{i=1}^{N} {e_i}^2 \sum_{i=1}^{N} {X_i}^2}}.$$

Assume all fourth moments exist and are finite; denoted by

$$\mu_{mn} = E[(e - \mu_e)^m (X - \mu_X)^n] = E[e^m X^n] < \infty,$$

for $m + n \leq 4$. Applying the multivariate central limit theorem and multivariate delta method gives the limiting distribution as

$$\sqrt{N}(\hat{\rho}_{ex} - \rho_{ex}) \overset{d}{\rightarrow} \mathcal{N}\left(0, \mu_{22} - \rho_{ex}(\mu_{13} + \mu_{31}) + \frac{{\rho_{ex}}^2}{4}(\mu_{40} + 2\mu_{22} + \mu_{04})\right),$$

where μ_{mn} has as an estimate

$$\hat{\mu}_{mn} = \frac{1}{N}\sum_{i=1}^{N}(e_i - \bar{e})^m (X_i - \bar{X})^n = \frac{1}{N}\sum_{i=1}^{N} {e_i}^m {X_i}^n,$$

for $m + n \leq 4$. Under the assumption of normality, a variance-stabilizing transformation gives Fisher's Z-transformation:

$$\frac{\sqrt{N}}{2}\left(\ln\left(\frac{1+\hat{\rho}_{ex}}{1-\hat{\rho}_{ex}}\right) - \ln\left(\frac{1+\rho_{ex}}{1-\rho_{ex}}\right)\right) \overset{d}{\rightarrow} \mathcal{N}(0,1).$$

GMM ESTIMATOR

Once we have identified the set of valid equation we need to obtain the estimate for the coefficient. For subject i, let $\mathbf{y}_i = (y_{i1}, \dots, y_{iT})'$ be a T x 1 vector of outcomes associated with matrix $\mathbf{X}_i = \begin{bmatrix} x_{i11} & \cdots & x_{i1J} \\ \vdots & \ddots & \vdots \\ x_{iT1} & \cdots & x_{iTJ} \end{bmatrix}$, where at time t the row vector, $\mathbf{x}_{it.} = (x_{it1}, \dots, x_{itJ})$ and for the j^th^ covariate the column vector $\mathbf{x}_{i.j} = (x_{i1j}, \dots, x_{iTj})'$ such that $t = 1, \dots, T$; and $j = 1, \dots, J$. Arrange $\mathbf{X}_i$ such that the 1st column of $\mathbf{X}_i$ is the intercept term, a T x 1 vector consisting of value 1, and the last $T - 1$ columns are indicator variables for the set of times.

Let $\boldsymbol{\beta} = (\beta_1, \dots \beta_J)'$ be the J x 1 vector of parameters. The optimal GMM estimator $\widehat{\boldsymbol{\beta}}_{\mathrm{GMM}}$ minimizes a quadratic objective function $\mathbf{G}_\mathbf{n}' \mathbf{W}_\mathbf{n} \mathbf{G}_\mathbf{n}$ where $\mathbf{G}_\mathbf{n}$ is a N_v x 1 vector consists of all valid moment conditions, and $\mathbf{W}_\mathbf{n}$ is a N_v x N_v weight matrix, where N_v denotes the total number of valid moment conditions.

Let $\mathbf{T}_{\mathbf{v}\mathrm{j}}$ be a T x T matrix that specifies valid moment conditions for the j^{th} covariate. Elements in $\mathbf{T}_{\mathbf{v}\mathrm{j}}$ take two values only: 0 and 1. If the element in row s, column t of $\mathbf{T}_{\mathbf{v}\mathrm{j}}$ takes value 1, it indicates that the moment condition

$$E_{\boldsymbol{\beta}_0}\left[\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}\{y_{it} - \mu_{it}(\boldsymbol{\beta}_0)\}\right] = 0,$$

is valid for the j^{th} covariate. Reshape $\mathbf{T}_{\mathbf{v}\mathrm{j}}$ into a 1 x (TxT) row vector for $j = 1, \dots, J$ and concatenate the rows for all covariates to form $\mathbf{T}_{\mathbf{shape}}$, a J x (TxT) matrix. The number of 1's in $\mathbf{T}_{\mathbf{shape}}$ represents the total number of valid moment conditions, denoted by N_v.

Let $\mathbf{g}_i$ be a N_v x 1 vector containing the computed value of all valid moment condition for subject i, as a function of initial value $\boldsymbol{\beta}_0$. The elements in $\mathbf{g}_i$ takes the form $\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}[y_{it} - \mu_{it}(\boldsymbol{\beta}_0)]$ such that the element in row s, column t of $\mathbf{T}_{\mathbf{v}\mathrm{j}}$ takes value 1. Empirically the N_v x 1 vector $\mathbf{G}_\mathbf{n}$ is computed by

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{g}_i = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial \mu_{is}(\boldsymbol{\beta}_0)}{\partial \beta_j}[y_{it} - \mu_{it}(\boldsymbol{\beta}_0)].$$

The N_v x N_v weight matrix $\mathbf{W}_\mathbf{n}$ is computed by $\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{g}_i\mathbf{g}_i^T\right)^{-1}$.

