ON IDENTIFICATION OF TRANSFER FUNCTION MODELS

by

Per-Olov Edlund

1987:2

University of Stockholm
Department of Statistics

ISBN 91-7474-096-2
ISSN 0280-7564

March 1987

ON IDENTIFICATION OF TRANSFER FUNCTION MODELS

by P-O Edlund

SUMMARY

Different ways to identify (preliminarily estimate) the impulse response function of the Box & Jenkins transfer function model are discussed. The discussion is based on the situation when there are several input variables that are correlated with each other. It is found that most of the methods proposed are unsuitable, some are not reliable when there are correlated input variables, and some are expensive or difficult to use. Therefore an extension of a regression approach used by Pukkila (1980) is proposed. The new approach is based on the solution of some problems connected with the application of the regression method in our particular situation, namely the multicollinearity problem and the problem of autocorrelated residuals. It is found that the use of biased regression estimators on variables transformed with respect to the noise model should give better estimates than the usual ordinary regression estimator. To test the new approach a simulation experiment has been designed and performed . The results from the simulations indicate that the proposed method may be of value to the practitioner. It gives estimates with smaller mean squared error and lower estimated standard error.

This is a revised version of a paper presented at the workshop on "Time Series Analysis in Management" at the European Institute for Advanced Studies in Management, Brussels, November 26-27, 1981, and at the second "International Symposium on Forecasting" held by the International Institute of Forecasters, Istanbul, 6-9 July, 1982. An earlier version of the paper titled "Identification of the Multi-input Box-Jenkins Transfer Function Model" has also been published in the Journal of Forecasting, p 297-308, vol 3, 1984. The main additions here are found in Sections 5.4 and 5.5.

KEY WORDS: Time series Transfer function model Identification procedure Biased regression Monte Carlo

CONTENTS Page

1. INTRODUCTION

In economic time series analysis we frequently work with models where it is necessary to include lagged values of the independent variables to make the model realistic. If we e.g. study the relationship between advertising and sales it is reasonable to assume that this month's sales will be effected not only by this month's advertising but also by the advertising expenditures during previous months. In econometrics this possible relationship between past and present advertising and present sales could be represented by means of e.g. Almon-lags, Koyck schemes or rational lag distributions (see e.g. Johnston, 1972, ch 10). (These schemes are called "distributed lags").

If we do not have a relevant theory that explicitly tells us the shape of the lag structure it has to be estimated by a trial-and-error method.

Even if an acceptable lag scheme is eventually found there will probably be autocorrelated residuals which will violate the basic assumptions of the regression model.

To avoid these difficulties the transfer function models described by Box and Jenkins (1976) could be used. Their approach gives us a technique for identification of a proper model and allows us to include a model for the residual structure.

Unfortunately their preferred identification procedure was developed for the case when there is only one independent variable (or when the independent variables are mutually uncorrelated). In economic time series analysis this is unlikely to be the case and for that reason it is interesting to try to find another method for identification which will work well when there are several independent variables which are intercorrelated.

The purpose of this study is to investigate an extension of a regression method, proposed by Pukkila (1980), for preliminary estimation of the

impulse response function of the transfer function model in the identification phase of building a transfer function model. This extended method will be useful when the input variables are correlated. (Note: In the following, unless otherwise stated, the term identification will be used as a synonym of preliminary estimation of the impulse response function.)

To investigate the possible benefits of the extended method a simulation experiment has been performed. The model used is from Pukkila (1980).

Outline of the paper:
In Chapter 2 the Box-Jenkins transfer function approach is presented and some methods for impulse response function identification are discussed. Chapter 3 deals with the regression method and the solution of the multicollinearity problem by biased regression, and how to adjust for autocorrelated residuals. A suggestion for an extension of Pukkila's regression method, which could be used when the input variables are correlated will be described in Chapter 4. To explore the possible benefits of the new method a simulation experiment has been performed. The outline of the experiment and the results will be reported in Chapter 5. Some concluding remarks will finally be given in Chapter 6.

2. TRANSFER FUNCTION MODELS

The use of univariate models for description and prediction of economic time series has been criticized by econometricians because the ARMA models could be used without understanding of the underlying economic system and that ARMA models could not be used to predict future values of a series when the system has been "shocked" (e.g. when an extreme value of an exogenous variable has occured).

To overcome this criticism, Box and Jenkins transfer function models (1976, Ch. 11) may be used instead. These models resemble ordinary regression models but have the advantage of an explicit noise model which allows the residuals to be autocorrelated.

A transfer function model with one input variable, x_t, may be split into two components following Jenkins (1979),

$$y_t = u_t + n_t$$

where y_t is the dependent variable (suitably differenced/transformed to be mean and variance stationary), u_t contains that part of y_t which can be explained exactly in terms of x_t (suitably differenced/transformed to be mean and variance stationary) and n_t is an error term which represents all "missing" x variables plus the pure noise.

The relationship between x_t and u_t can be expressed by a linear dynamic relationship of the kind,

$$u_t - \delta_1 u_{t-1} - \ldots - \delta_r u_{t-r} = \omega_0 x_{t-b} - \omega_1 x_{t-b-1} - \ldots - \omega_s x_{t-b-s}$$

i.e.,

$$u_t = \frac{\omega_0 - \omega_1 B - \ldots - \omega_s B^s}{1 - \delta_1 B - \ldots - \delta_r B^r} x_{t-b} = \frac{\omega(B)}{\delta(B)} x_{t-b} = \nu(B) x_t$$

where $\nu(B) = \frac{\omega(B)}{\delta(B)} B^b$, B is the ordinary lag operator, $\omega(B)$ is a "moving average" operator, $\delta(B)$ is an "autoregressive" operator and b is a pure delay parameter which represents the number of complete time intervals before a change in x_t begins to have an effect on y_t.

The transfer function $\nu(B)$ is a rational lag structure which may represent any linear dynamic relationship between x_t and y_t to any specified degree of accuracy. This formulation of the transfer function weights has also been used in econometrics, see e.g. Jorgenson (1966).

In general differently differenced input and output variables may be used.

n_t may be replaced by an ARMA(p,q) model of the form

$$n_t = c + \frac{\theta(B)}{\phi(B)} a_t$$

where $\phi(B)$ is an autoregressive operator, $\theta(B)$ a moving average operator and a_t a white noise series. If n_t is eliminated between the two expressions above a transfer function-noise model is obtained,

$$(2.1) \qquad y_t = c + \frac{\omega(B)}{\delta(B)} x_{t-b} + \frac{\theta(B)}{\phi(B)} a_t$$

When there are more than one input variable, say m variables x_{1t}, x_{2t}, ..., x_{mt}, the expression above is easily generalized to

$$(2.2) \qquad y_t = c + \sum_{j=1}^{m} \frac{\omega_j(B)}{\delta_j(B)} x_{j,t-b_j} + \frac{\theta(B)}{\phi(B)} a_t$$

It is also possible to allow the series to be seasonal.

It is easy to show that the ordinary regression model is a special case of the more general model (2.2) above.

2.1 Identification of transfer function models

Box and Jenkins (1976, p. 378) suggest the following identification procedure:

(1) Derive rough estimates $\hat{v}_j$ of the impulse response weights.

(2) Use the estimates $\hat{v}_j$ to make guesses of the orders s and r of the right-hand and left-hand operators, $\omega(B)$ and $\delta(B)$, and of the delay parameter b.

(3) Substitute the estimates $\hat{v}_j$ in the equations

$$
\begin{aligned}
v_j &= 0 && j < b \\
v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \ldots + \delta_r v_{j-r} + \omega_0 && j = b \\
v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \ldots + \delta_r v_{j-r} - \omega_{j-b} && j = b+1, b+2, \ldots, b+s \\
v_j &= \sigma_1 v_{j-1} + \delta_2 v_{j-2} + \ldots + \delta_r v_{j-r} && j > b+s
\end{aligned}
\qquad (2.3)
$$

with values of r, s and b obtained from (2) to obtain initial estimates of the parameters δ and ω in $\delta(B)$ and $\omega(B)$.

If the true $\hat{v}_j$ values were known, b, r and s may be guessed using the following facts. The response weights consist of

(1) b zero values $v_0, v_1, \ldots, v_{b-1}$

(2) a further s-r+1 values $v_b, v_{b+1}, \ldots, v_{b+s-r}$ following no fixed pattern (only if $s \geq r$)

(3) values v_j with $j \geq b+s-r+1$ which follow the pattern of an r:th order difference equation which has r starting values $v_{b+s}, \ldots, v_{b+s-r+1}$.

In Fig. 2.1 some common forms of $\delta(B)$ and $\omega(B)$ and the corresponding $v(B) = \omega(B)/\delta(B)$ are shown. Usually the orders of $\delta(B)$ and $\omega(B)$ are 0, 1

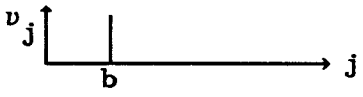
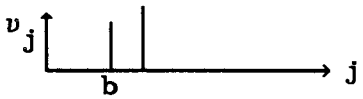
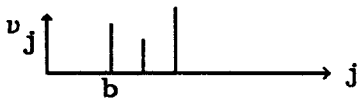
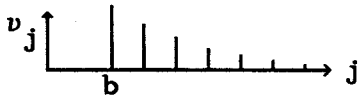
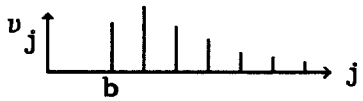
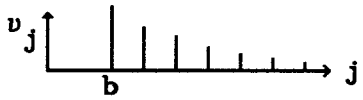
r,s,b	Transfer function	Impulse response, v_j	v_j j
0,0,b	$Y_t = \omega_0 B^b X_t$		0 $j < b$ ω_0 $j = b$ 0 $j > b$
0,1,b	$Y_t = (\omega_0 - \omega_1 B) B^b X_t$		0 $j < b$ ω_0 $j = b$ $-\omega_1$ $j = b+1$ 0 $j > b+1$
0,2,b	$Y_t = (\omega_0 - \omega_1 B - \omega_2 B^2) B^b X_t$		0 $j < b$ ω_0 $j = b$ $-\omega_1$ $j = b+1$ $-\omega_2$ $j = b+2$ 0 $j > b+2$
1,0,b	$(1-\delta_1 B) Y_t = \omega_0 B^b X_t$		0 $j < b$ ω_0 $j = b$ $\delta_1 v_{j-1}$ $j > b$
1,1,b	$(1-\delta_1 B) Y_t = (\omega_0 - \omega_1 B) B^b X_t$		0 $j < b$ ω_0 $j = b$ $\delta_1\omega_0 - \omega_1$ $j = b+1$ $\delta_1 v_{j-1}$ $j > b+1$
2,0,b	$(1-\delta_1 B - \delta_2 B^2) Y_t = \omega_0 B^b X_t$		0 $j < b$ ω_0 $j = b$ $\delta_1 v_{j-1} + \delta_2 v_{j-2}$ $j > b$
2,2,b	$(1-\delta_1 B - \delta_2 B^2) Y_t = (\omega_0 - \omega_1 B - \omega_2 B^2) B^b X_t$		0 $j < b$ ω_0 $j = b$ $\delta_1\omega_0 - \omega_1$ $j = b+1$ $(\delta_1^2 + \omega_2)\omega_0 - \delta_1\omega_1$ $-\omega_2$ $j = b+2$ $\delta_1 v_{j-1} + \delta_2 v_{j-2}$ $j > b+2$

Fig. 2.1. Examples of impulse response functions from transfer functions of order (r,s,b), $(1-\delta_1 B - \ldots - \delta_r B^r) Y_t = (\omega_0 - \omega_1 B - \ldots - \omega_s B^s) B^b X_t$.

or 2.

When the noise model is included in the transfer function model, a combined transfer function-noise model is obtained. The noise will of course "disturb" the empirical response function in Fig. 2.1. We therefore need a "good" method for estimating $v(B)$, a method that will give efficient estimates of $v(B)$. The method should also be easy to use and not too expensive in computer time. The aim of this study is to investigate some methods for estimating the impulse response function (transfer function) that have been proposed. Of special interest is the case when there are several correlated input (x) variables in the model.

The "true" identification problem of step 2 and 3 above can be solved in different ways. That part of the identification phase will not be discussed further in this study.

2.2 Methods for estimating the impulse response function

In their book, Box and Jenkins discuss three methods for estimating the transfer function weights. Two of these methods are time domain methods, the regression method (see Chapter 3) and the prewhitening-cross-correlation method (Section 2.2.1). The third method is a frequency domain method, the cross spectral analysis method (Section 2.2.3). They found that the regression method had several disadvantages and that the prewhitening-crosscorrelation method was to be preferred. However, they only discuss the case when there is only one input variable. There have also been some other methods proposed or used in practical applications.

Priestley (1971) proposed a method, the covariance contraction method (Section 2.2.2) that is similar to Box and Jenkins' prewhitening-crosscorrelation method. In practical applications the transfer function model could be identified by fitting a model that contains too many (or too few) parameters (Section 2.2.4). By trial-and-error the "right" model will eventually be found.

2.2.1 The prewhitening-crosscorrelation method

In order to study the relationship between x and y, the cross covariance or the cross correlation coefficients at different lags may be computed. The theoretical cross covariance coefficient between x and y at lag k is

$$\gamma_{xy}(k) = E((x_t-\mu_x)(y_{t+k}-\mu_y)) \qquad k = 0, \pm1, \pm2, \ldots \tag{2.4}$$

and between y and x at lag k

$$\gamma_{yx}(k) = E((y_t-\mu_y)(x_{t+k}-\mu_x)) \qquad k = 0, \pm1, \pm2, \ldots \tag{2.5}$$

In general, $\gamma_{xy}(k) \neq \gamma_{yx}(k)$. However, since $\gamma_{xy}(k) = \gamma_{yx}(-k)$ only one function $\gamma_{xy}(k)$ for $k = 0, \pm1, \pm2, \ldots$ is needed. This cross covariance function is not in general symmetric about $k = 0$. The function

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \cdot \sigma_y} \qquad k = 0, \pm1, \pm2, \ldots \tag{2.6}$$

is called the cross correlation function. In practice these functions are estimated from

$$c_{xy}(k) = \begin{cases} \frac{1}{n}\sum_{t=1}^{n-k}(x_t-\bar{x})(y_{t+k}-\bar{y}) & k = 0, 1, 2, \ldots, n-1 \\ \\ \frac{1}{n}\sum_{t=1}^{n+k}(y_t-\bar{y})(x_{t-k}-\bar{x}) & k = 0, -1, -2, \ldots, -n+1 \end{cases} \tag{2.7}$$

where $\bar{x}$ and $\bar{y}$ are the means of the x series and y series, respectively. The cross correlation function is then estimated by

$$r_{xy}(k) = \frac{c_{xy}(k)}{s_x s_y} \qquad k = 0, \pm1, \pm2, \ldots \tag{2.8}$$

where $s_x = c_{xx}(0)^{1/2}$ and $s_y = c_{yy}(0)^{1/2}$.

If x and y are ergodic processes, $r_{xy}(k)$ dies out fairly rapidly.

To obtain an estimate of υ_k we could compute the regression coefficient of y_t on x_{t-k}. This coefficient may be a poor estimate of υ_k, partly because of the autocorrelation in the x variable. To overcome this problem Box and Jenkins proposed the following method.

If the input follows a white noise process, the regression coefficient would be a fairly good estimate of υ_k. When the input follows some other process it could be transformed to white noise by a linear transformation. If the same transformation is applied to the output series both variables have been prewhitened.

It is assumed that the input process has been suitably differenced to be stationary. Then the differenced series can be represented by an ARMA(p,q) model

$$(2.9) \qquad \phi_x(B)(x_t-\mu_x) = \theta_x(B)\alpha_t$$

or $$\theta_x^{-1}(B)\phi_x(B)(x_t-\mu_x) = \alpha_t$$

The noise series α_t is then a close approximation to an uncorrelated white noise series. Applying the same transformation to the stationary series y_t gives

$$(2.10) \qquad \beta_t = \theta_x^{-1}(B)\phi_x(B)(y_t-\mu_y).$$

The transfer function-noise model may then be written as

$$(2.11) \qquad \beta_t = \upsilon(B)\alpha_t + \epsilon_t$$

where $\epsilon_t = \theta_x^{-1}(B)\phi_x(B)n_t$ is the transformed noise series. Since α_t is white noise and n_t is assumed to be independent of the input process, it is possible to obtain the coefficients υ_j from

$$(2.12) \qquad \nu_k = \frac{\gamma_{\alpha\beta}(k)}{\sigma_\alpha^2} \qquad k = 0, 1, 2, \ldots$$

where $\gamma_{\alpha\beta}(k)$ is the cross covariance at lag k between α and β. Alternatively (2.12) can be written as

$$(2.13) \qquad \nu_k = \frac{\rho_{\alpha\beta}(k)\sigma_\beta}{\sigma_\alpha} \qquad k = 0, 1, 2, \ldots$$

In practice ν_k is estimated by

$$(2.14) \qquad \hat{\nu}_k = \frac{s_\beta}{s_\alpha} \cdot r_{\alpha\beta}(k) \qquad k = 0, 1, 2, \ldots$$

When there are more than one input variable, the prewhitening technique can again be applied, if the input processes are not cross correlated, to give the estimates $\hat{\nu}_{kj}$, for k = 1, 2, ..., m and j = 0, 1, 2, ... If some or all input processes are cross correlated, the prewhitening technique is not directly applicable. For an example, see Damsleth (1979).

In the case of one input variable, the estimate $\hat{\nu}_k$ could be thought of as a regression coefficient of y_t on the variable x_{t-k}. If x_t is autocorrelated the x variables x_t, x_{t-1}, ..., x_{t-k}, ... will be correlated. This means that we will have multicollinearity between regressors. On the other hand, if the x variable is prewhitened as described above, the regression variables x_t, x_{t-1}, ..., x_{t-k}, ... will be made orthogonal to each other.

2.2.2 The covariance contraction method.

This method suggested by Priestley (1971) is an approach similar to that of the prewhitening method described above.

Priestley suggests that both x_t and y_t are prewhitened by fitting univariate ARMA models to each process. This leads to

$$(2.15) \quad \alpha_t = \theta_x^{-1}(B)\phi_x(B)(x_t-\mu_x)$$

and

$$(2.16) \quad \eta_t = \theta_y^{-1}(B)\phi_y(B)(y_t-\mu_y)$$

where α_t and η_t are white noise processes. Now a transfer function model can be fitted to the residuals

$$(2.17) \quad \eta_t + p_1\eta_{t-1} + \ldots + p_n\eta_{t-n} = q_0\alpha_t + q_1\alpha_{t-1} + \ldots + q_m\alpha_{t-m} + \epsilon_t$$

or $$P(B)\eta_t = Q(B)\alpha_t + \epsilon_t$$

The corresponding transfer function model for x_t and y_t is then given by

$$(2.18) \quad P(B)\theta_y^{-1}(B)\phi_y(B)(y_t-\mu_y) = Q(B)\theta_x^{-1}(B)\phi_x(B)(x_t-\mu_x) + \epsilon_t$$

or $$B(B)(y_t-\mu_t) = A(B)(x_t-\mu_x) + N_t'$$

where $$B(B) = P(B)\phi_y(B)\theta_x(B)$$
$$A(B) = Q(B)\phi_x(B)\theta_y(B)$$
$$N_t' = \theta_x(B)\theta_y(B)\epsilon_t$$

The main reason for this approach is that the structures of the operators A(B), B(B) depend on both the autocorrelation and cross correlation structure of x_t and y_t. When individual models are fitted to x_t and y_t the autocorrelation structures are removed and, therefore, it is reasonable to assume that the form of the operators, Q(B) and P(B), will be much simpler than the form of A(B) and B(B).

