

OVERDISPERSION IN POISSON REGRESSION

By

W. Brad McNeney

B. Sc. (Mathematics) University of British Columbia

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF STATISTICS

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

1992

© W. Brad McNeney, 1992

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of STATISTICS

The University of British Columbia
Vancouver, Canada

Date April 29, 1992

Abstract

Investigation of a possible relationship between air quality and human health in the community of Prince George, British Columbia was undertaken after a public opinion poll in 1972 discovered that poor air quality was the number one concern of the residents of Prince George. An analysis which attempted to identify such relationships using a data set including air quality measurements and hospital admissions for the period April 1, 1984 to March 31, 1986 is discussed in Knight, Leroux, Millar, and Petkau (1988). A similar analysis using emergency room visits during the same period rather than hospital admissions is described in Knight, Leroux, Millar, and Petkau (1989). The data set described here was collected to carry out a follow-up study to the emergency room visits analysis.

The main part of the analyses carried out involved the use of Poisson regression models with a minor extension to account for over-dispersion in the data. The results of the analysis were not consistent with either the earlier study or with the expectations of the investigators. For example, higher levels of one of the air quality variables was found to be associated with a *decrease* in the number of emergency room visits for respiratory disease in the winter, but an *increase* in emergency room visits for respiratory disease in the fall. A mechanism to explain such effects is difficult to imagine. These counter-intuitive results motivated a simulation study to assess the methods used in the analysis and to compare these to other possible estimators and test statistics that can be employed in the analysis of over-dispersed Poisson data.

Table of Contents

Abstract	ii
List of Tables	v
List of Figures	vii
Acknowledgement	viii
1 The Prince George Study	1
1.1 Data Set Description	2
1.1.1 Emergency Room Visits Data	2
1.1.2 Air Pollution Data	2
1.1.3 Meteorological Data	3
1.2 Methodology for the Prince George Study	3
1.2.1 Poisson Regression	3
1.2.2 Accounting for Over-Dispersion	8
1.3 Results of the Data Analysis	10
1.3.1 Temporal and Meteorological Effects	11
1.3.2 Respiratory Visits and TRS	12
1.3.3 Other Analyses of Pollution Covariates	15

1.4	Questions Raised by the Analysis	15
2	Methodology	17
2.1	Quasi-Likelihood in Generalized Linear Models	17
2.1.1	The Estimating Equations	19
2.1.2	Test Statistics for Hypotheses Regarding Regression Parameters	24
2.2	Test Statistics for Hypotheses in Over-Dispersed Poisson Regression	27
3	A Simulation Study Related to the Prince George Study	31
3.1	The General Simulation Procedure	34
3.1.1	Generating the Data	34
3.1.2	Fitting the Full Model	35
3.1.3	Fitting the Reduced Model	36
3.2	The Specifics of the Simulation Study	36
4	Results of the Simulation Study	42
4.1	Estimating the Dispersion Parameter ϕ	42
4.2	The Simple Case with a Single Series	50
4.3	The Simple Case with Three Series	60
4.4	The More Complicated Case with a Single Series	71
4.5	The More Complicated Case with Three Series	92
5	Discussion	106
	Bibliography	111

List of Tables

Table 1. Mean Values of Estimates (\pm Standard Deviation) in 1000 Simulated Data Sets. .	43
Table 2. Mean Values of Estimates of β	50
Table 3. Standard Deviations of Estimates of β and Mean Values of Estimated Standard Errors (\pm s.d.).	51
Table 4. Observed Rejection Probabilities Under Reduced Model.	54
Table 5. Mean Values of Estimated Parameters (\pm s.d.) Under Full Model.	57
Table 6. Observed Rejection Probabilities Under Full Model.	59
Table 7. Mean Values of Parameter Estimates.	60
Table 8. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	62
Table 9. Observed Rejection Probabilities: Common Pollution to No Pollution Model.	63
Table 10. Mean values of Parameter Estimates.	64
Table 11. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	65
Table 12. Observed Rejection Probabilities: Separate to Common.	66
Table 13. Observed Rejection Probabilities: Common to Null.	67
Table 14. Mean Values of Parameter Estimates.	68
Table 15. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	69

Table 16. Observed Rejection Probabilities: Separate to Common.	70
Table 17. Mean Values of Estimated Parameters Under Reduced Model.	71
Table 18. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	75
Table 19. Observed Rejection Probabilities Under Null Hypothesis.	78
Table 20. Mean Values of Estimated Parameters Under Alternative Hypothesis.	80
Table 21. Standard Deviations of Estimates Under Alternative Hypothesis.	83
Table 22. Observed Rejection Probabilities Under Alternative Hypothesis.	89
Table 23. Mean Values of Parameter Estimates.	92
Table 24. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	94
Table 25. Observed Rejection Probabilities: Common to Null Model.	95
Table 26. Mean Values of Parameter Estimates.	97
Table 27. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	98
Table 28. Observed Rejection Probabilities: Separate to Common.	100
Table 29. Observed Rejection Probabilities: Common to Null.	100
Table 30. Mean Values of Parameter Estimates.	101
Table 31. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).	103
Table 32. Observed Rejection Probabilities: Separate to Common.	104

List of Figures

Figure 1. Comparison of Estimators of Φ : Correct Variance Simulations ($\Phi=1$ or 1.4)	45
Figure 2. Comparison of Estimators of Φ : Correct Variance Simulations ($\Phi=2$ or 3)	46
Figure 3. Comparison of Estimators of Φ : Misspecified Variance Simulations ($\Phi=1$ or 1.4)	47
Figure 4. Comparison of Estimators of Φ : Misspecified Variance Simulations ($\Phi=2$ or 3)	48

Acknowledgement

I would like to express my gratitude to my advisor, Professor A. John Petkau, for his support, advice and patience throughout the writing of this thesis. I would also like to thank Professor Nancy E. Heckman for reading the manuscript and providing many helpful comments.

Chapter 1

The Prince George Study

Motivation for this work was provided by the analysis of a data set collected in Prince George, British Columbia, to study the possibility of associations between the ambient levels of air pollution and human health as measured by the rates of emergency room visits for respiratory illnesses. Investigation of such a possible relationship was undertaken after a public opinion poll in 1972 discovered that poor air quality was the number one concern of the residents of Prince George, ahead of such issues as crime, alcohol abuse and recreation facilities. Although monitoring of air quality parameters has been carried out at several monitoring stations in Prince George since 1980, most attempts to study this issue have compared hospital admissions rates and/or mortality rates in Prince George to those in other communities in British Columbia rather than attempt to identify an association between ambient levels of air pollution in Prince George and human health.

An analysis which did attempt to identify such relationships using a data set including air quality measurements and hospital admissions for respiratory illnesses for the period April 1, 1984 to March 31, 1986 is discussed in Knight, Leroux, Millar, and Petkau (1988). A similar analysis for the same period based on emergency room visits rather than hospital admissions is described in Knight, Leroux, Millar, and Petkau (1989). The data set described here were collected to carry out a follow-up study to the emergency room visits analysis.

1.1 Data Set Description

Three measures of air quality were monitored and the daily numbers of visits to the emergency room (ER) for respiratory related illnesses were recorded during the period April 1, 1986 to March 31, 1988. In addition, meteorological data in the form of temperature and maximum relative humidity was collected.

Description of the relationship between the pollution variables and the ER visits was the goal of the study. However, meteorological and temporal information was also included as potential explanatory factors so that any relationships discovered with pollution variables could reasonably be attributed to the pollution itself and not to some other factor that might underly both pollution and ER visits.

1.1.1 Emergency Room Visits Data

Each visit to the emergency room at Prince George Regional Hospital by a resident of Prince George over the two year period from April 1, 1986 to March 31, 1988 was originally classified according to 35 diagnostic categories potentially related to respiratory illness. These 35 categories were grouped into four broad categories which we will refer to as Asthma, Bronchitis+, Ear, and Others. Only the classification according to these four broad categories was available in the data set provided to us. A total of 8,079 visits were included in the study and there was one visit in at least one of the four respiratory categories on every day except January 15, 1987.

1.1.2 Air Pollution Data

Data on three air quality parameters were available:

TRS – measured via a continuous monitor with readings subsequently converted to hourly averages. The average of these hourly averages becomes the *daily average* and this was the summary used.

TSP – measured via a 24 hour vacuum sample once every six days.

SO₂ – measured via a continuous monitor in the same fashion as TRS. The *daily average* was the summary used.

1.1.3 Meteorological Data

The meteorological data consisted of daily summaries for each of the following parameters:

Temperature – the average of 24 hourly readings in degrees Celsius.

Maximum Relative Humidity – the largest of 24 hourly readings.

1.2 Methodology for the Prince George Study

The main part of the analyses carried out involved the use of Poisson regression models (or log-linear models). We give a brief description of these models and then describe an extension of Poisson regression models, which was required for the modelling of emergency room visits.

1.2.1 Poisson Regression

The Poisson regression model is a special case of a larger class of models referred to as generalized linear models (GLMs), the theory of which is derived largely from the properties of the exponential family of distributions. Thus before we describe the Poisson regression model in

detail, a brief discussion of the exponential family and GLMs seems appropriate (see McCullagh and Nelder, 1989, for a more detailed description).

For a random variable Y in the exponential family with canonical parameter θ and (known) nuisance parameter ϕ , we may write its density function as $f(y; \theta) = \exp\{(y\theta - b(\theta))/a(\phi) + c(y; \phi)\}$. Denote the mean value of Y as μ , which is related to the parameter θ by the function $\theta = g(\mu)$. We will refer to the function $g(\cdot)$ as the link function.

For making inference about θ we introduce the log-likelihood function, the log of the above density function, which we denote $l(\theta)$. In the case of n independent observations Y_i ($i = 1, \dots, n$), with possibly different canonical parameters, the joint log-likelihood is given by $l(\theta) = \sum_{i=1}^n l_i(\theta_i)$. In GLMs a linear model is specified for the parameters θ_i of the form $\theta_i = g(\mu_i) = X_i\beta$ (where β is a vector of $p \leq n$ parameters and X_i is the i^{th} row of an $n \times p$ design matrix). It is then possible to estimate β via maximum likelihood by taking the derivative of the joint log-likelihood with respect to each of the parameters, equating the resulting p equations to zero and solving for the parameters of interest. These estimators have several nice characteristics including a limiting multivariate normal distribution with mean vector equal to β^* , the true value of the vector of parameters, and covariance matrix equal to the inverse of the Fisher Information matrix corresponding to the joint distribution of the Y_i ($i = 1, \dots, n$).

Now consider a Poisson random variable Y . We write the density function of Y as $f(y; \mu) = (e^{-\mu}\mu^y)/y!$ or equivalently as $f(y; \mu) = \exp\{y \log \mu - \mu - \log y!\}$. The latter form allows us to identify the Poisson distribution as a member of the exponential family with $\theta = \log \mu$ (and therefore link function given by $g(\cdot) = \log(\cdot)$), $b(\theta) = \mu$, $a(\phi) = 1$, and $c(y; \phi) = -\log y!$.

The motivation behind using the Poisson regression model for the Prince George study lies

in the intuitively reasonable assumption that the daily numbers of emergency room visits follow Poisson distributions with possibly different means. The mean number of visits on a given day will be described by a function of the available covariate information up to and including that day, which includes temporal variables such as day of the week and month, the meteorological variables, daily temperature and maximum humidity, and the pollution variables, TSP, TRS and SO₂. The use of Poisson regression depends on the following basic assumptions:

- (a) the expected number of visits on a given day depends explicitly only on time, meteorology and pollution and not on the number of visits on past days.
- (b) daily visits are conditionally independent given the temporal, meteorological and pollution covariates.

The assumption that the expected number of visits does not depend explicitly on past visits seems to be a reasonable starting point in this situation. However, more general models in which the mean depends explicitly on past visits could be considered as well.

Suppose now that t_i is a vector of temporal covariates, m_i is a vector of meteorological covariates and s_i is a vector of pollution covariates for day i ; s_i and m_i can include values of pollution or meteorology prior to day i as well as those values on day i . If y_i is the number of emergency room visits on day i ($i = 1, 2, \dots, 731$) for a particular diagnostic class, the log-linear model assumes that y_i has a Poisson distribution with mean μ_i where

$$\log \mu_i = \alpha + t_i' \tau + m_i' \delta + s_i' \gamma$$

with α , τ , δ and γ unknown parameters (τ , δ and γ are vectors having the same length as t_i , m_i and s_i respectively). The components of the parameter γ give the relative changes in expected

visits on a given day (that is, with temporal and meteorological covariates fixed) associated with changes in the pollution variables. In particular, for small changes in the pollution variables, from s_i to $s_i + \Delta_i$, $\exp(\Delta_i' \gamma) - 1 \approx \Delta_i' \gamma$ gives the percentage change in the corresponding expected visits.

In order to account for possibly different patterns of emergency room visits for the different diagnostic categories, the basic model is generalized by allowing different parameters α , τ , δ and γ for the different diagnostic categories, and thus describing the expected number of visits for category j ($j = 1, 2, 3, 4$) on day i by the equation

$$\log \mu_{ij} = \alpha_j + t_i' \tau_j + m_i' \delta_j + s_i' \gamma_j.$$

To fit the above model into the framework described for GLMs let $\mu = (\mu_1^t, \mu_2^t, \mu_3^t, \mu_4^t)^t$ and $\beta = (\alpha^t, \tau^t, \delta^t, \gamma^t)^t$ where here $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)^t$, $\tau = (\tau_1^t, \tau_2^t, \tau_3^t, \tau_4^t)^t$, and so on. Then, if X is the appropriate design matrix containing all of the covariate information, our model can be expressed as $\log \mu = X\beta$.

In addition to assuming the observations to be conditionally independent given the covariates, we assume that data from different diagnostic categories is independent, so that the joint likelihood is simply the product of the marginal Poisson distributions. Thus the joint log-likelihood is:

$$l(\beta) = \sum_{j=1}^4 \sum_{i=1}^{731} [y_{ij} \log \mu_{ij} - \mu_{ij} - \log y_{ij}!].$$

For ease of notation, it will be more convenient to write the above log-likelihood using one subscript. We order the y_{ij} 's as y_1, \dots, y_{2924} , with the μ_i 's numbered correspondingly. Now

express the joint log-likelihood as:

$$l(\beta) = \sum_{i=1}^{2924} [y_i \log \mu_i - \mu_i - \log y_i!]. \quad (1.1)$$

To obtain maximum likelihood estimating equations for the parameters in β (suppose there are p parameters in β), we differentiate the expression in (1.1) with respect to each of the p components of β and equate the resulting derivatives to zero to arrive at:

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{2924} \left(\frac{y_i - \mu_i}{\mu_i} \right) \frac{\partial \mu_i}{\partial \beta_j} = 0; \quad j = 1, \dots, p. \quad (1.2)$$

Note that our model assumes $\mu_i = \exp\{X_i\beta\}$, or $\log \mu_i = X_i\beta$. Thus, (1.2) can be rewritten as

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{2924} (y_i - \mu_i) X_{ij} = 0; \quad j = 1, \dots, p. \quad (1.3)$$

The resulting estimators, $\hat{\beta}$, are asymptotically normal, with mean vector given by β^* , the true value of the vector of parameters, and covariance matrix equal to

$$\left[-E \left(\frac{\partial^2 l}{\partial \beta^2} \right) \right]^{-1}.$$

In this case

$$-\frac{\partial^2 l}{\partial \beta^2} = \sum_1^{2924} X_i^t X_i \mu_i = -E \left(\frac{\partial^2 l}{\partial \beta^2} \right).$$

Given the solution to (1.3), the fitted values, $\hat{\mu}_i = \exp\{X_i\hat{\beta}\}$, may be substituted in the above expression to obtain an estimate of the asymptotic covariance matrix. Using these fitted values, the *deviance* is then defined to be -2 times the log-likelihood, or

$$\text{dev.} = -2 \sum_{i=1}^{2924} [y_i \log \hat{\mu}_i - \hat{\mu}_i - \log y_i!].$$

For a fitted model with p parameters and a joint likelihood from n independent observations, the degrees of freedom associated with the deviance is $n - p$. By taking the difference in the

deviances of two fitted models, one of which is nested within the other, we may form a likelihood ratio statistic that can be compared to its limiting chi-square distribution to examine if the smaller model constitutes a significant deterioration in fit compared to the larger model.

The Pearson residuals are defined to be

$$r_i = \frac{(y_i - \hat{\mu}_i)}{\hat{\mu}_i}.$$

These quantities are often used for diagnostic purposes in the same way the residuals from an ordinary regression (assuming normality) are used. We may also use these residuals to form a goodness-of-fit statistic, the familiar Pearson X^2 , by summing the squares of these Pearson residuals. Another potential use of the Pearson X^2 emerges in the following section dealing with over-dispersion.

1.2.2 Accounting for Over-Dispersion

The extension of the Poisson regression model to be described concerns accounting for over-dispersion. Suppose that the variance of the daily visits given the temporal, meteorological and pollution variables is greater than the mean; in this situation the emergency room visits data are said to be “over-dispersed”. Over-dispersion represents a violation of the Poisson assumption underlying the Poisson regression model, which requires that the variance and mean be equal. One way to handle over-dispersed data is to introduce a dispersion parameter $\phi > 1$ and assume that the variance of the visits on day i is $\phi\mu_i$. We now describe the procedure for statistical inference in this new model.

The parameter vector β which appears in the specification of the daily mean visits rates is estimated exactly as for the model without a dispersion parameter, i.e. by maximizing the

Poisson log-likelihood function of (1.1). However, when assessing the statistical significance of estimated parameters it is necessary to take into account the dispersion parameter. To describe how this is done in the context of model selection (the dispersion parameter must also be taken into account when evaluating the standard errors of estimated parameters), assume that we begin with a model containing a number of parameters (the full model) and wish to test the viability of another model which is nested within the full model (the reduced model). Under our independence assumptions, the difference in the deviances for the two models is approximately distributed as ϕ times a χ^2 random variable with degrees of freedom given by the difference in the degrees of freedom associated with the two deviances. Thus the difference in deviances must be divided by an estimate of the dispersion parameter before it can be compared to the χ^2 distribution to assess whether reduction from the full model to the reduced model is permissible.

One choice of an estimate of the dispersion parameter is the ratio of the sum of squared Pearson residuals, the Pearson X^2 statistic, to its degrees of freedom (McCullagh and Nelder, 1989, p.200), both determined under the full model. Another estimate, which is often very close to the previous one, is the ratio of the deviance to its degrees of freedom, also under the full model. For this data the estimate based on the deviance was used; this leads to the following statistic for testing whether reduction from the full model to the reduced model is permissible:

$$\Delta'_{\text{dev.}} = \frac{\text{deviance (reduced)} - \text{deviance (full)}}{\text{deviance (full)/df (full)}};$$

this statistic is compared to the χ^2 distribution with degrees of freedom given by $\Delta \text{df} = \text{df (reduced)} - \text{df (full)}$.

The Poisson regression model with unspecified dispersion parameter is a particular case of a

class of models called “quasi-likelihood” models, so-named because the likelihood which is used to obtain estimates, the Poisson likelihood in our case, is not the true likelihood for the data but plays the role of a likelihood and so is called a quasi-likelihood. Quasi-likelihood models are discussed in more detail in Chapter 2.

We now describe the results of the analysis of the Prince George data using the methods described above.

1.3 Results of the Data Analysis

The possibility of over-dispersion in a series of counts can be judged by the index of dispersion:

$$X^2 = \frac{(n-1)\text{variance}}{\text{mean}},$$

where n is the sample size, 731 in this case. Under sampling from a Poisson distribution, X^2 has an approximate chi-squared distribution with $n-1$ degrees of freedom and can be converted to an approximate standard normal variable by $z = \sqrt{2X^2} - \sqrt{2(n-1)-1}$.

The results for the emergency room visits data are presented the following table:

category	mean	variance	X^2	z
Asthma	0.76	0.80	763	0.87
Bronchitis+	1.54	3.36	1593	18.2
Ear	3.62	6.92	1395	14.6
Other	5.13	14.2	2027	25.5

Except for the Asthma category, these exhibit a large degree of over-dispersion relative to the Poisson distribution; the values of z given in the above table clearly reflect extreme over-dispersion for the categories Bronchitis+, Ear and Other. In fact, these distributions exhibit the pattern of a mixture of Poisson distributions, having higher than expected numbers of very

small and very large counts, at the expense of moderate counts. This over-dispersion must be taken into account in the data analysis.

As mentioned previously, we must allow for the possibility that emergency room visits depend on factors other than just pollution, such as temporal and meteorological effects. These covariates were considered first, and model reductions were attempted to find a model which parsimoniously describes these effects. The pollution covariates were then added to the model and model reductions were attempted to identify the pollution covariates which had a substantial effect on emergency room visits for respiratory illness.

1.3.1 Temporal and Meteorological Effects

The first step of the analysis involved the incorporation of adjustments for temporal effects. The full model for the log of the visits rate included separate effects for each of the 12 months and each of the 7 days of the week for each diagnostic category. Attempted model reductions to common month or common day-of-the-week effects, that is common across all four diagnostic categories, resulted in unacceptably large changes in adjusted deviance and therefore the full collection of separate effects for each diagnosis was retained as the temporal model.

Next, the meteorological covariates (temperature and maximum relative humidity) were added to the temporal model at lags of 0 to 3 days; that is for any $i = 4, \dots, 731$ we allow for the possibility that meteorology at times i , $i - 1$, $i - 2$ and $i - 3$ might affect the rate of emergency room visits at time i . We first allow for separate effects at each lag for each of the four seasons, winter (December to February), spring (March to May), summer (June to August) and fall (September to November), but with each of these effects common to all four diagnostic

categories. Based on our model reduction criterion, these separate season effects were found to be unimportant compared to common effects across seasons for each lag.

When separate meteorological effects at each lag for each diagnostic category were considered, we found unacceptable increases in adjusted deviance when attempting to simplify the separate temperature effects to common effects. On the other hand, only common maximum humidity effects were needed. Thus our model which adjusts for possible temporal and meteorological effects included separate diagnosis effects for month, day-of-the-week and temperature, and common maximum humidity effects.

1.3.2 Respiratory Visits and TRS

We next investigated a model which expresses the daily emergency room visits rates for any particular diagnosis category as

$$\begin{aligned} \log \text{ visits rate on day } i = & \text{temporal effects} + \text{meteorological effects} + \\ & \gamma_0 \log \text{TRS}_i + \gamma_1 \log \text{TRS}_{i-1} + \gamma_2 \log \text{TRS}_{i-2} + \gamma_3 \log \text{TRS}_{i-3}. \end{aligned}$$

Here the effects of each of the TRS values are common across diagnosis categories; we also considered models involving separate effects for the four diagnosis categories or separate effects for the four seasons. For all models fitted, the temporal and meteorological effects had the same structure as previously identified. In what follows, we will abbreviate this structure as the tem/met model.

We started with separate effects for each diagnostic category but found the reduction to common effects acceptable and furthermore removal of the TRS effects altogether resulted in

only a negligible deterioration in the fit. The following table examines separate TRS effects for each season, where as before the separate season effects are common to all diagnostic categories:

model terms	deviance	d.f.	Δ' dev.	Δ d.f.
tem/met effects + separate season TRS effects	3827	2804	18.4	12
tem/met effects + common TRS effects	3852	2816	1.7	4
tem/met effects only	3855	2820		

Reduction to common effects leads to an increase in adjusted deviance of 18.4 versus a decrease of 12 in the number of parameters which is not so easily dismissed ($p=0.11$). Unfortunately the estimated effects from the full model present a confusing picture as we see in the following table which presents those estimates which are greater than one standard error in magnitude:

effect	coefficient	s.e.
Winter Lag 0	-0.026	0.020
Spring Lag 0	-0.027	0.026
Summer Lag 0	0.043	0.028
Winter Lag 2	-0.035	0.023
Summer Lag 2	-0.035	0.029

Looking at the effects at each lag we get the suggestion there may be some TRS effect at lags of 0 and 2 days. Unfortunately the effects from one lag to the other are not consistent and furthermore, it is difficult to understand why higher levels of TRS should be associated with higher numbers of emergency room visits at some times of the year and lower numbers at other times of the year. These results may not be particularly surprising since the overall reduction in adjusted deviance for going from the full model to one with the tem/met effects only ($p=0.20$) indicates that the separate effects in the above table could be large simply by chance.

However, we could also examine the effects grouped by season as in the following table:

Lag	Winter	Spring	Summer	Fall
0	-0.026	-0.027	0.043	-
1	-	-	-	-
2	-0.035	-	-0.035	-
3	-	-	-	-

where the dashes represent estimated effects which are less than their standard error. Looking down the columns of this table we see the possibility of a negative cumulative effect in Winter. Accordingly we now present results for a model which includes a single cumulative effect (an average of the logarithm of the lag 0 to lag 3 pollution measurements) for each of the four seasons:

model terms	deviance	d.f.	Δ' dev.	Δ d.f.
tem/met effects + separate season TRS effects	3827	2804	6.7	12
tem/met effects + cumulative TRS effects	3837	2816	13.4	4
tem/met effects only	3855	2820		

While there is little difference in the fit of the separate season effects model and the cumulative effects model, the cumulative TRS effects constitute a substantial improvement in the fit over the tem/met only model (Δ' dev.=13.4, Δ d.f.=4, $p=0.001$). The estimates and standard errors of these cumulative effects are presented in the following table:

effect	coefficient	s.e.
Winter	-0.060	0.023
Spring	-0.002	0.033
Summer	0.009	0.034
Fall	0.075	0.030

These effects suggest that an increase in the level of TRS is associated with a *decrease* in the number of emergency room visits for respiratory disease in the winter, but an *increase* in the number of emergency room visits for respiratory disease in the fall. This model provides a more

parsimonious description of the effect of TRS on emergency room visits for respiratory disease than the season effects model, yet a mechanism to explain such effects is difficult to imagine.

1.3.3 Other Analyses of Pollution Covariates

Similar analyses were carried out for SO₂ (as well as TRS and SO₂ simultaneously) and for TSP. Details of the model reductions and results can be found in McNeney and Petkau (1991). These analyses proceeded in the same fashion as the TRS analysis described above, and similar problems were encountered; that is, the parameter estimates obtained in the final models were such that no plausible explanation for their structure came to mind.