The GMM estimator $\widehat{\boldsymbol{\beta}}_{\mathrm{GMM}}$ is the argument to minimize the quadratic objective function $\mathbf{G}_\mathbf{n}(\boldsymbol{\beta}_0)'\, \mathbf{W}_\mathbf{n}(\boldsymbol{\beta}_0)\ \mathbf{G}_\mathbf{n}(\boldsymbol{\beta}_0)$,

$$\widehat{\boldsymbol{\beta}}_{\mathrm{GMM}} = \mathrm{argmin}_{\boldsymbol{\beta}_0}\, \mathbf{G}_\mathbf{n}(\boldsymbol{\beta}_0)'\, \mathbf{W}_\mathbf{n}(\boldsymbol{\beta}_0)\ \mathbf{G}_\mathbf{n}(\boldsymbol{\beta}_0).$$

The asymptotic variance of $\widehat{\boldsymbol{\beta}}_{\mathrm{GMM}}$ is computed by

$$\left[\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)^T \mathbf{W}_\mathbf{n}(\boldsymbol{\beta})\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)\right]^{-1},$$

evaluated at $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}_{\mathrm{GMM}}$.

In the case of logistic regression, the elements in $\mathbf{g}_{\mathrm{i}}$ take the form

$$\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}[y_{it} - \mu_{it}(\boldsymbol{\beta_0})] = x_{isj}\mu_{is}(\boldsymbol{\beta_0})[1 - \mu_{is}(\boldsymbol{\beta_0})][y_{it} - \mu_{it}(\boldsymbol{\beta_0})],$$

where

$$\mu_{it}(\boldsymbol{\beta_0}) = \frac{\exp\,(\mathbf{x}_{it.}\boldsymbol{\beta})}{1+\exp\,(\mathbf{x}_{it.}\boldsymbol{\beta})},$$

such that the element in row s, column t of $\mathbf{T}_{\mathbf{v}j}$ takes value 1.

For the N_v x J matrix $\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \left[\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \beta_1}, \quad \dots, \quad \frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \beta_J}\right]$, the N_v x 1 column vector $\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \beta_k}$, for $k = 1, \dots, J$, for logistic regression can be computed by

$$\frac{\partial\left\{\left[\frac{\partial \mu_{is}(\boldsymbol{\beta})}{\partial \beta_j}\right][y_{it} - \mu_{it}(\boldsymbol{\beta})]\right\}}{\partial \beta_k}$$
$$= x_{isj}\mu_{is}(\boldsymbol{\beta})[1 - \mu_{is}(\boldsymbol{\beta})]\left\{x_{isk}[1 - 2\mu_{is}(\boldsymbol{\beta})][y_{it} - \mu_{it}(\boldsymbol{\beta})] - x_{itj}\mu_{it}(\boldsymbol{\beta})[1 - \mu_{it}(\boldsymbol{\beta})]\right\}.$$

In the case of normal regression, the elements in $\mathbf{g}_{\mathrm{i}}$ take the form

$$\frac{\partial \mu_{is}(\boldsymbol{\beta_0})}{\partial \beta_j}[y_{it} - \mu_{it}(\boldsymbol{\beta_0})] = x_{isj}[y_{it} - \mu_{it}(\boldsymbol{\beta_0})],$$

such that the element in row s, column t of $\mathbf{T}_{\mathbf{v}j}$ takes value 1.

For the N_v x J matrix $\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \left[\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \beta_1}, \quad \dots, \quad \frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \beta_J}\right]$, the N_v x 1 column vector $\frac{\partial \mathbf{g}_i(\boldsymbol{\beta})}{\partial \beta_k}$, for $k = 1, \dots, J$, for normal regression can be computed by

$$\frac{\partial\left\{\left[\frac{\partial \mu_{is}(\boldsymbol{\beta})}{\partial \beta_j}\right][y_{it} - \mu_{it}(\boldsymbol{\beta})]\right\}}{\partial \beta_k} = -x_{isj}x_{isk}\,.$$

CHAPTER 4

ILLUSTRATIVE EXAMPLES

MODELING PROBABILITY OF REHOSPITALIZATION

Medicare is a social insurance program administered by the United States government, providing health insurance coverage to people who are aged 65 and over, or who meet other special criteria. Medicare currently pays for all rehospitalizations, except those in which patients are rehospitalized within 24 hours after discharge for the same condition for which they had initially been hospitalized (Jencks, Williams and Coleman, 2009).

I use data from the Arizona State Inpatient Database (SID). This dataset contains patient information from Arizona hospital discharges for 3-year period from 2003 through 2005. This dataset contains information of those who were admitted to a hospital exactly 4 times. There were 1625 patients in the dataset with complete information; each has three observations indicating three different times to rehospitalizations. I classify those who returned to the hospital within 30-days as one opposed to zero for those who did not. My list of chosen covariates is initiated by the findings of Jencks et al. (2009) and includes multitude of diseases (NDX), number of procedures (NPR), length of stay (LOS), coronary atherosclerosis (DX101) and time dummies (T2 and T3).

In fitting a logistic regression model for the probability of rehospitalization I include the effects of the covariates due to the time varying aspect. I use the GEE model and the GMM model with the extended method as presented in Chapter 3. In the GMM model with the extended method I first determine the type of each covariate through a logistic model.