Then the fact that η_t and α_t are white noise processes is used when the cross covariance function is used to indicate the forms of P(B) and Q(B)

as follows. If

$$\xi_t = P(B)\eta_t \tag{2.19}$$

then

$$\xi_t = q_0\alpha_t + q_1\alpha_{t-1} + \ldots + q_m\alpha_{t-m} + \epsilon_t \tag{2.20}$$

The cross covariance function between ξ_t and α_t is given by

$$\rho_{\xi\alpha}(k) = E(\alpha_t\xi_{t+k}) = \begin{cases} q_k & k = 0, 1, \ldots, m \\ 0 & \text{otherwise} \end{cases} \tag{2.21}$$

This means that the cross covariance function at lag k is simply the coefficient q_k and when Q(B) contains a finite number of terms, $\rho_{\xi\alpha}(k)$ will be zero, except for lags k = 0, 1, 2, ..., m. On the other hand the cross covariance function between η_t and α_t, $\rho_{\eta\alpha}(k)$, will not in general vanish after a finite number of terms since the operator $(P^{-1}(B)Q(B))$, in general, will produce an infinite series in powers of B. Luckily, there is a simple relationship between $\rho_{\eta\alpha}(k)$ and $\rho_{\xi\alpha}(k)$,

$$\rho_{\xi\alpha}(k) = P(B)\rho_{\eta\alpha}(k) \tag{2.22}$$

(where the shift operator B acts on the variable k). Therefore, P(B) may be regarded as the operator which "contracts" the cross covariance function, $\rho_{\eta\alpha}(k)$, into the function $\rho_{\xi\alpha}(k)$.

In practice the estimated cross covariance function, $r_{\eta\alpha}(k)$, is used and a suitable form of P(B) can be found by seeking the filter which causes the function $r_{\eta\alpha}(k)$ to decay quickly to zero.

When the form of P(B) is found, the form of Q(B) may be determined by inspection of the contracted cross covariance function, $r_{\xi\alpha}(k)$. (Only lags for which $r_{\xi\alpha}(k)$ differs significantly from zero are of interest.)

If there is a pure delay between the two series, some of the first q_i are

zero.

When P(B) and Q(B) have been determined, the structure of B(B) and A(B) can be found by polynomial multiplication.

This method does not seem to be superior to the prewhitening-cross correlation method. In fact, the identification of P(B) seems to be a difficult task in practice (see Liu and Hanssens, 1982).

Haugh and Box (1977) used a similar approach. They estimated a_t and η_t as Priestley, but then they used the Box and Jenkins cross correlation method to obtain an estimate of the impulse response function. The identified model was then combined with the models for x_t and y_t to obtain the transfer function model.

Fask and Robinson (1977) generalized Priestley's approach to multivariate dynamic models.

2.2.3 The cross spectral analysis method

Box & Jenkins (1976, Appendix A11.1) also give an identification method that does not require prewhitening of the input. This method is based on spectral analysis. It could also be extended to multiple (cross correlated) inputs.

They redefine the transfer function $\upsilon(B)$ so that it could have non-zero impulse response weights υ_k for k a negative integer, so that

$$(2.23) \qquad \upsilon(B) = \sum_{k=-\infty}^{\infty} \upsilon_k B^k$$

Then if the transfer function-noise model is

$$(2.24) \qquad y_t = \upsilon(B)x_t + n_t$$

the theoretical autocovariance function between x_t and y_t is

$$(2.25) \quad \gamma_{xy}(k) = \sum_{j=-\infty}^{\infty} v_j \gamma_{xx}(k-j) \qquad k = 0, \pm 1, \pm 2, \ldots$$

Let

$$(2.26) \quad \gamma^{xy}(B) = \sum_{k=-\infty}^{\infty} \gamma_{xy}(k) B^k$$

denote the cross covariance generating function. Then, multiplying throughout in (2.25) by B^k and summing gives

$$(2.27) \quad \gamma^{xy}(B) = v(B)\gamma^{xx}(B)$$

where γ^{xx} is the autocovariance generating function. Substituting $B = e^{-i2\pi f}$ into (2.26) and (2.27) gives

$$(2.28) \quad v\left[e^{-i2\pi f}\right] = \frac{P_{xy}(f)}{P_{xx}(f)} \qquad -\frac{1}{2} \leq f < \frac{1}{2}$$

where

$$(2.29) \quad v\left[e^{-i2\pi f}\right] = G(f)e^{i2\pi\phi(f)} = \sum_{k=-\infty}^{\infty} v_k e^{-i2\pi fk}$$

$p_{xy}(f)$ is the cross spectrum between input and output. (2.29) is called the frequency response function of the system and is the Fourier transform of the impulse response function. Since $v\left[e^{-i2\pi f}\right]$ is complex it can be written as a product involving a gain function $G(f)$ and a phase function $\phi(f)$. If $v\left[e^{-i2\pi f}\right]$ was known, it would be possible to obtain the impulse response function v_k from

$$(2.30) \qquad \upsilon_k = \int_{-1/2}^{1/2} \upsilon\left[e^{-i2\pi f}\right] e^{i2\pi f} df$$

In practice, $\upsilon\left[e^{-i2\pi f}\right]$ has to be estimated. The integral in (2.30) can be replaced by a finite sum.

It is also possible to estimate the noise autocovariance function $\gamma_{nn}(k)$.

For multiple input transfer function models with m input variables it is possible to extend the method above. Let us assume, that after differencing the transfer function-noise model may be written as

$$(2.31) \qquad y_t = \upsilon_1(B)x_{1,t} + \ldots + \upsilon_m(B)x_{m,t} + n_t$$

Multiplying throughout by $x_{1,t-k}$, $x_{2,t-k}$, $\ldots$, $x_{m,t-k}$ in turn, taking expectations and forming the generating functions, the following system of equations is obtained:

$$(2.32) \qquad \begin{aligned} \gamma^{x_1 y}(B) &= \upsilon_1(B)\gamma^{x_1 x_1}(B) + \ldots + \upsilon_m(B)\gamma^{x_1 x_m}(B) \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ \gamma^{x_m y}(B) &= \upsilon_1(B)\gamma^{x_m x_1}(B) + \ldots + \upsilon_m(B)\gamma^{x_m x_m}(B) \end{aligned}$$

Substituting $B = e^{-i2\pi f}$ into (2.32) the spectral equations are obtained

$$(2.33) \qquad \begin{aligned} p_{x_1 y}(f) &= H_1(f)p_{x_1 x_1}(f) + \ldots + H_m(f)p_{x_1 x_m}(f) \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ p_{x_m y}(f) &= H_1(f)p_{x_m x_1}(f) + \ldots + H_m(f)p_{x_m x_m}(f) \end{aligned}$$

where $H_j(f) = \upsilon_j\left[e^{-i2\pi f}\right]$ can be estimated and substituted into (2.30) to give the υ_{jk}-weights. This method has been described in Pukkila (1979). He has also performed some simulations to investigate the properties of

this method. From his results it seems as if this method will work well even when the input processes are cross correlated. One disadvantage with this method is the computational effort needed. For the practitioner it could also be difficult to understand the method since we are partly working in the frequency domain.

2.2.4 The under-/overfitting method

If we have a good theory it may be possible to specify a tentative model. We may then estimate the model and test the coefficients and residuals to see if the model is adequate. If not, it can be modified according to the result of the diagnostic checking.

A slightly different approach would be to overparametrize the model and then delete the parameters that are non-significant. This approach may lead to a non-parsimonious model, there can be common factors in $\delta(B)$ and $\omega(B)$ which may not be detected.

If, on the other hand, the model is underparametrized, parameters are added as they are needed. Also in this case it is possible to stop before the "best" model is reached.

If the number of input variables is large, there may be many models to be estimated and we are not even sure we got the best one.

3. THE REGRESSION METHOD

Box and Jenkins (1976, p 379) also discuss a simple regression method for identification of the impulse response function without prewhitening. They write the model (2.1) without a constant as

$$(3.1) \qquad y_t = \upsilon_0 x_t + \upsilon_1 x_{t-1} + \upsilon_2 x_{t-2} + \ldots + n_t$$

where y_t, x_t and n_t are stationary processes with zero means. Then, multiplying throughout in (3.1) by x_{t-k} for k = 0, 1, 2, ... gives the following equations

$$(3.2) \qquad x_{t-k}y_t = \upsilon_0 x_{t-k}x_t + \upsilon_1 x_{t-k}x_{t-1} + \ldots + x_{t-k}n_t \qquad k = 0,1,2,\ldots$$

Taking expectations in (3.2), on the assumption that x_{t-k} is uncorrelated with n_t for all k the following set of equations is obtained

$$(3.3) \qquad \gamma_{xy}(k) = \upsilon_0\gamma_{xx}(k) + \upsilon_1\gamma_{xx}(k-1) + \ldots \qquad k = 0,1,2,\ldots$$

Assuming that $\upsilon_j \approx 0$ for $k > K$, then it is possible to write the first K+1 equations of (3.3) as

$$(3.4) \qquad \gamma_{xy} = \Gamma_{xx}\upsilon$$

where

$$\gamma_{xy} = \begin{bmatrix} \gamma_{xy}(0) \\ \gamma_{xy}(1) \\ \vdots \\ \gamma_{xy}(K) \end{bmatrix} \qquad \Gamma_{xx} = \begin{bmatrix} \gamma_{xx}(0) & \gamma_{xx}(1) & \ldots & \gamma_{xx}(K) \\ \gamma_{xx}(1) & \gamma_{xx}(0) & \ldots & \gamma_{xx}(K-1) \\ \vdots & \vdots & & \vdots \\ \gamma_{xx}(K) & \gamma_{xx}(K-1) & \ldots & \gamma_{xx}(0) \end{bmatrix} \qquad \upsilon = \begin{bmatrix} \upsilon_0 \\ \upsilon_1 \\ \vdots \\ \upsilon_K \end{bmatrix}$$

To estimate υ, $\gamma_{xx}(k)$ is replaced by $c_{xx}(k)$ and $\gamma_{xy}(k)$ by $c_{xy}(k)$.

Box and Jenkins point out that these equations,

a) do not in general provide efficient estimates

b) are cumbersome to solve

c) in any case require knowledge of the point K beyond which υ_j is effectively zero.

If there are more than one input variable equations (3.3) and (3.4) could easily be extended to include several input variables.

In addition to Box and Jenkins' remarks on this method it could be seen that if the input variables are autocorrelated and/or cross correlated then the covariance matrix Γ_{xx} will be multicollinear, i.e. if we view the lagged variables $x_{1,t}, \ldots, x_{1,t-K}, x_{2,t}, \ldots, x_{2,t-K}, \ldots, x_{m,t}, \ldots, x_{m,t-K}$ as different variables these "independent" variables will be correlated. Therefore the estimate $\hat{\upsilon}$ will have larger variance than if the inputs were white noise and not cross correlated. Even if Γ_{xx} is orthogonal there would still be a problem with autocorrelated residuals from n_t.

To avoid these problems, which may be of great importance when input variables are auto- and cross correlated, the x variables could be transformed with respect to the noise model and then some form of biased regression could be used to reduce the effects of multicollinearity. The problem of multicollinearity and its solution is discussed in the two following Sections. The problem of autocorrelated residuals is discussed in Section 3.3.

From knowledge of the underlying system it may be possible to specify which lags υ_k may be non-zero. This may decrease the order of the Γ_{xx} matrix.

Pukkila (1980) investigated this regression method and found that the estimates were surprisingly good for systems where the input variables were moderately cross correlated. Therefore, it is of interest to study how this method works when the input variables are more seriously cross correlated.

Liu and Hanssens (1982) solved the multicollinearity problem by transforming the y and x variables by a common filter. This filter was constructed to eliminate AR factors with roots close to one in the ARMA processes for the x variables. To avoid the effects of autocorrelated residuals the transfer function weights were estimated with generalized least squares (GLS) rather than ordinary least squares (OLS).

Erickson (1981) used the ridge estimator to estimate the transfer function weights in a "direct lag model" for the famous Lydia Pinkham Data. He used only one input variable and did not correct for autocorrelation in the resdiuals.

3.1 Multicollinearity

In general, there will be correlation between the variables $x_{1,t}$, $x_{1,t-1}$, ..., $x_{2,t}$, $x_{2,t-1}$, ..., $x_{m,t}$, $x_{m,t-1}$, ... The correlations will be of two different kinds,

a) autocorrelation between $x_{i,t}$ and $x_{i,t+k}$ $\quad i = 1, 2, \ldots, m$
b) cross correlation between $x_{i,t}$ and $x_{i+l,t+k}$ $\quad i = 1, 2, \ldots, m;\ l \neq 0.$

This means that the independent variables are not independent, i.e. there is a multicollinearity problem.

The main consequences of this are (Johnston, 172, p 160):

- The precision of estimation falls. This means that the specific estimates may have very large errors; these errors may be highly correlated, one with another, and finally, the sampling variances of the coefficients will be very large. (it is also possible that the estimated coefficients have the wrong signs.)
- We may be led to drop variables incorrectly because their coefficients are non-significant.
- Estimates of coefficients become very sensitive to particular sets of sample data, the addition of more data may produce large shifts in

some of the coefficients.

Several ways to detect and test for multicollinearity in the data matrix have been discussed in literature, see e.g. Webster et al (1974) and Haitovsky (1969).

Several "remedies" have also been proposed to decrease the bad effects of multicollinearity, see e.g. Intriligator (1978, Ch. 6), Silvey (1969) and Farrar and Glauber (1967). Some of them are not applicable when economic data are used, and in particular not possible to use when estimation of the impulse response function is of interest.

The author therefore proposes the use of biased regression to reduce the effects of multicollinearity.

The basic idea of biased regression is that if a small bias is introduced in the estimate it is possible to reduce the variance of the estimate considerably. Then the mean squared error, MSE, will be lower for the biased estimator than for the OLS (Ordinary Least-Squares) estimator. ($MSE = E\{(\hat{\upsilon}-\upsilon)^2\} = E\{((\hat{\upsilon}-E\hat{\upsilon})+(E\hat{\upsilon}-\upsilon))^2\}$ = Variance + Square of bias.)

Various biased estimators have been described in literature. Among them the principal component estimator and the ridge estimator seem to be of particular interest. Other biased regression estimators have been proposed by Marquardt (1970) (Generalized Inverse Estimator), James and Stein (1961), and Webster, Gunst and Mason (1974) (Latent Root Estimator).

3.2 Biased regression methods

In this Section two of the biased estimators that have been proposed and used to decrease the effects of multicollinearity will be briefly described.

3.2.1 The principal component estimator

If the idependent variables are correlated then it is possible to describe most of the variation in the dependent variable by a subset of the independent variables. In principal component regression linear combinations of the x variables are constructed, principal components, and then these new variables are used in the regression. The principal components are chosen so that they are pairwise uncorrelated and that the first component will have the maximum possible variance, the second the maximum possible variance among those uncorrelated with the first, and so forth. The first component can be written as,

$$(3.5) \qquad z_{1t} = a_{11}x_{1t} + a_{21}x_{1,t-1} + \ldots + a_{h1}x_{m,t-k} \qquad t = 1, \ldots, n$$

where $h = m\cdot(k+1)$ (h is the number of independent variables, including lags of the original variables). In matrix form,

$$(3.6) \qquad z_1 = Xa_1$$

Usually a_1 is normalized by setting $a_1'a_1 = 1$. This means that $z_1'z_1 = \lambda_1 a_1'a_1 = \lambda_1$ where λ_1 is the largest eigenvalue of the X'X matrix. Continuing in this way gives, in matrix form,

$$(3.7) \qquad \underset{(nxh)}{Z} = \underset{(nxh)}{X} \cdot \underset{(hxh)}{A}$$

where $A = (a_1, a_2, \ldots, a_h)$ is a matrix of eigenvectors and Z is a nxh matrix with h principal components. Transformation of the x variables by the A matrix gives the least squares solution of γ, the transformed coefficients, as

$$(3.8) \qquad \hat{\gamma} = (Z'Z)^{-1}Z'y \quad \text{or} \qquad \hat{\gamma} = \Lambda^{-1}Z'y$$

where Λ is a diagonal matrix of size hxh with $\lambda_1, \lambda_2, \ldots, \lambda_h$ on its diagonal. To obtain $\hat{\upsilon}$ the transformation is reversed and $\hat{\upsilon}$ is given as $\hat{\upsilon} = A\hat{\gamma}$. If X'X is orthogonal all λ_i are equal to one. Even though the data matrix Λ is orthogonal the estimates in $\hat{\upsilon}$ are as imprecise as before. To improve on the OLS estimates some of the principal components have to be deleted (meaning that the effective rank of X'X is reduced). This will introduce a bias, but if the data are highly collinear the reduction in the variance of the estimate $\hat{\upsilon}$ will be larger than the effect of the bias. There are several ways to determine the number of and which of the principal components to delete:

a) The $\hat{\gamma}$ values could be plotted for different numbers of deleted variables. This plot is called a Principal Component Trace by Vinod (1974). From this trace it may be possible to find the point where the $\hat{\gamma}$'s are stabilized and choose the corresponding number of principal components.

b) Marquardt (1970) discusses the principal component estimator and its generalization to non-integer ranks. He calls his estimator the Generalized Inverse estimator. He proposes a criterion for choosing an integer rank that will include "substantially all" of the variation in the x variables. The criterion is that the smallest value of r for which

$$(3.9) \qquad \frac{\sum_{j=h-r}^{h} \lambda_j}{\sum_{j=1}^{h} \lambda_j} < \omega$$

is chosen, where $\lambda_1 > \lambda_2 > \ldots > \lambda_h$. Typically ω is selected to be 10^{-5}, or in the interval 10^{-1} to 10^{-7}.

c) Massy (1965) gives two alternative criteria for deleting components,
(i) Delete the components with the smallest eigenvalues.

(ii) Delete the components that are relatively unimportant as predictors of y, i.e. the components with the smallest value of $\hat{\gamma}$ in equation (3.8) are deleted.

There is no reason why the two criteria should give the same result because y need not be highly correlated with components having large eigenvalues. Greenberg (1975) summarizes this, "including components with small eigenvectors increases variance, while including such components, if correlated with y reduces bias".

The principal component method may be very useful in our case since there are a large number of variables. If the rank of the X'X matrix is reduced it will be easier to solve the equation (3.8) for $\hat{\gamma}$.

3.2.2 The ridge estimator

Hoerl and Kennard (1970 a,b) introduced a biased regression method called Ridge Regression. Their estimator may be written as,

$$(3.10) \quad \hat{\upsilon}_R = (X'X+kI_h)^{-1}X'y$$

i.e. a small quantity $k > 0$ is added to the diagonal elements of X'X before inversion of the matrix. They showed that the sum of MSE's for individual parameters of the ridge estimator is always lower than the corresponding MSE for the OLS estimator for some $k < \sigma^2/\gamma^2_{max}$ (where γ_{max} is the largest γ from regression on principal components as described in the previous Section). The summed MSE for the OLS estimator is,

$$(3.11) \quad E(L_1^2(\hat{\upsilon})) = \sigma^2 \sum_{i=1}^{h} \frac{1}{\lambda_i}$$

and for the ridge estimator

$$(3.12)\quad E(L_1^2(\hat{\upsilon}_R)) = \sigma^2\Sigma \frac{\lambda_i}{(\lambda_i+k)^2} + k^2\Sigma \frac{\lambda_i^2}{(\lambda_i+k)^2}$$

where the first term on the right is the variance and the second is the squared bias. As $k \to \infty$ the variance $\to 0$ and the bias $\to \upsilon'\upsilon$. The effect of increasing k is to force the $\hat{\upsilon}_i$ towards zero. k "shrinks" the $\hat{\upsilon}$ vector.