1.4 Questions Raised by the Analysis

The results of the data analysis described above make it difficult to come up with a reasonable and coherent explanation of the effects of air pollution on human health based on this data set. There are many possible reasons for our results, such as:

- (1) the air pollution data do not accurately represent the exposure of the residents of Prince George to pollutants,
- (2) emergency room visits for respiratory illnesses are poor indicators of the effects of air pollution on human health,
- (3) some other important factor was not considered in the analysis, or
- (4) the method of analysis is potentially misleading.

With respect to the last item, the following questions come to mind:

- (a) Is the estimate of the scale parameter, ϕ , used in this analysis, namely $\frac{\text{deviance}(\text{full})}{\text{d.f.}(\text{full})}$, a good estimate, or should we have used $\frac{\text{Pearson } X^2(\text{full})}{\text{d.f.}(\text{full})}$?
- (b) Is the change in adjusted deviance criterion appropriate for model reduction in this situation?
- (c) Are statistics other than the adjusted deviance better for testing hypotheses concerning the regression parameters when over-dispersion is present?
- (d) Given the amounts of over-dispersion in the emergency room visits series, are the estimates of the model parameters and standard errors accurate?

These questions motivated the simulation study we describe in Chapters 3 and 4. Before proceeding to the simulation study, Chapter 2 presents a more detailed description of quasi-likelihood theory and the estimating equations obtained from this theory for estimating regression and dispersion parameters. We also discuss some of the test statistics available for testing the significance of regression parameters. Finally, over-dispersed Poisson regression is discussed as a special case of quasi-likelihood.

Chapter 2

Methodology

In this section we describe the aspects of maximum likelihood theory (in the context of GLMs) which motivate the theory of quasi-likelihood first proposed by Wedderburn (1974) and also discussed by McCullagh and Nelder (1989, Chapter 9). In particular, we are interested in the estimating equations and test statistics derived from this theory. We also discuss the application of quasi-likelihood to the analysis of the Prince George data.

2.1 Quasi-Likelihood in Generalized Linear Models

Recall the log-likelihood for a single random variable in the exponential family

$$l(\theta) = (y\theta - b(\theta))/a(\phi) + c(y; \phi).$$

Maximum likelihood theory assumes sufficient regularity conditions to ensure that the following relations hold

$$E \left[\frac{\partial l}{\partial \theta} \right] = 0; \text{ and } E \left[\frac{\partial^2 l}{\partial \theta^2} \right] + E \left[\left(\frac{\partial l}{\partial \theta} \right)^2 \right] = 0.$$

In this discussion we will assume these conditions hold. The derivative of the log-likelihood is $\partial l / \partial \theta = (y - b'(\theta))/a(\phi)$, therefore $E(\partial l / \partial \theta) = 0$ implies that $E(Y) = b'(\theta)$. We also have

$$-\frac{\partial^2 l}{\partial \theta^2} = \frac{b''(\theta)}{a(\phi)} \text{ so that } -E \left[\frac{\partial^2 l}{\partial \theta^2} \right] = \frac{b''(\theta)}{a(\phi)}.$$

Finally

$$E \left[\left(\frac{\partial l}{\partial \theta} \right)^2 \right] = E \left[\left(\frac{(Y - b'(\theta))}{a(\phi)} \right)^2 \right] = \frac{1}{a^2(\phi)} \text{Var } Y,$$

which leads to the expression $\text{Var}(Y) = b''(\theta)a(\phi)$.

Rewriting the log-likelihood in terms of the mean μ , we have $l(\theta) = l(g(\mu))$ and

$$\frac{\partial l}{\partial \mu} = \frac{\partial l}{\partial \theta} \frac{\partial \theta}{\partial \mu},$$

so that

$$\frac{\partial l}{\partial \mu} = \left(\frac{y - b'(\theta)}{a(\phi)} \right) \frac{\partial \theta}{\partial \mu} = \frac{y - b'(\theta)}{a(\phi) \frac{\partial \mu}{\partial \theta}}.$$

Since

$$\mu = b'(\theta) \text{ and } b''(\theta) = \frac{\partial b'(\theta)}{\partial \theta} = \frac{\partial \mu}{\partial \theta},$$

the score function $S(\mu) = \frac{\partial l}{\partial \mu}$ can be rewritten as $S(\mu) = (y - E(Y))/\text{Var}(Y)$. Note the following properties of the score function:

1. $E(S) = 0$,
2. $\text{Var}(S) = E(S^2) = \frac{1}{(\text{Var} Y)^2} E(Y - \mu)^2 = \frac{1}{\text{Var}(Y)}$,
3. $-E\left(\frac{\partial S}{\partial \mu}\right) = \frac{1}{\text{Var}(Y)}$.

Now consider a random variable Y for which we do not specify a distribution but only suppose it has mean value μ and a variance that is related to μ by some function $V(\mu; \phi)$ where ϕ is an unknown parameter. Then if $V(\mu; \phi)$ correctly specifies the variance, the function

$$U^*(Y; \mu, \phi) = \frac{(Y - \mu)}{V(\mu; \phi)}$$

has similar properties to a score function resulting from a log-likelihood in that:

1. $E(U^*) = 0$,
2. $\text{Var}(U^*) = E(U^{*2}) = \frac{1}{V^2(\mu; \phi)} E(Y - \mu)^2 = \frac{1}{V(\mu; \phi)}$,
3. $-E\left(\frac{\partial U^*}{\partial \mu}\right) = E\left[\frac{1}{V(\mu; \phi)} + (Y - \mu) \frac{\frac{\partial}{\partial \mu} V(\mu; \phi)}{V^2(\mu; \phi)}\right] = \frac{1}{V(\mu; \phi)}$

Since most asymptotic likelihood theory is based on the above three properties, we might expect the integral of U^* to behave like a log-likelihood. We refer to the integral

$$Q(y; \mu, \phi) = \int_{\mu}^y U^*(y; t, \phi) dt = \int_{\mu}^y \frac{(y - t)}{V(t; \phi)} dt,$$

if it exists, as the quasi-log-likelihood or simply the quasi-likelihood. For n independent observations having the same unknown parameter ϕ we have the quasi-likelihood

$$Q_n(y; \mu, \phi) = \sum_{i=1}^n Q(y_i; \mu_i, \phi)$$

where now y and μ are vectors of length n . As in GLMs, the mean vector μ is assumed to be related to the regression parameters by the link function $g(\mu) = X\beta$.

2.1.1 The Estimating Equations

As in the case where we have a true likelihood, the approach with a quasi-likelihood is to differentiate Q_n with respect to each of the components of β and equate the resulting p derivatives to zero to obtain maximum quasi-likelihood estimators for the parameters. If ϕ were known, then under the assumption that $\mu_i = \exp\{X_i\beta\}$, we would solve the following estimating equations:

$$U_j(\beta, \phi) = \frac{\partial Q_n}{\partial \beta_j} = \sum_{i=1}^n \frac{(y_i - \mu_i)}{V(\mu_i; \phi)} \frac{\partial \mu_i}{\partial \beta_j} = 0; \quad j = 1, \dots, p. \quad (2.4)$$

We will refer to the p -dimensional vector $(U_1, \dots, U_p)'$ as U^β .

Because ϕ is unknown, we must introduce an additional estimating equation. This equation is based on the obvious notion that the variance function $V(\mu_i; \phi)$ should reflect the true variances of the observed y_i (denote these σ_i^2). Thus, ϕ should be estimated so that $V(\mu_i; \phi)$, the model-based variance, agrees well, in some sense, with the true variance σ_i^2 . Only estimates of the true variances are available from the data, so a sensible estimate of ϕ is provided by the solution to the “moment” equation

$$U^\phi(\beta, \phi) = \sum_{i=1}^n \left[\frac{(y_i - \mu_i)^2}{V(\mu_i; \phi)} - 1 \right] = 0. \quad (2.5)$$

To correct for the bias resulting from the simultaneous estimation of the p regression parameters, Breslow (1984) suggests modifying this equation slightly to

$$U^\phi(\beta, \phi) = \sum_{i=1}^n \left[\frac{(y_i - \mu_i)^2}{V(\mu_i; \phi)} - \frac{(n-p)}{n} \right] = 0 \quad (2.6)$$

in place of (2.5).

Given an initial estimate of ϕ we may use (2.4) to obtain an estimate of β . This estimate of β can then be used in (2.6) to estimate ϕ . This procedure is iterated to obtain the joint solution $(\tilde{\beta}, \tilde{\phi})$ to the estimating equations (2.4) and (2.6). Subject to regularity conditions (Inagaki, 1973; White, 1982) this joint solution converges in probability, as the number of observations increases, to (β^*, ϕ^*) where β^* is the true value of β , and ϕ^* is the value of ϕ which satisfies the limit equation

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \frac{\sigma_i^2}{V(\mu_i; \phi)} = 1.$$

To motivate the limiting distribution of the estimator $(\tilde{\beta}, \tilde{\phi})$ consider the following. Let $\theta = (\beta^t, \phi)^t$ and let $U(\theta) = (U^{\beta^t}, U^\phi)^t$, where U^β is the vector form of the equations given by (2.4) and U^ϕ is given in (2.6). While (2.6) is not a maximum quasi-likelihood equation,

we may consider (2.4) and (2.6) as moment-type estimating equations. The derivation of the asymptotic properties of $(\tilde{\beta}, \tilde{\phi})$ will proceed along the same lines as the derivation of the properties of maximum likelihood estimators, although some modifications are required.

The fact that $\tilde{\theta}$ is the solution to $U(\tilde{\theta}) = 0$ along with the Mean Value Theorem allows us to write the Taylor series expansion of $U(\tilde{\theta})$ about the point θ^* as

$$0 = U(\tilde{\theta}) = U(\theta^*) + \left. \frac{\partial U}{\partial \theta^t} \right|_{\theta^m} (\tilde{\theta} - \theta^*),$$

where each of the $p + 1$ elements of θ^m are between the corresponding elements of $\tilde{\theta}$ and θ^* .

Re-arrangement allows us to write this as

$$\sqrt{n}(\tilde{\theta} - \theta^*) = \left(-\frac{1}{n} \frac{\partial U}{\partial \theta^t} \right|_{\theta^m} \right)^{-1} \frac{1}{\sqrt{n}} U(\theta^*). \quad (2.7)$$

Under mild regularity conditions, $\tilde{\theta}$ will converge in probability to θ^* (see Moore 1986, Theorem 1, p.586). Then, since θ^m is between $\tilde{\theta}$ and θ^* and $\partial U / \partial \theta^t$ is a sum of n independent components, we have

$$-\frac{1}{n} \frac{\partial U}{\partial \theta^t} \Big|_{\theta^m} = -\frac{1}{n} \sum_{i=1}^n \frac{\partial U_i}{\partial \theta^t} \Big|_{\theta^m} \xrightarrow{p} \lim_{n \rightarrow \infty} -\frac{1}{n} \sum_{i=1}^n E \frac{\partial U_i(\theta^*)}{\partial \theta^t} = \lim_{n \rightarrow \infty} -\frac{1}{n} E \frac{\partial U(\theta^*)}{\partial \theta^t},$$

where here the subscript i refers to the contribution of the i^{th} observation; $i = 1, \dots, n$ and the expectation is taken with respect to the true value θ^* under the assumption that $\mu = \exp\{X\beta\}$ correctly specifies the mean. Now $-E[\partial U(\theta^*) / \partial \theta^t]$ is a $(p + 1) \times (p + 1)$ matrix of the form

$$\begin{bmatrix} A_n & b_n \\ c_n^t & d_n \end{bmatrix},$$

where

$$A_n = -E \frac{\partial U^\beta}{\partial \beta^t} = -E \frac{\partial}{\partial \beta^t} \left[\sum_{i=1}^n \frac{(Y_i - \mu_i)}{V(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \right]_{\beta^*}$$

$$\begin{aligned}
&= E \sum_{i=1}^n \frac{1}{V(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta^t} \Big|_{\beta^*} + E \sum_{i=1}^n \frac{(Y_i - \mu_i)}{V^2(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \frac{\partial V}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta^t} \Big|_{\beta^*} \\
&= \sum_{i=1}^n \frac{1}{V(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta^t} \Big|_{\beta^*};
\end{aligned}$$

$$b_n = -E \frac{\partial}{\partial \phi} \left[\sum_{i=1}^n \frac{(Y_i - \mu_i)}{V(\mu_i; \phi)} \frac{\partial \mu_i}{\partial \beta} \Big|_{\beta^*} \right]_{\phi^*} = E \left[\sum_{i=1}^n \frac{(Y_i - \mu_i)}{V^2(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \frac{\partial V}{\partial \phi} \Big|_{\theta^*} \right] = 0;$$

$$c_n = -E \frac{\partial}{\partial \beta} \left[\sum_{i=1}^n \left\{ \frac{(Y_i - \mu_i)^2}{V(\mu_i; \phi^*)} - \frac{(n-p)}{n} \right\} \right]_{\beta^*};$$

and

$$d_n = -E \frac{\partial}{\partial \phi} \left[\sum_{i=1}^n \left\{ \frac{(Y_i - \mu_i)^2}{V(\mu_i; \phi)} - \frac{(n-p)}{n} \right\} \right]_{\phi^*} = E \sum_{i=1}^n \frac{(Y_i - \mu_i)^2}{V^2(\mu_i; \phi)} \frac{\partial V}{\partial \phi} \Big|_{\phi^*} = \sum_{i=1}^n \frac{\sigma_i^2}{V^2(\mu_i; \phi)} \frac{\partial V}{\partial \phi} \Big|_{\phi^*}.$$

Now consider the other random vector entering (2.4):

$$\frac{1}{\sqrt{n}} U(\theta^*) = \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{(Y_i - \mu_i)}{V(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \Big|_{\beta^*} \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{(Y_i - \mu_i)^2}{V(\mu_i; \phi^*)} - \frac{(n-p)}{n} \right] \end{pmatrix}.$$

Assuming the regression model correctly specifies the mean, we have

$$E \frac{1}{\sqrt{n}} U(\theta^*) = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{\sigma_i^2}{V(\mu_i; \phi^*)} - \frac{(n-p)}{n} \right] \end{pmatrix}.$$

By the definition of ϕ^* and provided regularity conditions hold (see Moore 1986, Lemma 1, p.585), it follows that this random vector has a limiting multivariate normal distribution

$$N \left(0, \begin{bmatrix} B & e \\ e^t & f \end{bmatrix} \right)$$

where $B = \lim_{n \rightarrow \infty} \frac{1}{n} B_n$ is the normed limit of

$$B_n = E(U^\beta U^{\beta^t}) = E \left[\sum_{i=1}^n \left[\frac{(Y_i - \mu_i)}{V(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \right] \sum_{i=1}^n \left[\frac{(Y_i - \mu_i)}{V(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \right]^t \right]_{\theta^*}.$$

Because the observations are assumed to be independent, this reduces to

$$B_n = E \left[\sum_{i=1}^n \frac{(Y_i - \mu_i)^2}{V^2(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta^t} \right]_{\theta^*} = \sum_{i=1}^n \frac{\sigma_i^2}{V^2(\mu_i; \phi^*)} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta^t} \Big|_{\theta^*}.$$

For our purposes, the forms of e and f are of no consequence.

Combining these results with Slutsky's theorem we have

$$\sqrt{n}(\tilde{\theta} - \theta^*) \xrightarrow{d} N \left(0, \begin{bmatrix} A & 0 \\ c^t & d \end{bmatrix}^{-1} \begin{bmatrix} B & e \\ e^t & f \end{bmatrix} \left(\begin{bmatrix} A & 0 \\ c^t & d \end{bmatrix}^{-1} \right)^t \right),$$

where $A = \lim_{n \rightarrow \infty} \frac{1}{n} A_n$, $c = \lim_{n \rightarrow \infty} \frac{1}{n} c_n$, and $d = \lim_{n \rightarrow \infty} \frac{1}{n} d_n$. As

$$\begin{bmatrix} A & 0 \\ c^t & d \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ -\frac{c^t}{d} A^{-1} & d^{-1} \end{bmatrix},$$

the asymptotic variance of $\sqrt{n}(\tilde{\theta} - \theta^*)$ is

$$\begin{bmatrix} A^{-1} & 0 \\ -\frac{c^t}{d} A^{-1} & d^{-1} \end{bmatrix} \begin{bmatrix} B & e \\ e^t & f \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ -\frac{c^t}{d} A^{-1} & d^{-1} \end{bmatrix}^t.$$

The asymptotic covariance matrix for $\sqrt{n}(\tilde{\beta} - \beta^*)$ is the upper left-hand $p \times p$ sub-matrix of the above matrix, namely $A^{-1} B A^{-1}$. Note that this is exactly the covariance matrix one would obtain for the estimation of β based on the estimating equations (2.4), with ϕ fixed at ϕ^* rather than estimated; the asymptotic covariance for $\tilde{\beta}$ is unaffected by the estimation of ϕ . This occurs despite the fact that $\tilde{\beta}$ and $\tilde{\phi}$ are, in general, correlated (the entries of the upper right-hand $p \times 1$ vector of the above matrix need not be zero); it is a result of the fact that the $p \times 1$ vector b_n has all p elements identically equal to 0, that is, $E[\frac{\partial}{\partial \phi} U^\beta] = 0$. Note further that if the function $V(\mu_i; \phi)$ correctly specifies the variances σ_i^2 , then $A = B$ and $\sqrt{n}(\tilde{\beta} - \beta^*)$ has asymptotic covariance matrix given by A^{-1} .

Thus we can consider two estimates of the asymptotic covariance of $\tilde{\beta}$; the first is model-based and assumes that $V(\mu_i; \phi)$ correctly specifies the variances. Denote this estimate $\tilde{\Sigma}_M =$

\tilde{A}_n^{-1} where \tilde{A}_n is A_n with μ_i replaced by the fitted values $\tilde{\mu}_i = \mu_i(X_i; \tilde{\beta}_i)$ and ϕ^* replaced by $\tilde{\phi}$. The second estimate we refer to as the empirical estimate. Denote this $\tilde{\Sigma}_E = \tilde{A}_n^{-1} \tilde{B}_n \tilde{A}_n^{-1}$ where

$$\tilde{B}_n = \sum_{i=1}^n \frac{(y_i - \tilde{\mu}_i)^2}{V^2(\tilde{\mu}_i; \tilde{\phi})} \frac{\partial \tilde{\mu}_i}{\partial \tilde{\beta}} \frac{\partial \tilde{\mu}_i}{\partial \tilde{\beta}^t}$$

The idea of using an empirical covariance matrix in connection with misspecified models has been developed by numerous authors, including White (1982), Royall (1986), Liang and Zeger (1986), and Carroll and Rupert (1988, Section 4.3.2).

We now have estimating equations for β and ϕ which yield consistent estimators and we have an expression for the asymptotic covariance of $\sqrt{n}(\tilde{\beta} - \beta^*)$. In the next sub-section, we use this information to form statistics to test the significance of the regression parameters.

2.1.2 Test Statistics for Hypotheses Regarding Regression Parameters

Consider the p -dimensional vector β partitioned into β_1 and β_2 of dimensions p_1 and p_2 respectively. Here the point is to think of β_2 as parameters which have fixed values β_2^o (usually zero) under some null hypothesis H_o . We will refer to the model where all of the parameters $\beta = (\beta_1^t, \beta_2^t)^t$ are estimated as the full model, and the model where β_1 is estimated when β_2 is set to its hypothesized value β_2^o as the reduced model.

Under H_o , the random vector $(\tilde{\beta}_2 - \beta_2^o)$ has a limiting multivariate normal distribution with mean 0 and covariance Σ_{22} which is estimated by $\tilde{\Sigma}_{22}$, the lower right-hand $p_2 \times p_2$ sub-matrix of either $\tilde{\Sigma}_M$ or $\tilde{\Sigma}_E$. We therefore have model-based and empirical versions of a test statistic we refer to as the Wald test where we compare the test statistic $(\tilde{\beta}_2 - \beta_2^o)^t \tilde{\Sigma}_{22}^{-1} (\tilde{\beta}_2 - \beta_2^o)$ to the $\chi_{(p_2)}^2$ distribution.

Another possible type of test, one that does not require fitting of the full model, is the score test. Recall from Section 2.1.1 that $E[\frac{\partial U_j}{\partial \phi}] = 0$; $j = 1, \dots, p$. With sufficient regularity conditions, this implies that for any sequence of estimates $\tilde{\beta}$ converging to β^* the statistics $U^\beta(\tilde{\beta}, \tilde{\phi})$ and $U^\beta(\tilde{\beta}, \phi^*)$ are asymptotically equivalent. This can be seen by comparing the following two asymptotic expansions:

$$\frac{1}{\sqrt{n}}U^\beta(\tilde{\beta}, \tilde{\phi}) = \frac{1}{\sqrt{n}}U^\beta(\beta^*, \phi^*) + \frac{1}{n} \frac{\partial U^\beta}{\partial \beta} \bigg|_{\beta^{m_1}, \phi^m} \sqrt{n}(\tilde{\beta} - \beta^*) + \frac{1}{n} \frac{\partial U^\beta}{\partial \phi} \bigg|_{\beta^{m_1}, \phi^m} \sqrt{n}(\tilde{\phi} - \phi^*), \quad (2.8)$$

and

$$\frac{1}{\sqrt{n}}U^\beta(\tilde{\beta}, \phi^*) = \frac{1}{\sqrt{n}}U^\beta(\beta^*, \phi^*) + \frac{1}{n} \frac{\partial U^\beta}{\partial \beta} \bigg|_{\beta^{m_2}} \sqrt{n}(\tilde{\beta} - \beta^*) \quad (2.9)$$

for some β^{m_i} ($i=1,2$) between $\tilde{\beta}$ and β^* and similarly for some ϕ^m . The right-hand side of these equations are of the same form except for the final term in (2.8). But, as n gets large, $\sqrt{n}(\tilde{\phi} - \phi^*)$ converges in distribution while $\frac{1}{n} \frac{\partial U^\beta}{\partial \phi}$, which is an average of independent components, converges to its expectation which is 0. Also, recall from Section 2.1.1 that the asymptotic covariance matrix of $\sqrt{n}(\tilde{\beta} - \beta^*)$ does not depend on how, or even if, ϕ is estimated so that the distributions of the random vectors $\sqrt{n}(\tilde{\beta} - \beta^*)$ are the same in both (2.8) and (2.9). This shows that the random vectors defined by the left-hand sides of (2.8) and (2.9) have the same asymptotic mean and variance. Therefore, in what follows we may suppose that the estimator of β is based on the estimating equations (2.4) alone, with ϕ fixed at its limiting value ϕ^* .

Now we consider the maximum quasi-likelihood estimates under the reduced model. With U^β partitioned into components U^{β_1} and U^{β_2} of dimension p_1 and p_2 respectively, let $\tilde{\beta}_{red}$ denote this estimator, namely the solution to $U^{\beta_1}[(\beta_1^t, \beta_2^{ot})^t, \phi^*] = 0$. We wish to determine the

asymptotic distribution under the null hypothesis of the score statistic for testing $H_o: \beta_2 = \beta_2^o$.

To accomplish this, first examine the expansion

$$0 = U^{\beta_1}(\tilde{\beta}_{red}) = U^{\beta_1}(\beta^*) + \left. \frac{\partial U^{\beta_1}}{\partial \beta_1} \right|_{\beta_1^{m_1}} (\tilde{\beta}_1 - \beta_1^*) \quad (2.10)$$

and the corresponding expression for the score statistic

$$U^{\beta_2}(\tilde{\beta}_{red}) = U^{\beta_2}(\beta^*) + \left. \frac{\partial U^{\beta_2}}{\partial \beta_1} \right|_{\beta_1^{m_2}} (\tilde{\beta}_1 - \beta_1^*) \quad (2.11)$$

where $\beta_1^{m_i}$ ($i=1,2$) are between $\tilde{\beta}_1$ and β_1^* . Rearranging (2.10) we get

$$(\tilde{\beta}_1 - \beta_1^*) = - \left[\left. \frac{\partial U^{\beta_1}}{\partial \beta_1} \right|_{\beta_1^{m_1}} \right]^{-1} U^{\beta_1}(\beta^*),$$

which, substituted into (2.11), allows us to write

$$\frac{1}{\sqrt{n}} U^{\beta_2}(\tilde{\beta}_{red}) = \frac{1}{\sqrt{n}} \left[U^{\beta_2}(\beta^*) - \left. \frac{\partial U^{\beta_2}}{\partial \beta_1} \right|_{\beta_1^{m_2}} \left[\left. \frac{\partial U^{\beta_1}}{\partial \beta_1} \right|_{\beta_1^{m_1}} \right]^{-1} U^{\beta_1}(\beta^*) \right].$$

As $n \rightarrow \infty$ we have

$$\frac{1}{\sqrt{n}} U^{\beta_2}(\beta^*) \xrightarrow{d} N(0, B_{22}), \quad \frac{1}{\sqrt{n}} U^{\beta_1}(\beta^*) \xrightarrow{d} N(0, B_{11}),$$

and

$$\frac{1}{n} \left. \frac{\partial U^{\beta_2}}{\partial \beta_1} \right|_{\beta_1^{m_2}} \xrightarrow{p} A_{21}, \quad \frac{1}{n} \left. \frac{\partial U^{\beta_1}}{\partial \beta_1} \right|_{\beta_1^{m_1}} \xrightarrow{p} A_{11}.$$

where the matrices A_{11} , A_{21} , B_{11} , and B_{22} are the appropriate submatrices of the matrices A and B . It follows that the asymptotic distribution of this score statistic is multivariate normal with mean 0 and covariance matrix

$$\text{Var}_\infty \left[\frac{1}{\sqrt{n}} U^{\beta_2}(\tilde{\beta}_{red}) \right] = B_{22} - A_{21} A_{11}^{-1} B_{12} - B_{21} A_{11}^{-1} A_{12} + A_{21} A_{11}^{-1} B_{11} A_{11}^{-1} A_{12}. \quad (2.12)$$

If the variance function correctly specifies the variances, then $A = B$ and (2.12) reduces to the usual expression

$$\text{Var}_\infty \left[\frac{1}{\sqrt{n}} U^{\beta_2}(\tilde{\beta}_{red}) \right] = A_{22} - A_{21} A_{11}^{-1} A_{12}. \quad (2.13)$$

Therefore we can obtain a model-based estimate of the asymptotic covariance matrix by substituting \tilde{A}_n for A in (2.13) or an empirical estimate by substituting \tilde{A}_n for A and \tilde{B}_n for B in (2.12). The score test for testing $H_o: \beta_2 = \beta_2^o$ is given by comparing $U^{\beta_2^o}(\tilde{\beta}_{red}) \tilde{\Sigma}^{-1} U^{\beta_2^o}(\tilde{\beta}_{red})$ to the $\chi_{(p_2)}^2$ distribution where $\tilde{\Sigma}$ is our particular choice of the covariance matrix estimate.