The results are summarized in Table 1. In this table we have the correlations, p-values based on the asymptotic distribution of correlations and the validity of estimating equations for each covariate, NDX, NPR, LOS and DX101. In the logistic regression case

Table 1

Correlation Tests for Estimating Equations with Medicare Data

	NDX			**NPR**		
CORRELATION	TIME1	TIME 2	TIME 3	TIME 1	TIME 2	TIME 3
RSD1	0.005	-0.010	-0.099	0.004	-0.029	0.018
RSD2	0.035	0.002	0.049	0.004	0.000	-0.012
RSD3	0.008	0.005	0.012	-0.039	0.006	0.006
P-VALUE						
RSD1	0.849	0.714	0.000	0.886	0.228	0.447
RSD2	0.147	0.943	0.030	0.878	0.999	0.615
RSD3	0.755	0.847	0.624	0.148	0.817	0.815
VALIDITY						
RSD1	1	1	0	1	1	1
RSD2	1	1	0	1	1	1
RSD3	1	1	1	1	1	1
	LOS			**DX101**		
CORRELATION	TIME 1	TIME 2	TIME 3	TIME 1	TIME 2	TIME 3
RSD1	0.028	0.087	0.026	0.000	0.062	0.051
RSD2	0.040	0.017	0.125	-0.015	-0.002	-0.012
RSD3	0.058	0.069	0.032	-0.032	0.009	-0.001
P-VALUE						
RSD1	0.249	0.000	0.331	0.991	0.010	0.015
RSD2	0.102	0.473	0.000	0.603	0.937	0.572
RSD3	0.014	0.004	0.231	0.269	0.721	0.957
VALIDITY						
RSD1	1	0	1	1	0	0
RSD2	1	1	0	1	1	1
RSD3	0	0	1	1	1	1

we need to examine $\mathrm{Corr}\left(x_{sj}\,\mu_s\left(1-\mu_s\right),\,y_t-\mu_t\right)$ to check the validity of estimating equations for each covariate. A small p-value suggests that the estimating equation $E_{\boldsymbol{\beta_0}}\left[x_{sj}\,\mu_s(\boldsymbol{\beta_0})\left\{1-\mu_s\left((\boldsymbol{\beta_0})\right)\right\}\left\{y_t-\mu_t(\boldsymbol{\beta_0})\right\}\right]=0$ fails to hold for the jth covariate for the particular combination of s and t. First we need to fit logistic regression based on all the covariates (except time indicators) for each time and obtain the predicted probability

μ_{it} for t=1,2,3. For NDX, we examine the correlations between the residuals from the logistic regression at time t, i.e. $y_t - \mu_t$, t =1,2,3, denoted by rsd, and $NDX_s\mu_s(1-\mu_s)$, the weighted covariate NDX at time s, s=1,2,3. The small p-value for the correlation when t=1, s=3 suggests that the estimating equation for t=1, s=3 should not be included, corresponding to value 0 for validity. Likewise, for NDX we should not include the estimating equation for t=2, s=3 either. Thus we have the rest 7 estimating equations for NDX, corresponding to value 1 for validity. Similarly we can use all of the equations for NPR. For LOS we leave out the equations for t=1, s=2; t=2, s=3; t=3, s=1; and t=3, s=2. For DX101 we leave out the estimating equations for t=1, s=2 and t=1, s=3.

I fit the logistic regression model with the covariates NDX, NPR, LOS, and DX101 in addition to time dummies T2 and T3. The GEE results along with the GMM results using the extended method are given in Table 2. The GEE model ignores the time varying among the responses and the covariates while the GMM model do not ignore. Both models show that NDX, LOS, and time have an impact on probability of rehospitalization. Unlike the GEE model, the GMM model finds that NPR had some significance of an impact on the probability of rehospitalization.

Table 2

Comparison of GEE and GMM with the extended method for Medicare Data

	GEE		**GMM**	
PARAMETER	**EST**	**P-VALUE**	**EST**	**P-VALUE**
INTERCEPT	-0.3675	0.0035	-0.4076	0.0009
NDX	0.0648	**<.0001**	0.0642	**0.0000**
NPR	-0.0306	0.11	-0.0315	0.0922
LOS	0.0344	**<.0001**	0.0396	**0.0000**
DX101	-0.1143	0.2224	-0.0517	0.5776
T2	-0.3876	**<.0001**	-0.3840	**0.0000**
T3	-0.2412	**0.0005**	-0.2686	**0.0001**

MODELING MEAN MORBIDITY

As an illustrative example for non-binary response data with the extended method of fitting GMM, I choose to revisit the data analyzed by Lai and Small (2007). They consider a dataset that was collected by the International Food Policy Research Institute in the Bukidnon Province in the Philippines and focus on quantifying the association between body mass index (BMI) and morbidity four months into the future. Data were collected at four time points, separated by 4-month intervals (Bhargava, 1994). There were 370 children with three observations. The predictors are BMI, age, gender, and time dummies. Following Lai and Small (2007), I model the sickness intensity measured by adding the duration of sicknesses and taking a logistic transformation of the proportion of time for which a child is sick (with a continuity correction for extreme values; Cox, 1970). I fit the GEE model with the independent correlation structure, the GMM model with Lai and Small's three-type classification, and the GMM model with the extended method proposed in this paper but adjusted for non-binary data.

Table 3

Correlation Tests for Estimating Equations with Philippine Data

	BMI			AGE		
CORRELATION	TIME1	TIME 2	TIME 3	TIME1	TIME 2	TIME 3
RSD1	0.000	-0.042	0.023	0	0.003	0.002
RSD2	-0.067	0.000	-0.104	0.001	0	0.000
RSD3	-0.036	0.022	0.000	0.001	-0.001	0
P-VALUE						
RSD1	1.000	0.551	0.732	1	0.962	0.964
RSD2	0.159	1.000	**0.037**	0.991	1	1.000
RSD3	0.444	0.663	1.000	0.980	0.986	1
VALIDITY						
RSD1	1	1	1	1	1	1
RSD2	1	1	0	1	1	1
RSD3	1	1	1	1	1	1