As can be seen from equation (3.11) the effect of multicollinearity is that the MSE is greatly increased. For an orthogonal X'X $\lambda_1 = \ldots = \lambda_h = 1$ and $E(L_1^2(\hat{\upsilon})) = h\cdot\sigma^2$. When the multicollinearity is strong at least some $0 < \lambda_i << 1$. Then $E(L_1^2(\hat{\upsilon})) >> h\cdot\sigma^2$. It may then be assumed that at least some $\hat{\upsilon}_i$ are too large. This may also be seen from the expected sum of squared coefficients

$$(3.13)\quad E(\hat{\upsilon}'\hat{\upsilon}) = \upsilon'\upsilon + \sigma^2 tr(X'X)^{-1} = \upsilon'\upsilon + \sigma^2 \sum_{i=1}^{h} \frac{1}{\lambda_i}$$

which will be larger than $\upsilon'\upsilon$ on the average.

This estimator may then be used in our case. However, there is one problem that has to be solved, the value of k has to be determined. Since its optimal value depends on the unknown parameters σ and τ (or υ) k has to be estimated from our data. Unfortunately there is no guarantee that $L_1^2(\hat{\upsilon}_R) < L_1^2(\hat{\upsilon})$ for our estimated k. Therefore many techniques for estimating k have been proposed and a number of simulation studies have been performed to investigate the relative merits of ridge regression v. OLS regression and of the methods for choosing k. (See e.g. Dempster, Schatzoff and Wermuth (1977), Wichern and Churchill (1978), Hocking, Speed and Lynn (1976), Gunst and Mason (1977), Lawless and Wang (1976), Hoerl and Kennard (1976), McDonald and Galarneau (1975)). Here only some of the methods are given:

i) Hoerl, Kennard and Baldwin (1975)

$$(3.14) \quad k = \frac{h \cdot \hat{\sigma}^2}{\Sigma \hat{\gamma}_i^2} \qquad \hat{\gamma} \text{ is the OLS estimate of } \gamma$$

ii) Lawless and Wang (1976)

$$(3.15) \quad k = \frac{h \cdot \hat{\sigma}^2}{\Sigma \lambda_i \hat{\gamma}_i^2}$$

iii) Bulcock, Lee and Luk (1981)

Choose k to satisfy

$$(3.16) \quad \frac{1}{h} \Sigma \frac{\lambda_i}{(\lambda_i + k)^2} = 1$$

iv) Hocking, Speed and Lynn (1976)

$$(3.17) \quad k = \hat{\sigma}^2 \frac{\Sigma \lambda_i^2 \hat{\gamma}_i^2}{\Sigma \lambda_i^2 \hat{\gamma}_i^4}$$

v) Hoerl and Kennard (1976)

An iterative version of i) above

$$(3.18) \quad k_0 = \frac{h \cdot \hat{\sigma}^2}{\Sigma \hat{\gamma}_i^2}$$

$$k_t = \frac{h \cdot \hat{\sigma}^2}{\Sigma \hat{\gamma}_{it}^{*2}} \qquad \hat{\gamma}_{it}^{*} = \frac{\lambda_i}{(\lambda_i + k_{t-1})} \hat{\gamma}_i$$

$$\text{Stop when } \frac{k_{t+1} - k_t}{k_t} \leq \delta = 20(\text{tr}(X'X)^{-1}/h)^{-1.30}$$

vi) Hoerl and Kennard (1970 b)

The Ridge Trace: $\hat{\gamma}_R$ is computed for different values of k and plotted in a diagram with k on the x-axis and $\hat{\gamma}_R(k)$ on the y-axis. The optimal value of k is then determined by inspection. The value of k for which the ridge trace has stabilized is chosen.

In Leskinen (1980) the results from a simulation study are reported. He makes the following conclusions (p 78-9): The ridge estimator is more favourable to the OLS estimator when,

a) the number of explanatory variables increases
b) the degree of multicollinearity increases
c) the signal-to-noise ratio $\gamma'\gamma/\sigma^2$ decreases
d) the direction of the parameter vector changes from the eigenvector corresponding to the smallest eigenvalue of the X'X matrix to the eigenvector corresponding to the largest eigenvalue of the X'X matrix.

The ridge method described above may be called the Ordinary ridge estimator. The same value k is added to all diagonal elements of X'X. It is also possible as pointed out in Hoerl and Kennard (1970a p 63) to have a more general form of ridge regression by replacing $k \cdot I_h$ by K where K is a hxh diagonal matrix with diagonal elements equal to k_i, i.e. a k value is determined for each of the explanatory variables. It is then possible to adjust the bias for each variable. This estimator is called the Generalized ridge estimator. The optimal value of k_i is $k_i = \sigma^2/\gamma_i^2$. As above σ^2 and γ_i are unknown so k_i has to be estimated. Hemmerle (1975) and Goldstein and Smith (1974) have proposed a non-iterative and an iterative method for estimating the k_i respectively.

Since some of the $\hat{\upsilon}_i$ may be relatively unaffected by the multicollinearity it may be wise to add a value k_i only to those diagonal

elements corresponding to variables affected by the collinearity. Then the total bias will be reduced. This estimator, the Directed ridge estimator has been proposed by Guilkey and Murphy (1975).

3.3 Autocorrelated residuals

A basic assumption in regression analysis is that the residuals are uncorrelated. If this is not true the OLS estimates will not be efficient (minimum variance) and ordinary tests of significance can not be used. There will also be a bias in the estimation of the variance of the stochastic disturbance term. The OLS estimates are still defined, linear, unbiased and consistent. Since we want to test the significance of individual regression coefficients it is important that tests for autocorrelation are performed and if autocorrelation is found the estimates are corrected for this.

In regression there have been several methods proposed for dealing with autocorrelation. These methods, e.g. Durbin's method and Cochrane-Orcutt's method (see e.g. Johnston, 1972), assume that the residuals can be described by an AR model of low order (often an AR(1) model). The technique is to assume and estimate the noise model and then transform the x and y variables according to the noise model. The difficulty is to identify and estimate the noise model. The two methods mentioned above are of two different kinds. Durbin's method is a two-step procedure and Cochrane-Orcutt's method is an iterative method where the iteration is on the estimation of υ and the noise model.

Box and Tiao (1975) used the transformation technique in a case study where they wanted to estimate an intervention function model. In the case of transfer function-noise models the variables can be transformed (if the true model was known) as,

$$\phi(B)y_t = \theta(B)y'_t \qquad \text{all } t$$

(3.19)

$$\phi(B)x_{i,t-j} = \theta(B)x'_{i,t-j} \qquad \begin{matrix} i = 1, \ldots, m \\ j = 0, \ldots, k \end{matrix}$$

where y'_t and $x'_{i,t-j}$ are the transformed variables. Equation (2.2) can then be written as

$$(3.20) \qquad y'_t = c' + \sum_{j=1} \frac{\omega_j(B)}{\delta_j(B)} x'_{j,t-b_j} + a_t$$

This model may then be estimated by one of the biased regression methods of the previous Section.

4. A TWO-STEP PROCEDURE FOR THE IDENTIFICATION OF THE IMPULSE RESPONSE FUNCTION WHEN THE INPUT VARIABLES ARE CORRELATED.

In this Chapter a biased regression method is presented. The method can be used to identify the impulse response function when the input variables are correlated and when the noise model can be described by a (seasonal) ARMA model. As has been shown in the previous Chapter it is possible to deal with multicollinearity and autocorrelated resisuals by biased regression and transformation respectively. Therefore in the first step the noise model is estimated and the x and y variables are transformed, in the second step the impulse response function is estimated by a biased regression estimator as e.g. the principal component estimator or the ridge estimator.

4.1 Step one: Identification, estimation and checking of the noise model and transformation of the input and output variables

To identify the noise model, Box and Jenkins (1976, p 384) suggest that the noise series n_t is estimated by

$$(4.1) \qquad \hat{n}_t = y_t - \hat{\nu}(B)x_t = y_t - \hat{\nu}_0 x_t - \hat{\nu}_1 x_{t-1} - \hat{\nu}_2 x_{t-2} - \ldots$$

where the $\hat{\nu}(B)$ are preliminary estimates of the impulse response function. Formula (4.1) may easily be modified to include several input variables. From the noise series $\hat{n}_t$ an ARMA model (or a seasonal ARMA model) can be identified by the standard procedure of Box and Jenkins. To estimate the impulse response function it is suggested that one of the biased regression methods described in Section 3.2 is used. Those estimators do in general give better estimates of ν and better predictions of future values than the ordinary least-squares estimator. The identified model is then estimated by e.g. Marquardt's (1963) non-linear estimation method. The results are checked and if the model

is inadequate it has to be modified and re-estimated till it is found acceptable. The so obtained model

$$(4.2) \qquad \hat{n}_t = \frac{\hat{\theta}(B)}{\hat{\phi}(B)} \hat{a}_t$$

is then used to transform the original variables y_t, x_{1t}, ..., x_{mt}. This gives

$$(4.3) \qquad \begin{aligned} \hat{\theta}(B)y'_t &= \hat{\phi}(B)y_t \qquad \text{all } t \\ \hat{\theta}(B)x'_{jt} &= \hat{\phi}(B)x_{jt} \qquad j = 1, \ldots, m \quad \text{all } t \end{aligned}$$

4.2 Step two: Estimation of the impulse response function from the transformed variables y'_t and x'_{jt}

In the second step

$$(4.4) \qquad y'_t = c' + \nu_{10}x'_{1,t} + \nu_{11}x'_{1,t-1} + \ldots + \nu_{mk}x'_{m,t-k} + a_t$$

is estimated by a biased regression method. In (4.4) the residuals a_t will be (almost) white noise and by using biased regression the bad effects of multicollinearity are decreased. The estimates $\hat{\nu}_{ij}$ will then hopefully be good estimates of ν_{ij} and the significance of the coefficients can be tested by the standard t-test. From the estimated values $\hat{\nu}_{ij}$ the transfer function model can be identified as described in Section 2.1.

If it is believed that the estimated noise model is inadequate, i.e the estimated residuals, $\hat{a}_t$, in (4.4) are not white noise, step one can be repeated with the estimates $\hat{\nu}_{ij}$ from (4.4) substituted into (4.1). The noise model is then re-identified and re-estimated, the variables transformed and (4.4) re-estimated. This procedure can be repeated until the estimated residuals, $\hat{a}_t$, in (4.4) are white noise.

5. EVALUATION BY SIMULATION

Since it is not possible to evaluate the proposed identification method of Chapter 4 by analytical methods the author has chosen to perform a simulation study.

5.1 The model

Of course it is an impossible task to make a complete evaluation of the proposed procedure by the use of simulation techniques. The results may be dependent on the model chosen and its parameter values. Therefore the aim is not at a complete evaluation, but merely at showing the effects for one model. The model chosen is from Pukkila (1979) and has also been used in Damsleth (1979) and Pukkila (1980). Pukkila considers two similar models, here the two input processes model has been chosen

$$(5.1)\quad y_t = (2+4B+B^2)x_{1,t} + (1-.6B)^{-1}B^3 x_{2,t} + n_t$$

where
$$n_t = a_t + .75a_{t-1}$$
$$x_{1,t} = .7x_{1,t-1} + a_{1t}$$
$$x_{2,t} = 1.25x_{2,t-1} - .75x_{2,t-2} + a_{2t}$$

The processes a_{1t}, a_{2t} and a_t are normal white noise processes such that a_t is uncorrelated with a_{1t} and a_{2t}, and a_{1t} and a_{2t} are correlated with the covariance matrix Σ.

If we expand the model we obtain

$$(5.2)\quad y_t = 2x_{1,t} + 4x_{1,t-1} + x_{1,t-2} + x_{2,t-3} + .6x_{2,t-4} + .36x_{2,t-5} + \ldots + .0279936x_{2,t-10} + \ldots + n_t$$

We are interested in trying to estimate the coefficients in eq. (5.2) by the use of the regression methods presented in Chapter 3.

5.2 Criteria for comparing estimators

The most commonly used criterion is the MSE criterion which may be computed for each coefficient or for the whole vector of coefficients.

We have

$$(5.3)\quad MSE(\hat{\upsilon}) = E(\hat{\upsilon}-\upsilon)'(\hat{\upsilon}-\upsilon) = \sum_{i=1}^{h} E(\hat{\upsilon}_i-\upsilon_i)^2$$

where $\hat{\upsilon}$ is the estimated vector of coefficients. The mean squared error may be thought of as the sum of two components, bias^2 + variance. In this study main focus has been on the total $MSE(\hat{\upsilon})$ as in (5.3). The MSE has also been computed for each coefficient.

Another possible criterion is the generalized mean squared error

$$(5.4)\quad GMSE(\hat{\upsilon}) = E(\hat{\upsilon}-\upsilon)'G(\hat{\upsilon}-\upsilon)$$

where G is a symmetric positive semi-definite matrix of order hxh. If $G = I_h$ we obtain the $MSE(\hat{\upsilon})$ in (5.3). Usually G = X'X so that the $GMSE(\hat{\upsilon})$ is a measure of the predictive ability of the $\hat{\upsilon}_i$:s. Since the primary interest is not in prediction this criterion will not be used.

When the impulse response weights are estimated by regression methods not only the true values of the υ_i:s are of interest, but also their standard errors (for testing their significance). Therefore estimators which give low standard errors for the estimated coefficients are of interest. As a measure of the ability to identify the significant coefficients the standard error of the $\hat{\upsilon}_i$:s have been computed.

5.3 Selected estimators

In Chapter 3 several biased regression estimators were discussed. For this study two of the more well known estimators have been selected ; the principal component estimator and the ridge estimator.

For the ridge estimator the k-value has been computed by the method of Lawless & Wang (1976), i.e.

$$(5.5) \qquad k = \frac{h \cdot \hat{\sigma}^2}{\Sigma \lambda_i \hat{\tau}_i^2}$$

where $\hat{\tau}_i$ is the OLS estimate of τ_i, λ_i is the i:th eigenvalue and $\hat{\sigma}^2$ is the OLS estimate of the standardized σ^2.

The principal component estimator uses the first r principal components corresponding to the r largest eigenvalues. The value of r has been determined so that the r components contain at least 99.5 % of the variance of the 22 standardized x-variables. Another way of choosing the principal components is to choose the components which contributes significantly to the explanation of y. There may be some of the components with small eigenvalues that have significant $\hat{\tau}$:s. This would have been more time consuming in terms of CPU time.

As a standard of comparison the OLS estimates have also been computed .

For these three estimators, estimates are computed both for the original data and for the transformed data.

5.4 The experimental design

As was noted in Section 3.2.2 the ridge estimator is more favourable to the OLS estimator when

a) the number of explanatory variables increases,
b) the degree of multicollinearity increases,
c) the signal-to-noise ratio $\tau'\tau/\sigma^2$ decreases,
d) the direction of the parameter vector τ changes from the eigenvector corresponding to the smallest eigenvalue of the X'X matrix to the eigenvector corresponding to the largest eigenvalue of the X'X.

In this experiment the effects of changes in b) and c) have been studied by varying the degree of multicollinearity by using different values of the covariance between a_{1t} and a_{2t} and by varying the signal-to-noise ratio by using different values of σ^2.

The number of explanatory variables is 22, since lags up to 10 for each variable have been used. (As can be seen from eq (5.2) the value of $\upsilon_{2,10} \approx 0$, i.e. the cut-off is after lag 10.). More x variables could have been included by using more explanatory variables or by using a larger lag for each variable. Since the ridge estimator is known to be more superior to the OLS estimator when the number of variables increases, there is no need to include more variables at this preliminary stage.

It would be more interesting to experiment with different orientations of the parameter vector τ but that is left for a later study. (The number of possible combinations of levels on the multicollinearity, signal-to-noise ratio and parameter vector orientation would soon become very large.)

The degree of multicollinearity may be varied in two different ways. First, the intra correlation (autocorrelation) may be changed by changing the parameters of the two input processes. The coefficient of $x_{1,t-1}$ could e.g. be changed to, say, .9. Then the autocorrelation for the x_{1t} series would be increased, i.e. the correlation between $x_{1,t}$ and $x_{1,t-k}$ would increase etc..

The second way is to change the inter correlation (cross correlation) between the two variables. This is done by changing the off-diagonal

element of the covariance matrix Σ, i.e. the correlation between a_{1t} and a_{2t} is changed.

Since the prewhitening technique seems to be less efficient when the cross correlations are strong, it is interesting to see if the proposed method may be an improvement in such situations, and therefore only the latter way to change the multicollinearity in the X'X matrix has been used.

In the experiment the following Σ matrices were used

$$\Sigma = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} 3 & 3 \\ 3 & 4 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} 3 & 3.4 \\ 3.4 & 4 \end{bmatrix}$$

which gives the following approximate correlations between a_{1t} and a_{2t}: .58, .87 and .98 respectively. (The variance for a_1 and the covariances are not the same as in Pukkila (1979). He used only one Σ matrix with the correlation coefficient .612.)

The second factor considered in this experiment is the signal-to-noise ratio. The ratio has been changed by using different values of σ^2. σ^2 = .01, 1 and 25 were used.

Fifty replications were made for each combination of Σ and σ^2.