The final type of test statistic we will consider is a deviance or likelihood ratio statistic. Recall the quasi-log-likelihood $Q_n(y; \mu, \phi) = \sum_{i=1}^n Q(y_i; \mu_i, \phi)$. Define the *deviance* under the full and reduced models as $-2Q_n(y; \mu(X, \tilde{\beta}_{full}), \tilde{\phi})$ and $-2Q_n(y; \mu(X, \tilde{\beta}_{red}), \tilde{\phi})$ respectively, where the estimate $\tilde{\phi}$ is obtained under the full model in both cases. The deviance statistic is then

$$\Delta \text{dev.} = 2[Q_n(y; \mu(X, \tilde{\beta}_{red}), \tilde{\phi}) - Q_n(y; \mu(X, \tilde{\beta}_{full}), \tilde{\phi})]$$

which should behave like a log-likelihood ratio and thus have an approximate $\chi_{(p_2)}^2$ distribution (McCullagh and Nelder, 1989, p. 471).

In the next section, we will examine the specifics of applying these tests for the significance of regression parameters to the case where we have over-dispersed Poisson data.

2.2 Test Statistics for Hypotheses in Over-Dispersed Poisson Regression

In Section 1.2 the estimating equations and scaled deviance test statistic for over-dispersed Poisson regression were presented as an extension of Poisson regression. This extension was

necessary because all inference which requires estimates of variability, such as tests of hypotheses, will be in error if the over-dispersion is not taken into account, even though reasonable amounts of over-dispersion have very little effect on estimation of the regression parameters by ordinary Poisson regression (Cox, 1983).

We saw that an easy way to accomplish this was to introduce a dispersion parameter ϕ to model this extra variability. In quasi-likelihood this is equivalent to specifying the variance of the observed data to be described by $V(\mu_i, \phi) = \phi\mu_i$ for the i^{th} observation. In this case, for the link function $g(\mu) = \log \mu$ equations (2.4) and (2.6) of Section 2.1.1 become

$$U_j(\beta, \phi) = \sum_{i=1}^n \frac{(y_i - \mu_i)}{\phi} X_{ij} = 0; \quad \text{for } j = 1, \dots, p, \quad (2.14)$$

$$U^\phi(\beta, \phi) = \sum_{i=1}^n \left[\frac{(y_i - \mu_i)^2}{\phi\mu_i} - \frac{n-p}{n} \right] = 0, \quad (2.15)$$

where $\mu_i = \exp\{X_i\beta\}$. Note that ϕ may be factored out of (2.14) and that the estimate of ϕ obtained from (2.15) is Pearson's X^2 statistic (using the Pearson residuals defined in Section 1.2) divided by its degrees of freedom. An alternate estimate of ϕ which is thought to be very similar is to replace X^2 by the deviance given by the joint Poisson likelihood under the full model, which for this discussion we denote G^2 . For this formulation, it is immediate that the estimates of β do not depend on the estimated value of ϕ ; in other words, we may use the usual Poisson estimating equations for β and estimate ϕ afterwards with either of the two methods described above. Clearly this approach has considerable appeal over an approach involving specification of a variance function in which ϕ cannot be factored out.

With the choices $V(\mu_i; \phi) = \phi \mu_i$ and $\mu_i = \exp\{X_i \beta\}$, then from Section 2.1.1 the model-based estimate of the covariance matrix of $\tilde{\beta}$ is given by $\tilde{\Sigma}_M = \tilde{A}_n^{-1}$ where

$$\tilde{A}_n = \sum_{i=1}^n \frac{1}{V(\tilde{\mu}_i; \tilde{\phi})} \frac{\partial \tilde{\mu}_i}{\partial \tilde{\beta}} \frac{\partial \tilde{\mu}_i}{\partial \tilde{\beta}^t} = \sum_{i=1}^n \frac{1}{\tilde{\phi} \tilde{\mu}_i} X_i^t e^{X_i \tilde{\beta}} X_i e^{X_i \tilde{\beta}} = \tilde{\phi}^{-1} \sum_{i=1}^n X_i^t X_i \tilde{\mu}_i$$

where the X_i are the row vectors corresponding to the i^{th} row of X . Note that $\tilde{\phi}$ may be factored out of \tilde{A}_n and the resulting covariance matrix is exactly $\tilde{\phi}$ times that obtained from Poisson regression.

Alternately, the empirical estimate of the covariance matrix of $\tilde{\beta}$ is given by $\tilde{\Sigma}_E = \tilde{A}_n^{-1} \tilde{B}_n \tilde{A}_n^{-1}$ where

$$\tilde{B}_n = \sum_{i=1}^n \frac{(y_i - \tilde{\mu}_i)^2}{V^2(\tilde{\mu}_i; \tilde{\phi})} \frac{\partial \tilde{\mu}_i}{\partial \tilde{\beta}} \frac{\partial \tilde{\mu}_i}{\partial \tilde{\beta}^t} = \sum_{i=1}^n \frac{(y_i - \tilde{\mu}_i)^2}{\tilde{\phi}^2 \tilde{\mu}_i^2} X_i^t e^{X_i \tilde{\beta}} X_i e^{X_i \tilde{\beta}} = \tilde{\phi}^{-2} \sum_{i=1}^n (y_i - \tilde{\mu}_i)^2 X_i^t X_i.$$

Note that the factors $\tilde{\phi}$ appearing in \tilde{A}_n and \tilde{B}_n cancel out in the expression for $\tilde{\Sigma}_E$.

Considering these two estimates of the covariance matrix and using either $\tilde{\phi} = X^2/\text{d.f.}_{full}$ or $\tilde{\phi} = G_{full}^2/\text{d.f.}_{full}$ leads to three possible versions of the Wald test; two model-based and one empirical.

We can similarly examine three versions of the score test. The two model-based versions use the expression given in equation (2.13) for the variance with \tilde{A}_n appropriately partitioned and the score vector as given by the p_2 equations of the form (2.14) which correspond to parameters with some hypothesized value under H_0 . The resulting statistic could use either of the estimates of ϕ described above. The empirical score test, on the other hand, uses the empirical covariance matrix of the form described by (2.12) with \tilde{A}_n and \tilde{B}_n appropriately partitioned. Examination of the expressions involved in this statistic reveals that the $\tilde{\phi}$ terms disappear here also due to cancellation.

Finally, we may consider two versions of the deviance test. The expression for the quasi-likelihood is simply the Poisson log-likelihood divided by an estimate of ϕ . Therefore our deviance tests, often referred to as scaled deviance tests, are given by

$$\Delta_{\text{dev.}} = \frac{1}{\hat{\phi}} \sum_{i=1}^n \left(y_i \log\left(\frac{\tilde{\mu}_i}{\tilde{\mu}_i^o}\right) - (\tilde{\mu}_i - \tilde{\mu}_i^o) \right)$$

where $\tilde{\mu}_i$ are the fitted values under the full model and $\tilde{\mu}_i^o$ are the fitted values under the reduced model. Note that the best estimate of ϕ (using either $X^2/\text{d.f.}$ or $G^2/\text{d.f.}$) will be the one obtained under the full model since this model has the best estimates of μ_i .

Now that we have discussed some of the methods available for modelling and making inference about parameters in over-dispersed Poisson data, we wish to address the questions raised in Section 1.4 and examine the performance of these procedures in a context similar to that encountered in the Prince George study. A simulation study designed to answer these questions is described in Chapter 3.

Chapter 3

A Simulation Study Related to the Prince George Study

Recall the methodological issues which arose in the Prince George study as described in Section 1.4:

- (a) Is the estimate of the scale parameter, ϕ , used in this analysis, namely $\frac{\text{deviance}(\text{full})}{\text{d.f.}(\text{full})}$, a good estimate, or should we have used $\frac{\text{Pearson } X^2(\text{full})}{\text{d.f.}(\text{full})}$?
- (b) Is the change in adjusted deviance criterion appropriate for model reduction in this situation?
- (c) Are statistics other than the adjusted deviance statistic better for testing hypotheses concerning the regression parameters when over-dispersion is present?
- (d) Given the amounts of over-dispersion in the emergency room visits series, are the estimates of the model parameters and standard errors accurate?

In this chapter we outline the purpose and procedures of a simulation study designed to address these questions. The primary goal of the simulation study was to examine the behavior of the model reduction test statistics described in Section 2.2 in contexts similar to the Prince George study (items (b) and (c)). Secondary goals were to examine the performance of estimates of over-dispersion parameters as well as of estimates of regression coefficients and their standard

errors (item (d)). Of particular concern were the effects of varying amounts of over-dispersion and misspecification of the variance function on this performance.

In particular, we will:

1. Compare the scaled deviance, score and Wald tests for model reduction.
2. Compare the parameter estimates to the simulated values.
3. Compare empirical and model-based estimated standard errors to the standard deviations of the estimated parameters.

However, as an essential preliminary to the main thrust of our simulation study, we must first consider the question of how the over-dispersion parameter should be estimated (item (a)). As already mentioned, two possibilities are via $G^2/\text{d.f.}$ (as was done in the Prince George study) and $X^2/\text{d.f.}$ Results presented in Section 4.1 provide a clear indication that the latter is preferred, and this estimate is therefore employed in all subsequent work. In particular, the only version of the scaled deviance model reduction test considered is the one where ϕ is estimated by $X^2/\text{d.f.}$ Both empirical and model-based versions of the Wald statistic will be evaluated, but only the empirical score test will be evaluated; the latter test is reported to perform well even in the presence of over-dispersion and with misspecified variances (see Breslow, 1990). To carry out these comparisons, a variety of data configurations, all intended to be qualitatively similar to the context of the Prince George study, will be simulated. Different data configurations will correspond to different log-linear models for the mean levels of the observed counts and different amounts of over-dispersion. In all cases, the model fit to the simulated data will be that of over-dispersed Poisson regression; that is, estimation of the regression coefficients will

be based on the estimating equations (2.14) which result from an assumed variance function of the form $V(\mu_i; \phi) = \phi\mu_i$.

The simulation study will not be concerned with the possible misspecification of the log-linear model for μ_i (the regression function), but will examine the effects of misspecification of the variance function. For this purpose we define the “correct” variance to be $V(\mu_i; \phi) = \phi\mu_i$, the model to be used in the fitting, and the “misspecified” variance to be $V(\mu_i; \phi) = \mu_i + (\phi - 1)\mu_i^2$. The latter variance function arises from thinking of the datum y_i as being sampled from a distribution which is, conditional on the value of an unobserved variable λ_i , Poisson with mean λ_i ; if λ_i is considered to be sampled from a distribution with mean μ_i and variance $(\phi - 1)\mu_i^2$, then the marginal distribution of y_i has mean μ_i and variance $\mu_i + (\phi - 1)\mu_i^2$.

This simulation study differs in several respects from a study carried out in a different context by Breslow (1990). Firstly, Breslow’s study only considered simulated data with mean values no smaller than 2. In the Prince George study the Asthma series had many zero values and a mean value of only about 0.75. In the current study we wish to investigate the possibility that such a series could lead to poor performance of the estimators or test statistics. Another difference is that Breslow was concerned with a 3×4 factorial design with n replicates per cell (and a single continuous covariate) whereas our study will be concerned with one or more long data series similar to the series available for the Prince George study. This difference is perhaps not as great as it first appears because all the data in Breslow’s simulations were generated under the hypothesis that the column factor of the 3×4 design had no effect. The result is that, in effect, three data series of length $4n$ are being considered and for $n = 144$ (the largest value Breslow considered) the length of the series is comparable with the length

we will employ (700). The last key difference is that the current study is only concerned with test statistics and estimating equations based on an assumed variance function $V(\mu_i; \phi) = \phi\mu_i$. We will examine the performance of these test statistics and estimating equations using data generated under “correct” and “misspecified” variance functions, both for a single data series and for three series with differing amounts of over-dispersion (to represent the separate diagnostic categories), keeping in mind that the methods used will provide only a single overall estimate of ϕ . In contrast, Breslow generated each of his data sets under the variance function $V(\mu_i; \phi) = \mu_i + (\phi - 1)\mu_i^2$ and then examined the performance of the estimating equations and test statistics based on this variance function and also based on $V(\mu_i; \phi) = \phi\mu_i$.

3.1 The General Simulation Procedure

The simulation procedure, or sampling experiment, consists of three phases: generating the data, fitting a full model, and fitting one or more reduced models. All calculations are carried out in the programming language C.

3.1.1 Generating the Data

The data generated are to represent the main qualitative features of the data analyzed in the Prince George study. To accomplish this we will generate time series of 700 independent over-dispersed Poisson observations. We may choose to generate a single series (Sections 4.2 and 4.4) or we may generate three separate series (Sections 4.3 and 4.5).

Each data set to be generated requires a choice of a log-linear model for the mean μ_i of the form $\log \mu_i = X_i\beta$, a choice of the variance function to be used in the simulation, and a choice

of the value for the over-dispersion parameter ϕ .

Over-dispersion in the data is introduced through gamma-mixtures of Poisson random variables. To generate the datum y_i , it is most convenient to think of our situation as follows. Let $\log \mu_i = X_i \beta$ and let there be a random variable ν_i which has a gamma distribution with mean 1 and variance V_i . Then for each $i=1, \dots, 700$ we first generate a realization of the gamma random variable ν_i and then generate a realization of a Poisson random variable with parameter $\nu_i \mu_i$. The resulting observations are negative binomial, with mean μ_i and variance $\mu_i + \mu_i^2 V_i$. The choice $V_i = (\phi - 1)/\mu_i$ leads to simulated data with the “correct” variance function $\phi \mu_i$, whereas the choice $V_i = (\phi - 1)$ leads to simulated data with the “misspecified” variance function $\mu_i + (\phi - 1)\mu_i^2$. Note that both variance functions reduce to the usual Poisson variance when $\phi = 1$, and in this case the Poisson variates were generated directly.

3.1.2 Fitting the Full Model

We fit the full model to obtain estimates of the parameters $\tilde{\beta}_{full}$ using an iterative method for solving the estimating equations (2.14) of Chapter 2. We then estimate ϕ using the Pearson X^2 statistic divided by its degrees of freedom. In contrast to Breslow (1990), when we obtain an estimate of ϕ which is less than one, we do not adjust that estimate to be equal to one. In our situation the dispersion parameter is only a nuisance parameter used to account for dispersion. It is not necessarily thought to arise from the variance of a mixing distribution (which would require values of $\phi > 1$); that is to say, *under-dispersion* is not ruled out. Using this estimate, both the empirical and the model-based estimates of the covariance matrix of $\tilde{\beta}_{full}$ are calculated (allowing the empirical and model-based Wald tests to be evaluated), as

well as the scaled deviance under the full model.

3.1.3 Fitting the Reduced Model

Next, the reduced model is fit to obtain estimates of the parameters $\tilde{\beta}_{red}$. The empirical estimate of the covariance matrix of $\tilde{\beta}_{red}$ is calculated and used in the calculation of the score statistic. We also calculate the scaled deviance using the estimate of ϕ from the full model so that the scaled deviance statistic for testing the model reduction may be computed.

The general procedure described above of generating data, fitting full and reduced models, calculating estimates and standard errors, and calculating the model Wald, empirical Wald, empirical score and scaled deviance test statistics is repeated 1000 times for each particular specification of the log-linear model, choice of $V(\mu_i; \phi)$ and value of ϕ .

In the following section the specific log-linear models, and the choices for the parameters in these models, as well as the levels of over-dispersion utilized in this simulation study are discussed.

3.2 The Specifics of the Simulation Study

As mentioned in Section 3.1.1, we may choose to simulate either a single series of data, or three separate series to be modelled simultaneously. The single series simulations are small enough, in terms of the computing time required to generate the data and fit a model, to allow an investigation of the effects on the estimators and test statistics of varying both the level of dispersion and the simulated mean values. The three series simulations recreate an important feature of the data analysed in the Prince George study, namely separate series with possibly

different amounts of dispersion (we chose to use three series here instead of four as in the Prince George study simply to reduce the computing time).

The overall mean values of the series and the levels of over-dispersion to be considered in these simulations (whether in a single series or in three) was based on the data series from the Prince George study. To recap the approximate mean values of those series and to get a rough idea of how much dispersion was present, we have the following table:

	mean	variance	variance/mean
Asthma	0.76	0.80	1.05
Bronchitis+	1.54	3.36	2.18
Ear	3.62	6.92	1.91
Other	5.13	14.2	2.77

From the above, the values 0.75 , $\exp\{0.7\}(\approx 2.01)$ and $\exp\{1.6\}(\approx 4.95)$ were chosen as possible overall mean values for the series to be simulated. In the single series simulations we will carry out sampling experiments using each of these values, and we will use the three values together as the overall mean levels in the three series simulations.

Based on the third column of the above table, the values 1, 2 and 3 seem to be reasonable levels of over-dispersion to consider, although the estimates of ϕ for each series would be smaller than the ratio of the variance to the mean ($\tilde{\phi} = \frac{1}{\text{d.f.}} \sum_i (y_i - \tilde{\mu}_i)^2 / \tilde{\mu}_i$ is less than variance/mean $= \frac{1}{n-1} \sum_i (y_i - \bar{y})^2 / \bar{y}$ in general). We also wish to simulate the value 1.4 since this is very close to the overall estimate of the dispersion parameter from the Prince George study, and a simulated value $\phi = 5$ is of possible interest as an “extreme” amount of over-dispersion.

In the single series simulations, sampling experiments using each of these levels of dispersion are carried out at each possible mean level, with the exception that $\phi = 5$ is only simulated when the mean level is $\exp\{1.6\}$ (using $\phi = 5$ with lower mean values would result in series with

mostly zeros and only a few large values; this situation is *too* extreme and is not of particular interest in this thesis). In the three series simulations we chose two different combinations of dispersion parameters. The first is the triplet (1,2,3) which we refer to as combination I. These values seem like a plausible representation of the level of over-dispersion encountered in the Prince George study. We also consider the triplet (1.4,2,5), which we refer to as combination II, in order to examine the performance of the estimators and test statistics in a situation with a large amount of dispersion (relative to the amounts encountered in the Prince George study).

Thus sampling experiments will be carried out using the following choices of overall mean values and dispersion parameters:

Single Series Simulations					
Mean Value	ϕ				
	1	1.4	2	3	5
0.75	✓	✓	✓	✓	×
$\exp\{0.7\}$	✓	✓	✓	✓	×
$\exp\{1.6\}$	✓	✓	✓	✓	✓
× = no sampling experiment carried out					

Three Series Simulations		
Mean Values	Combination of ϕ	
	I=(1,2,3)	II=(1.4,2,5)
(0.75, $\exp\{0.7\}$, $\exp\{1.6\}$)	✓	✓

The models used for generating the data should, in some sense, reproduce the main features of the data encountered in the Prince George study. We first describe some very simple models that are used to generate single series of data before we attempt to reproduce the more complicated features of the Prince George study data. The simplest models to be considered are for a single series with the mean value of the j^{th} observation given by

$$\mu_j = \exp\{\beta + \gamma x_j\}, \quad j = 1, \dots, 700 \quad (3.16)$$

where β is the overall effect and γ is the coefficient corresponding to the covariate $x_j = \sin(4\pi j/700)$, used to simulate a pollution variable with a regular annual cycle. To simulate data under the reduced model, one with no pollution, the choices for the simulated coefficients are $\beta = \log 0.75$, 0.7 , or 1.6 (with $\gamma = 0$), corresponding to series with overall mean levels 0.75 , $\exp\{0.7\}$, and $\exp\{1.6\}$ respectively. To simulate data under the full model, one with pollution included, only $\beta = \log 0.75$ is used but three different values of γ are considered ($\gamma = 0.05$, 0.10 or 0.15). The simulations using models of the form (3.16) to generate data are referred to as the one series simple case.

These simple models are also extended to a three series simple case with the mean of the j^{th} observation in the i^{th} series given by

$$\mu_{ij} = \exp\{\beta_i + \gamma_0 x_j + \gamma_i x_j\} \quad i = 1, 2, 3; \quad j = 1, \dots, 700. \quad (3.17)$$

where β_i is the overall effect for the i^{th} diagnostic category, γ_0 is the common pollution effect and γ_i is the pollution effect specific to the i^{th} diagnosis (with $\gamma_3 \equiv 0$ for identifiability). This parameterization with a common and separate pollution effects is utilized to facilitate the score test by allowing the hypothesis of common pollution effects to be stated as $H: \gamma_i = 0$ for $i = 1, 2, 3$. For the alternate parameterization $\gamma'_i = \gamma_0 + \gamma_i$ for $i = 1, 2, 3$, the same hypothesis would take the more usual form of $H: \gamma'_1 = \gamma'_2 = \gamma'_3$ (which can be tested by the Wald or scaled deviance tests). In these simulations the simulated values of the β_i are $\beta_1 = \log 0.75$, $\beta_2 = 0.7$ and $\beta_3 = 1.6$. The null model (no pollution effects) follows from setting $\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = 0$. We may also simulate data sets with a common pollution effect (we use $\gamma_0 = 0.1$, $\gamma_1 = \gamma_2 = \gamma_3 = 0$), or with separate pollution effects (we use $\gamma_0 = 0.1$, $\gamma_1 = 0.1$,

$\gamma_2 = -0.2$ and $\gamma_3 = 0$).

The j^{th} observation in the more complicated models with a single series has mean described by

$$\mu_j = \exp\{\beta + \sum_{k=1}^4 \tau_k s_k + \gamma x_j\} \quad (3.18)$$

where the s_k are indicators designed to represent four “seasons” of equal length (± 1 day), within each of our two “years” of 350 observations and x_j is the sinusoidal pollution covariate just as in the simple case. The seasonal structure is the only attempt made in this thesis to recreate the temporal structure of the data in the Prince George study and no attempt is made to recreate the meteorological structure. For the purpose of simulating data under the reduced (or null) model, the values $\tau_1 = 0.25$, $\tau_2 = 0.25$ and $\tau_3 = -0.5$ (with $\tau_4 = 0$ for identifiability) were used with one of $\beta = \log 0.75$, 0.7 , or 1.6 . For the full model (one with pollution included) only $\beta = \log 0.75$ is used with one of $\gamma = 0.05$, 0.10 or 0.15 . The simulations based on models described by (3.18) include seasonal effects and are referred to as the one series complicated case.

As with the simple simulations, the one series complicated models are extended to a three series complicated case with the mean for the j^{th} observation in the i^{th} series given by

$$\mu_{ij} = \exp\{\beta_i + \sum_{k=1}^4 \tau_{ik} s_k + \gamma_0 x_j + \gamma_i x_j\} \quad (3.19)$$

where β_i is the overall effect for the i^{th} diagnostic category, τ_{ik} is the effect of “season” k in diagnosis i , γ_0 is the common pollution effect, and γ_i is the pollution effect specific to the i^{th} diagnostic category. The values of the coefficients used in the simulations are:

$$\begin{array}{lll}
\beta_1 = \log 0.75 & \beta_2 = .70 & \beta_3 = 1.6 \\
\tau_{11} = 0.2 & \tau_{21} = 0.2 & \tau_{31} = 0.2 \\
\tau_{12} = 0.2 & \tau_{22} = 0.2 & \tau_{32} = 0.3 \\
\tau_{13} = -0.2 & \tau_{23} = -0.4 & \tau_{33} = -0.4 \\
\tau_{14} = 0 & \tau_{24} = 0 & \tau_{34} = 0
\end{array}$$

There were no compelling reasons for the above choices for the season effects other than to make them separate effects; that is, different for each series. As in the three series simple case, the null model (no pollution) corresponds to $\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = 0$ while the common pollution effect model uses simulated values $\gamma_0 = 0.1$, $\gamma_1 = \gamma_2 = \gamma_3 = 0$, and the separate pollution effects model uses $\gamma_0 = 0.1$, $\gamma_1 = 0.1$, $\gamma_2 = -0.2$ and $\gamma_3 = 0$.

The results of fitting the models from each of these four cases (with the various possible choices of simulated coefficients, variance function, and dispersion parameter ϕ) are presented in Chapter 4. We begin Chapter 4 with a comparison of the possible estimators of ϕ (as the remainder of the study depends on a choice of estimator of ϕ). For this comparison, only the one series simple case models were needed as these provide a strong indication that the estimator $X^2/\text{d.f.}$ is preferable to $G^2/\text{d.f.}$

Chapter 4

Results of the Simulation Study

In this chapter we describe the results of the investigation of the performance of the estimators and test statistics. We begin with the comparison of possible estimators for ϕ .

4.1 Estimating the Dispersion Parameter ϕ

Of particular interest in this comparison are any shortcomings in the performance of the estimator $\tilde{\phi}_G = G^2/d.f.$ used in the Prince George study and the possibility that use of the alternate estimator $\tilde{\phi}_X = X^2/d.f.$ in the scaled deviance model reduction test statistic might have led to different conclusions than those sketched in Chapter 1.

For this investigation single series of data were generated according to the simple model given by (3.16) using one of the three possible values of β . Only the reduced model (set $\gamma = 0$) was used to generate data, so that μ_i does not depend on i . As we shall see, the results in even this most simple of cases indicates the estimator $\tilde{\phi}_X = X^2/d.f.$ is preferable. Table 1 presents the mean values and standard deviations of the estimates of ϕ obtained from the two estimators based on the results of fitting the reduced model. Part A of the table contains the results when the correct variance function is used to simulate the data, while part B presents the results for the misspecified variance function, along with the values ϕ^s where $\phi^s \mu_i = \mu_i + (\phi - 1)\mu_i^2$. The values ϕ^s are the estimated values of ϕ we should see if our incorrectly specified model is still

correctly accounting for the extra variability; that is, ϕ^s is the amount of dispersion in the data in terms of the variance function $V(\mu_i; \phi) = \phi\mu_i$ used in the estimating equations.