Table 3 provides the correlation tests and the selection of estimating equations in the use of the extended method. Recall that in the normal regression case we examine $\mathrm{Corr}\left(x_{sj},\ y_t - \mu_t\right)$ to check the validity of estimating equations for each covariate. A small p-value suggests that the estimating equation $E_{\boldsymbol{\beta_0}}\left[x_{sj}\{y_t - \mu_t(\boldsymbol{\beta_0})\}\right] = 0$ fails to hold for the j^{th} covariate for the particular combination of s and t. First we need to fit normal regression based on all the covariates (except time indicators) for each time and obtaine the predicted value μ_{it} for t=1,2,3. For BMI, we examine the correlations between the residuals from the normal regression at time t, i.e. $y_t - \mu_t$, t =1,2,3, denoted by rsd, and BMI_s, the covariate at time s, s=1,2,3. The small p-value for the correlation when t=2, s=3 suggests that estimating equation for t=2, s=3 should not be included, corresponding to value 0 for validity. We can use the rest 8 estimating equations for BMI, corresponding to value 1 for validity. Similarly we have all of the equations valid for age and gender.

Table 4

Comparison of GEE, GMM with the Three-Type Method and GMM with the Extended Method for Philippine Data

	GEE		GMM					
			Lai and Small			Extended		
	EST	P	TYPE	EST	P	TYPE	EST	P
Intercept	-0.972	0.215	III	-0.888	0.178	All	-0.625	0.326
BMI	-0.062	0.176	II	-0.072	0.061	Exclude (s=3, t=2)	-0.087	**0.019**
Age	-0.013	**0.000**	I	-0.012	**0.000**	All	-0.012	**0.000**
Gender	0.145	0.183	III	0.087	0.387	All	0.073	0.464
T2	-0.28	**0.012**	I	-0.277	**0.007**	All	-0.272	**0.008**
T3	0.024	0.847	I	-0.018	0.876	All	-0.034	0.772

Table 4 provides the results of modeling the mean sickness intensity using GEE, GMM with Lai and Small's three-type method, and GMM with the extended method. The GEE model which ignores the correlations on account of time varying covariates gives age and period 2 as significant. I use the results in Lai and Small (2007) and classify age as type I (that means all the equations are used) and BMI as type II (that means the estimating equations for s=1, t=2; s=1, t=3; and s=2, t=3 are omitted). The GMM model with Lai and Small's classification gives age and period 2 as significant and BMI as marginally insignificant. The GMM model with the extended method gives age, BMI, and period 2 as significant.

In this case Lai and Small's method relies on more estimating equations than the GEE method but two less than the extended method. However, those extra set of equations are enough to have BMI shown to be significant with the extended method but not with Lai and Small's method and the GEE method (Table 5).

Table 5

Change in P-Values as Estimating Equations Increase for BMI

	GEE	**GMM**	
		LAI & SMALL	**EXTENDED**
BMI	3	6	8
P-VALUE	0.176	0.061	0.019

Results of Number of Estimating Equations in BMI

Although this is not a simulation study I examined the effects of the increasing number of estimating equations when estimating the time-varying covariate, BMI on the mean sickness intensity for Filipino children. This was undertaken to get a sense of the penalty involved when estimating equations are left out. In Table 6 I provide the estimates and the standard errors for the effect of BMI while controlling for age and gender. In this study I used all the estimating equations for age and gender. The standard error seems to get larger as fewer equations are allowed. We see that when all equations are considered BMI gave an estimate of -0.0715 with standard error equal to 0.0367, while when we only allow the same cases as GEE we get an estimate of -0.0972 with a standard error of 0.0418.

Table 6

Change in Estimates and Standard Errors as the Estimating Equations Allowed for BMI Decrease

SET	EQUATIONS	BMI	STDERR	AGE	STDERR	GENDER	STDERR
I	ALL EQUATIONS	-0.0715	0.0367	-0.0110	0.0031	0.0810	0.1000
II	I WITHOUT Y_2BMI_3	-0.0802	0.0368	-0.0116	0.0031	0.0740	0.0999
III	II WITHOUT Y_2BMI_1	-0.1019	0.0386	-0.0123	0.0031	0.0537	0.1004
IV	III WITHOUTY_3BMI_1	-0.1000	0.0386	-0.0126	0.0031	0.0530	0.1004
V	IV WITHOUTY_1BMI_2	-0.1026	0.0386	-0.0129	0.0032	0.0449	0.1006
VI	V WITHOUT Y_3BMI_2	-0.1017	0.0392	-0.0129	0.0032	0.0426	0.1013
VII	VI WITHOUT Y_1BMI_3	-0.0972	0.0418	-0.0127	0.0033	0.0433	0.1013

CHAPTER 5

CONCLUSIONS

Researchers are aware that in the analysis of repeated measures binary data the correlation present on account of the repeated measures in the responses must be addressed. However, until recent times the dependency also present in the covariates that change over time due to factors other than the natural growth have been ignored. Thus the modeling of repeated measures data must address two sets of correlation inherent; one due to the responses and the other due to the covariates. While the generalized method of moments is an improved choice over GEE with independent working correlation, it is not at present available in statistical software packages such as SAS, or SPSS though can be done in R (Lalonde and Wilson, 2010). However, I provide a procedure in SAS through PROC IML as I compare to existing methods.

I develope a new approach to marginal models for time-dependent covariates both for binary and non-binary responses. Unlike Lai and Small (2007)'s approach of classifying variables into three types I take a different approach. The advantage of my approach is that I do not assume any feedback will be consistent or significant over time. As such I postulate that there is an advantage to my approach when the period followed are longer as one would expect associations to change as time increases. I use a correlation technique to determine which estimating equation should be considered valid.

REFERENCES

Anderson TW. An Introduction to Multivariate Statistical Analysis. New York: Wiley, 1966.