The practical work has been carried out as follows:

(The program was written in FORTRAN-77 and run on a PRIME 750 computer. All real variables have been declared in double precision (REAL*8). NAG-routines have been used to generate random normal deviates and to compute correlations, eigenvalues and eigenvectors.)

a) 150 random normal deviates were generated for each of the two x-variables and for 50 series of a_t.
b) The x_1-series was computed from the normal deviates. Only the last 100+k values have been used where k is the lag.

c) For each covariance matrix
   1) the x_2-series was computed
   2) the correlation matrix was computed
   3) the eigenvalues and eigenvectors were computed
   4) the number of principal components to be used was determined so that the components correspond to 99.5 % of the variation in the x-variables.

d) For each value of the residual variance and for each replication
   1) y_t was computed
   2) the coefficent vectors and the corresponding variances were estimated. In matrix form the following relations were used

OLS:

$$\hat{\gamma}_{OLS} = \Lambda^{-1}Z'y, \qquad \hat{\upsilon}_{OLS} = T\hat{A}\hat{\gamma}_{OLS}$$

where T is a diagonal matrix with the scale factors s_y/s_{x_j} on its diagonal

(5.6)

$$V(\hat{\upsilon}_{OLS}) = \hat{\sigma}^2_{OLS}TA\Lambda^{-1}A'T' \quad \text{where} \quad \hat{\sigma}^2_{OLS} = \frac{1-\hat{\gamma}'_{OLS}Z'y}{100-22}$$

Ridge:

$$\hat{\gamma}_{RR} = K\hat{\gamma}_{OLS}, \qquad \hat{\upsilon}_{RR} = T\hat{A}\hat{\gamma}_{RR}$$

where K is a diagonal matrix with λ_i/λ_i+k on its diagonal

(5.7)

$$V(\hat{\upsilon}_{RR}) = \hat{\sigma}^2_{RR}TAK\Lambda^{-1}K'A'T' \quad \text{where} \quad \hat{\sigma}^2_{RR} = \frac{1-\hat{\gamma}'_{RR}Z'y-k\hat{\gamma}'_{RR}\hat{\gamma}_{RR}}{100-22}$$

Principal component:

$$\hat{\gamma}_{PC} = \Lambda_r^{-1}A_r'X'y, \qquad \hat{\upsilon}_{PC} = TA_r\hat{\gamma}_{PC}$$

where Λ_r^{-1} and A_r contain the first r (largest) eigenvalues and eigenvectors respectively

(5.8)

$$V(\hat{\upsilon}_{PC}) = \hat{\sigma}^2_{PC}TA_r\Lambda_r^{-1}A_r'T' \quad \text{where} \quad \hat{\sigma}^2_{PC} = \frac{1-\hat{\gamma}'_{PC}Z_r'y}{100-22}$$

e) The x variables and the y variable were transformed with the true, rather than the estimated, noise model

$$y'_t + .75y'_{t-1} = y_t \quad \text{or} \tag{5.9}$$

$$y'_t = y_t - .75y'_{t-1}$$

and

$$x'_{it} + .75x'_{it} = x_i \quad \text{or} \tag{5.10}$$

$$x'_{it} = x_{it} - .75x'_{it-1} \quad i = 1, 2$$

This was done in order to simplify the computations. Of course, in practice the noise model has to be estimated as described in Chapter 4.

f) For each estimator and for each coefficient, j, the following statistics were computed

$$\text{MSE}(\hat{v})_j = \sum_{i=1}^{50} (\hat{v}_{ij}-v_j)^2/50 \qquad j = 1, \ldots, 22 \tag{5.11}$$

$$\text{BIAS}(\hat{v})_j = \sum_{i=1}^{50} (\hat{v}_{ij}-v_j)/50 \qquad j = 1, \ldots, 22 \tag{5.12}$$

$$S(\hat{v})_j = \sum_{i=1}^{50} s_{ij}/50 \qquad j = 1, \ldots, 22 \tag{5.13}$$

where s_{ij} is the estimated standard error of regression coefficient j in the i:th replication.

Then the following statistics were computed

$$\overline{\text{MSE}}(\hat{v}) = \sum_{j=1}^{22} \text{MSE}(\hat{v})_j/22 \tag{5.14}$$

$$(5.15) \qquad \overline{BIAS}(\hat{\upsilon}) = \sum_{j=1}^{22} BIAS(\hat{\upsilon})_j/22$$

$$(5.16) \qquad \overline{S}(\hat{\upsilon}) = \sum_{j=1}^{22} S(\hat{\upsilon})_j/22$$

$\overline{MSE}(\hat{\upsilon})$ measures the average MSE over all coefficients and replications, $\overline{BIAS}(\hat{\upsilon})$ measures the average bias over all coefficients and replications, and $\overline{S}(\hat{\upsilon})$ measures the average standard error of the estimate for all coefficients and over all replications.

In order to compare the suggested procedure and its estimators the ratios

$$(5.17) \qquad MSE_m = \frac{\overline{MSE}(\hat{\upsilon})_m}{\overline{MSE}(\hat{\upsilon})_{OLS}} \quad \text{and} \quad S_m = \frac{\overline{S}(\hat{\upsilon})_m}{\overline{S}(\hat{\upsilon})_{OLS}}$$

where m = RR, PC, OLS(t), RR(t) and PC(t) were computed (t = transformed).

By using the same time series for variable x_1 for all simulations and the same time series for variable x_2 for all replications with a given combination of values of Σ and σ^2 the variance of the estimators have been reduced. This means that the comparisons between different combinations of Σ and σ^2 will be less influenced by the fact that only 50 replications were made. By this technique we will also avoid the problem of having stochastic x variables.

5.5 Results of simulations

Before presenting the summarized results from the simulations the results for one replication where $\rho_{a_1a_2} = .87$ and $\sigma^2 = 1$ are presented.

For this replication the y variable was computed from the two x series (x_{1t} and x_{2t}) and from a residual series. The two x series remained the same for all replications with a given combination of $\rho_{a_1a_2}$ and σ^2. Different sets of residuals where used for each replication.

All series were transformed by the true residual model (see Section 5.4) to reduce the autocorrelation in the residual series and to make the estimated residual variance from OLS regression an estimate of the true variance of the residual series. The true models for the transformed variables, x'_{1t} and x'_{2t} are,

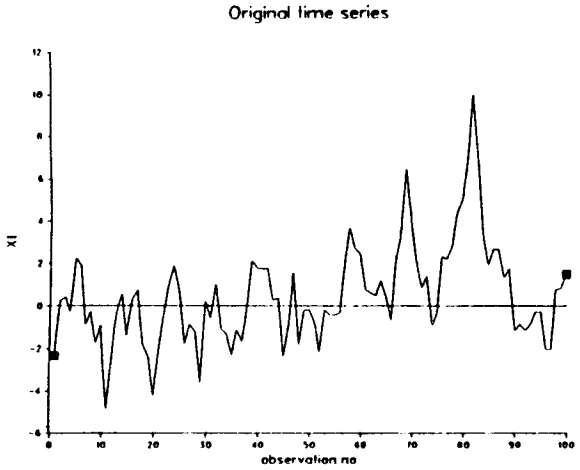
$$x'_{1t} = -0.05x'_{1t-1} + 0.525x'_{1t-2} + a'_{1t}$$

$$x'_{2t} = 0.5x'_{2t-1} + 0.1875x'_{2t-2} - 0.5625x'_{2t-3} + a'_{2t}$$

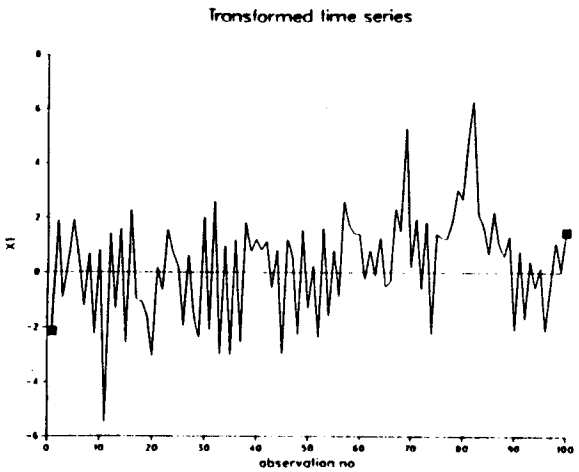
i.e. the orders of the two AR processes have increased from AR(1) to AR(2) and from AR(2) to AR(3) respectively.

The simulated series are shown in Fig. 5.1. Apart from the x_1 series the original and transformed series look very much the same. Please note that the scales are different, the transformed series have lower variances. The series are given in Appendix 1.

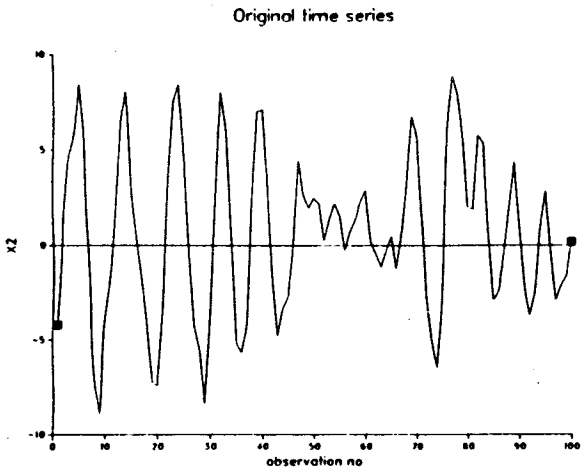
In Table 5.1 the autocorrelation and crosscorrelation functions are given for x_{1t}, x_{2t}, x'_{1t} and x'_{2t}. They correspond fairly well to their theoretical values. (The autocorrelation and cross correlation functions for other values of $\rho_{a_1a_2}$ are given in Appendix 2 (simulated) and Appendix 4 (theoretical).) When estimating the impulse response weights these correlation functions determine the correlation matrix (the X'X



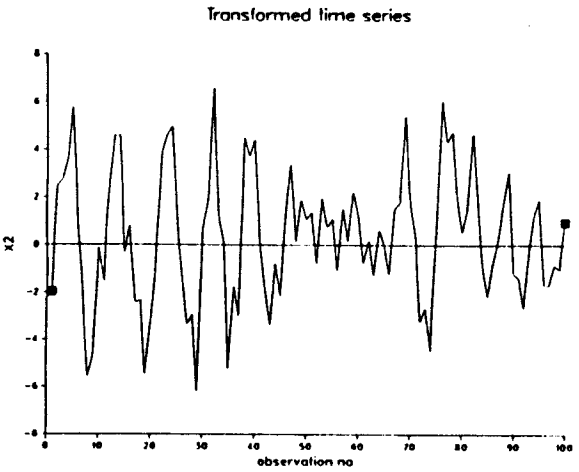
x_1 series, original



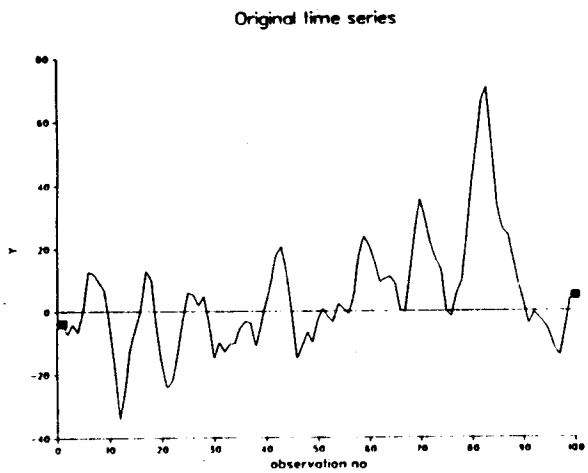
x_1 series, transformed



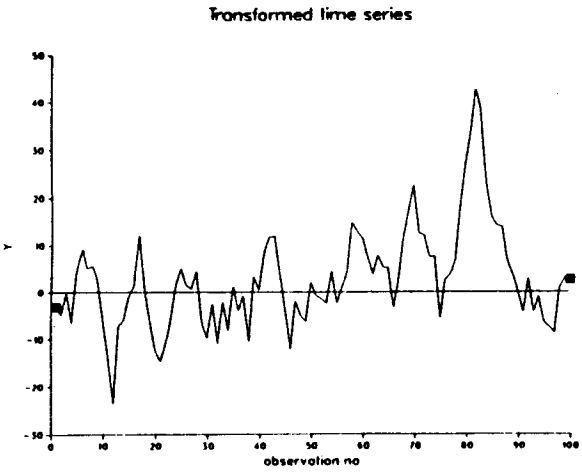
x_2 series, original



x_2 series, transformed



y series, original



y series, transformed

Fig 5.1. Simulated series, x_1, x_2 and y, original and transformed.

Table 5.1. Simulated (r) and theoretical (ρ) autocorrelation and cross correlation functions for original and transformed variables.

	Original variables				Transformed variables			
	x_{1t}		x_{2t}		x'_{1t}		x'_{2t}	
lag k	r(k)	ρ(k)	r(k)	ρ(k)	r(k)	ρ(k)	r(k)	ρ(k)
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.74	0.70	0.71	0.71	-0.03	-0.11	0.52	0.51
2	0.53	0.49	0.11	0.14	0.55	0.53	0.13	0.16
3	0.36	0.34	-0.44	-0.36	-0.04	-0.08	-0.45	-0.39
4	0.23	0.24	-0.68	-0.55	0.26	0.28	-0.56	-0.45
5	0.14	0.17	-0.56	-0.42	-0.05	-0.06	-0.52	-0.39
6	0.14	0.12	-0.20	-0.11	0.12	0.15	-0.12	-0.06
7	0.18	0.08	0.20	0.17	0.08	-0.04	0.16	0.15
8	0.22	0.06	0.44	0.30	0.07	0.08	0.41	0.28
9	0.25	0.04	0.46	0.25	0.20	-0.02	0.36	0.20
10	0.22	0.03	0.29	0.08	0.01	0.04	0.25	0.07

a) Simulated (r) and theoretical (ρ) autocorrelation functions.

lag k	Original variables		Transformed variables	
	$r_{x_1x_2}$(k)	$\rho_{x_1x_2}$(k)	$r_{x_1x_2}$(k)	$\rho_{x_1x_2}$(k)
-10	0.19	0.02	0.13	0.02
-9	0.14	0.02	0.08	-0.01
-8	0.05	0.03	0.04	0.05
-7	-0.04	0.05	-0.06	-0.01
-6	-0.08	0.07	-0.02	0.09
-5	-0.05	0.10	-0.09	-0.01
-4	0.05	0.14	0.09	0.16
-3	0.18	0.20	0.04	0.00
-2	0.29	0.28	0.29	0.31
-1	0.40	0.41	0.09	0.04
0	0.52	0.58	0.55	0.59
1	0.33	0.42	0.09	0.13
2	0.01	0.09	0.09	0.15
3	-0.27	-0.20	-0.28	-0.23
4	-0.36	-0.32	-0.19	-0.16
5	-0.27	-0.25	-0.24	-0.21
6	-0.01	-0.07	0.02	0.00
7	0.27	0.10	0.16	0.05
8	0.41	0.18	0.29	0.14
9	0.34	0.15	0.23	0.08
10	0.12	0.05	0.06	0.04

b) Simulated (r) and theoretical (ρ) cross correlation functions.

matrix).

This can be seen in Fig. 5.2 and 5.3 where the correlation matrices for the original and transformed variables are shown. Fig. 5.2 shows the correlation between the original 22 variables ($x_{1t}, \ldots, x_{1t-10}, x_{2t}, \ldots, x_{2t-10}$) and Fig. 5.3 the corresponding matrix for the transformed variables.

	x_1											x_2										
t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9	t-10	t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9	t-10	
1.00	.74	.53	.36	.23	.14	.14	.18	.22	.25	.22	.52	.40	.29	.18	.05	-.05	-.08	-.04	.05	.14	.19	$x_{1,t}$
	1.00	.74	.53	.36	.23	.15	.15	.20	.23	.26	.33	.53	.41	.29	.17	.04	-.06	-.07	-.03	.06	.14	$x_{1,t-1}$
		1.00	.73	.52	.35	.23	.16	.17	.21	.25	.01	.34	.53	.41	.27	.16	.03	-.05	-.06	-.02	.06	$x_{1,t-2}$
			1.00	.74	.53	.35	.23	.15	.16	.21	-.27	.00	.34	.54	.41	.27	.16	.03	-.05	-.06	-.02	$x_{1,t-3}$
				1.00	.74	.52	.34	.20	.11	.13	-.36	-.28	-.01	.34	.54	.41	.29	.16	.02	-.07	-.07	$x_{1,t-4}$
					1.00	.73	.51	.31	.17	.09	-.27	-.37	-.29	-.02	.34	.54	.43	.29	.14	.00	-.08	$x_{1,t-5}$
						1.00	.73	.50	.30	.17	-.01	-.26	-.37	-.30	-.02	.33	.54	.43	.29	.14	.00	$x_{1,t-6}$
							1.00	.73	.51	.32	.27	.01	-.25	-.38	-.31	-.04	.32	.54	.43	.29	.14	$x_{1,t-7}$
								1.00	.75	.53	.41	.29	.03	-.25	-.40	-.33	-.06	.32	.54	.43	.28	$x_{1,t-8}$
									1.00	.76	.34	.43	.30	.01	-.29	-.42	-.34	-.04	.32	.53	.40	$x_{1,t-9}$
										1.00	.12	.36	.44	.29	-.02	-.31	-.43	-.33	-.03	.32	.51	$x_{1,t-10}$
											1.00	.71	.11	-.44	-.68	-.56	-.20	.20	.44	.46	.29	$x_{2,t}$
												1.00	.71	.10	-.45	-.69	-.57	-.19	.21	.45	.45	$x_{2,t-1}$
													1.00	.70	.08	-.46	-.69	-.56	-.18	.21	.44	$x_{2,t-2}$
														1.00	.70	.09	-.45	-.69	-.57	-.19	.21	$x_{2,t-3}$
															1.00	.70	.10	-.45	-.71	-.59	-.20	$x_{2,t-4}$
																1.00	.71	.10	-.46	-.71	-.58	$x_{2,t-5}$
																	1.00	.70	.09	-.46	-.71	$x_{2,t-6}$
																		1.00	.71	.10	-.46	$x_{2,t-7}$
																			1.00	.70	.09	$x_{2,t-8}$
																				1.00	.70	$x_{2,t-9}$
																					1.00	$x_{2,t-10}$

Fig 5.2. The correlation matrix for the original variables when $\rho_{a_1 a_2} \approx .87$

	x_1											x_2										
t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9	t-10	t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9	t-10	
1.00	-.03	.55	-.04	.26	-.05	.12	.08	.07	.20	.01	.55	.09	.29	.04	.09	-.09	-.02	-.06	.04	.08	.13	$x_{1,t}$
	1.00	-.02	.55	-.04	.27	-.05	.12	.08	.09	.20	.09	.54	.09	.29	.04	.09	-.10	-.02	-.05	.05	.09	$x_{1,t-1}$
		1.00	-.02	.54	-.04	.27	-.04	.13	.10	.10	.09	.10	.55	.10	.28	.03	.08	-.09	-.01	-.05	.05	$x_{1,t-2}$
			1.00	-.02	.55	-.05	.27	-.04	.14	.11	-.28	.09	.11	.55	.09	.27	.02	.07	-.09	-.01	-.05	$x_{1,t-3}$
				1.00	-.02	.55	-.06	.27	-.06	.13	-.19	-.28	.08	.11	.56	.10	.29	.02	.07	-.10	-.01	$x_{1,t-4}$
					1.00	-.02	.54	-.06	.24	-.06	-.24	-.19	-.29	.07	.10	.55	.11	.29	.01	.05	-.11	$x_{1,t-5}$
						1.00	-.03	.54	-.08	.23	.02	-.25	-.19	-.30	.08	.11	.56	.11	.29	.01	.05	$x_{1,t-6}$
							1.00	-.02	.54	-.06	.16	.04	-.23	-.19	-.31	.06	.09	.55	.11	.29	.00	$x_{1,t-7}$
								1.00	.00	.55	.29	.17	.05	-.23	-.20	-.32	.05	.09	.56	.12	.29	$x_{1,t-8}$
									1.00	.03	.23	.32	.19	.05	-.26	-.23	-.33	.05	.10	.55	.11	$x_{1,t-9}$
										1.00	.06	.24	.33	.20	.04	-.27	-.24	-.33	.06	.12	.55	$x_{1,t-10}$
											1.00	.52	.13	-.45	-.56	-.52	-.12	.16	.41	.36	.25	$x_{2,t}$
												1.00	.53	.13	-.45	-.56	-.53	-.12	.17	.42	.36	$x_{2,t-1}$
													1.00	.52	.11	-.47	-.57	-.52	-.11	.17	.41	$x_{2,t-2}$
														1.00	.51	.10	-.47	-.57	-.52	-.11	.17	$x_{2,t-3}$
															1.00	.52	.12	-.46	-.57	-.54	-.12	$x_{2,t-4}$
																1.00	.53	.12	-.47	-.59	-.54	$x_{2,t-5}$
																	1.00	.53	.11	-.47	-.58	$x_{2,t-6}$
																		1.00	.53	.12	-.47	$x_{2,t-7}$
																			1.00	.53	.11	$x_{2,t-8}$
																				1.00	.52	$x_{2,t-9}$
																					1.00	$x_{2,t-10}$

Fig 5.3. The correlation matrix for the transformed variables when $\rho_{a_1 a_2} \approx .87$.