Table 1. Mean Values of Estimates (\pm Standard Deviation) in 1000 Simulated Data Sets.

A. Correct Variance Function.

Simulated β	Estimator	Simulated ϕ				
		1	1.4	2	3	5
log 0.75	$\tilde{\phi}_X$	1.002 (.055)	1.401 (.096)	2.001 (.174)	2.996 (.337)	-
	$\tilde{\phi}_G$	1.108 (.039)	1.349 (.057)	1.642 (.086)	2.018 (.136)	-
0.70	$\tilde{\phi}_X$	0.998 (.056)	1.400 (.084)	2.002 (.134)	2.983 (.235)	-
	$\tilde{\phi}_G$	1.137 (.059)	1.515 (.074)	2.010 (.097)	2.700 (.139)	-
1.60	$\tilde{\phi}_X$	0.998 (.052)	1.399 (.076)	2.006 (.118)	3.001 (.208)	4.985 (.378)
	$\tilde{\phi}_G$	1.046 (.056)	1.447 (.075)	2.035 (.108)	2.938 (.162)	4.537 (.246)

B. Misspecified Variance Function.

Simulated β	Estimator	Simulated ϕ				
		1	1.4	2	3	5
log 0.75	ϕ^s	1	1.3	1.75	2.5	-
	$\tilde{\phi}_X$	1.000 (.053)	1.300 (.086)	1.749 (.135)	2.497 (.255)	-
	$\tilde{\phi}_G$	1.107 (.038)	1.292 (.051)	1.526 (.072)	1.844 (.112)	-
0.70	ϕ^s	1	1.806	3.014	5.028	-
	$\tilde{\phi}_X$	1.005 (.054)	1.799 (.114)	2.992 (.248)	5.041 (.535)	-
	$\tilde{\phi}_G$	1.143 (.057)	1.854 (.088)	2.705 (.147)	3.823 (.244)	-
1.60	ϕ^s	1	2.981	5.953	10.906	20.812
	$\tilde{\phi}_X$	1.002 (.054)	2.979 (.201)	5.948 (.503)	10.867 (1.14)	20.66 (2.95)
	$\tilde{\phi}_G$	1.050 (.058)	2.919 (.159)	5.226 (.310)	8.150 (.552)	12.17 (1.08)

In part A of the table the estimator $\tilde{\phi}_X$ provides good estimates of ϕ , on average, at all levels of dispersion. The estimator $\tilde{\phi}_G$ appears to slightly overestimate ϕ at low levels of dispersion and underestimate ϕ in the presence of large amounts of dispersion. The largest differences in the estimators occur when the level of dispersion is very high relative to the mean value (see, for example, the simulations with $\beta = \log 0.75$ and $\phi = 3$). In part B the estimator $\tilde{\phi}_X$ yields

good estimates of ϕ^* , on average, whereas $\tilde{\phi}_G$ produces very poor estimates when the level of over-dispersion is high. Note that in both parts of Table 1, except perhaps when $\phi=1$, the estimator $\tilde{\phi}_X$ generally appears substantially more variable than $\tilde{\phi}_G$. This suggests that for some of the simulated data sets, $\tilde{\phi}_X$ may differ noticeably from the simulated value of ϕ (and may be worse than $\tilde{\phi}_G$) although on average it is clearly superior to $\tilde{\phi}_G$.

To answer the key question of how these differences affect the model reduction test statistics we use the above data sets (generated under the simple model $\log \mu_i = \beta$) to evaluate the scaled deviance statistic using either $\tilde{\phi}_G$ or $\tilde{\phi}_X$ based on the full model given by (3.16). This model reduction, a test of the hypothesis $H: \gamma = 0$, is intended to be an over-simplified representation of the model reductions tested in the Prince George study. The plots in Figures 1, 2, 3 and 4 present the values of three test statistics; the scaled deviance using $\tilde{\phi}_G$, the scaled deviance using $\tilde{\phi}_X$ and the empirical score test (which does not depend on an estimate of ϕ , see Section 2.2), versus the quantiles of the chi-squared distribution with 1 degree of freedom. Results are presented only for a simulated mean of 0.75 and for $\phi = 1, 1.4, 2$ and 3; the results for the other combinations of β and ϕ are qualitatively similar. These plots are intended to judge how well the test statistics agree with their predicted asymptotic distribution. The score test is included to allow comparison with a test statistic that has been reported (in a somewhat different context) to perform reasonably well even in the presence of over-dispersion (Breslow 1990). This statistic is also useful for comparison in this situation because it does not involve an estimate of the dispersion parameter.

We see that for data sets generated under either the correct variance function (Figures 1 and 2) or the misspecified variance function (Figures 3 and 4), the use of the estimator $\tilde{\phi}_G$

Figure 1. Comparison of Estimators of Phi: Correct Variance Simulations (Phi=1 or 1.4)

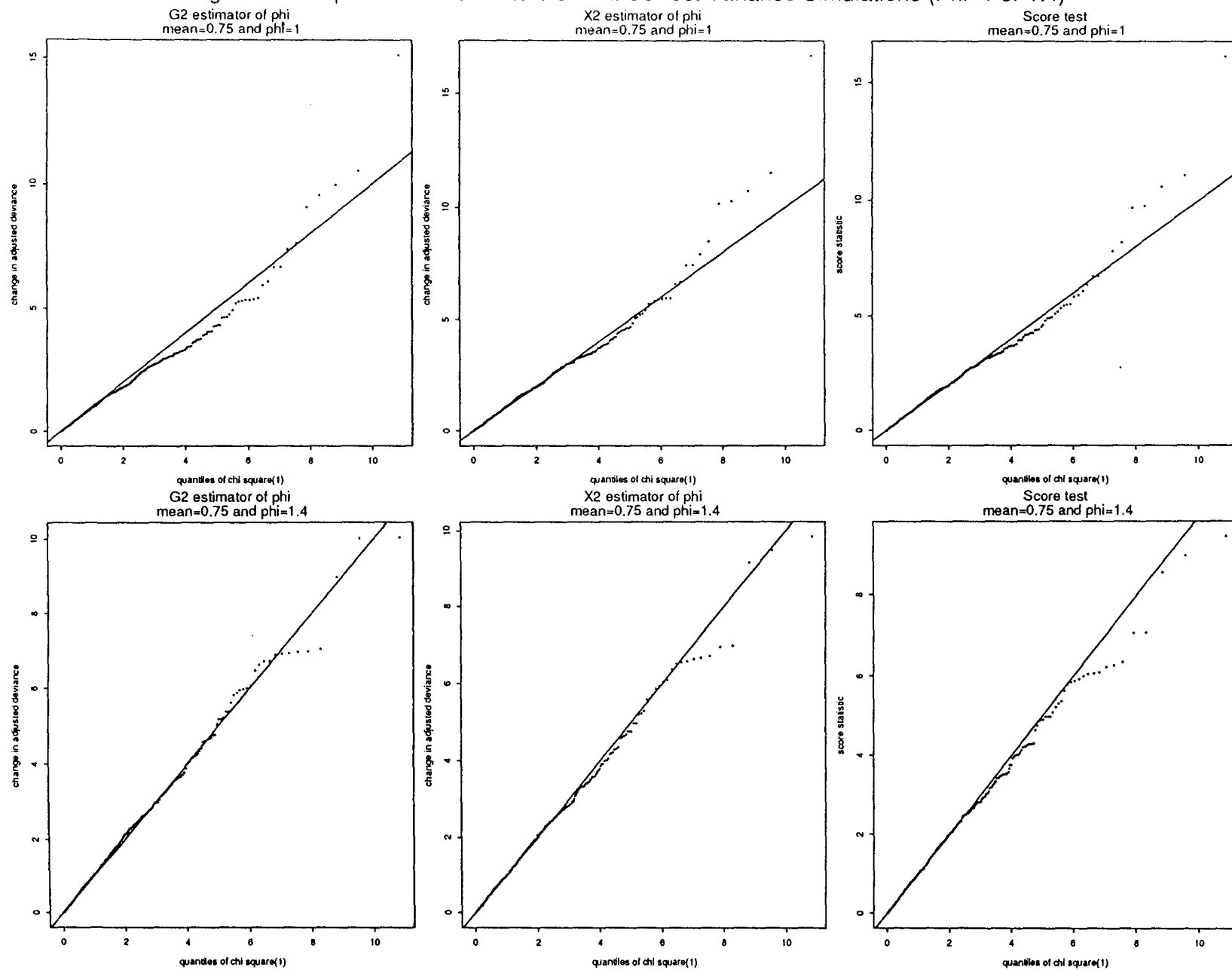


Figure 2. Comparison of Estimators of Phi: Correct Variance Simulations (Phi=2 or 3)

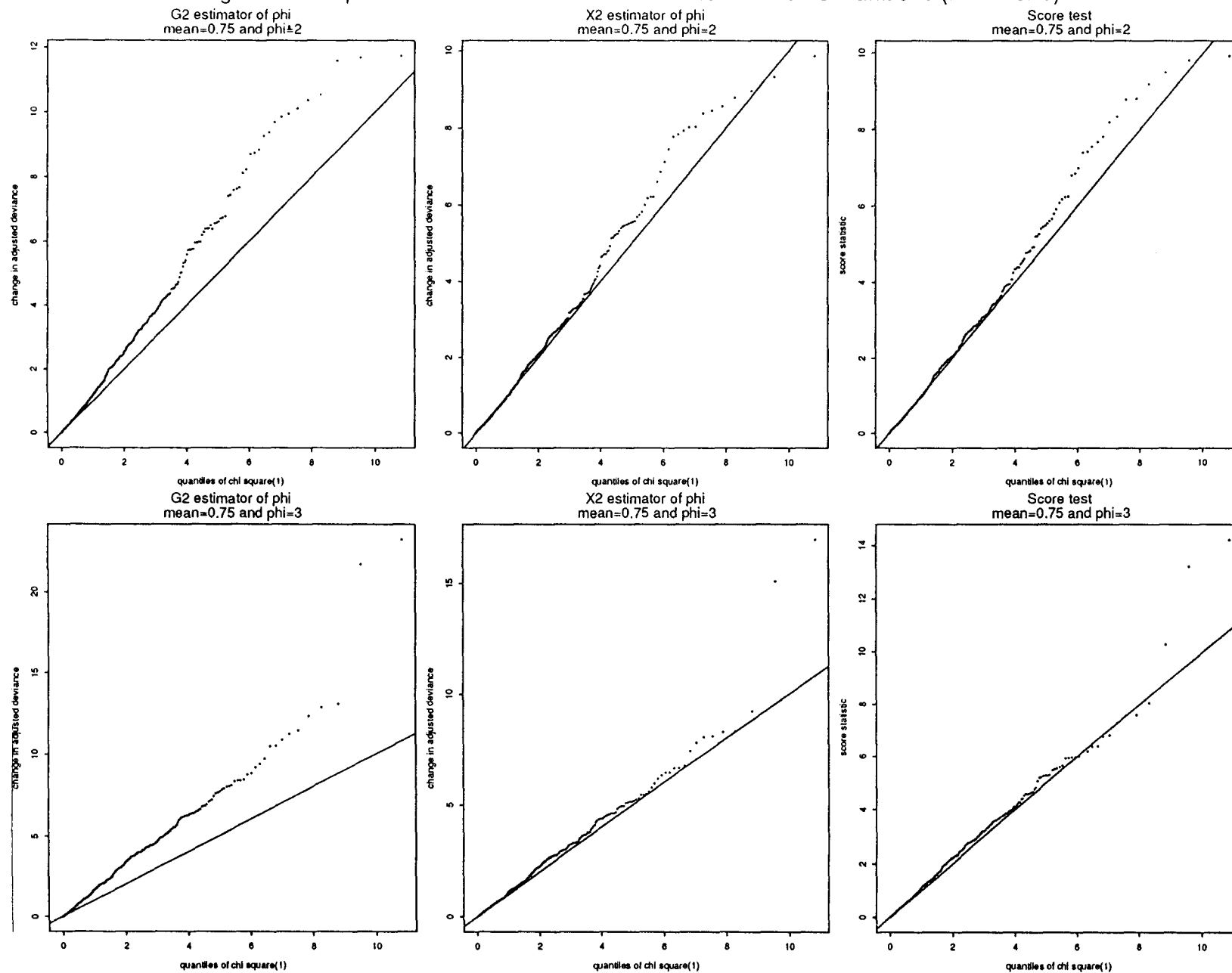


Figure 3. Comparison of Estimators of Phi: Misspecified Variance Simulations (Phi=1 or 1.4)

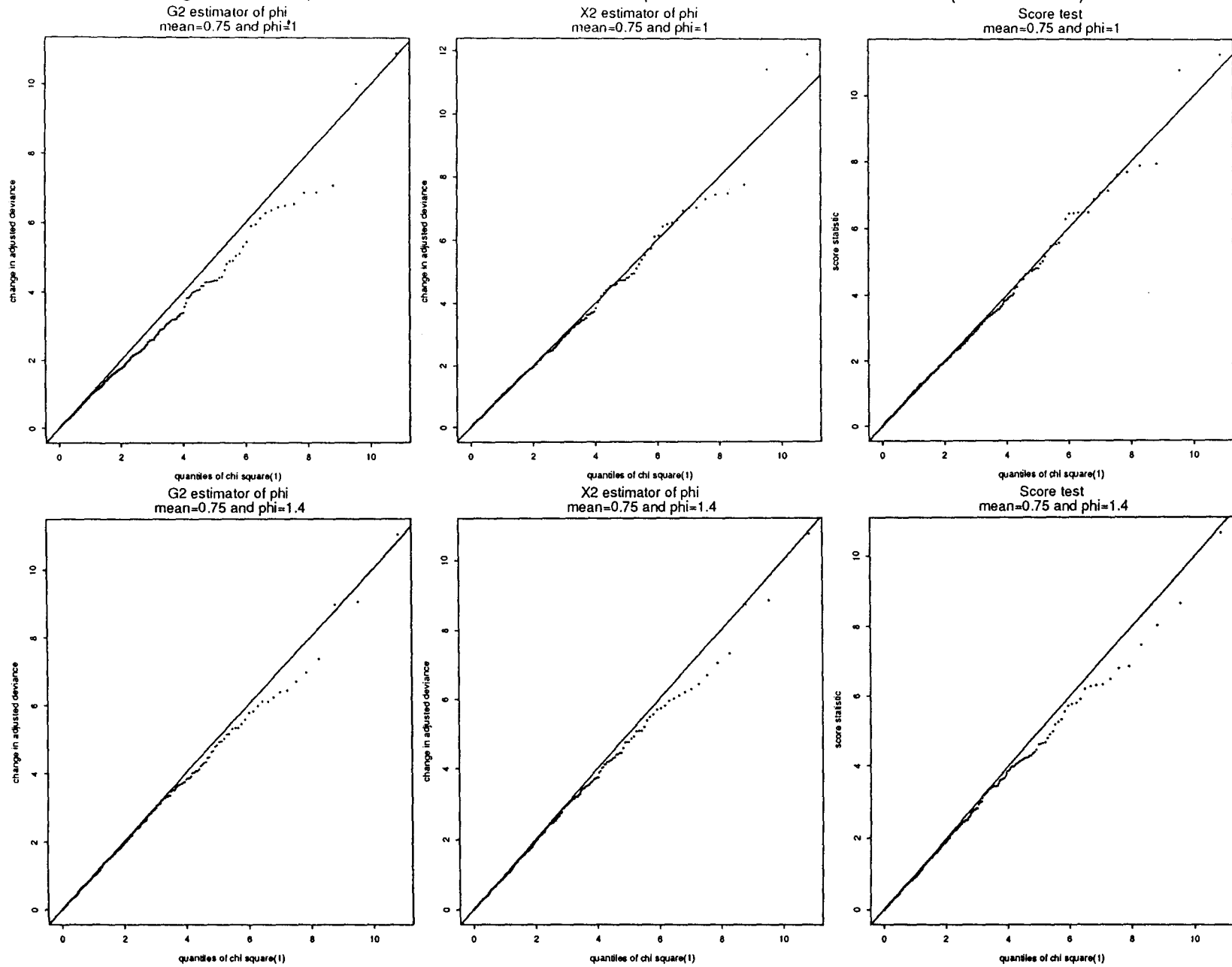
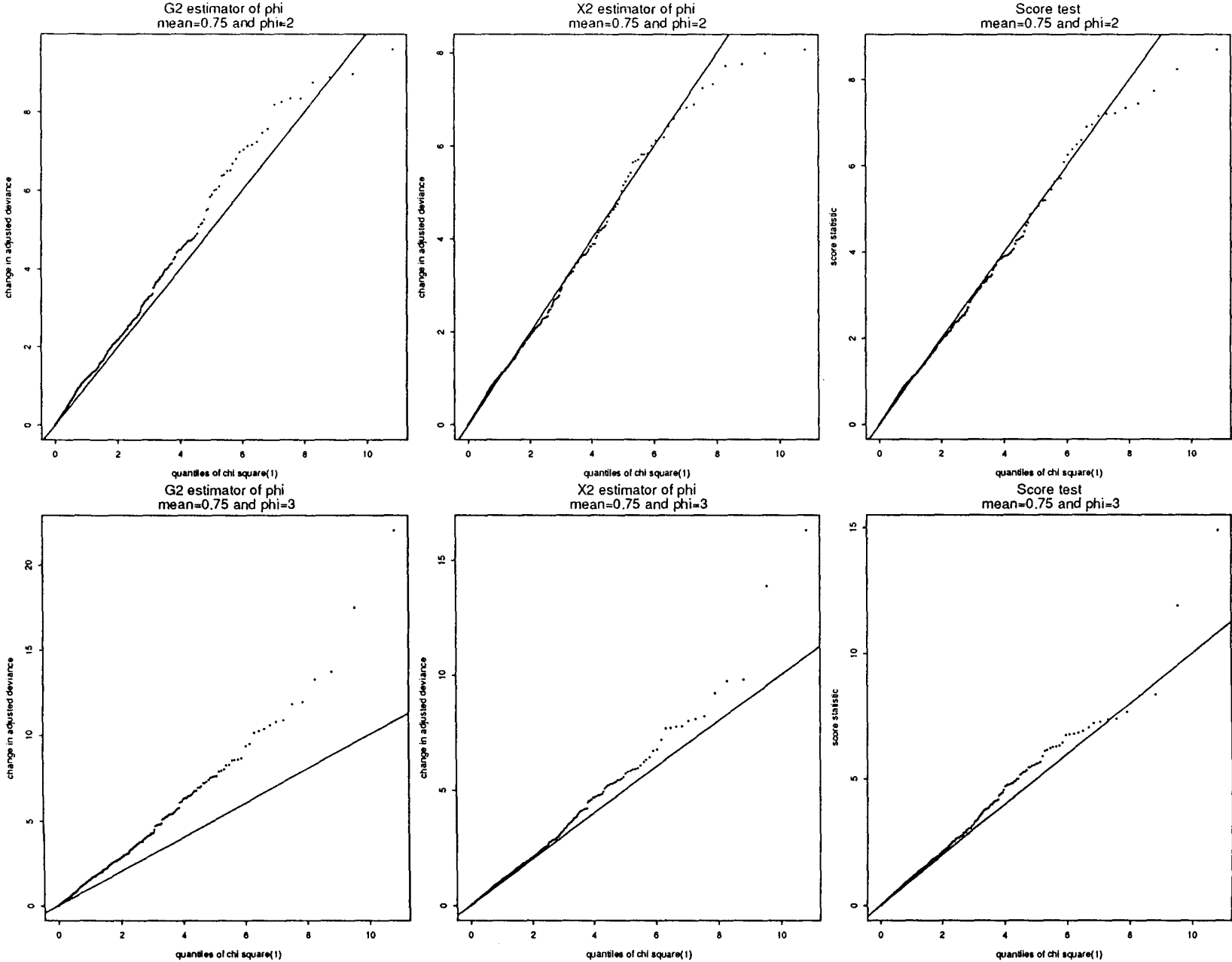


Figure 4. Comparison of Estimators of Phi: Misspecified Variance Simulations (Phi=2 or 3)



causes the scaled deviance test statistic to reject too often at higher levels of over-dispersion. The scaled deviance test using the estimator $\tilde{\phi}_X$ and the empirical score tests agree reasonably well with their predicted asymptotic distributions, except for obvious departures at the most extreme values. Similar plots for the other values of β show the same departures of the scaled deviance statistic using $\tilde{\phi}_G$ from its predicted asymptotic distribution while the scaled deviance using $\tilde{\phi}_X$ and the empirical score tests perform well.

Recall that the estimator $\tilde{\phi}_X$ is the most natural estimator of the two, arising as the moment estimate of ϕ , whereas $\tilde{\phi}_G$ was used because it was more convenient and thought to generally agree closely with $\tilde{\phi}_X$. The limited results presented above indicate that $\tilde{\phi}_G$ may *not* agree well with $\tilde{\phi}_X$ and that the latter is preferable, especially when substantial amounts of dispersion are present in the data. Therefore, in the simulations which follow, when an estimate of ϕ is required we will always use $\tilde{\phi}_X$.

With reference to the Prince George study, the question that arises is: Would this improved estimator have led to different conclusions? It appears that use of $\tilde{\phi}_X$ would not have changed the nature of the results outlined in Chapter 1. For the pollutant TRS, in the situations where we rejected the null hypothesis in favor of a more complex alternative, $\tilde{\phi}_X$ was evaluated and found to be smaller than $\tilde{\phi}_G$ (although the results of Table 1 would suggest that $\tilde{\phi}_X$ is generally larger than $\tilde{\phi}_G$), thus the null hypothesis would still be rejected. It appears quite likely the model reduction process would have led to the same final models and very similar final sets of estimates.

We now begin to examine the performance of the estimators for parameters and their standard errors, and the test statistics outlined in Section 2.2.

4.2 The Simple Case with a Single Series

For a single series we generate data under the reduced model ($\gamma = 0$) corresponding to equation (3.16) at the different possible mean levels and amounts of over-dispersion to be considered. Table 2 presents the mean values of the estimated regression coefficient under the reduced model.

Table 2. Mean Values of Estimates of β .

A. Correct Variance Function.

Simulated β	Simulated ϕ				
	1	1.4	2	3	5
$\log 0.75^1$	-0.287	-0.288	-0.288	-0.289	-
0.7	0.699	0.700	0.699	0.696	-
1.6	1.601	1.600	1.600	1.599	1.598

B. Misspecified Variance Function.

Simulated β	Simulated ϕ				
	1	1.4	2	3	5
$\log 0.75^1$	-0.288	-0.289	-0.290	-0.287	-
0.7	0.700	0.698	0.695	0.702	-
1.6	1.600	1.601	1.600	1.598	1.592

¹ $\log 0.75 \approx -0.288$

For the simulations under the correct variance function, the estimates of β are very close to the simulated values, even in the presence of substantial amounts of dispersion; the same is true in the misspecified variance simulations. This may not be surprising since in this situation the estimate of $\mu = \exp\{\beta\}$ is just an average of the observed data.

Table 3 summarizes the estimates of the variability of the estimates of β under the reduced model. In this table, “Simulated” standard error refers to the standard deviation of the 1000 parameter estimates. The details of parts A and B of Table 3 are very similar. In all cases the empirical and model-based estimates are identical to three decimal places – even for extreme

amounts of over-dispersion. Nor is any difference in the variability of these estimates apparent. Further, these estimated standard errors accurately reflect the actual standard deviations, even when the variance function is misspecified. The agreement seen here between the simulated, empirical and model-based estimates of variability is somewhat surprising. It was expected that the model-based estimators of the covariance matrices would begin to yield poor estimates as the dispersion increased, but this is not the case. It appears that in this simple situation the over-dispersion in the data is being adequately accounted for, even when the variance function is misspecified; the quantity ϕ^s presented in part B of Table 1 was estimated very well and serves the purpose of accounting for the extra variability.

Table 3. Standard Deviations of Estimates of β
and Mean Values of Estimated Standard Errors (\pm s.d.).

A. Correct Variance Function		Simulated ϕ				
Simulated β	Method	1	1.4	2	3	5
log 0.75	Simulated	.043	.051	.062	.077	-
	Model	.044 (.002)	.052 (.002)	.062 (.003)	.075 (.004)	-
	Empirical	.044 (.002)	.052 (.002)	.062 (.003)	.075 (.004)	-
0.70	Simulated	.026	.031	.037	.046	-
	Model	.027 (.001)	.032 (.001)	.038 (.001)	.046 (.002)	-
	Empirical	.027 (.001)	.032 (.001)	.038 (.001)	.046 (.002)	-
1.60	Simulated	.017	.021	.024	.030	.039
	Model	.017 (.001)	.020 (.001)	.024 (.001)	.029 (.001)	.038(.001)
	Empirical	.017 (.001)	.020 (.001)	.024 (.001)	.029 (.001)	.038(.001)

Table 3 B. Misspecified Variance Function

Simulated β	Method	Simulated ϕ				
		1	1.4	2	3	5
log 0.75	Simulated	.042	.050	.059	.069	-
	Model	.044 (.002)	.050 (.002)	.058 (.002)	.069 (.003)	-
	Empirical	.044 (.002)	.050 (.002)	.058 (.002)	.069 (.003)	-
0.70	Simulated	.026	.036	.048	.057	-
	Model	.027 (.008)	.036 (.001)	.046 (.002)	.060 (.003)	-
	Empirical	.027 (.008)	.036 (.001)	.046 (.002)	.060 (.003)	-
1.60	Simulated	.017	.031	.041	.056	.081
	Model	.017 (.001)	.029 (.001)	.041 (.002)	.056 (.003)	.077 (.005)
	Empirical	.017 (.001)	.029 (.001)	.041 (.002)	.056 (.003)	.077 (.005)

We now turn to the performance of the test statistics for the hypothesis $H: \gamma = 0$, summarized in Table 4. The observed rejection probabilities presented in this table can be considered as averages of 1000 Bernoulli trials with success probability p , which we may approximate by the nominal level. Thus the standard errors of the observed probabilities are approximately $\sqrt{p(1-p)/n}$; for nominal levels of 0.10, 0.05 and 0.01 we may expect standard errors of approximately .009, .007 and .003. For the results in part A of Table 4, keeping the precision of these observed rejection probabilities in mind, we find that they agree quite well with the nominal levels at all mean values and even at higher levels of dispersion. We also note the observed rejection probabilities are very similar across the four tests for each choice of mean and dispersion suggesting that any deviations of the observed rejection probabilities from the nominal levels are a reflection of the particular 1000 data sets generated. The agreement between model and empirical Wald tests was to be expected given the agreement of the two estimators of the variances seen in Table 3. In general, the agreement among the four statistics may not be particularly surprising either. Conventional wisdom would suggest that, over-dispersion aside,

each of the tests should perform well with a sample size of 700 (although there was some question if this was true when the series consist of a large number of zeros such as in the simulations using $\beta = \log 0.75$). Thus, in light of how effectively the over-dispersion is accounted for, we should not expect any of the test statistics to do poorly.