Bhargava A. “Modelling the Health of Filipino Children.” Journal of the Royal Statistical Society, Series A 157, no.3 (1994): 417-432

Cox DR. Analysis of Binary Data. London: Chapman and Hall, 1970.

Dobson AJ. An Introduction to Generalized Linear Models. Chapman and Hall, 2002.

Diggle P, Heagerty P, Liang K, Zeger S. Analysis of Longitudinal Data, Oxford University Press, 2002.

Fitzmaurice GM. “A Caveat Concerning Independence Estimating Equations with Multivariate Binary Data.” Biometrics 51, no.1 (1995):309-317.

Hastie T, Tibshirani R. Generalized Additive Models. Chapman and Hall, 1990.

Hedeker D, Gibbons RD. Longitudinal Data Analysis. New York: Wiley-Interscience, 2006.

Hu FC. “A Statistical Methodology for Analyzing the Causal Health Effect of A Time Dependent Exposure From Longitudinal Data.” ScD dissertation, Harvard School of Public Health, 1993.

Jencks SF, Williams MV and Coleman EA. “Rehospitalizations among Patients in the Medicare Fee-for-Service Program.” The New England Journal of Medcine 360, no.14 (2009): 1418-1428.

Jha AK, Orav EJ, Epstein AM. “Public Reporting of Discharge Planning and Rates of Readmissions.” The New England Journal of Medicine 361, no.27 (2009):2637-2645.

Lai TL, Small D. “Marginal Regression Analysis of Longitudinal Data with Time-Dependent Covariates: A Generalized Method-of-Moments Approach.” Journal of the Royal Statistical Society, Series B 69, no.1 (2007):79-99.

Lalonde T, Wilson JR. “A Generalized Method of Moments Approach for Binary Data with Time-Dependent Covariates.” Proceedings of ASA Meetings Section on Statistical Computing, Vancouver, Canada 2010.

Liang KY, Zeger SL. “Longitudinal Data Analysis Using Generalized Linear Models.” Biometrika 73, no.1 (1986):13-22.

McCullagh PJ, Nelder JA. Generalized Linear Models. London: Chapman and Hall, 1989.

Medelsee M. "Estimating Pearson's Correlation Coefficient with Bootstrap Confidence Interval from Serially Dependent Time Series." Mathematical Geology 35, no.6 (2003):651-665.

Nelder JA, Wedderburn RWM. "Generalized Linear Models." Journal of the Royal Statistical Society, Series A 135, no.3 (1972):370-384.

Pan W, Connett JE. "Selecting the Working Correlation Structure in Generalized Estimating Equations with Application to the Lung Health Study." Statistica Sinica 12, no.2 (2002):475-490.

Pepe MS, Anderson GL. "A Cautionary Note on Inference for Marginal Regression Models with Longitudinal Data and General Correlated Response Data." Communications in Statistics-Simulation and Computation 23, no.4 (1994):939-951.

Tate RF. "Correlation between a Discrete and a Continuous Variable." The Annual of Mathematical Statistics 25 (1954):603-607.

Tate RF. "Applications of Correlation Models for Biserial Data." The Journal of American Statistical Association 50 (1955):1078-1095.

Zeger SL, Liang KY. "Longitudinal Data Analysis for Discrete and Continuous Outcomes." Biometrics 42, no.1 (1986):121-130.

Zeger SL, Liang KY, Albert PS. "Models for Longitudinal Data: A Generalized Estimating Equation Approach." Biometrics 44, no.4 (1988):1049-1060.

Zeger SL, Liang KY. "An Overview of Methods for the Analysis of Longitudinal Data." Statistics in Medicine 11, no.14-15 (1992):1825-1839.