As can be seen in Fig. 5.2 the "upper" triangle contains the autocorrelations of variable x_{1t} and the "lower" triangle the autocorrelations of x_{2t}. The "square" shows the cross correlations between the two variables.

The effects of transforming x_{1t} and x_{2t} may be seen in Fig. 5.3 where some of the correlations drop significantly. The transformation may be viewed as a sort of differencing. The effect in this case is a reduction of the multicollinearity. For the two correlation matrices the eigenvalues and some other interesting measures of multicollinearity are shown in Table 5.2. (All correlation matrices and eigenvalues for the simulated series are given in Appendix 3. Their theoretical values are given in Appendix 5). From the results in Table 5.2 the following may be noted:

a) The sum of variance inflation factors are much larger than for an orthogonal matrix of the same order, about 29 and 8 times resp..
b) The spectral condition number is 1 if the correlation matrix is orthogonal. This measure also shows that the X'X matrices are ill-conditioned, especially the first one.
c) Haitovskys test variable is effectively zero for both matrices. This means that there are a $\approx$ 100 % multicollinearity in the matrices.
d) A relatively low number of principal components may represent most of the variation in the x variables. In order not to introduce too much bias in the principal components estimator the number of retained components were chosen so that they correspond to at least 99.5 % of the variation in the x variables.

The estimated regression coefficients for all estimators together with their theoretical values are shown in Table 5.3.

Table 5.2. Some measures of multicollinearity for the two correlation matrices of Fig 5.2 and 5.3.

Original variables				Transformed variables			
i	λ_i	%	Σ%	i	λ_i	%	Σ%
1	5.964	27.1	27.1	1	4.758	21.6	21.6
2	4.935	22.4	49.5	2	3.893	17.7	39.3
3	4.796	21.8	71.3	3	2.942	13.4	52.7
4	1.997	9.1	80.4	4	2.589	11.8	64.5
5	1.109	5.0	85.5	5	1.808	8.2	72.7
6	0.875	4.0	89.4	6	1.343	6.1	78.8
7	0.462	2.1	91.5	7	0.930	4.2	83.0
8	0.364	1.7	93.2	8	0.660	3.0	86.0
9	0.301	1.4	94.6	9	0.648	2.9	89.0
10	0.233	1.1	95.6	10	0.521	2.4	91.3
11	0.186	0.8	96.5	11	0.492	2.2	93.6
12	0.173	0.8	97.2	12	0.468	2.1	95.7
13	0.164	0.7	98.0	13	0.271	1.2	96.9
14	0.143	0.7	98.6	14	0.183	0.8	97.8
15	0.129	0.6	99.2	15	0.121	0.5	98.3
16	0.091	0.4	99.6	16	0.096	0.4	98.7
17	0.034	0.2	99.8	17	0.077	0.4	99.1
18	0.012	0.1	99.9	18	0.074	0.3	99.4
19	0.009	0.0	99.9	19	0.045	0.2	99.6
20	0.009	0.0	99.9	20	0.042	0.2	99.8
21	0.007	0.0	100.0	21	0.021	0.1	99.9
22	0.007	0.0	100.0	22	0.018	0.1	100.0
	22.000	100.0			22.000	100.0	

Sum of variance inflation factors:

$\Sigma\ 1/\lambda_i = 672.40$ $\Sigma\ 1/\lambda_i = 217.86$

The spectral condition number:

$\lambda_1/\lambda_{22} = 848.52$ $\lambda_1/\lambda_{22} = 267.82$

Determinant of the correlation matrix:

$|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 4.90423 \cdot 10^{-17}$ $|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 3.73064 \cdot 10^{-11}$

Haitovskys test variable:

$\chi^2_H(231) = 0.000$ $\chi^2_H(231) = 0.000$

Table 5.3. Estimates of impulse response weights for all estimators and the corresponding theoretical weights.

	Untransformed variables			Transformed variables			Theoretical
	OLS	Ridge	PC	OLS	Ridge	PC	weights
$\hat{v}_{10}$	2.15	2.16	2.35	2.07	2.07	2.27	2.00
$\hat{v}_{11}$	3.90	3.88	3.76	3.97	3.93	3.81	4.00
$\hat{v}_{12}$	1.03	1.01	0.79	1.03	1.01	0.78	1.00
$\hat{v}_{13}$	-0.09	-0.05	0.20	-0.12	-0.08	0.21	0.00
$\hat{v}_{14}$	-0.12	-0.12	-0.07	-0.09	-0.06	0.00	0.00
$\hat{v}_{15}$	0.11	0.10	-0.02	0.15	0.13	-0.16	0.00
$\hat{v}_{16}$	0.24	0.22	0.09	0.17	0.16	0.23	0.00
$\hat{v}_{17}$	-0.03	-0.02	0.20	0.00	0.02	0.19	0.00
$\hat{v}_{18}$	-0.22	-0.20	-0.01	-0.08	-0.08	-0.12	0.00
$\hat{v}_{19}$	0.03	0.01	-0.25	-0.08	-0.09	-0.16	0.00
$\hat{v}_{110}$	0.08	0.08	0.12	0.05	0.05	0.06	0.00
$\hat{v}_{20}$	-0.07	-0.08	-0.24	-0.07	-0.07	-0.23	0.00
$\hat{v}_{21}$	0.01	0.02	0.17	-0.01	0.02	0.19	0.00
$\hat{v}_{22}$	-0.07	-0.05	0.08	-0.06	-0.06	0.04	0.00
$\hat{v}_{23}$	1.02	0.97	0.61	1.04	1.01	0.67	1.00
$\hat{v}_{24}$	0.57	0.59	0.75	0.52	0.54	0.66	0.60
$\hat{v}_{25}$	0.32	0.32	0.39	0.30	0.32	0.53	0.36
$\hat{v}_{26}$	0.05	0.05	0.07	0.12	0.11	-0.10	0.22
$\hat{v}_{27}$	0.25	0.23	0.04	0.18	0.16	0.11	0.13
$\hat{v}_{28}$	0.15	0.14	0.07	0.04	0.04	0.14	0.08
$\hat{v}_{29}$	-0.22	-0.20	0.06	-0.05	-0.04	-0.05	0.05
$\hat{v}_{210}$	0.14	0.13	-0.06	0.03	0.03	-0.01	0.03
MSE	0.015	0.013	0.032	0.006	0.005	0.032	
Bias	-0.010	-0.012	-0.016	-0.016	-0.015	-0.018	

We can see that the ridge estimators are better than the corresponding OLS estimator in terms of MSE and that the principal components estimators are rather poor compared to the other estimators even though the bias is of the same order. Both the OLS and ridge estimators are much better on the transformed variables than the original variables. OLS on untransformed variables overestimated the true residual standard deviation (1.28 vs 1.00). On transformed variables the residual standard deviation was 1.01 which is very close to the true value. The principal component estimator used 16 components for the untransformed variables and 19 components for the transformed variables.

The estimated standard errors are shown in Table 5.4. We can again

Table 5.4. Estimates of standard errors for the estimated impulse response weights for all estimators.

$\hat{s}_{v_{ij}}$	Untransformed variables			Transformed variables		
	OLS	Ridge	PC	OLS	Ridge	PC
1,0	0.16	0.15	0.09	0.13	0.12	0.09
1	0.20	0.18	0.11	0.13	0.12	0.11
2	0.20	0.18	0.10	0.16	0.15	0.09
3	0.20	0.19	0.11	0.16	0.15	0.10
4	0.20	0.19	0.11	0.16	0.15	0.11
5	0.20	0.19	0.11	0.16	0.15	0.10
6	0.20	0.18	0.11	0.16	0.15	0.11
7	0.20	0.18	0.10	0.16	0.15	0.08
8	0.20	0.18	0.10	0.13	0.13	0.11
9	0.16	0.15	0.10	0.11	0.11	0.09
10	0.10	0.10	0.09	0.08	0.08	0.08
2,0	0.14	0.13	0.06	0.11	0.11	0.08
1	0.21	0.19	0.04	0.12	0.11	0.10
2	0.22	0.20	0.04	0.13	0.12	0.08
3	0.23	0.20	0.04	0.14	0.13	0.06
4	0.23	0.20	0.04	0.14	0.13	0.09
5	0.23	0.20	0.05	0.14	0.13	0.09
6	0.22	0.20	0.04	0.14	0.13	0.09
7	0.22	0.19	0.03	0.14	0.13	0.06
8	0.22	0.19	0.03	0.11	0.10	0.08
9	0.19	0.17	0.03	0.11	0.10	0.10
10	0.12	0.11	0.05	0.09	0.09	0.08

observe that the ridge estimator is at least as good as the OLS estimator. We can also notice that the principal components estimator now has much lower values that the other estimators.

Now the results for all simulations will be presented. The results have been summarized in tables to ease comparisons.

Table 5.5. Estimated ratios of $\overline{MSE}(v)_m$ to $\overline{MSE}(v)_{OLS}$.

MSE_m	ρ = .58			ρ = .87			ρ = .98		
Estimator	σ =.1	σ = 1	σ = 5	σ =.1	σ = 1	σ = 5	σ =.1	σ = 1	σ = 5
OLS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
RR	1.00	0.96	0.57	0.99	0.89	0.33	0.99	0.53	0.08
PC	25.21	1.21	0.82	81.04	1.49	0.29	152.48	1.71	0.09
OLS(t)	0.64	0.82	0.82	0.72	0.83	0.83	0.83	0.85	0.85
RR(t)	0.64	0.79	0.50	0.72	0.76	0.32	0.82	0.52	0.08
PC(t)	76.54	2.00	0.73	86.50	1.70	0.44	188.79	2.11	0.10
$MSE(\hat{v})_{OLS}$	0.0001	0.009	0.217	0.0003	0.021	0.530	0.001	0.133	3.320

In Table 5.5 the average MSE values are presented. It may be noted that:

a) The MSE for the OLS estimator increases roughly proportional to the residual variance for a given value of ρ.
b) The ridge estimator (RR) used is always better than (or equal to) the corresponding OLS estimator. The principal component estimator (PC) is better than the corresponding OLS estimator when the signal-to-noise ratio is low (σ_a = 5). The RR estimator is with one exception always better than the PC estimator. The largest reductions in MSE from using RR are obtained when the average MSE is large (signal-to-noise ratio is low) and when the correlation between a_{1t} and a_{2t} is high (strong multicollinearity).

c) The estimates on transformed variables are always better (or equal to) estimates on original variables for OLS and RR.

d) The PC estimator performes very poorly except when the signal-to-noise is low. This may be an effect of the criteria used for selecting principal components. In Table 5.6 the number of principal components used for different correlations between a_{1t} and a_{2t} are shown. In Table 5.7 the average number of deleted significant principal components are shown for different combinations of correlation and residual standard deviation.

Table 5.6. The number of principal components, r, used for different correlations between a_{1t} and a_{2t}.

Correlation	Data series Untransformed	Transformed
ρ = .58	19	21
ρ = .87	16	19
ρ = .98	13	13

Table 5.7. Average number of deleted significant principal components for different combinations of correlations between a_{1t} and a_{2t} and residual standard deviations.

Correlation between a_{1t} and a_{2t}	True standard deviation	Data series Untransformed no	%	Transformed no	%
ρ = .58	σ = .1	3.00	100	1.00	100
(3 and 1 pc	σ = 1	0.10	3	0.76	76
deleted)	σ = 5	0.02	1	0.08	8
ρ = .87	σ = .1	4.34	72	3.00	100
(6 and 3 pc	σ = 1	0.82	14	1.12	37
deleted)	σ = 5	0.20	3	0.20	7
ρ = .98	σ = .1	5.86	65	7.88	88
(9 pc dele-	σ = 1	2.34	26	2.56	28
ted)	σ = 5	0.86	10	0.76	8

From Table 5.7 it is quite clear that the effect of deleting components is most serious when the residual standard deviation is low. We may therefore expect that the PC estimator would do much better if the components that are significant are selected rather than the r components with largest eigenvalues (which is based only on the correlation matrix and not on the residual standard deviation).

Table 5.8. Estimated average bias for the estimators ($x10^{-4}$).

BIAS	ρ = .58			ρ = .87			ρ = .98		
Estimator	σ =.1	σ = 1	σ = 5	σ =.1	σ = 1	σ = 5	σ =.1	σ = 1	σ = 5
OLS	-3	8	55	4	17	75	12	38	154
RR	-3	4	-16	4	13	7	11	25	-4
PC	-5	6	53	-24	-9	57	-79	-78	-75
OLS(t)	-7	6	66	1	17	87	10	42	184
RR(t)	-7	-2	-108	1	12	-29	10	38	26
PC(t)	-31	-18	41	-19	-3	68	82	82	81

Table 5.8 shows the average bias for the estimators. As may be expected, on the average the absolute value of the bias increases as the multicollinearity increases and the signal-to-noise ratio decreases. Still, the largest average bias is not more than 0.0154 which is small compared to the values of most of the coefficients. For individual coefficients the bias may be much larger than the averages given above.

From Table 5.9 below it may be noted that:

a) The average standard error for the OLS estimator increases roughly proportional to the residual standard deviation for a given value of ρ.

b) When we look at the average standard error for the coefficients we find the same relationship between OLS and RR as before, but that the PC estimator nearly always is the best estimator in its class (original vs transformed data). It is interesting to notice that the reduction in $\bar{S}(\hat{\upsilon})$ is largest when the multicollinearity is strong and when the signal-to-noise ratio is small. As may be seen from the bottom row of the Table it is almost impossible to obtain any significant coefficient from the OLS estimator when ρ = .98 and σ = 5. The average s.e. is then about 2.3 as compared with the coefficients of which some are less than 1 and the largest 4. If we then use the RR estimator on the transformed variables the average s.e. decreases to about .3. The RR is of course a much better estimator in this situation since we are interested in identifying significant coefficients.

Table 5.9. Estimated ratios of $\bar{S}(\hat{\upsilon})_m$ to $\bar{S}(\hat{\upsilon})_{OLS}$.

S_m	ρ = .58			ρ = .87			ρ = .98		
Estimator	σ =.1	σ = 1	σ = 5	σ =.1	σ = 1	σ = 5	σ =.1	σ = 1	σ = 5
OLS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
RR	1.00	0.97	0.66	1.00	0.91	0.44	0.99	0.61	0.17
PC	1.29	0.77	0.76	0.96	0.38	0.37	0.90	0.15	0.12
OLS(t)	0.68	0.71	0.71	0.70	0.71	0.71	0.70	0.70	0.70
RR(t)	0.68	0.69	0.53	0.70	0.67	0.38	0.70	0.52	0.14
PC(t)	1.81	0.66	0.63	1.47	0.49	0.47	0.60	0.11	0.09
$\bar{S}(\hat{\upsilon})_{OLS}$	0.013	0.117	0.587	0.019	0.183	0.915	0.046	0.457	2.286

c) The estimated s.e. from the transformed data are in general lower than the corresponding estimate from the original data when the multicollinearity is large and/or the signal-to-noise ratio is low. This is to be expected since when we have autocorrelated residuals

the estimate of the variance of a_t will be biased (too large). The transformation procedure will then give a practically unbiased estimate of σ_a (and unbiased estimates of the s.e. of $\hat{v}_i$). This also gives us the possibility to use the standard t-test to test the significance of the estimated coefficients, see Table 5.10.

Table 5.10. The estimated residual standard deviation for OLS on untransformed and transformed data.

True standard deviation	Correlation between a_{1t} and a_{2t}	Data series Untransformed	Transformed
$\sigma = .1$	$\rho = .58$	0.136	0.108
	$\rho = .87$	0.129	0.103
	$\rho = .98$	0.123	0.100
$\sigma = 1$	$\rho = .58$	1.214	0.992
	$\rho = .87$	1.213	0.992
	$\rho = .98$	1.214	0.991
$\sigma = 5$	$\rho = .58$	6.062	4.957
	$\rho = .87$	6.063	4.956
	$\rho = .98$	6.071	4.958

In Table 5.11 the values of the shrinkage parameter k for the ridge estimator are given. The values are proportional to the residual variance of the simulated series as can be seen from eq. (5.5). Except when $\sigma = 5$ the value of k is very small.

Table 5.11 The average estimated value of k in ridge regression on untransformed and transformed data.

Correlation between a_{1t} and a_{2t}	True standard deviation	Data series Untransformed	Transformed
ρ = .58	σ = .1	0.00001	0.00003
	σ = 1	0.00117	0.00214
	σ = 5	0.02813	0.05055
ρ = .87	σ = .1	0.00001	0.00002
	σ = 1	0.00101	0.00186
	σ = 5	0.02449	0.04438
ρ = .98	σ = .1	0.00001	0.00002
	σ = 1	0.00093	0.00172
	σ = 5	0.02259	0.04108

If any recommendations are to be made from this study it would be to advice the practitioner to start by studying the eigenvalues of the correlation matrix. If these show clear signs of multicollinearity or if it is believed that the signal-to-noise ratio is low, then the ridge estimator should be used after the transformation of the original variables. If the signal-to-noise ratio is high it may still be wise to use the ridge estimator on the transformed variables, even though the gain in terms of MSE may not be so dramatic.

6. CONCLUDING REMARKS

In this paper different ways to estimate the impulse response function weights in the Box-Jenkins transfer function model have been discussed (Chapter 2). The discussion was based on the case when there are several input variables that are correlated with each other. It was found that most of the methods proposed are unsuitable, some are not reliable when there are correlated input variables, and some are expensive or difficult to use. Therefore an extension of a simple regression approach used by Pukkila (1980) was proposed. The new approach is based on the solution of some problems in connection with the application of the regression method in our particular situation, namely the multicollinearity problem and the problem of autocorrelated residuals. It was found that the use of biased regression estimators on variables transformed with respect to the noise model should give better estimates than the ordinary regression estimator (Chapter 4).

To test the new approach a simulation experiment was designed and performed. The results from the simulations indicate that the proposed method may be of value to the practitioner (Chapter 5).

It seems as if the greatest benefit of the proposed method is the possible reduction in the s.e. of $\hat{\nu}_i$. The use of biased regression decreases the risk of obtaining too large estimates of the coefficients. This reduces the risk of over-parametrization of the identified transfer function model. The method should be easy to apply since almost all computer systems have programs for biased regression.

It was noted that the two criteria used gave somewhat different results. This indicates that the results may be dependent on the criteria chosen. One other criterion that may be of interest is the proportion of times that the different estimators were better than the OLS estimator. In Bulcock et al (1981) some other criteria are discussed.

Since the results of the simulation study are very promising further studies will be made in the following areas:

a) Lawless and Wangs method of determining k was used. Several other methods have been proposed and some of them may be particulary suitable for this type of data. Further research is needed in this area. It is also interesting to investigate other critera for choosing r, the number of components in the PC estimator. This may have changed the conclusions regarding the benefits of the PC estimator, especially in terms of the MSE.
b) A comparative simulation study including some of the other methods discussed in Chapter 2. Then different models (different orientations of the parameter vector) would be used. Other relevant aspects to study are the effects of varying the number of lags included (including too few lags), the length of the time series, effects of seasonality, other distributions for the independent variables and the residual model.
c) Some real world applications to study the practical use of the proposed methods.

When drawing conclusions from this study it should be remembered that the conclusions are based on simulations of one model (under different conditions). Even though the results are in line with what may be expected one should be very careful to say anything about the gain in MSE from using the proposed method. It is likely that the proposed method gives smaller or equal MSE than ordinary least-squares but that the gain in MSE may vary between models.