The results in the misspecified variance simulations are similar. In these simulations the agreement between the four tests may seem particularly surprising at first; we might expect the model-based Wald and the scaled deviance tests to do poorly since they rely on the particular choice of variance function used to fit the data. However, considering how the choice of variance function does not seem to matter in this simple case (as noted for Table 3), the above agreement should not be surprising.

Table 4. Observed Rejection Probabilities Under Reduced Model.

A. Correct Variance Function

Simulated β	Test/Method	$\phi = 1$			$\phi = 1.4$			$\phi = 2$		
		0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.105	.041	.010	.096	.045	.007	.106	.053	.015
	Wald/Empirical	.105	.040	.010	.097	.046	.005	.114	.056	.016
	Score/Empirical	.102	.040	.010	.095	.045	.005	.111	.056	.016
	Deviance/Scaled	.106	.041	.010	.096	.045	.008	.108	.054	.015
0.70	Wald/Model	.103	.056	.005	.082	.046	.012	.110	.062	.017
	Wald/Empirical	.104	.055	.006	.083	.046	.011	.111	.061	.017
	Score/Empirical	.102	.054	.005	.082	.046	.011	.110	.060	.014
	Deviance/Scaled	.103	.056	.005	.082	.046	.013	.110	.062	.017
1.60	Wald/Model	.112	.056	.011	.095	.043	.004	.087	.039	.007
	Wald/Empirical	.110	.057	.012	.097	.043	.004	.091	.038	.008
	Score/Empirical	.108	.057	.012	.090	.038	.007	.090	.038	.007
	Deviance/Scaled	.112	.056	.011	.087	.039	.007	.087	.039	.007

Simulated β	Test/Method	$\phi = 3$			$\phi = 5$		
		0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.119	.056	.009	-	-	-
	Wald/Empirical	.122	.059	.011	-	-	-
	Score/Empirical	.117	.054	.009	-	-	-
	Deviance/Scaled	.121	.056	.011	-	-	-
0.70	Wald/Model	.114	.061	.013	-	-	-
	Wald/Empirical	.122	.065	.012	-	-	-
	Score/Empirical	.119	.063	.011	-	-	-
	Deviance/Scaled	.114	.061	.013	-	-	-
1.60	Wald/Model	.103	.043	.010	.092	.043	.009
	Wald/Empirical	.096	.048	.013	.090	.043	.009
	Score/Empirical	.095	.045	.010	.088	.043	.008
	Deviance/Scaled	.103	.043	.010	.092	.043	.009

Table 4 B. Misspecified Variance Function

Simulated β	Test/Method	$\phi = 1$			$\phi = 1.4$			$\phi = 2$		
		0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.094	.045	.009	.094	.044	.006	.091	.045	.009
	Wald/Empirical	.095	.047	.009	.095	.046	.007	.092	.050	.012
	Score/Empirical	.095	.046	.009	.094	.046	.006	.091	.047	.010
	Deviance/Scaled	.094	.045	.009	.095	.044	.006	.093	.046	.009
0.70	Wald/Model	.091	.050	.016	.095	.048	.012	.107	.046	.010
	Wald/Empirical	.091	.051	.016	.091	.047	.012	.108	.046	.009
	Score/Empirical	.092	.051	.015	.091	.046	.011	.108	.043	.007
	Deviance/Scaled	.091	.050	.017	.095	.048	.012	.109	.047	.010
1.60	Wald/Model	.097	.041	.007	.103	.052	.009	.083	.033	.004
	Wald/Empirical	.097	.039	.009	.105	.051	.010	.081	.034	.005
	Score/Empirical	.096	.039	.007	.103	.050	.009	.079	.033	.004
	Deviance/Scaled	.097	.041	.007	.103	.052	.010	.083	.034	.004

Simulated β	Test/Method	$\phi = 3$			$\phi = 5$		
		0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.112	.063	.014	-	-	-
	Wald/Empirical	.114	.066	.017	-	-	-
	Score/Empirical	.110	.062	.015	-	-	-
	Deviance/Scaled	.112	.064	.015	-	-	-
0.70	Wald/Model	.101	.044	.010	-	-	-
	Wald/Empirical	.102	.048	.011	-	-	-
	Score/Empirical	.099	.042	.009	-	-	-
	Deviance/Scaled	.101	.045	.010	-	-	-
1.60	Wald/Model	.111	.059	.014	.102	.056	.011
	Wald/Empirical	.114	.066	.016	.104	.053	.012
	Score/Empirical	.108	.059	.013	.100	.047	.006
	Deviance/Scaled	.111	.059	.014	.102	.056	.012

In summary, for the simulation results when the data is generated under the reduced model there is no clear difference between the empirical and model-based estimates of variances. Nor are there any differences among the four model reduction test statistics considered. We will look to the three series case in Section 4.3 to provide more insight into possible differences in the estimators and test statistics in this simple situation.

A limited number of simulations were also carried out using the full model corresponding to (3.16) to generate the data. In particular, 1000 data sets were generated according to this model with $\beta = \log 0.75$ and one of $\gamma = 0.05, 0.10$, or 0.15 . The estimation of ϕ , β and γ were of interest as well as power for testing the hypothesis $H:\gamma = 0$. The mean values of the parameter estimates are summarized in Table 5. First consider part A of the table, corresponding to simulations where the data were generated with the correct variance function. In all cases (that is, all combinations of simulated ϕ and γ) the estimators yield good estimates. The mean values of both model and empirical versions of the estimated standard errors (not presented here) are in close agreement with the displayed standard deviations. The standard errors of the estimates increase as the dispersion increases as would be expected. Notice that the estimates of ϕ (and their standard deviations) are almost identical to the corresponding quantities for $\tilde{\phi}_X$ in part A of Table 1 (which were based on data generated using the reduced model).

In part B, we see that β and γ are estimated very well at all levels of dispersion considered. To determine what the estimates of ϕ should be, first note that in these misspecified variance simulations, the amount of extra variability varies as the covariate x_j , and thus the mean values μ_j , vary. In the present situation, the average of the ϕ_j^s (where $\mu_j \phi_j^s = \mu_j + (\phi - 1)\mu_j^2$) for each cell in part B of Table 5 are the same to two decimals as the values ϕ^s from part B of Table 1. It is therefore reasonable that the estimates of ϕ are close to the estimates in part B of Table 1, and again $\tilde{\phi}_X$ is accounting for the extra variability very well.

Table 5. Mean Values of Estimated Parameters (\pm s.d.) Under Full Model.

A. Correct Variance Function					
Simulated γ	Parameter	Simulated ϕ			
		1	1.4	2	3
0.05	ϕ	1.00 (.055)	1.40 (.095)	2.00 (.178)	2.97 (.335)
	β	-.288 (.044)	-.289 (.050)	-.292 (.064)	-.292 (.076)
	γ	.049 (.061)	.051 (.075)	.050 (.087)	.053 (.109)
0.10	ϕ	1.00 (.054)	1.40 (.093)	1.99 (.179)	2.97 (.323)
	β	-.288 (.042)	-.290 (.051)	-.291 (.063)	-.293 (.076)
	γ	.100 (.063)	.098 (.073)	.099 (.086)	.102 (.110)
0.15	ϕ	1.00 (.054)	1.40 (.096)	2.00 (.171)	2.98 (.323)
	β	-.288 (.043)	-.289 (.051)	-.288 (.063)	-.293 (.075)
	γ	.151 (.062)	.149 (.072)	.150 (.085)	.153 (.109)
B. Misspecified Variance Function					
Simulated γ	Parameter	Simulated ϕ			
		1	1.4	2	3
0.05	ϕ	1.00 (.055)	1.30 (.083)	1.75 (.133)	2.49 (.260)
	β	-.288 (.044)	-.290 (.050)	-.292 (.059)	-.292 (.069)
	γ	.049 (.061)	.051 (.069)	.055 (.084)	.053 (.098)
0.10	ϕ	1.00 (.054)	1.30 (.082)	1.75 (.134)	2.49 (.258)
	β	-.288 (.042)	-.290 (.050)	-.290 (.056)	-.294 (.069)
	γ	.100 (.063)	.102 (.068)	.099 (.083)	.104 (.102)
0.15	ϕ	1.00 (.054)	1.30 (.084)	1.75 (.137)	2.50 (.255)
	β	-.288 (.043)	-.291 (.049)	-.291 (.058)	-.291 (.070)
	γ	.151 (.062)	.151 (.071)	.150 (.081)	.151 (.100)

The observed rejection probabilities under the alternative hypothesis are shown in Table 6 below. The results presented in part A show no difference in the power of the four statistics. For a given value of γ , the observed rejection probabilities decrease as the amount of dispersion increases (the increased variability in the data makes it more difficult to detect an effect).

The results in part B are very similar. In general, for given values of γ and $\phi > 1$, the rejection probabilities are slightly higher in the misspecified variance simulations compared to the correct variance simulations. Note that for values of $\mu_j \leq 1$, such as in these simulations using $\beta = \log 0.75$ and γ no larger than 0.15, the value of $\mu_j + (\phi - 1)\mu_j^2$ (the variance under

the misspecified formulation) is less than the value of $\phi\mu_j$ (the variance under the correct formulation). Therefore from our earlier observation that the power decreases as the amount of variability in the data increases, we should expect the rejection probabilities in the misspecified variance to be higher than the correct variance simulations for these values of γ and ϕ .

In summary, from the results of the simulations under the full model (Tables 5 and 6) the estimates of ϕ appear reasonable and the regression parameter estimates are close, on average, to the simulated values. There are also no indications of any differences in the power of the four test statistics considered. Further, estimated standard errors from the model and empirical estimators (not presented) were similar, both being very close, on average, to the simulated values.

Table 6. Observed Rejection Probabilities Under Full Model.

A. Correct Variance Function													
γ	Test	$\phi = 1$			$\phi = 1.4$			$\phi = 2$			$\phi = 3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
.05	Wald/Mod.	.205	.117	.033	.194	.114	.041	.152	.090	.025	.152	.089	.025
	Wald/Emp.	.204	.117	.035	.197	.119	.040	.155	.090	.027	.151	.086	.025
	Score/Emp.	.202	.116	.034	.197	.118	.038	.154	.086	.024	.147	.079	.024
	Dev/Scaled	.205	.117	.033	.194	.117	.041	.153	.092	.026	.156	.092	.025
.10	Wald/Mod.	.490	.374	.164	.348	.247	.116	.295	.205	.070	.253	.165	.059
	Wald/Emp.	.489	.372	.171	.351	.253	.119	.298	.207	.076	.266	.159	.065
	Score/Emp.	.487	.371	.164	.353	.249	.111	.296	.203	.072	.260	.156	.060
	Dev/Scaled	.490	.374	.164	.349	.247	.117	.296	.207	.074	.253	.166	.060
.15	Wald/Mod.	.786	.685	.445	.661	.522	.296	.529	.404	.186	.391	.274	.135
	Wald/Emp.	.788	.687	.445	.663	.521	.297	.527	.411	.187	.402	.274	.136
	Score/Emp.	.788	.683	.439	.662	.518	.287	.527	.408	.182	.395	.269	.128
	Dev/Scaled	.786	.685	.448	.663	.523	.300	.531	.405	.190	.391	.277	.136
B. Misspecified Variance Function													
γ	Test	$\phi = 1$			$\phi = 1.4$			$\phi = 2$			$\phi = 3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
.05	Wald/Mod.	.205	.117	.033	.193	.111	.024	.176	.108	.030	.165	.090	.021
	Wald/Emp.	.204	.117	.035	.195	.111	.028	.174	.112	.032	.161	.087	.020
	Score/Emp.	.202	.116	.034	.189	.107	.026	.169	.109	.030	.157	.082	.018
	Dev/Scaled	.205	.117	.033	.194	.111	.025	.177	.110	.032	.165	.093	.021
.10	Wald/Mod.	.490	.374	.164	.407	.291	.123	.336	.236	.087	.291	.195	.073
	Wald/Emp.	.489	.372	.171	.410	.299	.131	.333	.241	.097	.294	.203	.074
	Score/Emp.	.487	.371	.164	.409	.300	.122	.328	.233	.088	.291	.192	.066
	Dev/Scaled	.490	.374	.164	.410	.292	.123	.336	.237	.088	.291	.195	.073
.15	Wald/Mod.	.786	.685	.445	.680	.579	.337	.585	.455	.218	.470	.340	.158
	Wald/Emp.	.788	.687	.445	.683	.574	.342	.587	.455	.222	.481	.342	.162
	Score/Emp.	.788	.683	.439	.681	.565	.333	.584	.450	.209	.473	.333	.148
	Dev/Scaled	.786	.685	.448	.680	.580	.339	.586	.456	.219	.472	.342	.160

4.3 The Simple Case with Three Series

The results discussed in this section are for three separate series of data generated according to log-linear models given by (3.17) with one of the two combinations of dispersion parameters described in Section 3.2 (combination I or II) and one of the two possible variance functions. For the null model ($\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = 0$) the parameter values $\beta_1 = \log 0.75 (= -0.288)$, $\beta_2 = 0.7$ and $\beta_3 = 1.6$ are used to generate the data. The resulting data series are then fit using this same log-linear model. The mean values of the estimated parameters from 1000 simulated data sets generated under each of the four possible scenarios (2 combinations of $\phi \times 2$ variance formulations) are presented in Table 7.

Table 7. Mean Values of Parameter Estimates.

Parameter	Variance Function			
	A. Correct		B. Misspecified	
	Combination of ϕ		Combination of ϕ	
	I	II	I	II
ϕ	1.997	2.799	4.964	8.333
β_1	-0.288	-0.289	-0.289	-0.287
β_2	0.699	0.699	0.701	0.700
β_3	1.599	1.599	1.598	1.594

Intuitively the estimates of ϕ should be approximately equal to the average of the three dispersion values associated with the three generated data series in each of the 1000 data sets leading to the results in each of the four scenarios. These averages are 2.0 (combination I, correct variance), 2.8 (combination II, correct variance), 4.97 (combination I, misspecified variance) and 8.37 (combination II, misspecified variance). We can see that in each of the four scenarios, the estimates of ϕ are very close to these corresponding averages. The estimates of the regression parameters are also excellent in all four scenarios.

Table 8 presents the standard deviations of the parameter estimates and the mean values of the estimated standard errors for the regression parameters. In each column of this table the empirical estimates agree very well with the simulated standard deviations but the model-based estimates are most often quite different. The latter are much too large for the parameter β_1 corresponding to the series where the simulated dispersion is lower than the average of the three series, and too small for β_3 corresponding to the series where the simulated dispersion is larger than the average of the three series. For the regression parameter β_2 , first consider the correct variance simulations. We find that with combination I of ϕ 's the standard errors for β_2 are about right; this appears to be because the simulated ϕ for this series is equal to the average of the ϕ 's for all three series. However, using combination II of ϕ 's the simulated ϕ for series 2 is smaller than the average of the ϕ 's for all three series and the model-based standard error is distinctly larger than the standard deviation of the parameter estimates. In the misspecified variance simulations, we also find the model-based standard error to be larger than the standard deviation of the parameter estimates because the simulated ϕ for series 2 is smaller than the average of the ϕ 's in these cases. Clearly the empirical estimator is better in each of these cases (for both combinations of ϕ and for both variance formulations) and accurately reflects the true variability of these parameter estimates, even with the misspecified variance function.

Table 8. Standard Deviations of Estimates
and Mean Values of Estimated Standard Errors (\pm s.d.).

Parameter	Method	Variance Function			
		A. Correct		B. Misspecified	
		Combination of ϕ		Combination of ϕ	
		I	II	I	II
ϕ	Simulated	.084	.144	.399	.976
β_1	Simulated	.042	.052	.042	.050
	Model	.062 (.002)	.073 (.003)	.097 (.004)	.126 (.008)
	Empirical	.044 (.002)	.052 (.002)	.044 (.002)	.050 (.002)
β_2	Simulated	.038	.038	.046	.047
	Model	.038 (.001)	.045 (.001)	.059 (.003)	.077 (.005)
	Empirical	.038 (.001)	.038 (.001)	.046 (.002)	.046 (.002)
β_3	Simulated	.029	.037	.054	.079
	Model	.024 (.001)	.028 (.001)	.038 (.001)	.049 (.002)
	Empirical	.029 (.001)	.038 (.001)	.056 (.003)	.077 (.004)

Using the data generated under the reduced model, the levels achieved by the four test statistics under study are examined by considering an alternative model that allows for a common “pollution” effect ($\gamma_1 = \gamma_2 = \gamma_3 = 0$, but γ_0 possibly not zero). The hypothesis to be tested is $H: \gamma_0=0$. Table 9 presents the observed rejection probabilities of each of the four test statistics. The precision of the observed rejection probabilities, which should be kept in mind when interpreting these results, are the same as those reported in Section 4.2. Consider first the correct variance simulations. In the model reductions for the correct variance simulations, the model-based tests (model Wald and scaled deviance) clearly reject too often while the rejection probabilities for the empirical tests (empirical score and empirical Wald) agree quite well with the nominal levels based on the predicted $\chi^2_{(1)}$ critical values. A similar pattern is apparent in the misspecified variance simulations, although the departure of the model-based tests from

the nominal levels is considerably more exaggerated than in the correct variance simulations.

Table 9. Observed Rejection Probabilities: Common Pollution to No Pollution Model.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.159	.099	.019	.173	.102	.039	.180	.104	.045	.226	.139	.060
Wald/Emp	.114	.057	.004	.110	.058	.014	.089	.054	.016	.108	.062	.012
Score/Emp	.114	.056	.004	.109	.058	.014	.088	.053	.012	.102	.057	.011
Dev/Scaled	.159	.099	.018	.172	.102	.039	.180	.104	.044	.227	.139	.060

Thus in this multiple series situation (without any pollution effects at this time) we find a clear difference between the empirical and model-based estimates of variance, with the empirical estimator being preferable. Correspondingly we find that the model-based test statistics are inferior to the empirical test statistics. Given the differences in the variance estimates, the fact that the empirical Wald is preferable to the model Wald test is not surprising. Nor is it surprising that the empirical score test performs well in all cases since it performed well even under misspecified variances in the study described in Breslow (1990). What is interesting is the very close agreement between the model Wald and scaled deviance tests and between the empirical score and empirical Wald tests. Conventional wisdom would suggest the Wald tests might be inferior to the other two tests, but we see that the model Wald gives observed rejection probabilities that are almost exactly the same as the scaled deviance test and the empirical Wald gives observed rejection probabilities that are almost exactly the same as the empirical score test. It seems that the test statistics which rely on a dispersion parameter to account for extra variability perform poorly, while those that estimate the extra variability

empirically perform much better.

A set of simulations with data generated according to the model with a common pollution effect ($\gamma_1 = \gamma_2 = \gamma_3 = 0$, but γ_0 non-zero in (3.17)) were also carried out. Recall that the value chosen for γ_0 is 0.10. Table 10 presents the mean values of the parameter estimates for the data sets generated in this fashion. The estimates of the β_i are very good in each column of the table as are the estimates of γ_0 . The estimates of ϕ are close to the estimates seen in Table 7 as we would expect given that in each case, the average of the dispersions is close to the averages encountered in the simulations under the null hypothesis.

Table 10. Mean values of Parameter Estimates.

Parameter	Variance Function			
	A. Correct		B. Misspecified	
	Combination of ϕ		Combination of ϕ	
	I	II	I	II
ϕ	1.998	2.799	4.970	8.325
β_1	-0.289	-0.292	-0.288	-0.292
β_2	0.699	0.701	0.699	0.699
β_3	1.599	1.599	1.597	1.591
γ_0	0.100	0.101	0.098	0.098

The standard deviations of the above estimates and the mean values of the estimated standard errors for the regression parameters are given in Table 11. The model-based estimates of the standard errors for the β_i show the same behavior as in Table 8 with the estimate being too large when the parameter corresponds to a series with simulated ϕ lower than $\tilde{\phi}$ and too small when the parameter corresponds to a series with simulated ϕ larger than $\tilde{\phi}$. The empirical estimator on the other hand, agrees very well with the simulated standard deviations. Notice that the model-based estimates of the standard error of γ_0 are too low throughout the table. Recall that the rejection probabilities were too high for the model-based test statistics in Table

9. This suggests that the standard errors of the estimates of γ_0 were lower than expected in the simulations under the null hypothesis as well.

Table 11. Standard Deviations of Estimates
and Mean Values of Estimated Standard Errors (\pm s.d.).

Parameter	Method	Variance Function			
		A. Correct		B. Misspecified	
		Combination of ϕ		Combination of ϕ	
		I	II	I	II
ϕ	Simulated	.086	.144	.395	.967
β_1	Simulated	.043	.053	.040	.052
	Model	.062 (.002)	.073 (.003)	.097 (.004)	.126 (.008)
	Empirical	.044 (.002)	.052 (.002)	.044 (.002)	.050 (.002)
β_2	Simulated	.038	.038	.046	.049
	Model	.038 (.001)	.045 (.001)	.059 (.003)	.077 (.005)
	Empirical	.038 (.001)	.038 (.001)	.046 (.002)	.046 (.002)
β_3	Simulated	.029	.037	.058	.079
	Model	.024 (.001)	.028 (.001)	.038 (.001)	.049 (.002)
	Empirical	.029 (.001)	.038 (.001)	.056 (.003)	.077 (.004)
γ_0	Simulated	.030	.038	.055	.076
	Model	.027 (.001)	.032 (.001)	.043 (.002)	.056 (.003)
	Empirical	.031 (.001)	.038 (.002)	.054 (.002)	.072 (.006)

We now examine the results of evaluating the four test statistics based on an alternative hypothesis that includes common and separate “pollution” effects for each series. Thus these model reductions, from the model with separate effects to the one with a common pollution effect (which was used to generate the data), represent the separate to common effects reductions described in the Prince George study. Note that these tests, the results of which are summarized in Table 12, are the first discussed in this thesis involving reduction by more than one parameter (and are thus the first that involve estimates of *covariance* in the Wald and score test statistics).

In the correct variance simulations the empirical test statistics achieve levels close to the nominal rejection probabilities with combination I of ϕ and still quite close with combination II (keeping in mind the precision of these observed rejection probabilities discussed above). The levels of the model-based test statistics are clearly too low under both combinations of ϕ . Even with the misspecified variance formulation, the empirical test statistics achieve levels close to nominal; while the rejection probabilities may be slightly too large at the 0.10 level under combination II, overall the agreement is good. The levels of the model-based test statistics are far too low at the 0.10 and 0.05 nominal levels under both combinations of ϕ , although it is not clear how they do at the 0.01 level.

Table 12. Observed Rejection Probabilities: Separate to Common.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.076	.035	.006	.062	.026	.001	.067	.036	.005	.056	.029	.009
Wald/Emp	.100	.051	.011	.092	.049	.010	.104	.057	.006	.117	.052	.014
Score/Emp	.099	.049	.008	.088	.049	.008	.100	.055	.006	.113	.051	.010
Dev/Scaled	.077	.035	.006	.062	.027	.001	.067	.036	.006	.056	.030	.009

We also examine the power of the four tests in this multiple series situation in Table 13 by considering the reduction from the common pollution effect to no pollution effects (the null model). The power for the two model-based tests are virtually identical throughout the table and the two empirical tests achieve similar power. However, the rejection probabilities for the model-based tests are higher than the empirical tests presumably as a result of the standard error of γ_0 being underestimated in the former tests (see Table 11).

Table 13. Observed Rejection Probabilities: Common to Null.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.958	.925	.834	.891	.827	.687	.695	.598	.404	.542	.447	.271
Wald/Emp	.941	.890	.751	.837	.767	.537	.570	.445	.226	.399	.293	.111
Score/Emp	.941	.891	.751	.837	.765	.535	.567	.436	.213	.394	.283	.100
Dev/Scaled	.958	.925	.835	.891	.828	.688	.695	.599	.406	.542	.447	.274

The final model to be considered for generating data in this three series simple case is the model with separate pollution effects. Recall the value chosen for γ_0 is again 0.10, while $\gamma_1 = 0.10$, $\gamma_2 = -0.20$ and $\gamma_3 = 0$. Results of fitting the full model to these data sets are summarized in Tables 14, 15 and 16 which follow.

The mean values of the parameter estimates (see Table 14 below) are very close to the correct values, even in the misspecified case with substantial amounts of dispersion. The estimates of ϕ are very close to the values seen in Tables 7 and 10, which is still to be expected because, even in the misspecified variance simulations, the average of the dispersion values over the entire 2100 observations is approximately equal to the averages in the corresponding simulations under either the null or the common pollution effect model.

Table 14. Mean Values of Parameter Estimates.

Parameter	Variance Function			
	A. Correct		B. Misspecified	
	Combination of ϕ		Combination of ϕ	
	I	II	I	II
ϕ	1.997	2.799	4.962	8.326
β_1	-0.290	-0.289	-0.290	-0.289
β_2	0.699	0.700	0.700	0.698
β_3	1.598	1.597	1.595	1.593
γ_0	0.098	0.102	0.102	0.096
γ_1	0.101	0.096	0.100	0.102
γ_2	-0.195	-0.200	-0.197	-0.195

The estimates of the standard errors are summarized in Table 15. In each of the columns of this table the empirical estimator is very close, on average, to the simulated standard deviations. Again the model-based estimator overestimates the standard errors for parameters which correspond to the series with dispersion lower than $\tilde{\phi}$ (most notably β_1 and γ_1) and underestimates the standard errors corresponding to the series with dispersion larger than $\tilde{\phi}$ (most notably β_3). It would seem that the low rejection probabilities in Table 12 for the model-based test statistic can be explained by the general overestimation of the standard error of γ_1 .