APPENDIX A

SAS CODE USING PROC IML FOR MEDICARE DATA

```
/*##########################################
* read data and create time dummies;
############################################*/

libname perm 'c:\SAS\perm';
data mydata; set perm.Medicare;
if time=1 then t1=1; else t1=0;
if time=2 then t2=1; else t2=0;
if time=3 then t3=1; else t3=0;
run;

/*##########################################
* obtain residuals from by-time regression;
############################################*/

title ' pooled logistic by time';
proc sort data=mydata out=mydatasorted;
by time; run;
proc logistic data=mydatasorted noprint;
by time;
model biRadmit (event='1') =NDX NPR LOS DX101 / aggregate
scale=none;
output out=outpool3 p=mu xbeta=xb RESCHI=rsdpsn
RESDEV=rsddev;
run;
data outpool3; set outpool3;
wt = mu*(1-mu);
rsdraw = biRadmit-mu;
run;
/*########################################
* examine corr by PROC IML;
##########################################*/
PROC SORT DATA=outpool3 OUT=outpool3 ;
  BY PNUM_R time; RUN;
proc iml;
use outpool3; * ## change ####;
read all VARIABLES {wt NDX NPR LOS DX101 t2 t3 PNUM_R time}
into Zmat;    * ## change ####;
read all var {rsdraw} into rsd;
close outpool3;

N=1625;T =3;  * ## change ####;

start rho(a,rsd) global(N,T);
abm = j(N,2*T,.);
abm[,1:T] = shape(rsd,N);
abm[,T+1:2*T] = shape(a,N);
corr = corr(abm);
rho = corr[1:T,T+1:2*T];
return(rho);
```

```
finish rho;

start stddev(a,rsd) global(N,T);
bm = shape(rsd,N);
bdev = bm-j(N,1,1)*bm[:,];
bdev2 = bdev#bdev;
am = shape(a,N);
adev = am-j(N,1,1)*am[:,];
adev2 = adev#adev;
stddev = sqrt( (1/N)*t(bdev2)*adev2 );
return(stddev);
finish stddev;

start stdzn(x);
N = nrow(x);
y = x-x[:,];
vcv = (1/(N-1))*t(y)*y;
v = diag(vcv);
sinv = diag(sqrt(1/vcv));
x2 = y*sinv;
return(x2);
finish stdzn;

print 'Corr Examination to Medicare Data';

print 'wt*NDX';
x1 = Zmat[,1]#Zmat[,2];
x2 = stdzn(x1);
rsd2 = stdzn(rsd);
r2 = rho(x2,rsd2); s2 = stddev(x2,rsd2);
z2 = sqrt(N)*(r2/s2);
p2 = 2*(1-cdf('normal',abs(z2)));
print r2, s2, z2, p2;

print 'wt*NPR';
x1 = Zmat[,1]#Zmat[,3];
x2 = stdzn(x1);
rsd2 = stdzn(rsd);
r2 = rho(x2,rsd2); s2 = stddev(x2,rsd2);
z2 = sqrt(N)*(r2/s2);
p2 = 2*(1-cdf('normal',abs(z2)));
print r2, s2, z2, p2;

print 'wt*LOS';
x1 = Zmat[,1]#Zmat[,4];
x2 = stdzn(x1);
rsd2 = stdzn(rsd);
r2 = rho(x2,rsd2); s2 = stddev(x2,rsd2);
z2 = sqrt(N)*(r2/s2);
p2 = 2*(1-cdf('normal',abs(z2)));
```

```
print r2, s2, z2, p2;


print 'wt*DX101';
x1 = Zmat[,1]#Zmat[,5];
x2 = stdzn(x1);
rsd2 = stdzn(rsd);
r2 = rho(x2,rsd2); s2 = stddev(x2,rsd2);
z2 = sqrt(N)*(r2/s2);
p2 = 2*(1-cdf('normal',abs(z2)));
print r2, s2, z2, p2;

 /*#####################################################
Obtain GEE Results Based on the Independent Correlation
Structure to Serve as Initial Values
#######################################################*/
proc genmod data=mydata descend;   * to model Prob(y=1);
      class PNUM_R time;
      model biRadmit=NDX NPR LOS DX101 t2 t3 / dist=bin ;
      repeated subject=PNUM_R /within=time corr=indep corrw;
       output out=GEEout xbeta=xb RESRAW = rraw;
run;



/*#####################################################
Obtain GMM Estimates
#######################################################*/
proc sort data=mydata;
by PNUM_R time; run;

proc iml;
use mydata;     * ## change variable list ####;
read all VARIABLES {NDX NPR LOS DX101 } into Zmat;
read all var {biRadmit} into yvec;
close mydata;
print '2SGMM with Extended Method to Medicare Data of
Binary Y';

N=1625; * number of observations;  * ## change ####;
Pn=7;   * number of parameters to estimate; * ## change ###;