When we evaluate the results of the simulation study we should keep in mind that in real world applications we face problems like how many, and which, x variables to be included (the missing variables problem) and measurement errors. It is also possible that the transfer function model is not providing good fit. It may therefore be better to start with a multivariate model to make sure of the direction of causality, before choosing the transfer function model.

Finally it is noted that even if we may identify a suitable transfer function it may still be difficult to obtain meaningful estimates if the

input processes are correlated. On the other hand, with the proposed method we may delete variables that are not to be included, and then reduce the effects of multicollinearity on the remaining variables. This may be the greatest advantage with the proposed method compared to OLS.

7 REFERENCES

BOX, G.E.P., JENKINS, G.M., 1976, Time Series Analysis: Forecasting and control, 2:nd ed., Holden-Day, San Francisco.

BOX, G.E.P., TIAO, G.C., 1975, Intervention Analysis with Applications to Economic and Environmental Problems, JASA, 70, p 70-9.

BULCOCK, J.W., LEE, W.F., LUK, W.S., 1981, Ordinary Least Squares, Ridge Regression and Normalization Ridge Regression: Some Comparisons, Paper presented at the Institute of Theology, University of Uppsala, May 1981.

DAMSLETH, E., 1979, Analysis of Multi-input Transfer Function Models when the Inputs are Correlated, Norwegian Computing Center, Oslo, publ. no 641.

DEMPSTER, A.P., SCHATZOFF, M., WERMUTH, N., 1977, A Simulation Study of Alternatives to Ordinary Least Squares, JASA, 72, p 77-91.

EDLUND, P.-O., 1984, Identification of the Multi-input Box-Jenkins Transfer Function Model, Journal of Forecasting, 3, p 297-308.

ERICKSON, G.M., 1981, Using Ridge Regression to Estimate Directly Lagged Effects in Marketing, JASA, 76, p 766-773.

FARRAR, D.E., GLAUBER, R.R., 1967, Multicollinearity in Regression Analysis: The problem revisited, Rev. of Economics and Statistics, 49, p 92-107.

FASK, A., ROBINSON, P.B., 1977, Identification of Multivariate Dynamic Econometric Models, American Statistical Association, 1977 Proceedings of the Business and Economic Statistics Section, p 284-289.

GOLDSTEIN, M, SMITH, A.F.M., 1974, Ridge-type Estimators for Regression Analysis, JRSS ser.B, 36, p 284-91.

GREENBERG, E., 1975, Minimum Variance Properties of Principal Component Regression, JASA, 70, p 194-7.

GUILKEY, D.K., MURPHY, J.L., 1975, Directed Ridge Regression Techniques in Cases of Multicollinearity, JASA, 70, p 769-75.

GUNST, R.F., MASON, R.L., 1977, Biased Estimation in Regression: An Evaluation Using Mean Squarred Error, JASA, 72, p 616-28.

HAITOVSKY, Y., 1969, Multicollinearity in Regression Analysis: Comment, Rev. of Economics and Statistics, 51, p 486-9.

HAUGH, L.D., BOX, G.E.P., 1977, Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series, JASA, 72, p 121-30.

HEMMERLE, W.J., 1975, An Explicit Solution for Generalized Ridge Regression, Technometrics, 17, p 309-14.

HOCKING, R.R., SPEED, F.M., LYNN, M.J., 1976, A Class of Biased Estimators in Linear Regression, Technometrics, 18, p 425-37.

HOERL, A.E., KENNARD, R.W., 1970a, Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics, 12, p 55-67.

HOERL, A.E., KENNARD, R.W., 1970b, Ridge Regression: Applications to Nonorthogonal Problems, Technometrics, 12, p 69-82.

HOERL, A.E., KENNARD, R.W., 1976, Ridge Regression: Iterative Estimation of the Biasing Parameter, Comm. in Statistics, A5, p 77-88.

HOERL, A.E., KENNARD, R.W., BALDWIN, K.F., 1975, Ridge Regression: Some Simulations, Comm. in Statistics, 4, p 105-23.

INTRILIGATOR, M.D., 1978, Econometric Models, Techniques and Applications, Prentice-Hall Inc, Englewood Cliffs, N.J.,

JAMES, W., STEIN, G., 1961, Estimation with Quadratic Loss, Proceedings of the Fourth Berkely Symposium on Mathematical Statistics and Probability, vol 1, ed. by Jerzy Neyman, Berkely and Los Angeles, University of California Press, p 361-79.

JENKINS, G.M., 1979, Practical Experiences with Modelling and Forecasting Time Series, G Jenkins & Partners Ltd, St Helier, Jersey.

JOHNSTON, J., 1972, Econometric Methods, 2:nd ed., McGraw-Hill Kogakusha Ltd, Tokyo.

JORGENSON, D.W., 1966, Rational Distributed Lag Functions, Econometrica, 32, p 135-48.

LAWLESS, J.F., WANG, P., 1976, A Simulation Study of Ridge and other Regression Estimators, Comm. in Statistics, A5, p 307-23.

LESKINEN, E., 1980, On the Biased Estimation in Regression: Studies on Ridge-type Estimation, Jyväskylän Yliopisto, Jyväskylä, Finland.

LIU, L.-M., HANSSENS, D., 1982, Identification of Multiple-input Transfer Function Models, Comm. in Statistics, B11, p 297-314.

MARQUARDT, D.W., 1963, An Algorithm for Least Squares Estimation of Nonlinear Parameters, Jour. Soc. Ind. Appl. Math., 11, p 431-41.

MARQUARDT, D.W., 1970, Generalized Inverses, Ridge Regression, Biased Linear and Nonlinear Estimation, Technometrics, 12, p 591-612.

MASSY, W.F., 1965, Principal Components Regression in Exploratory Statistical Research, JASA, 60, p 234-56.

McDONALD, G.C., GALARNEAU, D.I., 1975, A Monte Carlo Evaluation of Some Ridge-type Estimators, JASA, 70, p 407-16.

PRIESTLEY, M.B., 1971, Fitting Relationships between Time Series, Bull. Inst. Internat. Statist., 38, p 1-27, (paper presented at the 38:th session of the I.S.I., Washington, USA).

PUKKILA, T., 1979, On the Identification of Transfer Function Noise Models with Several Correlated Input Processes Using Frequency Domain Tools, Acta Universitatis Tamperensis, ser A, vol 110, Tampere, Finland.

PUKKILA, T., 1980, On Transfer Function Noise Model Identification, Dept. of Mathematical Sciences, Univ. of Tampere, Finland, (paper presented to the 8:th Nordic conference on Mathematical Statistics held on Ahvenanmaa 26-29.5 1980).

SILVEY, S.D., 1969, Multicollinearity and Imprecise Estimation, JRSS ser. B, 31, p 539-52.

VINOD, H.D., 1974, Ridge Estimation of a Trans-Log Production Function, 1974 Business and Economic Statistics Section Proceedings, ASA, p 596-601.

WEBSTER, J.T., GUNST, R.F., MASON, R.L., 1974, Latent Root Regression Analysis, Technometrics, 16, p 513-22.

WICHERN, D.W., CHURCHILL, G.A., 1978, A Comparison of Ridge Estimators, Technometrics, 20, p 301-11.

APPENDIX 1. Original and transformed time series used in the numerical example in Section 5.5.

t	y_t	x_{1t}	x_{2t}	t	y_t	x_{1t}	x_{2t}
1	-3.994	-2.368	-4.219	51	0.798	-0.743	2.193
2	-7.524	0.262	0.985	52	-1.996	-2.200	0.214
3	-4.072	0.447	4.604	53	-3.624	-0.181	1.353
4	-7.008	-0.299	5.643	54	2.462	-0.417	2.212
5	-0.364	2.245	8.451	55	0.773	-0.413	1.633
6	12.750	1.911	5.357	56	-0.984	-0.300	-0.280
7	11.913	-0.916	-1.559	57	4.788	1.925	0.724
8	9.321	-0.274	-7.359	58	17.881	3.665	1.277
9	6.828	-1.773	-8.954	59	23.849	2.713	2.315
10	-3.764	-0.902	-3.685	60	21.033	2.452	2.886
11	-18.346	-4.919	-1.668	61	16.014	0.770	0.133
12	-34.255	-2.717	1.403	62	9.166	0.607	-0.411
13	-25.222	-0.289	6.533	63	10.419	0.476	-1.150
14	-11.614	0.561	8.055	64	10.985	1.169	-0.345
15	-5.467	-1.429	3.093	65	8.966	0.424	0.461
16	0.553	0.320	0.511	66	0.392	-0.707	-1.282
17	13.010	0.750	-1.839	67	-0.422	2.119	0.555
18	9.921	-1.775	-4.145	68	13.236	3.242	2.925
19	-5.582	-2.350	-7.267	69	25.944	6.449	6.782
20	-17.292	-4.259	-7.459	70	35.488	4.145	5.801
21	-24.318	-2.130	-3.402	71	29.385	2.054	1.735
22	-22.114	-0.538	3.181	72	21.335	0.842	-2.927
23	-14.535	1.041	7.533	73	16.387	1.373	-5.145
24	-3.344	1.879	8.459	74	13.177	-0.930	-6.530
25	6.111	0.763	4.912	75	-0.026	-0.324	-3.058
26	5.309	-1.805	-0.314	76	-1.812	2.301	6.332
27	1.722	-0.860	-4.276	77	5.659	2.179	8.914
28	4.844	-1.198	-5.478	78	9.561	2.765	8.015
29	-3.628	-3.650	-8.443	79	23.406	4.423	5.584
30	-15.159	0.214	-3.959	80	40.201	4.948	2.028
31	-9.931	-0.618	2.435	81	52.154	6.903	1.863
32	-13.006	1.005	8.062	82	66.850	10.008	5.791
33	-10.441	-1.082	6.179	83	70.441	6.850	5.350
34	-10.060	-1.313	1.045	84	51.629	3.270	0.421
35	-5.178	-2.335	-5.179	85	33.036	1.881	-2.922
36	-3.138	-1.106	-5.717	86	26.055	2.663	-2.430
37	-3.821	-1.692	-4.336	87	24.341	2.628	-0.450
38	-11.325	-0.109	2.243	88	16.865	1.293	1.871
39	-4.796	2.109	7.080	89	8.510	1.730	4.379
40	2.793	1.768	7.169	90	2.947	-1.176	1.091
41	8.682	1.718	3.377	91	-4.248	-0.852	-2.293
42	18.001	1.748	-1.823	92	-0.445	-1.150	-3.731
43	20.673	0.275	-4.825	93	-2.227	-0.887	-2.531
44	13.190	0.379	-3.329	94	-4.161	-0.282	0.867
45	-0.610	-2.394	-2.755	95	-7.061	-0.274	2.827
46	-15.254	-1.034	-0.180	96	-12.270	-2.037	-0.279
47	-11.237	1.535	4.456	97	-14.402	-2.047	-2.954
48	-6.547	-1.843	2.633	98	-5.690	0.766	-2.109
49	-10.192	-0.181	1.975	99	3.689	0.855	-1.665
50	-2.728	-0.163	2.500	100	4.793	1.471	0.172

t	y'_t	x'_{1t}	x'_{2t}	t	y'_t	x'_{1t}	x'_{2t}
1	-3.234	-2.161	-1.994	51	-0.740	0.251	1.383
2	-5.099	1.883	2.480	52	-1.441	-2.388	-0.823
3	-0.248	-0.965	2.743	53	-2.543	1.611	1.971
4	-6.822	0.425	3.585	54	4.369	-1.625	0.734
5	4.752	1.926	5.763	55	-2.504	0.806	1.083
6	9.186	0.467	1.035	56	0.894	-0.904	-1.092
7	5.024	-1.267	-2.336	57	4.117	2.603	1.542
8	5.553	0.676	-5.607	58	14.793	1.713	0.120
9	2.663	-2.281	-4.749	59	12.755	1.428	2.225
10	-5.761	0.809	-0.123	60	11.467	1.381	1.217
11	-14.025	-5.525	-1.575	61	7.414	-0.266	-0.780
12	-23.736	1.427	2.585	62	3.605	0.807	0.175
13	-7.420	-1.359	4.594	63	7.715	-0.129	-1.281
14	-6.049	1.580	4.609	64	5.199	1.266	0.616
15	-0.930	-2.614	-0.364	65	5.066	-0.525	-0.001
16	1.251	2.281	0.783	66	-3.407	-0.313	-1.281
17	12.071	-0.961	-2.426	67	2.134	2.354	1.515
18	0.868	-1.055	-2.325	68	11.636	1.477	1.788
19	-6.233	-1.559	-5.523	69	17.217	5.342	5.441
20	-12.617	-3.090	-3.317	70	22.575	0.139	1.720
21	-14.855	0.187	-0.914	71	12.453	1.950	0.444
22	-10.973	-0.678	3.866	72	11.995	-0.620	-3.260
23	-6.305	1.550	4.634	73	7.391	1.838	-2.701
24	1.385	0.717	4.983	74	7.634	-2.309	-4.504
25	5.072	0.226	1.175	75	-5.751	1.407	0.320
26	1.505	-1.974	-1.195	76	2.502	1.246	6.092
27	0.594	0.620	-3.380	77	3.783	1.245	4.345
28	4.398	-1.663	-2.943	78	6.724	1.831	4.756
29	-6.927	-2.403	-6.235	79	18.363	3.050	2.017
30	-9.963	2.016	0.717	80	26.429	2.660	0.515
31	-2.458	-2.130	1.898	81	32.332	4.907	1.477
32	-11.162	2.602	6.639	82	42.601	6.327	4.683
33	-2.070	-3.033	1.200	83	38.491	2.105	1.838
34	-8.508	0.962	0.145	84	22.761	1.691	-0.958
35	1.203	-3.057	-5.288	85	15.965	0.613	-2.204
36	-4.041	1.186	-1.751	86	14.081	2.204	-0.777
37	-0.790	-2.581	-3.022	87	13.780	0.975	0.133
38	-10.733	1.827	4.510	88	6.530	0.562	1.771
39	3.254	0.739	3.698	89	3.612	1.308	3.051
40	0.353	1.214	4.395	90	0.238	-2.157	-1.198
41	8.417	0.808	0.080	91	-4.426	0.766	-1.395
42	11.688	1.142	-1.883	92	2.875	-1.725	-2.685
43	11.907	-0.581	-3.413	93	-4.383	0.407	-0.518
44	4.260	0.815	-0.769	94	-0.874	-0.587	1.255
45	-3.805	-3.005	-2.178	95	-6.406	0.166	1.886
46	-12.400	1.220	1.453	96	-7.465	-2.161	-1.693
47	-1.937	0.620	3.366	97	-8.803	-0.426	-1.684
48	-5.094	-2.309	0.109	98	0.913	1.085	-0.846
49	-6.371	1.550	1.893	99	3.005	0.041	-1.030
50	2.050	-1.326	1.080	100	2.539	1.440	0.945

APPENDIX 2. Autocorrelation and cross correlation functions for simulated processes.

	$\rho_{a_1a_2} \approx .58$				$\rho_{a_1a_2} \approx .98$			
	x_1	x_2	x_1'	x_2'	x_1	x_2	x_1'	x_2'
lag k	r(k)	r(k)	r(k)	r(k)	r(k)	r(k)	r(k)	r(k)
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.74	0.68	-0.03	0.43	0.74	0.73	-0.03	0.57
2	0.53	0.06	0.55	0.10	0.53	0.17	0.55	0.18
3	0.36	-0.45	-0.04	-0.47	0.36	-0.39	-0.04	-0.40
4	0.23	-0.63	0.26	-0.45	0.23	-0.68	0.26	-0.58
5	0.14	-0.45	-0.05	-0.41	0.14	-0.61	-0.05	-0.56
6	0.14	-0.10	0.12	0.00	0.14	-0.27	0.12	-0.21
7	0.18	0.19	0.08	0.12	0.18	0.16	0.08	0.15
8	0.22	0.31	0.07	0.30	0.22	0.47	0.07	0.42
9	0.25	0.27	0.20	0.15	0.25	0.53	0.20	0.47
10	0.22	0.16	0.01	0.16	0.22	0.37	0.01	0.31

	$r_{x_1x_2}$		$r_{x_1'x_2'}$		$r_{x_1x_2}$		$r_{x_1'x_2'}$	
lag k	r(k)	r(-k)	r(k)	r(-k)	r(k)	r(-k)	r(k)	r(-k)
0	0.27	0.27	0.32	0.32	0.63	0.63	0.64	0.64
1	0.16	0.17	0.01	-0.01	0.41	0.51	0.14	0.15
2	-0.05	0.10	0.03	0.11	0.05	0.38	0.12	0.38
3	-0.21	0.08	-0.21	0.01	-0.28	0.22	-0.29	0.05
4	-0.25	0.03	-0.10	0.05	-0.40	0.06	-0.23	0.10
5	-0.17	-0.01	-0.18	-0.05	-0.30	-0.06	-0.26	-0.11
6	0.03	-0.04	0.07	0.01	-0.04	-0.09	-0.01	-0.03
7	0.25	-0.04	0.12	-0.06	0.25	-0.04	0.17	-0.05
8	0.34	0.00	0.28	0.03	0.40	0.07	0.26	0.04
9	0.23	0.05	0.12	0.00	0.36	0.18	0.28	0.13
10	0.02	0.09	0.04	0.09	0.17	0.23	0.07	0.13

APPENDIX 3. Correlation matrices and eigenvalues for simulated processes.