Table 15. Standard Deviations of Estimates
and Mean Values of Estimated Standard Errors (\pm s.d.).

Parameter	Method	Variance Function			
		A. Correct		B. Misspecified	
		Combination of ϕ		Combination of ϕ	
		I	II	I	II
ϕ	Simulated	.085	.142	.393	.966
β_1	Simulated	.044	.054	.042	.050
	Model	.062 (.002)	.073 (.003)	.098 (.004)	.127 (.008)
	Empirical	.044 (.002)	.052 (.002)	.044 (.002)	.050 (.002)
β_2	Simulated	.039	.037	.047	.047
	Model	.038 (.001)	.045 (.001)	.059 (.003)	.077 (.005)
	Empirical	.038 (.002)	.038 (.001)	.046 (.002)	.046 (.002)
β_3	Simulated	.029	.039	.056	.076
	Model	.024 (.001)	.028 (.001)	.038 (.001)	.049 (.002)
	Empirical	.029 (.001)	.038 (.001)	.056 (.003)	.077 (.005)
γ_0	Simulated	.042	.054	.078	.113
	Model	.034 (.001)	.040 (.001)	.054 (.002)	.069 (.007)
	Empirical	.041 (.002)	.054 (.003)	.079 (.005)	.109 (.011)
γ_1	Simulated	.074	.090	.098	.133
	Model	.094 (.003)	.111 (.004)	.148 (.006)	.191 (.011)
	Empirical	.075 (.003)	.090 (.003)	.100 (.004)	.130 (.009)
γ_2	Simulated	.071	.076	.102	.131
	Model	.063 (.002)	.075 (.002)	.100 (.004)	.129 (.007)
	Empirical	.067 (.002)	.076 (.003)	.102 (.005)	.127 (.012)

Table 16. Observed Rejection Probabilities: Separate to Common.

Test	Variance Function											
	A. Correct Variance Function						B. Misspec. Variance Function					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.936	.890	.741	.870	.786	.558	.621	.465	.222	.340	.218	.088
Wald/Emp	.962	.932	.822	.944	.889	.747	.911	.858	.685	.880	.791	.593
Score/Emp	.962	.929	.818	.945	.888	.745	.909	.854	.671	.879	.786	.578
Dev/Scaled	.936	.891	.743	.870	.789	.560	.623	.468	.227	.341	.218	.088

In Table 16 we examine the power of the four test statistics under consideration in testing the reduction from the separate pollution effects simulated in these data sets to a model with a common pollution effect. The model-based test statistics have lower power than the empirical tests in all of the above cases. The difference is most noticeable in the cases where the standard errors of γ_1 were seriously overestimated, such as the misspecified variance simulations.

In all of the simulations in this three series simple case, in contrast to the one series simulations, it is clear that the empirical estimate of the covariance matrix for the regression parameters is superior to the model-based estimate. As for the observed levels of the four test statistics under consideration, we find that the model-based test statistics reject too often in some instances, and not as often as predicted by theory in other cases, whereas the empirical test statistics always achieve levels close to the nominal rates. Throughout we have found very close agreement between the model Wald and scaled deviance tests and between the empirical score and empirical Wald tests. It appears that, in these simple simulations, the two model-based test statistics that rely on a dispersion parameter to account for the extra variability perform poorly, while the empirical test statistics (which do not rely on the correctness of a specified model to estimate the variance-covariance matrix) perform much better.

4.4 The More Complicated Case with a Single Series

We now wish to study the performance of the estimators and test statistics when we generate data using the more complicated model for a single series, described by (3.18), that includes season effects. For the purpose of simulating data under the reduced model, the values $\tau_1 = \tau_2 = 0.25$, $\tau_3 = -0.5$ and $\tau_4 = 0$ (with $\gamma = 0$) were used. Table 17 presents the mean values of the parameter estimates from fitting the reduced model to each of the 1000 simulated data sets.

Table 17. Mean Values of Estimated Parameters Under Reduced Model.

A. Correct Variance		Simulated ϕ				
Simulated β	Parameter	1	1.4	2	3	5
log 0.75	ϕ	1.00	1.39	1.98	2.96	-
	β	-.291	-.286	-.295	-.292	-
	τ_1	.253	.245	.253	.247	-
	τ_2	.252	.238	.250	.253	-
	τ_3	-.498	-.513	-.513	-.520	-
0.70	ϕ	.997	1.40	1.99	2.98	-
	β	.699	.697	.698	.695	-
	τ_1	.246	.250	.247	.252	-
	τ_2	.250	.253	.250	.250	-
	τ_3	-.499	-.501	-.501	-.508	-
1.60	ϕ	.997	1.40	2.01	2.98	4.97
	β	1.60	1.60	1.60	1.60	1.60
	τ_1	.251	.249	.247	.251	.246
	τ_2	.250	.250	.248	.249	.251
	τ_3	-.503	-.501	-.502	-.504	-.499

Table 17 B. Misspecified Variance

Simulated β	Parameter	Simulated ϕ				
		1	1.4	2	3	5
log 0.75	ϕ	1.00	1.31	1.77	2.54	-
	β	-.287	-.292	-.297	-.296	-
	τ_1	.245	.246	.256	.248	-
	τ_2	.246	.251	.258	.256	-
	τ_3	-.507	-.503	-.503	-.507	-
0.70	ϕ	1.00	1.83	3.07	5.15	-
	β	.702	.695	.699	.698	-
	τ_1	.247	.253	.244	.246	-
	τ_2	.245	.254	.244	.248	-
	τ_3	-.505	-.501	-.506	-.502	-
1.60	ϕ	1.00	3.06	6.15	11.22	21.09
	β	1.60	1.60	1.60	1.59	1.58
	τ_1	.249	.252	.249	.252	.263
	τ_2	.248	.249	.248	.255	.258
	τ_3	-.500	-.500	-.505	-.497	-.495

In part A of the table, the correct variance simulations, the estimates $\tilde{\phi}$ are very close to the simulated values in each column of the table. The regression parameter estimates (β and τ_i) are also quite good (correct to 1 decimal) even when dispersion is high relative to the mean. There are cases where the estimates appear to be a bit off (for example, the simulations with $\beta = \log 0.75$ and $\phi \geq 1.4$), but the regression parameter estimators have substantially higher variances than in the simple simulations of the previous sections (compare the standard deviations of these estimates provided in Table 18 below with those in Table 3), so it is not unreasonable that the mean value from 1000 simulated data sets differs slightly from the simulated value.

We now consider the misspecified variance results in part B of the above table. Here the values of ϕ are roughly what they were in the one series simple case. This follows because the four seasons are like four short series with different mean levels and therefore different levels of dispersion. Results from the three series simple simulations (Section 4.3) suggest the average

of those levels of dispersion is what is being estimated estimated by the overall estimator $\tilde{\phi}$. Because the τ_i were chosen so that $\sum_i \tau_i = 0$, the average is approximately the same as the amount of dispersion in the one series simple simulations of Section 4.2 (slightly different because the mean levels of the four “series” don’t quite average out to the overall mean; $\sum_i \exp\{\beta + \tau_i\}$ does not equal $\exp\{\sum_i(\beta + \tau_i)\} = \exp\{\beta\}$). Thus the estimates of ϕ in the above table are as we would expect. More importantly, the mean values of the regression parameter estimates are quite close to the simulated values at all levels of dispersion and for all choices of β .

The standard deviations of the parameter estimates and the mean values of the standard errors of the regression parameters are summarized in Table 18. In the correct variance simulations (part A of the table), the empirical and model-based estimators are very close to each other in all cases, and agree well with the actual standard deviations of the parameter estimates. There is one instance ($\beta = \log 0.75$ and $\phi = 3$) when the mean values of the estimated standard errors for all the regression parameters are slightly lower than the standard deviations. This could be a suggestion that the estimators underestimate the standard errors when there is a large amount of dispersion and low overall mean level, but could also just be chance. In any case, there is no strong evidence that the empirical estimator particularly underestimates the standard errors in general as expected *a priori* (based on conventional wisdom), or at least not any more so than the model-based estimator. Overall there is no real difference in these estimated standard errors. Note however that the model-based estimated standard error is uniformly less variable than that based on the empirical estimator (although the difference is small in this particular situation).

In the misspecified variance simulations there may be a problem with the model-based estimator similar to the problems in the three series simple case simulations of Section 4.3. If we think of the four seasons as four short series, then under the misspecified variance formulation, the four series, with different mean levels, have different amounts of dispersion. In particular, the simulated dispersion for the data in the third season will be quite a bit lower than for the other seasons. The “overall” estimate $\tilde{\phi}$ will be larger than the amount of dispersion simulated for this season which might explain why the model-based estimates of standard errors for $\tilde{\tau}_3$ are too high on average. This same argument would suggest that the standard errors for $\tilde{\tau}_1$ and $\tilde{\tau}_2$ would be too low since these “series” have the highest means and therefore the most dispersion so that $\tilde{\phi}$ will be lower than the amount of dispersion simulated in these series. There is some suggestion of this as well in part B of Table 12.

As for the empirical estimator, overall it appears to estimate the standard deviations of the parameters much better than the model-based estimator. There is no clear evidence that this estimator consistently underestimates the standard errors of the regression parameters although whenever the average of the empirical estimates differs to any appreciable extent from a simulated standard deviation, this average does appear to be too low.

Table 18. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).

A. Correct Variance Function

Simulated β	Parameter	Method	Simulated ϕ				
			1	1.4	2	3	5
log 0.75	ϕ	Simulated	.056	.097	.175	.326	-
	β	Simulated	.086	.103	.124	.153	-
		Model	.088(.004)	.103(.006)	.123(.009)	.150(.012)	-
		Empirical	.087(.006)	.103(.008)	.123(.011)	.150(.015)	-
	τ_1	Simulated	.118	.135	.165	.206	-
		Model	.117(.005)	.138(.006)	.165(.009)	.201(.013)	-
		Empirical	.116(.006)	.137(.007)	.164(.010)	.200(.014)	-
	τ_2	Simulated	.115	.135	.164	.205	-
		Model	.117(.005)	.138(.006)	.165(.009)	.201(.013)	-
		Empirical	.116(.006)	.137(.007)	.164(.010)	.200(.014)	-
	τ_3	Simulated	.142	.166	.205	.258	-
		Model	.143(.007)	.168(.009)	.202(.012)	.247(.019)	-
		Empirical	.142(.008)	.168(.010)	.201(.014)	.244(.021)	-
0.70	ϕ	Simulated	.056	.087	.136	.245	-
	β	Simulated	.055	.064	.077	.093	-
		Model	.053(.002)	.063(.003)	.075(.004)	.092(.005)	-
		Empirical	.053(.003)	.063(.004)	.075(.005)	.092(.007)	-
	τ_1	Simulated	.074	.087	.097	.121	-
		Model	.071(.002)	.084(.003)	.100(.004)	.123(.005)	-
		Empirical	.071(.003)	.084(.004)	.100(.005)	.122(.006)	-
	τ_2	Simulated	.072	.085	.101	.123	-
		Model	.071(.002)	.084(.003)	.100(.004)	.123(.006)	-
		Empirical	.071(.003)	.084(.004)	.100(.005)	.122(.006)	-
	τ_3	Simulated	.089	.103	.123	.153	-
		Model	.087(.003)	.103(.004)	.123(.005)	.150(.007)	-
		Empirical	.086(.004)	.102(.005)	.122(.006)	.149(.009)	-
1.60	ϕ	Simulated	.054	.078	.122	.200	.388
	β	Simulated	.034	.042	.047	.060	.077
		Model	.034(.001)	.040(.001)	.048(.002)	.059(.002)	.076(.004)
		Empirical	.034(.002)	.040(.002)	.048(.003)	.059(.004)	.076(.005)
	τ_1	Simulated	.046	.056	.062	.081	.105
		Model	.045(.001)	.054(.002)	.064(.002)	.078(.003)	.101(.004)
		Empirical	.045(.002)	.053(.002)	.064(.003)	.078(.004)	.101(.005)
	τ_2	Simulated	.045	.054	.063	.079	.102
		Model	.045(.001)	.054(.002)	.064(.002)	.078(.003)	.101(.004)
		Empirical	.045(.002)	.053(.002)	.064(.003)	.078(.004)	.101(.005)
	τ_3	Simulated	.057	.067	.078	.100	.126
		Model	.055(.002)	.065(.002)	.078(.003)	.096(.004)	.123(.006)
		Empirical	.055(.002)	.065(.003)	.078(.003)	.095(.005)	.123(.007)

Table 18 B. Misspecified Variance Function

Simulated β	Parameter	Method	Simulated ϕ				
			1	1.4	2	3	5
log 0.75	ϕ	Simulated	.052	.087	.140	.253	-
	β	Simulated	.088	.101	.116	.139	-
		Model	.087(.005)	.100(.006)	.117(.008)	.140(.011)	-
		Empirical	.087(.006)	.099(.008)	.115(.009)	.137(.013)	-
	τ_1	Simulated	.119	.137	.162	.196	-
		Model	.117(.005)	.134(.006)	.156(.008)	.186(.011)	-
		Empirical	.116(.006)	.135(.007)	.157(.009)	.190(.012)	-
	τ_2	Simulated	.118	.136	.165	.192	-
		Model	.117(.005)	.134(.006)	.156(.008)	.186(.011)	-
		Empirical	.116(.006)	.134(.007)	.157(.009)	.189(.012)	-
	τ_3	Simulated	.140	.157	.185	.208	-
		Model	.143(.007)	.163(.009)	.191(.011)	.228(.016)	-
		Empirical	.142(.008)	.158(.010)	.178(.011)	.207(.015)	-
0.70	ϕ	Simulated	.054	.120	.261	.525	-
	β	Simulated	.055	.072	.094	.125	-
		Model	.053(.002)	.072(.003)	.093(.005)	.121(.009)	-
		Empirical	.053(.003)	.071(.005)	.091(.007)	.118(.011)	-
	τ_1	Simulated	.073	.100	.132	.172	-
		Model	.071(.002)	.096(.004)	.125(.006)	.161(.009)	-
		Empirical	.071(.003)	.097(.005)	.128(.007)	.165(.010)	-
	τ_2	Simulated	.071	.099	.130	.165	-
		Model	.071(.002)	.096(.004)	.125(.006)	.161(.009)	-
		Empirical	.071(.003)	.098(.004)	.127(.007)	.165(.011)	-
	τ_3	Simulated	.090	.111	.146	.177	-
		Model	.087(.003)	.118(.005)	.152(.008)	.197(.012)	-
		Empirical	.087(.004)	.109(.005)	.136(.008)	.173(.011)	-
1.60	ϕ	Simulated	.057	.209	.525	1.23	2.96
	β	Simulated	.034	.059	.084	.118	.155
		Model	.034(.001)	.059(.002)	.084(.005)	.114(.008)	.157(.014)
		Empirical	.034(.002)	.058(.004)	.082(.006)	.110(.010)	.152(.017)
	τ_1	Simulated	.047	.083	.118	.165	.217
		Model	.045(.001)	.079(.003)	.112(.005)	.152(.008)	.210(.015)
		Empirical	.045(.002)	.081(.004)	.115(.006)	.156(.010)	.215(.017)
	τ_1	Simulated	.046	.081	.119	.159	.214
		Model	.045(.001)	.079(.003)	.112(.005)	.152(.008)	.210(.015)
		Empirical	.045(.002)	.081(.004)	.115(.006)	.156(.010)	.215(.018)
	τ_3	Simulated	.055	.090	.122	.164	.221
		Model	.055(.002)	.097(.004)	.137(.007)	.186(.013)	.256(.020)
		Empirical	.055(.002)	.087(.004)	.120(.006)	.159(.012)	.218(.018)

In the next table we present the observed rejection probabilities for the four test statistics under study in each of the combinations of β and ϕ considered above. Here the alternate model (to be fit to the data sets generated under the reduced model) includes the sinusoidal “pollution” covariate. In the correct variance simulations the model Wald and scaled deviance tests give almost identical results in all simulations. The empirical tests are not as similar (either to each other or to the model-based tests), but overall there appears to be little to distinguish any of the tests. The observed rejection probabilities are reasonably close to the nominal levels in almost all cases. Possible exceptions might be the simulations with $\beta = 1.60$ and either of $\phi = 1$ or $\phi = 3$, but these are likely just chance deviations from the nominal levels.

In the misspecified variance simulations we also find the model Wald and scaled deviance tests to be almost identical in their observed rejection probabilities. In general the score test appears most conservative and it seems to get more conservative (relative to the others) as the amount of dispersion increases. This makes it look like it is performing better than the others when the others reject too often and it looks worse when the others reject at about the nominal rate.

As a side note, based on the model-based standard errors of Table 18, we might look carefully at the empirical Wald test to see if it rejects too often, since it seemed the empirical estimates of the standard errors of some of the $\tilde{\tau}_i$ were too small, which might imply that the standard errors of the coefficient of the pollution covariate could be underestimated. The empirical Wald test does reject more often than the score test in general, but it occasionally rejects less often than the model-based tests so that no general conclusion regarding its performance can be reached from these results.

Table 19. Observed Rejection Probabilities Under Null Hypothesis.

A. Correct Variance Function

Simulated β	Test/Method	$\phi = 1$			$\phi = 1.4$			$\phi = 2$		
		0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.093	.060	.011	.112	.056	.012	.110	.052	.014
	Wald/Empirical	.098	.063	.013	.114	.057	.010	.115	.053	.012
	Score/Empirical	.098	.062	.012	.112	.055	.009	.110	.047	.011
	Deviance/Scaled	.093	.060	.011	.111	.058	.012	.110	.053	.014
0.70	Wald/Model	.100	.052	.014	.099	.045	.008	.098	.052	.013
	Wald/Empirical	.101	.049	.012	.102	.044	.010	.101	.056	.011
	Score/Empirical	.095	.047	.012	.099	.044	.010	.097	.053	.011
	Deviance/Scaled	.100	.052	.014	.099	.044	.008	.098	.052	.013
1.60	Wald/Model	.115	.054	.011	.092	.040	.011	.109	.053	.009
	Wald/Empirical	.121	.055	.011	.091	.041	.012	.111	.053	.010
	Score/Empirical	.119	.053	.010	.091	.041	.010	.109	.053	.008
	Deviance/Scaled	.115	.055	.011	.092	.040	.011	.109	.054	.009

Simulated β	Test/Method	$\phi = 3$			$\phi = 5$		
		0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.112	.062	.012	-	-	-
	Wald/Empirical	.118	.066	.016	-	-	-
	Score/Empirical	.111	.059	.011	-	-	-
	Deviance/Scaled	.113	.062	.012	-	-	-
0.70	Wald/Model	.099	.048	.009	-	-	-
	Wald/Empirical	.107	.048	.013	-	-	-
	Score/Empirical	.101	.045	.007	-	-	-
	Deviance/Scaled	.100	.047	.009	-	-	-
1.60	Wald/Model	.121	.062	.014	.105	.048	.010
	Wald/Empirical	.125	.063	.015	.108	.057	.012
	Score/Empirical	.123	.057	.014	.106	.051	.009
	Deviance/Scaled	.121	.062	.014	.106	.049	.010

Table 19 B. Misspecified Variance Function

Simulated β	Test/Method	$\phi = 1$			$\phi = 1.4$			$\phi = 2$		
		0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.110	.051	.005	.112	.049	.004	.111	.058	.010
	Wald/Empirical	.113	.055	.005	.111	.047	.004	.112	.059	.007
	Score/Empirical	.111	.052	.005	.110	.042	.004	.109	.055	.005
	Deviance/Scaled	.110	.052	.005	.112	.049	.004	.111	.059	.010
0.70	Wald/Model	.095	.049	.009	.105	.049	.010	.109	.060	.013
	Wald/Empirical	.098	.051	.013	.103	.048	.008	.099	.055	.015
	Score/Empirical	.095	.051	.011	.100	.046	.007	.094	.057	.013
	Deviance/Scaled	.095	.049	.009	.105	.048	.009	.109	.060	.013
1.60	Wald/Model	.101	.051	.010	.102	.053	.013	.105	.052	.010
	Wald/Empirical	.102	.054	.011	.102	.046	.011	.098	.046	.008
	Score/Empirical	.102	.053	.010	.099	.045	.010	.092	.045	.008
	Deviance/Scaled	.101	.051	.010	.103	.053	.013	.104	.053	.011

Simulated β	Test/Method	$\phi = 3$			$\phi = 5$		
		0.10	0.05	0.01	0.10	0.05	0.01
log 0.75	Wald/Model	.113	.059	.010	-	-	-
	Wald/Empirical	.110	.056	.010	-	-	-
	Score/Empirical	.102	.052	.007	-	-	-
	Deviance/Scaled	.113	.060	.011	-	-	-
0.70	Wald/Model	.102	.052	.018	-	-	-
	Wald/Empirical	.093	.049	.014	-	-	-
	Score/Empirical	.087	.046	.013	-	-	-
	Deviance/Scaled	.103	.053	.018	-	-	-
1.60	Wald/Model	.108	.056	.016	.112	.056	.018
	Wald/Empirical	.104	.050	.017	.110	.056	.018
	Score/Empirical	.100	.046	.013	.103	.049	.015
	Deviance/Scaled	.108	.057	.016	.114	.057	.018

We now consider the simulations under the full model (with pollution included). As in the simple case simulations, we chose to simulate data using only $\beta = \log 0.75$ but with one of $\gamma = 0.05, 0.10$ and 0.15 . Table 20 presents the mean values of the estimated parameters from the simulated data sets generated under the above model.

Table 20. Mean Values of Estimated Parameters Under Alternative Hypothesis.

A. Correct Variance Function

Simulated γ	Parameter	Simulated ϕ			
		1	1.4	2	3
0.05	ϕ	1.00	1.40	1.97	2.94
	β	-.293	-.295	-.311	-.303
	τ_1	.254	.253	.274	.256
	τ_2	.249	.251	.276	.254
	τ_3	-.497	-.503	-.503	-.513
	γ	.054	.051	.038	.055
0.10	ϕ	1.00	1.39	1.97	2.94
	β	-.289	-.287	-.304	-.306
	τ_1	.246	.237	.259	.258
	τ_2	.247	.244	.260	.259
	τ_3	-.506	-.511	-.510	-.515
	γ	.103	.104	.097	.098
0.15	ϕ	1.00	1.40	1.98	2.94
	β	-.289	-.304	-.300	-.299
	τ_1	.248	.271	.260	.252
	τ_2	.247	.272	.259	.259
	τ_3	-.498	-.511	-.515	-.532
	γ	.154	.132	.147	.144

B. Misspecified Variance Function

Simulated γ	Parameter	Simulated ϕ			
		1	1.4	2	3
0.05	ϕ	1.00	1.31	1.77	2.54
	β	-.293	-.295	-.296	-.293
	τ_1	.254	.253	.250	.235
	τ_2	.249	.252	.254	.246
	τ_3	-.497	-.504	-.504	-.515
	γ	.054	.050	.050	.056
0.10	ϕ	1.00	1.31	1.79	2.56
	β	-.289	-.294	-.295	-.291
	τ_1	.246	.250	.248	.240
	τ_2	.247	.251	.255	.241
	τ_3	-.506	-.505	-.504	-.516
	γ	.103	.104	.101	.102
0.15	ϕ	1.00	1.32	1.79	2.57
	β	-.289	-.298	-.294	-.303
	τ_1	.248	.254	.246	.252
	τ_2	.247	.263	.246	.257
	τ_3	-.498	-.501	-.502	-.496
	γ	.154	.145	.149	.155

Before interpreting these results, note that there is non-negligible collinearity between the continuous pollution covariate and the seasonal indicators. If, in our simulations, two explanatory variables have a substantial positive correlation, one of the corresponding regression parameters could be estimated to be larger than its simulated (or true) value, while the other could be estimated to be smaller than its simulated value, with little effect on the goodness-of-fit. Similarly, both could be estimated to be larger than the simulated values if negatively correlated. Such deviations from the true values, beyond the normal variation we would expect, lead to more variability from one set of fitted parameters to another than if the predictors were not collinear. The sample correlations are approximately 0.52 between the pollution variable and each of the indicators for the first and second season, and -0.52 between the pollution variable and the indicator for the third season. While such correlations are not a major concern, they must be kept in mind when examining the results of these simulations.

For the correct variance simulations, the mean values of the regression parameter estimates are generally close to their simulated values, considering how big the standard deviations of the parameter estimates are (and thus the standard error of the average of the parameter estimates as an estimate of the true mean, see Table 21). However, in the cases $\gamma = 0.05$, $\phi = 2$ and $\gamma = 0.15$, $\phi = 1.4$ the averages of the estimates of τ_1 and τ_2 exceed the simulated value by more than two standard errors (while the estimate of γ is too small). In the case $\gamma = 0.15$, $\phi = 3$ the estimate of τ_3 is too small (a large negative value that differs from the true value by more than two standard errors) with the estimate of γ slightly too small, although it differs from the true value by less than one standard error. Note how the pattern of over and underestimation agrees with what we might expect when the pollution variable is positively correlated with the

first two season indicators and negatively correlated with the third. But there is also a hint of a pattern in the estimates of τ_3 which has no such plausible explanation; in general, as ϕ increases across the table, τ_3 appears to decrease (become a more negative value). The observations in the third season have the lowest mean values, and as the amount of dispersion increases, there will be more and more zero values among these observations. We might infer that for large enough amounts of over-dispersion, this could be causing the estimator to underestimate the season effect, but with the low degree of precision in the averages of the parameter estimates we can not conclude whether this is the case, or if the deviations from the simulated values are just chance.

In the misspecified variance simulations, the regression parameters estimates are very good; in fact, they appear somewhat better than in the correct variance case. The estimates of ϕ seen in parts A and B of this table are similar to those in Table 17.

The standard deviations of the parameter estimates in the 1000 simulations and the mean values of the standard errors of the regression parameters are summarized in Table 21. As in part A of Table 18, under the correct variance formulation the model and empirical estimates are very similar. Both estimators are very close to the standard deviations of the parameter estimates when the dispersion is 1 or 1.4. For simulated dispersions of 2 or 3, the averages of the estimates are not as close to the simulated values but the differences can be explained by the variability of the observed standard deviations and the estimated standard errors.