* Intercept   NDX      NPR     LOS       DX101   time2
time3;
*GEE  results as starting values; * ## change ####;
beta0 = {-0.3675 0.0648    -0.0306 0.0344 -0.1143 -0.3876
-0.2412};

nr = nrow(Zmat);
nc = ncol(Zmat);
int = j(nr,1,1);
```

```
Xmat =j(nr,nc+3,.);
Xmat[,1]=int; Xmat[,2:nc+1]=Zmat;
in = j(N,1,1);
tm2 = {0,1,0}; D2 = in@tm2;
tm3 = {0,0,1}; D3 = in@tm3;
Xmat[,nc+2]=D2; Xmat[,nc+3]=D3;

Tn=3; * number of periods per observation; * ## change ####;

* Intercept   NDX      NPR     LOS       DX101   time2
time3;
J={1,2,3,4,5,6,7}; * Type specification using raw residuals;
* ## change ####;

/*intercept * mu(1-mu) = */   * ## change ####;
T1 = {1 1 0,
       0 1 0,
      1 1 1};

/*NDX * mu(1-mu) = */
T2 = {1 1 0 ,
      1 1 0 ,
      1 1 1 };

/*NPR * mu(1-mu) = */
T3 = {1 1 1,
      1 1 1,
      1 1 1};

/*LOS * mu(1-mu) = */
T4 = {1 0 1,
      1 1 0,
      0 0 1};

/*DX101 * mu(1-mu) = */
T5 = {1 0 0,
      1 1 1,
      1 1 1};

/*t2 * mu(1-mu) = */
T6 = {0 1 0,
      0 1 0,
      0 1 0};

/*t3 * mu(1-mu) = */
T7 = {0 0  0,
      0 0  0,
      0 0  1};

T0 = {1 0  0,
      0 1  0,
      0 0  1};

Tshape = j(Pn,Tn*Tn,.);
neq = j(Pn,1,0);

do p =1 to Pn;
  if J[p]=1 then Tshape[p,] = shape(T1,1);
    else if  J[p]=2 then Tshape[p,] = shape(T2,1);
         else if J[p]=3 then Tshape[p,] = shape(T3,1);
           else if J[p]=4 then Tshape[p,] = shape(T4,1);
               else if J[p]=5 then Tshape[p,] = shape(T5,1);
                   else if J[p]=6 then Tshape[p,] =
shape(T6,1);
                      else if J[p]=7 then Tshape[p,] =
shape(T7,1);
                  else Tshape[p,] = shape(T0,1);
    neq[p] = ncol(loc(Tshape[p,]^=0));
end;

* nloc containing the starting/end positions of reg eq's
brought by each covariate  ;
nloc = j(1,Pn+1,0);
do p =1 to Pn;
  nloc[p+1] = sum(neq[1:p]);
end;
nv = sum(neq);

Wn = I(nv);                * initial weight matrix;
S = j(nv,nv,0);            * to compute covariance mtx ;

start TSGMM(beta)
global(Pn,Tn,N,Xmat,yvec,nv,Tshape,nloc,Wn,S);
Gn = j(nv,1,0);           * to collect valid mmt conditions;
S = j(nv,nv,0);            * to compute covariance mtx ;
eq = j(nv,N,0);

do i = 1 to N;
  x = Xmat[(i-1)*Tn+1:i*Tn,];
  y = yvec[(i-1)*Tn+1:i*Tn];
  mu = exp(x*t(beta)) / ( 1+exp(x*t(beta)) );
  Rsd = y - mu;
  do p = 1 to Pn;
    D = x[,p]#mu#(1- mu);
    Eqmtx = Rsd*t(D);
    eq[nloc[p]+1:nloc[p+1],i] = Eqmtx[loc(Tshape[p,]^=0)];
  end;
  S = S + eq[,i]*t(eq[,i]);
end;
Gn = eq[,:];
```

```
f = t(Gn)*Wn*Gn;               * the objective fn to be
minimized;
return(f);
finish TSGMM;

tc = {2000 2000}; optn = {0 2};
  call NLPNRA(rc, xres,"TSGMM", beta0,optn, , tc);
  beta0 = xres;
  Wn = ginv(S/N);

  call NLPNRA(rc, xres,"TSGMM", beta0,optn, , tc);
  beta = xres;
  Wn = ginv(S/N);

* ASYM VAR;
DG = j(nv,Pn,.);
do k = 1 to Pn;
  DGi = j(nv,N,0);
  do i = 1 to N;
    x = Xmat[(i-1)*Tn+1:i*Tn,];
    y = yvec[(i-1)*Tn+1:i*Tn];
    mu = exp(x*t(beta)) / ( 1+exp(x*t(beta)) );
    Rsd = y - mu;
    Dk =  x[,k]#mu#(1- mu);
     Dkz =  x[,k]#(1- 2*mu);
    do p = 1 to Pn;
      Dp = x[,p]#mu#(1- mu);
      Dkzp = Dkz#Dp;
       DGmtx = Rsd*t(Dkzp)-Dk*t(Dp);
      DGi[nloc[p]+1:nloc[p+1],i] =
DGmtx[loc(Tshape[p,]^=0)];
    end;
  end;
  DG[,k]= DGi[,:];
end;

AsymVar = (1/N)*ginv(t(DG)*Wn*DG);
AVvec = vecdiag(AsymVar);
StdDev = sqrt(AVvec);

zvalue = t(beta)/StdDev;
pvalue = 2*(1-cdf('normal',abs(zvalue)));

Outmtx = j(Pn,4,.);
Outtitle={'Estimate'  'StdDev'  'Zvalue'  'Pvalue'};
Outmtx[,1]=t(beta);
Outmtx[,2]=StdDev;
Outmtx[,3]=zvalue;
Outmtx[,4]=pvalue;
print Outtitle;
print Outmtx;

quit;
```