1.00	0.74	0.53	0.36	0.23	0.14	0.14	0.18	0.22	0.25	0.22	0.27	0.17	0.10	0.08	0.03	-0.01	-0.04	-0.04	0.00	0.05	0.09
0.74	1.00	0.74	0.53	0.36	0.23	0.15	0.15	0.20	0.23	0.26	0.16	0.28	0.18	0.10	0.08	0.03	-0.02	-0.04	-0.03	0.01	0.05
0.53	0.74	1.00	0.73	0.52	0.35	0.23	0.16	0.17	0.21	0.25	-0.05	0.16	0.29	0.18	0.10	0.07	0.02	-0.02	-0.04	-0.03	0.00
0.36	0.53	0.73	1.00	0.74	0.53	0.35	0.23	0.15	0.16	0.21	-0.21	-0.05	0.16	0.29	0.19	0.10	0.07	0.02	-0.02	-0.03	-0.03
0.23	0.36	0.52	0.74	1.00	0.74	0.52	0.34	0.20	0.11	0.13	-0.25	-0.22	-0.08	0.15	0.29	0.19	0.12	0.08	0.01	-0.03	-0.04
0.14	0.23	0.35	0.53	0.74	1.00	0.73	0.51	0.31	0.17	0.09	-0.17	-0.26	-0.25	-0.09	0.15	0.29	0.20	0.12	0.07	0.00	-0.04
0.14	0.15	0.23	0.35	0.52	0.73	1.00	0.73	0.50	0.30	0.17	0.03	-0.17	-0.26	-0.25	-0.10	0.14	0.29	0.20	0.12	0.06	-0.01
0.18	0.15	0.16	0.23	0.34	0.51	0.73	1.00	0.73	0.51	0.32	0.25	0.04	-0.16	-0.27	-0.26	-0.11	0.13	0.29	0.20	0.11	0.05
0.22	0.20	0.17	0.15	0.20	0.31	0.50	0.73	1.00	0.75	0.53	0.34	0.25	0.05	-0.17	-0.28	-0.27	-0.12	0.13	0.29	0.19	0.09
0.25	0.23	0.21	0.16	0.11	0.17	0.30	0.51	0.75	1.00	0.76	0.23	0.34	0.26	0.04	-0.18	-0.29	-0.28	-0.11	0.13	0.26	0.15
0.22	0.26	0.25	0.21	0.13	0.09	0.17	0.32	0.53	0.76	1.00	0.02	0.24	0.35	0.25	0.02	-0.20	-0.30	-0.27	-0.10	0.12	0.24
0.27	0.16	-0.05	-0.21	-0.25	-0.17	0.03	0.25	0.34	0.23	0.02	1.00	0.68	0.06	-0.45	-0.63	-0.45	-0.10	0.19	0.31	0.27	0.16
0.17	0.28	0.16	-0.05	-0.22	-0.26	-0.17	0.04	0.25	0.34	0.24	0.68	1.00	0.68	0.05	-0.47	-0.63	-0.45	-0.09	0.19	0.30	0.25
0.10	0.18	0.29	0.16	-0.08	-0.25	-0.26	-0.16	0.05	0.26	0.35	0.06	0.68	1.00	0.67	0.03	-0.48	-0.64	-0.44	-0.10	0.17	0.28
0.08	0.10	0.18	0.29	0.15	-0.09	-0.25	-0.27	-0.17	0.04	0.25	-0.45	0.05	0.67	1.00	0.67	0.03	-0.48	-0.64	-0.45	-0.11	0.16
0.03	0.08	0.10	0.19	0.29	0.15	-0.10	-0.26	-0.28	-0.18	0.02	-0.63	-0.47	0.03	0.67	1.00	0.67	0.04	-0.48	-0.65	-0.47	-0.12
-0.01	0.03	0.07	0.10	0.19	0.29	0.14	-0.11	-0.27	-0.29	-0.20	-0.45	-0.63	-0.48	0.03	0.67	1.00	0.67	0.04	-0.48	-0.65	-0.47
-0.04	-0.02	0.02	0.07	0.12	0.20	0.29	0.13	-0.12	-0.28	-0.30	-0.10	-0.45	-0.64	-0.48	0.04	0.67	1.00	0.67	0.04	-0.47	-0.64
-0.04	-0.04	-0.02	0.02	0.08	0.12	0.20	0.29	0.13	-0.11	-0.27	0.19	-0.09	-0.44	-0.64	-0.48	0.04	0.67	1.00	0.67	0.05	-0.47
0.00	-0.03	-0.04	-0.02	0.01	0.07	0.12	0.20	0.29	0.13	-0.10	0.31	0.19	-0.10	-0.45	-0.65	-0.48	0.04	0.67	1.00	0.67	0.04
0.05	0.01	-0.03	-0.03	-0.03	0.00	0.06	0.11	0.19	0.26	0.12	0.27	0.30	0.17	-0.11	-0.47	-0.65	-0.47	0.05	0.67	1.00	0.67
0.09	0.05	0.00	-0.03	-0.04	-0.04	-0.01	0.05	0.09	0.15	0.24	0.16	0.25	0.28	0.16	-0.12	-0.47	-0.64	-0.47	0.04	0.67	1.00

Correlation matrix for the original variables when $\rho_{a_1 a_2} \approx .58$.

1.00	-0.03	0.55	-0.04	0.26	-0.05	0.12	0.08	0.07	0.20	0.01	0.32	-0.01	0.11	0.01	0.05	-0.05	0.01	-0.06	0.03	0.00	0.09
-0.03	1.00	-0.02	0.55	-0.04	0.27	-0.05	0.12	0.08	0.09	0.20	0.01	0.32	0.00	0.12	0.01	0.06	-0.05	0.01	-0.06	0.04	0.00
0.55	-0.02	1.00	-0.02	0.54	-0.04	0.27	-0.04	0.13	0.10	0.10	0.03	0.02	0.32	0.00	0.11	0.01	0.05	-0.05	0.01	-0.06	0.03
-0.04	0.55	-0.02	1.00	-0.02	0.55	-0.05	0.27	-0.04	0.14	0.11	-0.21	0.03	0.03	0.33	0.00	0.11	0.00	0.04	-0.05	0.01	-0.06
0.26	-0.04	0.54	-0.02	1.00	-0.02	0.55	-0.06	0.27	-0.06	0.13	-0.10	-0.22	0.02	0.02	0.34	0.00	0.12	0.01	0.04	-0.05	0.02
-0.05	0.27	-0.04	0.55	-0.02	1.00	-0.02	0.54	-0.06	0.24	-0.06	-0.18	-0.10	-0.24	0.01	0.01	0.33	0.01	0.13	0.00	0.03	-0.06
0.12	-0.05	0.27	-0.05	0.55	-0.02	1.00	-0.03	0.54	-0.08	0.23	0.07	-0.18	-0.11	-0.24	0.01	0.01	0.33	0.01	0.13	0.00	0.03
0.08	0.12	-0.04	0.27	-0.06	0.54	-0.03	1.00	-0.02	0.54	-0.06	0.12	0.07	-0.18	-0.11	-0.25	0.00	0.00	0.33	0.02	0.12	-0.01
0.07	0.08	0.13	-0.04	0.27	-0.06	0.54	-0.02	1.00	0.00	0.55	0.28	0.12	0.08	-0.17	-0.11	-0.26	0.00	0.00	0.34	0.02	0.12
0.20	0.09	0.10	0.14	-0.06	0.24	-0.08	0.54	0.00	1.00	0.03	0.12	0.28	0.13	0.08	-0.20	-0.12	-0.27	0.00	0.01	0.31	0.00
0.01	0.20	0.10	0.11	0.13	-0.06	0.23	-0.06	0.55	0.03	1.00	0.04	0.13	0.29	0.14	0.08	-0.20	-0.14	-0.27	0.00	0.01	0.31
0.32	0.01	0.03	-0.21	-0.10	-0.18	0.07	0.12	0.28	0.12	0.04	1.00	0.43	0.10	-0.47	-0.45	-0.41	0.00	0.12	0.30	0.15	0.16
-0.01	0.32	0.02	0.03	-0.22	-0.10	-0.18	0.07	0.12	0.28	0.13	0.43	1.00	0.43	0.10	-0.47	-0.46	-0.41	0.01	0.12	0.30	0.15
0.11	0.00	0.32	0.03	0.02	-0.24	-0.11	-0.18	0.08	0.13	0.29	0.10	0.43	1.00	0.43	0.09	-0.49	-0.46	-0.41	0.01	0.11	0.29
0.01	0.12	0.00	0.33	0.02	-0.01	-0.24	-0.11	-0.17	0.08	0.14	-0.47	0.10	0.43	1.00	0.41	0.07	-0.49	-0.46	-0.41	0.00	0.10
0.05	0.01	0.11	0.00	0.34	0.01	0.01	-0.25	-0.11	-0.20	0.08	-0.45	-0.47	0.09	0.41	1.00	0.41	0.09	-0.49	-0.47	-0.42	-0.01
-0.05	0.06	0.01	0.11	0.00	0.33	0.01	0.00	-0.26	-0.12	-0.20	-0.41	-0.46	-0.49	0.07	0.41	1.00	0.42	0.10	-0.49	-0.48	-0.43
0.01	-0.05	0.05	0.00	0.12	0.01	0.33	0.00	0.00	-0.27	-0.14	0.00	-0.41	-0.46	-0.49	0.09	0.42	1.00	0.42	0.09	-0.48	-0.47
-0.06	0.01	-0.05	0.04	0.01	0.13	0.01	0.33	0.00	0.00	-0.27	0.12	0.01	-0.41	-0.46	-0.49	0.10	0.42	1.00	0.42	0.10	-0.48
0.03	-0.06	0.01	-0.05	0.04	0.00	0.13	0.02	0.34	0.01	0.00	0.30	0.12	0.01	-0.41	-0.47	-0.49	0.09	0.42	1.00	0.42	0.09
0.00	0.04	-0.06	0.01	-0.05	0.03	0.00	0.12	0.02	0.31	0.01	0.15	0.30	0.11	0.00	-0.42	-0.48	-0.48	0.10	0.42	1.00	0.41
0.09	0.00	0.03	-0.06	0.02	-0.06	0.03	-0.01	0.12	0.00	0.31	0.16	0.15	0.29	0.10	-0.01	-0.43	-0.47	-0.48	0.09	0.41	1.00

Correlation matrix for the transformed variables when $\rho_{a_1 a_2} \approx .58$.

Some measures of multicollinearity for the two simulated correlation matrices with $\rho_{a_1 a_2} \approx .58$.

Original variables				Transformed variables			
i	λ_i	%	Σ%	i	λ_i	%	Σ%
1	4.983	22.6	22.6	1	3.783	17.2	17.2
2	4.506	20.5	43.1	2	3.220	14.6	31.8
3	4.037	18.3	61.5	3	2.825	12.8	44.7
4	1.973	9.0	70.5	4	2.365	10.7	55.4
5	1.673	7.6	78.1	5	1.652	7.5	62.9
6	1.106	5.0	83.1	6	1.336	6.1	69.0
7	0.855	3.9	87.0	7	1.118	5.1	74.1
8	0.505	2.3	89.3	8	0.920	4.2	78.3
9	0.394	1.8	91.1	9	0.721	3.3	81.5
10	0.391	1.8	92.8	10	0.597	2.7	84.3
11	0.319	1.5	94.3	11	0.529	2.4	86.7
12	0.277	1.3	95.5	12	0.508	2.3	89.0
13	0.229	1.0	96.6	13	0.491	2.2	91.2
14	0.176	0.8	97.4	14	0.472	2.1	93.3
15	0.166	0.8	98.1	15	0.306	1.4	94.7
16	0.157	0.7	98.9	16	0.293	1.3	96.1
17	0.101	0.5	99.3	17	0.249	1.1	97.2
18	0.039	0.2	99.5	18	0.223	1.0	98.2
19	0.031	0.1	99.6	19	0.143	0.7	98.9
20	0.031	0.1	99.8	20	0.129	0.6	99.5
21	0.025	0.1	99.9	21	0.063	0.3	99.7
22	0.024	0.1	100.0	22	0.056	0.3	100.0
	22.000	100.0			22.000	100.0	

Sum of variance inflation factors:

$\Sigma\ 1/\lambda_i = 221.02$ $\Sigma\ 1/\lambda_i = 79.17$

The spectral condition number:

$\lambda_1/\lambda_{22} = 204.53$ $\lambda_1/\lambda_{22} = 67.22$

Determinant of the correlation matrix:

$|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 4.73826 \cdot 10^{-12}$ $|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 1.63570 \cdot 10^{-6}$

Haitovskys test variable:

$\chi^2_H(231) = 0.000$ $\chi^2_H(231) = 0.000$

1.00	0.74	0.53	0.36	0.23	0.14	0.14	0.18	0.22	0.25	0.22	0.63	0.51	0.38	0.22	0.06	-0.06	-0.09	-0.04	0.07	0.18	0.23
0.74	1.00	0.74	0.53	0.36	0.23	0.15	0.15	0.20	0.23	0.26	0.41	0.63	0.52	0.38	0.21	0.04	-0.07	-0.09	-0.02	0.08	0.18
0.53	0.74	1.00	0.73	0.52	0.35	0.23	0.16	0.17	0.21	0.25	0.05	0.42	0.63	0.51	0.35	0.19	0.04	-0.07	-0.07	-0.01	0.08
0.36	0.53	0.73	1.00	0.74	0.53	0.35	0.23	0.15	0.16	0.21	-0.28	0.04	0.41	0.63	0.50	0.35	0.19	0.04	-0.07	-0.07	-0.01
0.23	0.36	0.52	0.74	1.00	0.74	0.52	0.34	0.20	0.11	0.13	-0.40	-0.28	0.03	0.42	0.63	0.51	0.36	0.19	0.02	-0.08	-0.08
0.14	0.23	0.35	0.53	0.74	1.00	0.73	0.51	0.31	0.17	0.09	-0.30	-0.40	-0.29	0.03	0.42	0.64	0.52	0.36	0.17	0.00	-0.09
0.14	0.15	0.23	0.35	0.52	0.73	1.00	0.73	0.50	0.30	0.17	-0.04	-0.29	-0.40	-0.29	0.03	0.41	0.64	0.52	0.36	0.17	0.00
0.18	0.15	0.16	0.23	0.34	0.51	0.73	1.00	0.73	0.51	0.32	0.25	-0.01	-0.27	-0.40	-0.31	0.01	0.39	0.63	0.52	0.37	0.17
0.22	0.20	0.17	0.15	0.20	0.31	0.50	0.73	1.00	0.75	0.53	0.40	0.28	0.01	-0.28	-0.43	-0.33	-0.01	0.39	0.64	0.53	0.36
0.25	0.23	0.21	0.16	0.11	0.17	0.30	0.51	0.75	1.00	0.76	0.36	0.43	0.30	-0.01	-0.32	-0.45	-0.34	0.00	0.40	0.63	0.51
0.22	0.26	0.25	0.21	0.13	0.09	0.17	0.32	0.53	0.76	1.00	0.17	0.39	0.44	0.28	-0.05	-0.35	-0.46	-0.32	0.02	0.41	0.62
0.63	0.41	0.05	-0.28	-0.40	-0.30	-0.04	0.25	0.40	0.36	0.17	1.00	0.73	0.17	-0.39	-0.68	-0.61	-0.27	0.16	0.47	0.53	0.37
0.51	0.63	0.42	0.04	-0.28	-0.40	-0.29	-0.01	0.28	0.43	0.39	0.73	1.00	0.74	0.16	-0.40	-0.69	-0.62	-0.26	0.18	0.48	0.53
0.38	0.52	0.63	0.41	0.03	-0.29	-0.40	-0.27	0.01	0.30	0.44	0.17	0.74	1.00	0.72	0.14	-0.41	-0.69	-0.61	-0.24	0.19	0.48
0.22	0.38	0.51	0.63	0.42	0.03	-0.29	-0.40	-0.28	-0.01	0.28	-0.39	0.16	0.72	1.00	0.72	0.15	-0.41	-0.69	-0.61	-0.25	0.19
0.06	0.21	0.35	0.50	0.63	0.42	0.03	-0.31	-0.43	-0.32	-0.05	-0.68	-0.40	0.14	0.72	1.00	0.73	0.16	-0.41	-0.70	-0.62	-0.25
-0.06	0.04	0.19	0.35	0.51	0.64	0.41	0.01	-0.33	-0.45	-0.35	-0.61	-0.69	-0.41	0.15	0.73	1.00	0.73	0.15	-0.42	-0.71	-0.62
-0.09	-0.07	0.04	0.19	0.36	0.52	0.64	0.39	-0.01	-0.34	-0.46	-0.27	-0.62	-0.69	-0.41	0.16	0.73	1.00	0.73	0.14	-0.42	-0.71
-0.04	-0.09	-0.07	0.04	0.19	0.36	0.52	0.63	0.39	0.00	-0.32	0.16	-0.26	-0.61	-0.69	-0.41	0.15	0.73	1.00	0.73	0.14	-0.42
0.07	-0.02	-0.07	-0.07	0.02	0.17	0.36	0.52	0.64	0.40	0.02	0.47	0.18	-0.24	-0.61	-0.70	-0.42	0.14	0.73	1.00	0.73	0.14
0.18	0.08	-0.01	-0.07	-0.08	0.00	0.17	0.37	0.53	0.63	0.41	0.53	0.48	0.19	-0.25	-0.62	-0.71	-0.42	0.14	0.73	1.00	0.73
0.23	0.18	0.08	-0.01	-0.08	-0.09	0.00	0.17	0.36	0.51	0.62	0.37	0.53	0.48	0.19	-0.25	-0.62	-0.71	-0.42	0.14	0.73	1.00

Correlation matrix for the original variables when $\rho_{a_1 a_2} \approx .98$.

1.00	-0.03	0.55	-0.04	0.26	-0.05	0.12	0.08	0.07	0.20	0.01	0.64	0.15	0.38	0.05	0.10	-0.11	-0.03	-0.05	0.04	0.13	0.13
-0.03	1.00	-0.02	0.55	-0.04	0.27	-0.05	0.12	0.08	0.09	0.20	0.14	0.63	0.15	0.38	0.05	0.10	-0.11	-0.03	-0.05	0.05	0.13
0.55	-0.02	1.00	-0.02	0.54	-0.04	0.27	-0.04	0.13	0.10	0.10	0.12	0.15	0.64	0.15	0.36	0.03	0.09	-0.11	-0.03	-0.04	0.05
-0.04	0.55	-0.02	1.00	-0.02	0.55	-0.05	0.27	-0.04	0.14	0.11	-0.29	0.12	0.16	0.64	0.14	0.35	0.03	0.09	-0.11	-0.02	-0.03
0.26	-0.04	0.54	-0.02	1.00	-0.02	0.55	-0.06	0.27	-0.06	0.13	-0.23	-0.30	0.11	0.15	0.64	0.15	0.36	0.02	0.08	-0.12	-0.02
-0.05	0.27	-0.04	0.55	-0.02	1.00	-0.02	0.54	-0.06	0.24	-0.06	-0.26	-0.22	-0.30	0.11	0.15	0.63	0.16	0.36	0.02	0.06	-0.13
0.12	-0.05	0.27	-0.05	0.55	-0.02	1.00	-0.03	0.54	-0.08	0.23	-0.01	-0.26	-0.23	-0.30	0.12	0.16	0.64	0.16	0.35	0.01	0.06
0.08	0.12	-0.04	0.27	-0.06	0.54	-0.03	1.00	-0.02	0.54	-0.06	0.17	0.01	-0.24	-0.23	-0.31	0.09	0.14	0.63	0.17	0.36	0.01
0.07	0.08	0.13	-0.04	0.27	-0.06	0.54	-0.02	1.00	0.00	0.55	0.26	0.18	0.02	-0.24	-0.23	-0.32	0.08	0.14	0.64	0.17	0.36
0.20	0.09	0.10	0.14	-0.06	0.24	-0.08	0.54	0.00	1.00	0.03	0.28	0.30	0.21	0.02	-0.28	-0.27	-0.34	0.09	0.15	0.64	0.17
0.01	0.20	0.10	0.11	0.13	-0.06	0.23	-0.06	0.55	0.03	1.00	0.07	0.29	0.32	0.21	0.01	-0.29	-0.28	-0.33	0.09	0.17	0.64
0.64	0.14	0.12	-0.29	-0.23	-0.26	-0.01	0.17	0.26	0.28	0.07	1.00	0.57	0.18	-0.40	-0.58	-0.56	-0.21	0.15	0.42	0.47	0.31
0.15	0.63	0.15	0.12	-0.30	-0.22	-0.26	0.01	0.18	0.30	0.29	0.57	1.00	0.58	0.18	-0.41	-0.59	-0.58	-0.21	0.16	0.44	0.47
0.38	0.15	0.64	0.16	0.11	-0.30	-0.23	-0.24	0.02	0.21	0.32	0.18	0.58	1.00	0.58	0.15	-0.43	-0.60	-0.57	-0.19	0.18	0.44
0.05	0.38	0.15	0.64	0.15	0.11	-0.30	-0.23	-0.24	0.02	0.21	-0.40	0.18	0.58	1.00	0.57	0.14	-0.43	-0.60	-0.56	-0.19	0.18
0.10	0.05	0.36	0.14	0.64	0.15	0.12	-0.31	-0.23	-0.28	0.01	-0.58	-0.41	0.15	0.57	1.00	0.58	0.16	-0.42	-0.61	-0.58	-0.19
-0.11	0.10	0.03	0.35	0.15	0.63	0.16	0.09	-0.32	-0.27	-0.29	-0.56	-0.59	-0.43	0.14	0.58	1.00	0.59	0.16	-0.43	-0.62	-0.58
-0.03	-0.11	0.09	0.03	0.36	0.16	0.64	0.14	0.08	-0.34	-0.28	-0.21	-0.58	-0.60	-0.43	0.16	0.59	1.00	0.58	0.15	-0.44	-0.62
-0.05	-0.03	-0.11	0.09	0.02	0.36	0.16	0.63	0.14	0.09	-0.33	0.15	-0.21	-0.57	-0.60	-0.42	0.16	0.58	1.00	0.58	0.15	-0.44
0.04	-0.05	-0.03	-0.11	0.08	0.02	0.35	0.17	0.64	0.15	0.09	0.42	0.16	-0.19	-0.56	-0.61	-0.43	0.15	0.58	1.00	0.59	0.15
0.13	0.05	-0.04	-0.02	-0.12	0.06	0.01	0.36	0.17	0.64	0.17	0.47	0.44	0.18	-0.19	-0.58	-0.62	-0.44	0.15	0.59	1.00	0.58
0.13	0.13	0.05	-0.03	-0.02	-0.13	0.06	0.01	0.36	0.17	0.64	0.31	0.47	0.44	0.18	-0.19	-0.58	-0.62	-0.44	0.15	0.58	1.00

Correlation matrix for the transformed variables when $\rho_{a_1 a_2} \approx .98$.