Table 21. Standard Deviations of Estimates Under Alternative Hypothesis and Mean Values of Estimated Standard Errors (\pm s.d.).

A. Correct Variance Function

Simulated γ	Parameter	Method	Simulated ϕ			
			1	1.4	2	3
0.05	ϕ	Simulated	.054	.099	.166	.326
	β	Simulated	.122	.149	.177	.221
		Model	.125(.005)	.148(.007)	.176(.009)	.215(.014)
		Empirical	.124(.007)	.147(.010)	.175(.015)	.214(.024)
	τ_1	Simulated	.209	.255	.300	.374
		Model	.212(.007)	.251(.010)	.298(.013)	.364(.020)
		Empirical	.211(.012)	.249(.017)	.296(.026)	.362(.041)
	τ_2	Simulated	.210	.246	.305	.373
		Model	.213(.007)	.251(.010)	.298(.013)	.365(.020)
		Empirical	.211(.012)	.249(.018)	.296(.026)	.364(.042)
	τ_3	Simulated	.143	.171	.194	.261
		Model	.145(.007)	.172(.009)	.205(.012)	.251(.019)
		Empirical	.144(.008)	.171(.011)	.203(.014)	.247(.021)
	γ	Simulated	.138	.166	.203	.243
		Model	.139(.005)	.164(.007)	.205(.012)	.238(.013)
		Empirical	.138(.007)	.163(.009)	.203(.014)	.238(.021)
0.10	ϕ	Simulated	.054	.096	.173	.321
	β	Simulated	.126	.151	.172	.224
		Model	.126(.005)	.148(.007)	.177(.009)	.216(.014)
		Empirical	.125(.008)	.147(.011)	.175(.015)	.214(.023)
	τ_1	Simulated	.210	.255	.287	.380
		Model	.212(.008)	.250(.010)	.298(.013)	.364(.019)
		Empirical	.210(.012)	.248(.018)	.295(.026)	.361(.040)
	τ_2	Simulated	.217	.250	.289	.385
		Model	.212(.008)	.250(.010)	.299(.013)	.364(.019)
		Empirical	.211(.012)	.249(.018)	.295(.026)	.360(.041)
	τ_3	Simulated	.144	.174	.215	.261
		Model	.147(.007)	.174(.009)	.209(.013)	.255(.020)
		Empirical	.147(.008)	.173(.011)	.207(.015)	.251(.021)
	γ	Simulated	.139	.166	.188	.249
		Model	.138(.005)	.163(.006)	.195(.009)	.237(.012)
		Empirical	.138(.007)	.162(.010)	.193(.014)	.236(.022)
0.15	ϕ	Simulated	.054	.096	.178	.332
	β	Simulated	.128	.150	.174	.223
		Model	.126(.005)	.149(.007)	.177(.010)	.216(.015)
		Empirical	.125(.008)	.148(.011)	.176(.016)	.215(.024)
	τ_1	Simulated	.220	.251	.284	.372
		Model	.211(.007)	.250(.010)	.298(.013)	.363(.020)
		Empirical	.210(.012)	.248(.017)	.296(.027)	.361(.040)
	τ_2	Simulated	.216	.248	.292	.374
		Model	.211(.007)	.250(.010)	.298(.013)	.363(.020)
		Empirical	.211(.012)	.248(.018)	.297(.026)	.361(.040)
	τ_3	Simulated	.152	.175	.222	.266
		Model	.149(.007)	.177(.010)	.212(.014)	.259(.020)
		Empirical	.149(.008)	.176(.011)	.210(.015)	.255(.022)
	γ	Simulated	.141	.161	.186	.243
		Model	.138(.005)	.163(.006)	.194(.008)	.236(.013)
		Empirical	.137(.007)	.162(.010)	.193(.014)	.235(.021)

Table 21 B. Misspecified Variance Function

Simulated γ	Parameter	Method	Simulated ϕ			
			1	1.4	2	3
0.05	ϕ	Simulated	.053	.086	.144	.259
	β	Simulated	.122	.145	.169	.197
		Model	.125(.005)	.143(.007)	.167(.008)	.199(.012)
		Empirical	.124(.007)	.142(.010)	.166(.013)	.199(.020)
	τ_1	Simulated	.209	.245	.292	.351
		Model	.212(.007)	.243(.009)	.282(.012)	.338(.017)
		Empirical	.211(.012)	.243(.017)	.285(.023)	.343(.037)
	τ_2	Simulated	.210	.253	.290	.345
		Model	.213(.007)	.243(.010)	.283(.012)	.338(.017)
		Empirical	.211(.012)	.243(.017)	.286(.024)	.343(.037)
	τ_3	Simulated	.143	.163	.189	.212
		Model	.145(.007)	.166(.009)	.194(.011)	.232(.016)
		Empirical	.144(.008)	.159(.009)	.179(.011)	.208(.014)
	γ	Simulated	.138	.167	.191	.227
		Model	.139(.005)	.159(.006)	.185(.008)	.221(.011)
		Empirical	.138(.007)	.159(.009)	.186(.013)	.225(.020)
0.10	ϕ	Simulated	.054	.085	.143	.264
	β	Simulated	.126	.145	.172	.203
		Model	.126(.005)	.144(.007)	.168(.009)	.201(.012)
		Empirical	.125(.008)	.143(.010)	.167(.014)	.199(.020)
	τ_1	Simulated	.210	.244	.295	.350
		Model	.212(.008)	.243(.009)	.283(.012)	.339(.016)
		Empirical	.210(.012)	.244(.017)	.285(.024)	.343(.037)
	τ_2	Simulated	.217	.247	.295	.349
		Model	.212(.008)	.243(.009)	.283(.012)	.339(.017)
		Empirical	.211(.012)	.243(.017)	.286(.025)	.342(.037)
	τ_3	Simulated	.145	.171	.190	.220
		Model	.147(.007)	.169(.009)	.197(.012)	.237(.017)
		Empirical	.147(.008)	.162(.010)	.182(.01)	.210(.015)
	γ	Simulated	.139	.160	.193	.227
		Model	.138(.005)	.159(.006)	.184(.008)	.221(.010)
		Empirical	.138(.007)	.159(.009)	.187(.013)	.225(.019)
0.15	ϕ	Simulated	.054	.084	.146	.259
	β	Simulated	.120	.148	.170	.207
		Model	.126(.005)	.145(.007)	.168(.009)	.203(.012)
		Empirical	.125(.008)	.144(.010)	.166(.013)	.199(.019)
	τ_1	Simulated	.202	.252	.289	.348
		Model	.212(.008)	.243(.009)	.283(.012)	.340(.017)
		Empirical	.210(.012)	.244(.017)	.284(.023)	.342(.035)
	τ_2	Simulated	.204	.251	.296	.346
		Model	.212(.008)	.243(.009)	.283(.012)	.340(.017)
		Empirical	.211(.012)	.244(.017)	.285(.024)	.343(.035)
	τ_3	Simulated	.148	.168	.183	.226
		Model	.150(.007)	.172(.009)	.200(.013)	.241(.017)
		Empirical	.149(.008)	.164(.010)	.183(.012)	.211(.014)
	γ	Simulated	.136	.165	.194	.224
		Model	.138(.005)	.158(.006)	.184(.008)	.221(.011)
		Empirical	.137(.007)	.159(.009)	.187(.013)	.226(.019)

To estimate the variance of the observed standard deviations, note the following. Because regression parameter estimates are asymptotically normal, if $\tilde{\theta}_i$ is a regression parameter estimate from the i^{th} simulated data set and $\bar{\tilde{\theta}}$ is the average of these estimates, then $\sigma^{-2} \sum_i (\tilde{\theta}_i - \bar{\tilde{\theta}})^2$, where σ^2 represents $\text{Var}(\tilde{\theta}_i)$, should have approximately a χ^2_{n-1} distribution. It follows that the sample variance of these estimates should have variance roughly equal to $2\sigma^4/(n-1)$. Using the delta method we arrive at the approximation $\text{Var}(s) \approx \sigma^2/[2(n-1)]$.

Using this approximation, observed standard deviations of $s = 0.3$ and $s = 0.2$ would have standard errors of approximately 0.0067 and 0.0045 respectively. The mean values of the estimated standard errors will have approximate standard errors given by their standard deviations (enclosed in brackets in the table) divided by $\sqrt{1000}$. When ϕ is 2 this can be as large as $.027/\sqrt{1000} = 0.0009$. With these precisions in mind, the observed differences between the variance estimates and the standard deviations are not too large.

Another approach to determining whether the model-based and empirical standard errors approximate the simulated standard errors can be based on the increased precision provided by generating more data sets for a given choice of γ and ϕ . The following table summarizes the results of the simulations for the case $\gamma = 0.10$ and $\phi = 2$ as displayed in part A of Table 21, and the results of five repeats of the simulation experiment using the same γ and ϕ .

Parameter	Method	original	repeat 1	repeat 2	repeat 3	repeat 4	repeat 5	average
β	Simulated	.172	.169	.182	.180	.177	.179	.177
	Model	.177	.177	.177	.177	.177	.177	.177
	Empirical	.175	.175	.176	.176	.176	.175	.176
τ_1	Simulated	.287	.293	.311	.307	.296	.298	.299
	Model	.298	.299	.299	.298	.298	.298	.298
	Empirical	.295	.296	.296	.297	.295	.296	.296
τ_2	Simulated	.289	.289	.304	.299	.297	.300	.296
	Model	.299	.299	.299	.299	.298	.299	.299
	Empirical	.295	.296	.297	.297	.297	.296	.296
τ_3	Simulated	.215	.206	.209	.215	.207	.215	.211
	Model	.209	.209	.209	.209	.208	.209	.209
	Empirical	.207	.207	.207	.207	.207	.207	.207
γ	Simulated	.188	.196	.196	.197	.194	.192	.194
	Model	.195	.195	.195	.195	.194	.195	.195
	Empirical	.193	.194	.194	.193	.194	.194	.194

From the above table we note that in each simulation experiment, the mean values of the model and empirical estimates of the standard errors are very close to the average of the simulated standard deviations across the columns of the table. This confirms that the model and empirical estimators do a good job of estimating the standard deviations of the parameter estimates.

We now examine part B of Table 21 corresponding to the simulations using the misspecified variance formulation to generate the data. Here we seem to be seeing the same pattern as in the simulations using data generated with the reduced model (part B of Table 18). The model-based estimates appear too large for regression parameters which correspond to “seasons” with lower than average dispersion and too small for parameters which correspond to “seasons” with larger than average dispersion. In this situation, with a single pollution covariate, there is no clear picture of under or overestimation of the standard errors of γ . However, one might speculate that if the pollution covariate were associated with one of the seasons (such as a

season \times pollution interaction) the standard errors of the regression parameter might be under or overestimated, depending upon which season the covariate was related to, and this could lead to unexpected results. As for the empirical estimator, the results suggest the possibility that the standard errors of the regression parameters tend to be underestimated in these simulations. Of course, when making these observations we should keep in mind the possibility that the observed standard deviations of the parameter estimates may differ somewhat from the true standard deviations. A similar exercise to that carried out in the correct variance simulations was carried out for these simulations under the misspecified variance for the same combination $\gamma = 0.10$ and $\phi = 2$. The final column of the table which follows shows the averages of the 6 simulation experiments.

Parameter	Method	original	repeat 1	repeat 2	repeat 3	repeat 4	repeat 5	average
β	Simulated	.172	.165	.164	.171	.168	.167	.168
	Model	.168	.167	.168	.168	.168	.167	.168
	Empirical	.167	.166	.166	.167	.166	.166	.166
τ_1	Simulated	.295	.289	.287	.288	.282	.291	.289
	Model	.283	.282	.283	.282	.282	.282	.282
	Empirical	.285	.285	.284	.286	.285	.285	.285
τ_2	Simulated	.295	.288	.286	.293	.286	.289	.290
	Model	.283	.283	.283	.283	.283	.283	.283
	Empirical	.286	.285	.285	.286	.285	.285	.285
τ_3	Simulated	.190	.189	.185	.182	.188	.190	.187
	Model	.197	.196	.197	.197	.197	.197	.197
	Empirical	.182	.180	.181	.181	.181	.181	.181
γ	Simulated	.193	.193	.192	.185	.190	.195	.191
	Model	.185	.184	.184	.184	.184	.184	.184
	Empirical	.187	.187	.187	.187	.187	.187	.187

The final column of this table suggests the model-based estimator does underestimate the standard errors for τ_1 and τ_2 , and overestimate the standard error for τ_3 . We may also note that the standard error of γ is underestimated by the model-based estimator in these simulations.

This part of the table is also consistent with the observation that the empirical estimator appears to slightly underestimate the standard errors. Whether this is true for other combinations of γ and ϕ remains speculation.

In summary, for the correct variance simulations, both the model-based and the empirical estimators of standard errors perform well. However in the misspecified variance simulations, there is some evidence that the model-based estimator does not perform well and that the empirical estimator may underestimate the standard errors slightly.

Table 22. Observed Rejection Probabilities Under Alternative Hypothesis.

A. Correct Variance Function													
γ	Test	$\phi = 1$			$\phi = 1.4$			$\phi = 2$			$\phi = 3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
.05	Wald/Mod.	.121	.065	.014	.116	.062	.014	.121	.066	.012	.115	.063	.020
	Wald/Emp.	.124	.069	.015	.120	.062	.012	.127	.072	.014	.113	.066	.020
	Score/Emp.	.124	.070	.014	.116	.063	.012	.126	.067	.013	.110	.063	.017
	Dev/Scaled	.121	.065	.014	.116	.062	.014	.122	.067	.013	.115	.064	.022
.10	Wald/Mod.	.185	.104	.027	.170	.102	.030	.130	.060	.016	.150	.087	.018
	Wald/Emp.	.196	.110	.030	.171	.103	.034	.129	.068	.018	.159	.085	.019
	Score/Emp.	.196	.108	.028	.167	.100	.032	.130	.068	.016	.156	.081	.017
	Dev/Scaled	.185	.108	.027	.176	.104	.030	.131	.061	.017	.151	.089	.020
.15	Wald/Mod.	.298	.191	.076	.201	.126	.038	.195	.111	.019	.165	.087	.027
	Wald/Emp.	.300	.189	.081	.205	.134	.039	.188	.115	.027	.161	.095	.025
	Score/Emp.	.299	.190	.080	.206	.131	.040	.190	.114	.027	.159	.093	.023
	Dev/Scaled	.299	.192	.076	.206	.128	.040	.196	.113	.019	.167	.092	.028
B. Misspecified Variance Function													
γ	Test	$\phi = 1$			$\phi = 1.4$			$\phi = 2$			$\phi = 3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
.05	Wald/Mod.	.121	.065	.014	.136	.075	.019	.120	.062	.018	.122	.052	.012
	Wald/Emp.	.124	.069	.015	.129	.079	.023	.117	.062	.018	.118	.047	.011
	Score/Emp.	.124	.070	.014	.129	.076	.020	.116	.058	.016	.111	.042	.007
	Dev/Scaled	.121	.065	.014	.137	.076	.019	.122	.063	.018	.121	.053	.013
.10	Wald/Mod.	.185	.104	.027	.165	.092	.031	.167	.096	.024	.142	.078	.016
	Wald/Emp.	.196	.110	.030	.159	.090	.034	.165	.099	.020	.135	.074	.012
	Score/Emp.	.196	.108	.028	.159	.089	.032	.161	.097	.019	.129	.068	.012
	Dev/Scaled	.185	.108	.027	.165	.095	.031	.168	.097	.024	.145	.080	.018
.15	Wald/Mod.	.298	.191	.076	.253	.166	.055	.219	.136	.047	.190	.106	.027
	Wald/Emp.	.300	.189	.081	.253	.166	.055	.212	.137	.048	.184	.105	.024
	Score/Emp.	.299	.190	.080	.250	.163	.052	.207	.133	.046	.179	.104	.022
	Dev/Scaled	.299	.192	.076	.254	.167	.056	.221	.138	.048	.191	.109	.028

The power of the four test statistics under study to detect the simulated pollution effect is summarized in Table 22; this table is similar in nature to Table 6 for the one series simple simulations. The results of the correct variance simulations suggest very low power for all test statistics in this more complicated case (only marginally higher than the nominal levels for the smallest simulated value of γ), but this is not surprising given the variability of the estimates

of γ compared to their simulated values. As in Table 6 the power of the tests tends to decrease as the amount of dispersion increases. The test statistics show similar rejection probabilities in the misspecified variance simulations. As in Table 6, the rejection probabilities, in general, are slightly higher in the misspecified variance simulations than in the correct variance simulations because for each simulated value of ϕ , there is less dispersion in the data under the misspecified variance formulation. Thus, except for the fact that the rejection probabilities are lower in Table 21 than in Table 6 due to the larger standard errors of γ in these more complicated simulations, the test statistics perform similarly in the simple and complicated simulations with one series.

We now summarize the findings of these simulations using the more complicated model for a single series of data. Estimates of ϕ and of the regression parameters are relatively unaffected by the presence of over-dispersion although the estimators of these parameters became more variable as expected. Standard errors of the regression parameters were accurately estimated with either the model-based estimator or the empirical estimator when the data was generated using the correct variance formulation, but there is some evidence the model-based estimator performs poorly when the misspecified variance formulation was used to generate data. The empirical estimator appeared to do better in the misspecified variance simulations but it seemed to underestimate the standard errors when the level of dispersion was large relative to the overall mean level (this was not so apparent in the simulations under the reduced model reported in Table 19, but was more readily seen in the simulations under the full model reported in Table 21). In these simulations the problems with estimating the standard errors did not translate into problems with the model-based test statistics. This is in contrast to the simple case simulations with three series where poor performance of the model-based estimator of variances appeared

to result in poor performance of the model-based test statistics. There were no detectable differences among any of the four test statistics under study.

4.5 The More Complicated Case with Three Series

The last case to be considered is the simulations based on (3.19) describing the more complicated case with three series. Recall that the simulated values for the mean and temporal parameters were:

$$\begin{array}{lll} \beta_1 = \log 0.75 & \beta_2 = .70 & \beta_3 = 1.6 \\ \tau_{11} = 0.2 & \tau_{21} = 0.2 & \tau_{31} = 0.2 \\ \tau_{12} = 0.2 & \tau_{22} = 0.2 & \tau_{32} = 0.3 \\ \tau_{13} = -0.2 & \tau_{23} = -0.4 & \tau_{33} = -0.4 \\ \tau_{41} = 0 & \tau_{42} = 0 & \tau_{43} = 0 \end{array}$$

The first simulations to be considered use the null model ($\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = 0$) to generate data. The mean values of the parameter estimates from fitting the null model to these data sets are summarized in Table 23. For the correct variance simulations, we see that in both combinations, $\tilde{\phi}$ is again estimated to be the average of the dispersions. The mean values of the regression parameter estimates are close to the simulated values.

Table 23. Mean Values of Parameter Estimates.

Parameter	Variance Function			
	A. Correct		B. Misspecified	
	Combination of ϕ		Combination of ϕ	
	I	II	I	II
ϕ	1.998	2.785	5.157	8.676
β_1	-0.294	-0.292	-0.291	-0.300
β_2	0.695	0.697	0.692	0.700
β_3	1.597	1.594	1.595	1.575
τ_{11}	0.204	0.203	0.197	0.205
τ_{12}	0.204	0.198	0.197	0.211
τ_{13}	-0.198	-0.199	-0.195	-0.193
τ_{21}	0.205	0.198	0.205	0.203
τ_{22}	0.204	0.201	0.206	0.200
τ_{23}	-0.397	-0.397	-0.397	-0.407
τ_{31}	0.202	0.199	0.198	0.217
τ_{32}	0.305	0.304	0.302	0.321
τ_{33}	-0.396	-0.398	-0.395	-0.396

In the misspecified variance simulations, the estimates of $\tilde{\phi}$ are in the same neighborhood as the estimates in the three series simple simulations. This is expected because the average of the levels of dispersion over all seasons in all series in these simulations is very close to the average of the three levels of dispersion in the three series simple simulations. The mean values of the regression parameter estimates are fairly close to the simulated values. Any deviations from the simulated values that we see here are likely random error due to the substantial amount of variability in the data sets, especially in the simulations using combination II.

We now turn our attention to the standard deviations of the above estimates and the mean values of the estimated standard errors provided in Table 24. As in the three series simple case, in the correct variance simulations, the model-based standard errors for the parameters associated with the first series (β_1 and τ_{1j}) are too large. This is a result of using the estimate $\tilde{\phi}$ (estimated using data from all three series) to calculate the standard error for parameters corresponding to a series where the simulated dispersion is lower than $\tilde{\phi}$. Analogously, we see that the standard errors for β_3 and τ_{3j} are too small. For the regression parameters corresponding to series 2, we also find predictable results. Using combination I, the standard errors for β_2 and τ_{2j} are about right because the simulated ϕ for this series is equal to the average of the ϕ for all three series. However, using combination II, the simulated ϕ for series 2 is smaller than $\tilde{\phi}$ and we see that the model-based standard errors are larger than the standard deviations of the parameter estimates.

Table 24. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).

Parameter	Method	Variance Function			
		A. Correct		B. Misspecified	
		Combination of ϕ		Combination of ϕ	
		I	II	I	II
ϕ	Simulated	.084	.140	.407	.973
β_1	Simulated	.085	.105	.081	.103
	Model	.124 (.006)	.146 (.009)	.198 (.011)	.258 (.019)
	Empirical	.087 (.006)	.103 (.008)	.087 (.006)	.100 (.008)
β_2	Simulated	.076	.074	.092	.091
	Model	.075 (.003)	.089 (.004)	.121 (.007)	.156 (.011)
	Empirical	.075 (.005)	.075 (.005)	.092 (.007)	.092 (.007)
β_3	Simulated	.058	.074	.111	.157
	Model	.048 (.002)	.057 (.002)	.077 (.005)	.101 (.009)
	Empirical	.058 (.004)	.075 (.005)	.111 (.010)	.153 (.018)
τ_{11}	Simulated	.115	.141	.107	.136
	Model	.167 (.006)	.197 (.009)	.268 (.013)	.348 (.023)
	Empirical	.118 (.006)	.139 (.007)	.118 (.006)	.136 (.007)
τ_{12}	Simulated	.111	.141	.113	.140
	Model	.167 (.006)	.197 (.009)	.268 (.013)	.347 (.023)
	Empirical	.118 (.006)	.139 (.007)	.118 (.006)	.135 (.007)
τ_{13}	Simulated	.130	.154	.125	.150
	Model	.145 (.007)	.218 (.010)	.295 (.015)	.384 (.025)
	Empirical	.130 (.006)	.154 (.008)	.130 (.007)	.147 (.008)
τ_{21}	Simulated	.106	.105	.133	.124
	Model	.102 (.003)	.120 (.004)	.164 (.008)	.211 (.014)
	Empirical	.101 (.005)	.101 (.005)	.128 (.007)	.128 (.007)
τ_{22}	Simulated	.101	.102	.130	.125
	Model	.102 (.003)	.120 (.004)	.164 (.008)	.211 (.013)
	Empirical	.101 (.005)	.101 (.005)	.128 (.007)	.128 (.007)
τ_{23}	Simulated	.115	.114	.138	.138
	Model	.119 (.004)	.141 (.006)	.192 (.010)	.248 (.016)
	Empirical	.119 (.006)	.118 (.006)	.135 (.007)	.135 (.008)
τ_{31}	Simulated	.081	.099	.163	.216
	Model	.065 (.002)	.077 (.002)	.104 (.005)	.136 (.009)
	Empirical	.079 (.004)	.102 (.005)	.156 (.010)	.216 (.018)
τ_{32}	Simulated	.077	.098	.158	.212
	Model	.063 (.002)	.075 (.002)	.102 (.005)	.133 (.008)
	Empirical	.077 (.003)	.100 (.005)	.156 (.010)	.216 (.018)
τ_{33}	Simulated	.092	.119	.166	.225
	Model	.076 (.002)	.090 (.003)	.122 (.006)	.160 (.011)
	Empirical	.092 (.004)	.119 (.007)	.159 (.010)	.218 (.018)

The empirical estimates, on the other hand, are very good for all parameters (and at both levels of dispersion). There are only small differences between the empirical estimates and the observed standard deviations and there is no indication here that this estimator consistently underestimates the standard errors. Given the precision of the observed standard deviations and the average of the standard errors (discussed in the previous section), the observed differences can be explained by chance.

In the misspecified variance simulations, it is even more apparent that the model-based estimator of the standard errors performs poorly while the empirical estimates are quite good. One could not say that the empirical estimates are consistently too low in these simulations.

Table 25 summarizes the observed rejection probabilities of hypothesis tests for a common pollution covariate ($H: \gamma_0 = 0$) using the data generated under the reduced model.

Table 25. Observed Rejection Probabilities: Common to Null Model.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.135	.081	.016	.180	.108	.033	.206	.147	.056	.247	.153	.058
Wald/Emp	.101	.047	.006	.112	.055	.014	.116	.058	.021	.108	.056	.017
Score/Emp	.099	.045	.006	.112	.054	.010	.112	.057	.018	.102	.053	.016
Dev/Scaled	.135	.081	.016	.180	.108	.033	.206	.146	.056	.246	.153	.058

In the correct variance simulations we find that with both combinations of ϕ , the empirical test statistics do reasonably well, achieving levels close to the nominal rejection probabilities. The model-based tests reject too often in the simulations using the first combination of ϕ and even more so with the second combination. Notice again how very similar the results are for the

two model-based tests. As in Table 24, this overall pattern is even more clear in the misspecified variance simulations.

Thus in this three series complicated case we find that the parameter estimates are adequate as are the empirical estimates of the standard errors. However, as in the three series simple case, we find that the model-based estimator of standard errors performs poorly, presumably as a result of using a single estimate of ϕ when in fact different amounts of dispersion are simulated in each series. This poor performance of the model-based estimator appears to carry through to the model-based test statistics. This should be expected in the case of the model-based Wald test but it is interesting that it is also true of the scaled deviance test.

We now consider simulations using a model which includes the pollution covariate common to all three series; the simulated value is $\gamma_0=0.1$. The results summarized in Table 26 indicate that using either of the variance formulations, the overall effects and season effects are estimated just as well as they were in the null simulations. The mean values of the estimates of the pollution parameter are also close to the simulated value of 0.1.