APPENDIX B

SAS CODE USING PROC IML FOR PHILIPPINE DATA

```
/*############################################
* read data and create time dummies;
############################################*/
libname perm 'c:\SAS\perm';
data mydata; set perm.Philipppine;
if time=1 then t1=1; else t1=0;
if time=2 then t2=1; else t2=0;
if time=3 then t3=1; else t3=0;
run;

/*############################################
* obtain residuals from by-time regression;
############################################*/
title ' regression by time';
proc sort data=mydata out=mydatasorted;
by time; run;
proc reg data=mydatasorted noprint;
by time;
model y = bmi age gender;
output out=outpool3 p=pred r=rsd; * outpool3 is sorted by
time;
run;

/*##########################################
* examine corr by PROC IML;
##########################################*/
PROC SORT DATA=outpool3 OUT=outpool3 ;
  BY childid time; RUN;
proc iml;
use outpool3;    * ## change ####;
read all VARIABLES {bmi age gender t2 t3 childid time} into
Zmat;
* ## change ####;
read all var {rsd} into rsd;
close outpool3;

N=370;T =3;      * ## change ####;

start rho(a,rsd) global(N,T);
abm = j(N,2*T,.);
abm[,1:T] = shape(rsd,N);
abm[,T+1:2*T] = shape(a,N);
corr = corr(abm);
rho = corr[1:T,T+1:2*T];
return(rho);
finish rho;

start stddev(a,rsd) global(N,T);
bm = shape(rsd,N);
bdev = bm-j(N,1,1)*bm[:,];
```

```
bdev2 = bdev#bdev;
am = shape(a,N);
adev = am-j(N,1,1)*am[:,];
adev2 = adev#adev;
stddev = sqrt( (1/N)*t(bdev2)*adev2 );
return(stddev);
finish stddev;

start stdzn(x);
N = nrow(x);
y = x-x[:,];
vcv = (1/(N-1))*t(y)*y;
v = diag(vcv);
sinv = diag(sqrt(1/vcv));
x2 = y*sinv;
return(x2);
finish stdzn;

print 'Corr Examination to Philippine Data';

print 'bmi';
x1 = Zmat[,1];  * bmi;
x2 = stdzn(x1);
rsd2 = stdzn(rsd);
r2 = rho(x2,rsd2); s2 = stddev(x2,rsd2);
z2 = sqrt(N)*(r2/s2);
p2 = 2*(1-cdf('normal',abs(z2)));
print r2, s2, z2, p2;

print 'age';
x1 = Zmat[,2];  * age;
x2 = stdzn(x1);
rsd2 = stdzn(rsd);
r2 = rho(x2,rsd2); s2 = stddev(x2,rsd2);
z2 = sqrt(N)*(r2/s2);
p2 = 2*(1-cdf('normal',abs(z2)));
print r2, s2, z2, p2;
quit;

/*#####################################################
Obtain GEE Results Based on the Independent Correlation
Structure to Serve as Initial Values
#####################################################*/
proc genmod data=mydata;        class childid time;
      model y = bmi age gender t2 t3 / dist=normal
link=identity ;
      repeated subject=childid /within=time corr=indep
corrw;
       output out=GEEout pred = yhat xbeta=xb RESRAW = rraw;
run;
```

```
/*#####################################################
Obtain GMM Estimates
#####################################################*/
proc sort data=mydata;
by childid time; run;

proc iml;
use mydata;   * ## change variable list ####;
read all VARIABLES {bmi age gender } into Zmat;
read all var {y} into yvec;
close mydata;
print '2SGMM with Extended Method to Philippine Data of
Continuous Y';

N=370;   * number of individuals;
Pn=6;    * number of parameters to estimate; * ## change
###;
       * Intercept  bmi      age        gender      time2
time3 ;
beta0 = {-0.3173 -0.1006 -0.0136  0.1542  -0.2760  -
0.0092 };

nr = nrow(Zmat);
nc = ncol(Zmat);
int = j(nr,1,1);

Xmat =j(nr,nc+3,.);
Xmat[,1]=int; Xmat[,2:nc+1]=Zmat;

in = j(N,1,1);
tm2 = {0,1,0}; D2 = in@tm2;
tm3 = {0,0,1}; D3 = in@tm3;
Xmat[,nc+2]=D2; Xmat[,nc+3]=D3;

Tn=3;  * number of periods per observation; ## change ####;
 * Intercept  bmi age  gender t2 t3 ;
J = {1,2,1,1,1,1};   * ext_class bmi 8eq  CORRECT;


T0 = {1 0 0,
      0 1 0,
      0 0 1};

T1 ={1 1 1,
     1 1 1,
     1 1 1};

T2 = {1 1  1,
   1 1  0,
   1 1  1};
    * ## change ####;

Tshape = j(Pn,Tn*Tn,.);
neq = j(Pn,1,0);

do p =1 to Pn;
 if J[p]=1 then Tshape[p,] = shape(T1,1);
  else if  J[p]=2 then Tshape[p,] = shape(T2,1);
        else Tshape[p,] = shape(T0,1);
  neq[p] = ncol(loc(Tshape[p,]^=0));  /
end;   * ## change ####;

* nloc containing the starting/end positions of reg eq's
brought by each covariate  ;
nloc = j(1,Pn+1,0);
do p =1 to Pn;
 nloc[p+1] = sum(neq[1:p]);
end;
nv = sum(neq);

Wn = I(nv);              * initial weight matrix ;
S = j(nv,nv,0);        * to compute covariance mtx ;

start TSGMM(beta)
global(Pn,Tn,N,Xmat,yvec,nv,Tshape,nloc,Wn,S);
Gn = j(nv,1,0);         * to collect valid mmt
conditions;
S = j(nv,nv,0);        * to compute covariance mtx ;
eq = j(nv,N,0);

do i = 1 to N;
 x = Xmat[(i-1)*Tn+1:i*Tn,];
 y = yvec[(i-1)*Tn+1:i*Tn];
 mu =x*t(beta);
 Rsd = y - mu;
 do p = 1 to Pn;
  Eqmtx = Rsd*t(x[,p]);
  eq[nloc[p]+1:nloc[p+1],i] = Eqmtx[loc(Tshape[p,]^=0)];
 end;
 S = S + eq[,i]*t(eq[,i]);
end;
Gn = eq[,:];
f = t(Gn)*Wn*Gn;         * the objective fn to be
minimized;
return(f);
finish TSGMM;

tc = {2000 2000}; optn = {0 2};
```

```
  call NLPNRA(rc, xres,"TSGMM", beta0,optn, , tc);
  beta0 = xres;
  Wn = ginv(S/N);

  call NLPNRA(rc, xres,"TSGMM", beta0,optn, , tc);
  beta = xres;
  Wn = ginv(S/N);

* ASYM VAR;
DG = j(nv,Pn,.);
do k = 1 to Pn;
  DGi = j(nv,N,0);
  do i = 1 to N;
    x = Xmat[(i-1)*Tn+1:i*Tn,];
    do p = 1 to Pn;
        DGmtx = -x[,k]* t(x[,p]);
      DGi[nloc[p]+1:nloc[p+1],i] =
DGmtx[loc(Tshape[p,]^=0)];
    end;
  end;
  DG[,k]= DGi[,:];
end;

AsymVar = (1/N)*ginv(t(DG)*Wn*DG);
AVvec = vecdiag(AsymVar);
StdDev = sqrt(AVvec);

zvalue = t(beta)/StdDev;
pvalue = 2*(1-cdf('normal',abs(zvalue)));

Outmtx = j(Pn,4,.);
Outtitle={'Estimate'  'StdDev'  'Zvalue'  'Pvalue'};
Outmtx[,1]=t(beta);
Outmtx[,2]=StdDev;
Outmtx[,3]=zvalue;
Outmtx[,4]=pvalue;
print Outtitle;
print Outmtx;

quit;
```