Table 26. Mean Values of Parameter Estimates.

Parameter	Variance Function			
	A. Correct		B. Misspecified	
	Combination of ϕ		Combination of ϕ	
	I	II	I	II
ϕ	1.997	2.789	5.202	8.686
β_1	-0.293	-0.299	-0.294	-0.299
β_2	0.695	0.695	0.694	0.694
β_3	1.601	1.597	1.593	1.589
τ_{11}	0.203	0.212	0.205	0.202
τ_{12}	0.202	0.208	0.201	0.205
τ_{13}	-0.195	-0.193	-0.198	-0.192
τ_{21}	0.203	0.204	0.201	0.200
τ_{22}	0.200	0.199	0.201	0.200
τ_{23}	-0.397	-0.400	-0.402	-0.402
τ_{31}	0.196	0.198	0.200	0.192
τ_{32}	0.297	0.302	0.301	0.289
τ_{33}	-0.407	-0.402	-0.406	-0.400
γ_0	0.103	0.101	0.098	0.103

A summary of the standard deviations of the above parameter estimates as well as the mean values of the estimated standard errors of the regression parameters are presented in Table 27. For the correct variance simulations, as in Table 24, the model-based standard errors for β_1 and τ_{1j} are overestimated while those for β_3 and τ_{3j} are underestimated. Using combination I of ϕ , the model-based standard errors for β_2 and τ_{2j} are about right (as expected, see above). With combination II of ϕ , the model-based standard errors for β_2 are only slightly too large and somewhat surprisingly, only that for τ_{23} is too large (recall that in Table 24 τ_{21} , τ_{22} and τ_{23} were all overestimated on average). We also see that the model-based standard errors for γ_0 are underestimated. This is consistent with the larger than expected rejection probabilities in Table 25 for the model-based Wald test. The empirical estimates, on the other hand, are very good for all regression parameters.

Table 27. Standard Deviations of Estimates and Mean Values of Estimated Standard Errors (\pm s.d.).

Parameter	Method	Variance Function			
		A. Correct		B. Misspecified	
		Combination of ϕ		Combination of ϕ	
		I	II	I	II
ϕ	Simulated	.083	.140	.423	.987
β_1	Simulated	.100	.124	.119	.152
	Model	.133 (.006)	.158 (.009)	.215 (.012)	.279 (.021)
	Empirical	.100 (.006)	.119 (.007)	.121 (.006)	.148 (.011)
β_2	Simulated	.087	.097	.122	.149
	Model	.087 (.003)	.103 (.004)	.140 (.008)	.181 (.012)
	Empirical	.089 (.005)	.094 (.005)	.123 (.007)	.143 (.011)
β_3	Simulated	.076	.095	.036	.192
	Model	.063 (.002)	.074 (.002)	.102 (.005)	.132 (.008)
	Empirical	.075 (.005)	.095 (.006)	.138 (.012)	.188 (.022)
τ_{11}	Simulated	.156	.186	.200	.266
	Model	.185 (.006)	.218 (.008)	.297 (.014)	.385 (.024)
	Empirical	.147 (.005)	.176 (.007)	.201 (.011)	.257 (.020)
τ_{12}	Simulated	.150	.180	.196	.266
	Model	.185 (.006)	.217 (.008)	.298 (.014)	.385 (.024)
	Empirical	.147 (.005)	.176 (.007)	.201 (.011)	.257 (.020)
τ_{13}	Simulated	.130	.159	.130	.161
	Model	.190 (.007)	.226 (.010)	.307 (.016)	.397 (.027)
	Empirical	.134 (.007)	.159 (.009)	.134 (.007)	.150 (.008)
τ_{21}	Simulated	.133	.149	.206	.256
	Model	.128 (.003)	.151 (.004)	.207 (.009)	.267 (.015)
	Empirical	.134 (.005)	.148 (.006)	.207 (.012)	.253 (.021)
τ_{22}	Simulated	.132	.150	.209	.264
	Model	.129 (.003)	.152 (.004)	.208 (.009)	.268 (.015)
	Empirical	.134 (.005)	.149 (.006)	.208 (.012)	.254 (.021)
τ_{23}	Simulated	.119	.120	.136	.140
	Model	.123 (.004)	.145 (.006)	.199 (.010)	.257 (.017)
	Empirical	.122 (.006)	.122 (.006)	.137 (.007)	.137 (.007)
τ_{31}	Simulated	.120	.151	.219	.311
	Model	.101 (.002)	.120 (.003)	.163 (.006)	.212 (.011)
	Empirical	.118 (.007)	.148 (.010)	.224 (.020)	.303 (.038)
τ_{32}	Simulated	.121	.149	.225	.309
	Model	.101 (.002)	.120 (.003)	.163 (.006)	.212 (.011)
	Empirical	.118 (.007)	.148 (.010)	.225 (.021)	.306 (.040)
τ_{33}	Simulated	.094	.120	.162	.228
	Model	.078 (.002)	.093 (.003)	.127 (.006)	.164 (.011)
	Empirical	.096 (.005)	.123 (.007)	.160 (.010)	.217 (.017)
γ_0	Simulated	.070	.086	.127	.179
	Model	.061 (.001)	.072 (.002)	.098 (.004)	.127 (.006)
	Empirical	.068 (.003)	.084 (.004)	.127 (.010)	.171 (.018)

In the misspecified variance simulations, the model-based standard errors for β_1 and the τ_{1j} are overestimated while those for β_3 and the τ_{3j} are underestimated. The model-based standard errors for β_2 are overestimated with both combinations of ϕ while the standard errors for τ_{21} and τ_{22} appear quite good compared with the standard error of τ_{23} which is clearly overestimated. These results are similar to the correct variance results but are again different from the results of the simulations using the null model as summarized in Table 25. As for γ_0 , the standard error for this parameter appears to be underestimated on average, just as it was in the correct variance simulations.

The empirical estimates are close to the simulated values with combination I, but there is a hint the estimates are too low on average with combination II (although keeping the precisions of the estimates and standard deviations in mind, the differences seen here could be just chance). The empirical estimates seem to approximate the simulated standard deviations better here than they did in the single series complicated simulations (for data generated with the model including the pollution effect; see Table 21). The simplest explanation for the poorer results in the one series complicated case would be that in those simulations we considered much higher levels of dispersion relative to the overall mean value ($\beta = \log 0.75$) than in this three series case and these higher levels of dispersion contribute to underestimation of the standard errors.

The levels of the four test statistics in a model reduction from separate pollution effects for each series to a common effect using the data generated with a common pollution effect are summarized in Table 28. Under both variance formulations and for both combinations of ϕ , the empirical test statistics adhere reasonably well to the nominal levels, but the model-based tests do not reject as often as they should. For the model-based Wald test this would suggest

the standard errors for the separate pollution effects are overestimated (for at least one of the parameters). Again the model Wald and scaled deviance tests perform almost identically.

Table 28. Observed Rejection Probabilities: Separate to Common.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.074	.032	.008	.056	.027	.004	.054	.028	.010	.055	.027	.004
Wald/Emp	.096	.053	.012	.108	.050	.004	.117	.058	.016	.098	.057	.014
Score/Emp	.093	.050	.009	.106	.046	.004	.111	.058	.013	.097	.054	.011
Dev/Scaled	.074	.032	.008	.058	.028	.004	.053	.028	.010	.056	.028	.004

The powers of these tests to detect the common pollution effect represented by the simulated value of $\gamma_0=0.10$ are summarized in Table 29. In both variance formulations the model-based tests reject more often than the empirical tests (as expected given the underestimation of the standard errors of γ_0 and the consistent similarity of performance of the model Wald and scaled deviance tests). Predictably, the rejection probabilities decrease as ϕ increases.

Table 29. Observed Rejection Probabilities: Common to Null.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.522	.417	.224	.433	.334	.168	.335	.233	.102	.295	.219	.115
Wald/Emp	.451	.340	.147	.338	.233	.083	.205	.112	.043	.181	.118	.047
Score/Emp	.451	.340	.143	.338	.230	.082	.203	.108	.042	.171	.113	.038
Dev/Scaled	.522	.418	.225	.434	.334	.168	.336	.234	.103	.298	.219	.115

The final model we will consider for the more complicated simulations with three series will be the most general one which includes separate pollution effects for each series. The simulated

values were $\gamma_0=0.1$, $\gamma_1=0.1$, $\gamma_2=-0.2$ and $\gamma_3=0$. We begin as always with a table summarizing the mean values of the parameter estimates from 1000 simulated data sets generated with the above model.

Table 30. Mean Values of Parameter Estimates.

Parameter	Variance Function			
	A. Correct		B. Misspecified	
	Combination of ϕ		Combination of ϕ	
	I	II	I	II
ϕ	1.994	2.789	5.187	8.664
β_1	-0.289	-0.295	-0.299	-0.289
β_2	0.694	0.695	0.695	0.698
β_3	1.601	1.595	1.595	1.569
τ_{11}	0.198	0.203	0.204	0.188
τ_{12}	0.203	0.202	0.205	0.193
τ_{13}	-0.209	-0.201	-0.191	-0.202
τ_{21}	0.203	0.205	0.196	0.195
τ_{22}	0.203	0.203	0.203	0.192
τ_{23}	-0.393	-0.400	-0.392	-0.399
τ_{31}	0.200	0.200	0.193	0.228
τ_{32}	0.299	0.302	0.298	0.323
τ_{33}	-0.405	-0.401	-0.407	-0.402
γ_0	0.100	0.101	0.100	0.087
γ_1	0.099	0.103	0.106	0.124
γ_2	-0.200	-0.202	-0.195	-0.186

The estimates of β_i and τ_{ij} are very similar to those in Tables 23 and 26 for both correct and misspecified variance formulations. In the correct variance simulations, the estimates are very good for γ_0 , γ_1 and γ_2 , as they are in the misspecified variance simulations using combination I of ϕ . However the estimates of these pollution parameters differ noticeably from the true values in the simulations using combination II of ϕ , although the differences are not too great considering the amount of variability in the simulated data which is reflected in the standard deviations of these parameter estimates as shown in Table 31. Note that the predictor corresponding to

the overall pollution effect is not orthogonal to the predictors corresponding to the separate effects. The mean values of the parameter estimates in a reparameterized model with orthogonal predictors representing three separate pollution effects, as opposed to a main effect and two separate effects, would generally be closer to the simulated values (although the standard errors of those estimates would be smaller). In that case the interpretation of the table would remain the same as above; that is, the mean values of the estimates differ from the simulated values but are within acceptable limits considering the precisions of these mean values.

In the correct variance simulations the mean values of the model-based standard errors for β_i and τ_{ij} provided in Table 31 are very similar to those observed in Table 24 as one would expect. Recall that following the simulations with a common pollution effect (Table 27) there was some question about the model-based standard errors for τ_{21} and τ_{22} (they were noticeably overestimated for combination II in Table 24 but did not appear to be overestimated in Table 27). The results in the above table and Table 24 agree with the conclusion that standard errors are underestimated (overestimated) on average for parameters that correspond to data series with more (less) dispersion than the estimate $\tilde{\phi}$. We note that the estimated model-based standard error for the parameter γ_1 appears quite a bit larger than it should be. This would explain the large observed rejection probabilities for the model-based Wald tests (relative to the empirical) seen in Table 29. The empirical estimates of the standard errors for the pollution parameters look very good in these correct variance simulations.

Table 31. Standard Deviations of Estimates
and Mean Values of Estimated Standard Errors (\pm s.d.).

Parameter	Method	Variance Function			
		A. Correct		B. Misspecified	
		Combination of ϕ		Combination of ϕ	
		I	II	I	II
ϕ	Simulated	.084	.143	.416	1.05
β_1	Simulated	.127	.145	.122	.142
	Model	.178 (.006)	.211 (.010)	.287 (.015)	.370 (.026)
	Empirical	.125 (.008)	.148 (.010)	.126 (.008)	.143 (.010)
β_2	Simulated	.108	.108	.139	.133
	Model	.108 (.003)	.127 (.004)	.174 (.009)	.224 (.015)
	Empirical	.107 (.007)	.107 (.007)	.133 (.010)	.134 (.011)
β_3	Simulated	.084	.103	.165	.232
	Model	.069 (.002)	.082 (.003)	.112 (.005)	.145 (.009)
	Empirical	.084 (.005)	.108 (.007)	.164 (.016)	.224 (.026)
τ_{11}	Simulated	.213	.244	.207	.242
	Model	.297 (.009)	.352 (.013)	.480 (.022)	.620 (.040)
	Empirical	.210 (.013)	.247 (.017)	.210 (.012)	.242 (.016)
τ_{12}	Simulated	.212	.250	.207	.241
	Model	.298 (.009)	.353 (.013)	.481 (.022)	.621 (.040)
	Empirical	.210 (.012)	.248 (.017)	.210 (.012)	.242 (.016)
τ_{13}	Simulated	.134	.165	.139	.154
	Model	.196 (.008)	.233 (.013)	.317 (.016)	.409 (.029)
	Empirical	.138 (.007)	.164 (.010)	.139 (.007)	.154 (.009)
τ_{21}	Simulated	.185	.188	.236	.232
	Model	.185 (.005)	.218 (.007)	.298 (.014)	.385 (.024)
	Empirical	.184 (.012)	.183 (.013)	.230 (.019)	.231 (.020)
τ_{22}	Simulated	.189	.186	.238	.230
	Model	.186 (.005)	.220 (.007)	.300 (.013)	.387 (.025)
	Empirical	.184 (.012)	.184 (.013)	.231 (.019)	.232 (.020)
τ_{23}	Simulated	.109	.115	.140	.136
	Model	.115 (.004)	.136 (.005)	.186 (.010)	.240 (.017)
	Empirical	.115 (.006)	.114 (.006)	.134 (.007)	.134 (.007)
τ_{31}	Simulated	.144	.171	.284	.400
	Model	.116 (.003)	.137 (.004)	.188 (.007)	.243 (.012)
	Empirical	.141 (.009)	.182 (.013)	.316 (.032)	.387 (.052)
τ_{32}	Simulated	.142	.170	.289	.404
	Model	.116 (.003)	.138 (.004)	.188 (.007)	.244 (.012)
	Empirical	.141 (.009)	.182 (.013)	.287 (.031)	.392 (.051)
τ_{33}	Simulated	.095	.124	.159	.231
	Model	.078 (.002)	.093 (.003)	.127 (.006)	.165 (.012)
	Empirical	.095 (.005)	.123 (.007)	.160 (.010)	.217 (.017)
γ_0	Simulated	.092	.113	.181	.266
	Model	.075 (.002)	.089 (.002)	.122 (.004)	.157 (.008)
	Empirical	.092 (.005)	.119 (.007)	.187 (.016)	.256 (.029)
γ_1	Simulated	.165	.195	.220	.313
	Model	.207 (.005)	.245 (.008)	.334 (.014)	.431 (.026)
	Empirical	.164 (.006)	.199 (.008)	.231 (.014)	.301 (.025)
γ_2	Simulated	.156	.165	.235	.303
	Model	.143 (.003)	.169 (.004)	.230 (.009)	.298 (.017)
	Empirical	.152 (.006)	.169 (.007)	.240 (.014)	.297 (.026)

With the misspecified variance simulations we again find the model-based standard errors for β_i and τ_{ij} to be very similar to the corresponding standard errors in Table 24. As in the correct variance case, the model-based standard error for γ_1 is overestimated. The empirical estimates are also quite good even in the simulations using combination II of ϕ . There are some parameters for which the empirical estimator produces estimated standard errors which are on average too low but there is no clear indication of underestimation in this situation.

Finally, in Table 32, we present the observed rejection probabilities to evaluate the power of the four test statistics in the multiple parameter reduction from separate pollution effects to a common effect.

Table 32. Observed Rejection Probabilities: Separate to Common.

Test	Variance Function											
	A. Correct						B. Misspecified					
	Combination of ϕ						Combination of ϕ					
	I			II			I			II		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Wald/Mod	.358	.236	.102	.274	.168	.045	.120	.064	.012	.086	.047	.014
Wald/Emp	.441	.319	.144	.396	.290	.121	.350	.234	.076	.336	.226	.080
Score/Emp	.441	.315	.134	.390	.286	.115	.340	.224	.072	.325	.219	.071
Dev/Scaled	.359	.238	.102	.275	.168	.045	.120	.064	.012	.086	.047	.014

The model-based test statistics have lower power than the empirical test statistics in this table. Recall that the model-based estimator overestimated the standard errors for γ_1 (while the standard errors for γ_2 were about right). This overestimation would explain why the rejection probabilities for the model-based Wald test are lower than for the empirical tests and, as usual, the scaled deviance test gives rejection probabilities that are almost identical to those for the model-based Wald test.

The results of the simulations under the alternative hypotheses in this more complicated case with three series are qualitatively very similar to the results of the simple simulations with three series. We found that the estimator of ϕ performed adequately in estimating the average level of dispersion and that the regression parameter estimator did well for all parameters including the “pollution” parameters that were of most interest. We also found the empirical estimator of the variances to perform well in these simulations. The underestimation of standard errors apparent in the more complicated case with a single series was not so clearly apparent in these three series simulations, possibly because the levels of dispersion considered in this section were not as large relative to the overall mean levels as they were in the single series simulations where, in the simulations using the full model to generate data, only $\beta = \log 0.75$ was considered with values of ϕ as large as 2 or 3. The model-based estimator of variance performed very badly in these three series simulations because the models did not allow for the possibility that different series may have different amounts of dispersion. Thus in a sense, the variance was actually misspecified in all the simulations (even those labelled “correct variance”). The model-based test statistics considered also seemed to suffer from the misspecification of the variance function. When considering the levels of the model-based statistics, they rejected too often in some instances and less often than they should have in others. The empirical tests, even the empirical Wald test, achieved levels much closer to the nominal levels.

Chapter 5

Discussion

The analysis of relationships between air pollution and human health data from Prince George provided the motivation for carrying out a simulation study designed to evaluate the performance of possible variance estimators and test statistics available for making inference in over-dispersed Poisson models.

Moderate amounts of over-dispersion are reported to have little effect on the estimation of regression parameters (Cox, 1983), however the extra variability must be taken into account when estimating variances or testing hypotheses. In the Prince George study, over-dispersion was accounted for via a dispersion parameter ϕ (common to all of the data series) which was estimated by the deviance, G^2 , divided by its degrees of freedom. To adjust for this over-dispersion, standard errors of the regression parameters were multiplied by the square root of the estimate $\tilde{\phi}_G = G^2/\text{d.f.}$ and the usual likelihood ratio test (based on a Poisson likelihood), or deviance statistic, for testing the viability of a reduced model nested within a larger model, was divided by $\tilde{\phi}$. Proceeding with this methodology, models were fit to the emergency room visits data that included temporal, meteorological and the pollution parameters of primary interest. The final models, resulting from the model reduction procedures (for example, the final model for the pollutant TRS), included parameter estimates whose interpretation suggested that higher levels of pollution were related to *lower* numbers of emergency room visits for

respiratory illness. Such counter-intuitive results raised questions about the appropriateness of the methodology and motivated a simulation study. This study was designed to investigate the possibility that the amounts of over-dispersion encountered in the Prince George study were large enough (relative to the mean levels of the data series) to affect the estimation of regression parameters, and to examine the performance of other possible variance estimators and test statistics that could be employed in the analysis of such over-dispersed Poisson data.

An alternate estimator of the dispersion parameter involving the Pearson X^2 statistic, $\tilde{\phi}_X = X^2/\text{d.f.}$, was discussed and was found to be superior to the estimator $\tilde{\phi}_G$. In addition to the model-based estimator of the covariance matrix of the estimated regression parameters described in the methodology for the Prince George study, an empirical estimator, which does not rely on the correct specification of the variance function, was described. The two estimators of the covariance matrix give rise to two versions of the Wald test as alternatives to the scaled deviance test used in the Prince George study for determining if certain parameters (such as pollution effects) contribute to the fit of the model. An empirical score statistic, suggested in Breslow (1990), which uses the estimating equations themselves for inference in over-dispersed Poisson models, was also considered.

In the simple simulations described first in this thesis, the regression parameter estimates were all very close to the simulated values, even when data were simulated with large amounts of over-dispersion and with a misspecified variance function. The simple simulations with only one series of data did not show any clear differences between the estimators of the variances nor between the test statistics. With the three series simple simulations, the variance function used to fit the data assumed a common ϕ whereas a separate ϕ applies in the simulation of each

series. Thus the variance function was always misspecified, even for the simulations labelled “correct variance”. It appears that in this situation, the model-based estimator overestimates (underestimates) the standard errors for regression parameters corresponding to series with simulated dispersion smaller (larger) than the estimated dispersion (which is the average of the simulated values from the three series). We found that, when considering the levels of the test statistics, the observed rejection probabilities of the model-based test statistics often did not approximate the nominal rates well whereas the empirical test statistics always achieved levels close to nominal. We found an interesting agreement between the model Wald and scaled deviance tests, as well as between the empirical score and empirical Wald tests.

Using the more complicated model for a single series of data, estimates of ϕ and of the regression parameters were relatively unaffected by the presence of over-dispersion although the estimators of these parameters became more variable and may have been affected by collinearity between some of the predictors. Standard errors of the regression parameters were accurately estimated with either the model-based estimator or the empirical estimator when the data were generated using the correct variance formulation, but there was some evidence the model-based estimator performs poorly when the misspecified variance formulation was used to generate data. It was hypothesized that the problem with the model-based estimator was similar to the problem in the three series simple case. The observations in different seasons of the one series complicated case had different mean values and therefore, under the misspecified variance formulation, had different amounts of dispersion (a situation similar to four separate series generated with the correct variance function). The empirical estimator appeared to do better in the misspecified variance simulations but seemed to underestimate the standard errors when the

level of dispersion was large relative to the overall mean level. In these one series complicated simulations, the problems with the model-based standard errors did not translate into poor performance of the model-based test statistics; this is in contrast to the simple case simulations with three series. There were no detectable differences among any of the four test statistics under study in this one series complicated case.

The results of the simulations in the more complicated case with three series were very similar to the results of the simple simulations with three series. The estimator of ϕ performed adequately in estimating the average level of dispersion and the regression parameter estimators did well for all parameters including the “pollution” parameters that were of primary interest. The empirical estimator of the variances also performed well in these simulations. The under-estimation of standard errors apparent in the more complicated case with a single series was not so clearly apparent in the three series simulations, possibly because the levels of dispersion considered were not as large relative to the overall mean levels. The model-based estimator of variance performed very badly in the three series simulations because, again, the models used to fit the data did not allow for the possibility of differing amounts of over-dispersion in the different series. The model-based test statistics considered also seemed to suffer from the misspecification of the variance function. When considering the levels of the statistics, these tests did not achieve the nominal rates. The empirical tests, even the empirical Wald test, achieved levels much closer to the nominal levels.

Overall, it appears solutions of the score equations produce good estimates of regression parameters in a situation with levels of over-dispersion similar to those in the Prince George study. The empirical estimator of the covariance matrix of the regression parameters appears

preferable to the model-based estimator which relies on correct specification of the variance function. It appears that in situations where the variance function is misspecified, the two model-based test statistics which rely on a single estimated over-dispersion parameter may perform poorly. This is in sharp contrast to the empirical test statistics which always performed well in this study. An interesting side note to the investigation of the test statistics was a very close agreement between the two model-based tests throughout the simulation study.

Only brief speculations were made regarding the possibility that different, and perhaps more appropriate, methods would have led to different conclusions in the Prince George study. It appears that, on average, the estimators of the regression parameters would yield correct estimates of the effects of pollution on emergency room visits, but it is unclear whether empirical estimates of the standard errors or an empirical test statistic would have led to a different set of final parameters in the models.

Bibliography

- [1] Breslow, N. (1984). *Extra-Poisson Variation in Log-Linear Models*. Applied Statistics, 33, 38-44.
- [2] Breslow, N. (1990). *Tests of Hypotheses in Overdispersed Poisson Regression and Other Quasi-Likelihood Models*. Journal of the American Statistical Association, 85, 565-571.
- [3] Carroll, R. J. and Ruppert, D. (1988). *Transformation and Weighting in Regression*. New York: Chapman and Hall.
- [4] Cox, D. R. (1983). *Some Remarks on Over-dispersion*. Biometrika, 70, 269-274.
- [5] Inagaki, N. (1973). *Asymptotic Relations Between the Likelihood Estimating Function and the Maximum Likelihood Estimator*. Annals of the Institute of Statistical Mathematics, 265, 1-26.
- [6] Knight, K., Leroux, B., Millar, J., and Petkau, A.J. (1988). *Air Pollution and Human Health: A Study Based on Hospital Admissions Data from Prince George, British Columbia*. Report prepared under contract with the Health Protection Branch, Department of National Health and Welfare. Also issued as SIMS Technical Report 128, Department of Statistics, University of British Columbia (February 1989).
- [7] Knight, K., Leroux, B., Millar, J., and Petkau, A.J. (1989). *Air Pollution and Human Health: A Study Based on Emergency Room Visits Data from Prince George, British Columbia*. Report prepared under contract with the Health Protection Branch, Department of National Health and Welfare. Also issued as SIMS Technical Report 136, Department of Statistics, University of British Columbia (June 1989).
- [8] Liang, K. Y. and Zeger, S. L. (1986). *Longitudinal Data Analysis Using Generalized Linear Models*. Biometrika, 73, 13-22.
- [9] McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models, 2nd ed.* London: Chapman and Hall.
- [10] McNeney, B. and Petkau, A. J. (1991). *Air Pollution and Human Health: A Follow-up Study Based on Emergency Room Visits Data from Prince George, British Columbia*. Report prepared under contract with the Ministry of Health, Province of British Columbia.
- [11] Moore, D. F. (1986). *Asymptotic Properties of Moment Estimators for Overdispersed Counts and Proportions*. Biometrika, 73, 583-588.
- [12] Royall, R. M. (1986). *Model Robust Confidence Intervals*. International Statistical Review, 54, 199-214.
- [13] Wedderburn, R. W. (1974). *Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss-Newton Method*. Biometrika, 61, 439-447.
- [14] White, H. (1982). *Maximum Likelihood Estimation of Misspecified Models*. Econometrica, 50, 1-25.