Some measures of multicollinearity for the two simulated correlation matrices with $\rho_{a_1 a_2} \approx .98$.

Original variables				Transformed variables			
i	λ_i	%	Σ%	i	λ_i	%	Σ%
1	6.333	28.8	28.8	1	5.153	23.4	23.4
2	5.201	23.6	52.4	2	4.180	19.0	42.4
3	5.078	23.1	75.5	3	2.954	13.4	55.9
4	1.986	9.0	84.5	4	2.770	12.6	68.4
5	0.915	4.2	88.7	5	1.910	8.7	77.1
6	0.717	3.3	92.0	6	1.338	6.1	83.2
7	0.421	1.9	93.9	7	0.837	3.8	87.0
8	0.337	1.5	95.4	8	0.627	2.9	89.9
9	0.261	1.2	96.6	9	0.566	2.6	92.4
10	0.198	0.9	97.5	10	0.492	2.2	94.7
11	0.175	0.8	98.3	11	0.463	2.1	96.8
12	0.169	0.8	99.1	12	0.408	1.9	98.6
13	0.148	0.7	99.7	13	0.214	1.0	99.6
14	0.022	0.1	99.8	14	0.024	0.1	99.7
15	0.017	0.1	99.9	15	0.016	0.1	99.8
16	0.012	0.1	100.0	16	0.013	0.1	99.8
17	0.005	0.0	100.0	17	0.010	0.0	99.9
18	0.002	0.0	100.0	18	0.009	0.0	99.9
19	0.001	0.0	100.0	19	0.006	0.0	100.0
20	0.001	0.0	100.0	20	0.006	0.0	100.0
21	0.001	0.0	100.0	21	0.003	0.0	100.0
22	0.001	0.0	100.0	22	0.002	0.0	100.0
	22.000	100.0			22.000	100.0	

Sum of variance inflation factors:

$\Sigma\ 1/\lambda_i = 5069.50$ $\qquad$ $\Sigma\ 1/\lambda_i = 1516.92$

The spectral condition number:

$\lambda_1/\lambda_{22} = 7306.35$ $\qquad$ $\lambda_1/\lambda_{22} = 2120.83$

Determinant of the correlation matrix:

$|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 2.34242 \cdot 10^{-25}$ $\qquad$ $|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 2.73348 \cdot 10^{-19}$

Haitovskys test variable:

$\chi^2_H(231) = 0.000$ $\qquad$ $\chi^2_H(231) = 0.000$

APPENDIX 4. Autocorrelation and cross correlation functions for theoretical processes.

$\rho_{a_1 a_2} \approx .58$ | $\rho_{a_1 a_2} \approx .98$

lag k	x_1 $\rho(k)$	x_2 $\rho(k)$	x_1' $\rho(k)$	x_2' $\rho(k)$	x_1 $\rho(k)$	x_2 $\rho(k)$	x_1' $\rho(k)$	x_2' $\rho(k)$
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.70	0.71	-0.11	0.51	0.70	0.71	-0.11	0.51
2	0.49	0.14	0.53	0.16	0.49	0.14	0.53	0.16
3	0.34	-0.36	-0.08	-0.39	0.34	-0.36	-0.08	-0.39
4	0.24	-0.55	0.28	-0.45	0.24	-0.55	0.28	-0.45
5	0.17	-0.42	-0.06	-0.39	0.17	-0.42	-0.06	-0.39
6	0.12	-0.11	0.15	-0.06	0.12	-0.11	0.15	-0.06
7	0.08	0.17	-0.04	0.15	0.08	0.17	-0.04	0.15
8	0.06	0.30	0.08	0.28	0.06	0.30	0.08	0.28
9	0.04	0.25	-0.02	0.20	0.04	0.25	-0.02	0.20
10	0.03	0.08	0.04	0.07	0.03	0.08	0.04	0.07

	$\rho_{x_1 x_2}$		$\rho_{x_1' x_2'}$		$\rho_{x_1 x_2}$		$\rho_{x_1' x_2'}$	
lag k	$\rho(k)$	$\rho(-k)$	$\rho(k)$	$\rho(-k)$	$\rho(k)$	$\rho(-k)$	$\rho(k)$	$\rho(-k)$
0	0.39	0.39	0.39	0.39	0.66	0.66	0.67	0.67
1	0.28	0.27	0.09	0.03	0.48	0.46	0.15	0.04
2	0.06	0.19	0.10	0.21	0.10	0.32	0.17	0.35
3	-0.14	0.13	-0.15	0.00	-0.23	0.26	-0.26	0.01
4	-0.21	0.09	-0.11	0.11	-0.36	0.16	-0.18	0.18
5	-0.17	0.07	-0.14	-0.00	-0.28	0.11	-0.24	-0.01
6	-0.05	0.05	-0.00	0.06	-0.08	0.08	-0.01	0.10
7	0.07	0.03	0.03	-0.00	0.11	0.05	0.05	-0.01
8	0.12	0.02	0.09	0.03	0.20	0.04	0.16	0.05
9	0.10	0.02	0.05	-0.00	0.17	0.03	0.09	-0.01
10	0.03	0.01	0.03	0.02	0.06	0.02	0.05	0.03

APPENDIX 5. Correlation matrices and eigenvalues for theoretical processes.

1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.04	0.03	0.39	0.27	0.19	0.13	0.09	0.07	0.05	0.03	0.02	0.02	0.01
0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.04	0.28	0.39	0.27	0.19	0.13	0.09	0.07	0.05	0.03	0.02	0.02
0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.06	0.28	0.39	0.27	0.19	0.13	0.09	0.07	0.05	0.03	0.02
0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	-0.14	0.06	0.28	0.39	0.27	0.19	0.13	0.09	0.07	0.05	0.03
0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	-0.21	-0.14	0.06	0.28	0.39	0.27	0.19	0.13	0.09	0.07	0.05
0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	-0.17	-0.21	-0.14	0.06	0.28	0.39	0.27	0.19	0.13	0.09	0.07
0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	-0.05	-0.17	-0.21	-0.14	0.06	0.28	0.39	0.27	0.19	0.13	0.09
0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.07	-0.05	-0.17	-0.21	-0.14	0.06	0.28	0.39	0.27	0.19	0.13
0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.12	0.07	-0.05	-0.17	-0.21	-0.14	0.06	0.28	0.39	0.27	0.19
0.04	0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.10	0.12	0.07	-0.05	-0.17	-0.21	-0.14	0.06	0.28	0.39	0.27
0.03	0.04	0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.03	0.10	0.12	0.07	-0.05	-0.17	-0.21	-0.14	0.06	0.28	0.39
0.39	0.28	0.06	-0.14	-0.21	-0.17	-0.05	0.07	0.12	0.10	0.03	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30	0.25	0.08
0.27	0.39	0.28	0.06	-0.14	-0.21	-0.17	-0.05	0.07	0.12	0.10	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30	0.25
0.19	0.27	0.39	0.28	0.06	-0.14	-0.21	-0.17	-0.05	0.07	0.12	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30
0.13	0.19	0.27	0.39	0.28	0.06	-0.14	-0.21	-0.17	-0.05	0.07	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17
0.09	0.13	0.19	0.27	0.39	0.28	0.06	-0.14	-0.21	-0.17	-0.05	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11
0.07	0.09	0.13	0.19	0.27	0.39	0.28	0.06	-0.14	-0.21	-0.17	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42
0.05	0.07	0.09	0.13	0.19	0.27	0.39	0.28	0.06	-0.14	-0.21	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55
0.03	0.05	0.07	0.09	0.13	0.19	0.27	0.39	0.28	0.06	-0.14	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36
0.02	0.03	0.05	0.07	0.09	0.13	0.19	0.27	0.39	0.28	0.06	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14
0.02	0.02	0.03	0.05	0.07	0.09	0.13	0.19	0.27	0.39	0.28	0.25	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71
0.01	0.02	0.02	0.03	0.05	0.07	0.09	0.13	0.19	0.27	0.39	0.08	0.25	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00

Correlation matrix for the original variables when $\rho_{a_1 a_2} \approx .58$.

1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	-0.02	0.04	0.39	0.03	0.21	0.00	0.11	0.00	0.06	0.00	0.03	0.00	0.02
-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	-0.02	0.09	0.39	0.03	0.21	0.00	0.11	0.00	0.06	0.00	0.03	0.00
0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	0.10	0.09	0.39	0.03	0.21	0.00	0.11	0.00	0.06	0.00	0.03
-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	-0.15	0.10	0.09	0.39	0.03	0.21	0.00	0.11	0.00	0.06	0.00
0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.11	-0.15	0.10	0.09	0.39	0.03	0.21	0.00	0.11	0.00	0.06
-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	-0.14	-0.11	-0.15	0.10	0.09	0.39	0.03	0.21	0.00	0.11	0.00
0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	0.00	-0.14	-0.11	-0.15	0.10	0.09	0.39	0.03	0.21	0.00	0.11
-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.03	0.00	-0.14	-0.11	-0.15	0.10	0.09	0.39	0.03	0.21	0.00
0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	0.09	0.03	0.00	-0.14	-0.11	-0.15	0.10	0.09	0.39	0.03	0.21
-0.02	0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.05	0.09	0.03	0.00	-0.14	-0.11	-0.15	0.10	0.09	0.39	0.03
0.04	-0.02	0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	0.03	0.05	0.09	0.03	0.00	-0.14	-0.11	-0.15	0.10	0.09	0.39
0.39	0.09	0.10	-0.15	-0.11	-0.14	0.00	0.03	0.09	0.05	0.03	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28	0.20	0.07
0.03	0.39	0.09	0.10	-0.15	-0.11	-0.14	0.00	0.03	0.09	0.05	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28	0.20
0.21	0.03	0.39	0.09	0.10	-0.15	-0.11	-0.14	0.00	0.03	0.09	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28
0.00	0.21	0.03	0.39	0.09	0.10	-0.15	-0.11	-0.14	0.00	0.03	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15
0.11	0.00	0.21	0.03	0.39	0.09	0.10	-0.15	-0.11	-0.14	0.00	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06
0.00	0.11	0.00	0.21	0.03	0.39	0.09	0.10	-0.15	-0.11	-0.14	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39
0.06	0.00	0.11	0.00	0.21	0.03	0.39	0.09	0.10	-0.15	-0.11	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45
0.00	0.06	0.00	0.11	0.00	0.21	0.03	0.39	0.09	0.10	-0.15	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39
0.03	0.00	0.06	0.00	0.11	0.00	0.21	0.03	0.39	0.09	0.10	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16
0.00	0.03	0.00	0.06	0.00	0.11	0.00	0.21	0.03	0.39	0.09	0.20	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51
0.02	0.00	0.03	0.00	0.06	0.00	0.11	0.00	0.21	0.03	0.39	0.07	0.20	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00

Correlation matrix for the transformed variables when $\rho_{a_1 a_2} \approx .58$.

Some measures of multicollinearity for the two theoretical correlation matrices with $\rho_{a_1 a_2} \approx .58$.

Original variables				Transformed variables			
i	λ_i	%	Σ%	i	λ_i	%	Σ%
1	4.786	21.8	21.8	1	3.684	16.7	16.7
2	4.425	20.1	41.9	2	3.208	14.6	31.3
3	3.914	17.8	59.7	3	3.111	14.1	45.5
4	2.517	11.4	71.1	4	2.329	10.6	56.1
5	1.352	6.1	77.2	5	1.773	8.1	64.1
6	0.925	4.2	81.4	6	1.474	6.7	70.8
7	0.754	3.4	84.9	7	1.000	4.5	75.4
8	0.627	2.9	87.7	8	0.802	3.6	79.0
9	0.489	2.2	89.9	9	0.725	3.3	82.3
10	0.461	2.1	92.0	10	0.558	2.5	84.8
11	0.331	1.5	93.5	11	0.482	2.2	87.0
12	0.280	1.3	94.8	12	0.445	2.0	89.1
13	0.248	1.1	95.9	13	0.429	2.0	91.0
14	0.224	1.0	97.0	14	0.415	1.9	92.9
15	0.201	0.9	97.9	15	0.384	1.7	94.6
16	0.189	0.9	98.7	16	0.331	1.5	96.1
17	0.123	0.6	99.3	17	0.219	1.0	97.1
18	0.061	0.3	99.6	18	0.205	0.9	98.1
19	0.036	0.2	99.7	19	0.131	0.6	98.7
20	0.024	0.1	99.8	20	0.120	0.5	99.2
21	0.019	0.1	99.9	21	0.089	0.4	99.6
22	0.016	0.1	100.0	22	0.086	0.4	100.0
	22.000	100.0			22.000	100.0	

Sum of variance inflation factors:

$\Sigma\ 1/\lambda_i = 245.69$ $\Sigma\ 1/\lambda_i = 71.00$

The spectral condition number:

$\lambda_1/\lambda_{22} = 300.76$ $\lambda_1/\lambda_{22} = 42.75$

Determinant of the correlation matrix:

$|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 1.04877 \cdot 10^{-11}$ $|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 1.89621 \cdot 10^{-6}$

Haitovskys test variable:

$\chi^2_H(231) = 0.000$ $\chi^2_H(231) = 0.000$

1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.04	0.03	0.58	0.41	0.28	0.20	0.14	0.10	0.07	0.05	0.03	0.02	0.02
0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.04	0.42	0.58	0.41	0.28	0.20	0.14	0.10	0.07	0.05	0.03	0.02
0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.09	0.42	0.58	0.41	0.28	0.20	0.14	0.10	0.07	0.05	0.03
0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	-0.20	0.09	0.42	0.58	0.41	0.28	0.20	0.14	0.10	0.07	0.05
0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	-0.32	-0.20	0.09	0.42	0.58	0.41	0.28	0.20	0.14	0.10	0.07
0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	-0.25	-0.32	-0.20	0.09	0.42	0.58	0.41	0.28	0.20	0.14	0.10
0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	-0.07	-0.25	-0.32	-0.20	0.09	0.42	0.58	0.41	0.28	0.20	0.14
0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.10	-0.07	-0.25	-0.32	-0.20	0.09	0.42	0.58	0.41	0.28	0.20
0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.18	0.10	-0.07	-0.25	-0.32	-0.20	0.09	0.42	0.58	0.41	0.28
0.04	0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.15	0.18	0.10	-0.07	-0.25	-0.32	-0.20	0.09	0.42	0.58	0.41
0.03	0.04	0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.05	0.15	0.18	0.10	-0.07	-0.25	-0.32	-0.20	0.09	0.42	0.58
0.58	0.42	0.09	-0.20	-0.32	-0.25	-0.07	0.10	0.18	0.15	0.05	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30	0.25	0.08
0.41	0.58	0.42	0.09	-0.20	-0.32	-0.25	-0.07	0.10	0.18	0.15	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30	0.25
0.28	0.41	0.58	0.42	0.09	-0.20	-0.32	-0.25	-0.07	0.10	0.18	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30
0.20	0.28	0.41	0.58	0.42	0.09	-0.20	-0.32	-0.25	-0.07	0.10	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17
0.14	0.20	0.28	0.41	0.58	0.42	0.09	-0.20	-0.32	-0.25	-0.07	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11
0.10	0.14	0.20	0.28	0.41	0.58	0.42	0.09	-0.20	-0.32	-0.25	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42
0.07	0.10	0.14	0.20	0.28	0.41	0.58	0.42	0.09	-0.20	-0.32	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55
0.05	0.07	0.10	0.14	0.20	0.28	0.41	0.58	0.42	0.09	-0.20	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36
0.03	0.05	0.07	0.10	0.14	0.20	0.28	0.41	0.58	0.42	0.09	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14
0.02	0.03	0.05	0.07	0.10	0.14	0.20	0.28	0.41	0.58	0.42	0.25	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71
0.02	0.02	0.03	0.05	0.07	0.10	0.14	0.20	0.28	0.41	0.58	0.08	0.25	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00

Correlation matrix for the original variables when $\rho_{a_1 a_2} \approx .87$.

1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	-0.02	0.04	0.59	0.04	0.31	0.00	0.16	-0.01	0.09	-0.01	0.05	-0.01	0.02
-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	-0.02	0.13	0.59	0.04	0.31	0.00	0.16	-0.01	0.09	-0.01	0.05	-0.01
0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	0.15	0.13	0.59	0.04	0.31	0.00	0.16	-0.01	0.09	-0.01	0.05
-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	-0.23	0.15	0.13	0.59	0.04	0.31	0.00	0.16	-0.01	0.09	-0.01
0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.16	-0.23	0.15	0.13	0.59	0.04	0.31	0.00	0.16	-0.01	0.09
-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	-0.21	-0.16	-0.23	0.15	0.13	0.59	0.04	0.31	0.00	0.16	-0.01
0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	0.00	-0.21	-0.16	-0.23	0.15	0.13	0.59	0.04	0.31	0.00	0.16
-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.05	0.00	-0.21	-0.16	-0.23	0.15	0.13	0.59	0.04	0.31	0.00
0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	0.14	0.05	0.00	-0.21	-0.16	-0.23	0.15	0.13	0.59	0.04	0.31
-0.02	0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.08	0.14	0.05	0.00	-0.21	-0.16	-0.23	0.15	0.13	0.59	0.04
0.04	-0.02	0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	0.04	0.08	0.14	0.05	0.00	-0.21	-0.16	-0.23	0.15	0.13	0.59
0.59	0.13	0.15	-0.23	-0.16	-0.21	0.00	0.05	0.14	0.08	0.04	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28	0.20	0.07
0.04	0.59	0.13	0.15	-0.23	-0.16	-0.21	0.00	0.05	0.14	0.08	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28	0.20
0.31	0.04	0.59	0.13	0.15	-0.23	-0.16	-0.21	0.00	0.05	0.14	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28
0.00	0.31	0.04	0.59	0.13	0.15	-0.23	-0.16	-0.21	0.00	0.05	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15
0.16	0.00	0.31	0.04	0.59	0.13	0.15	-0.23	-0.16	-0.21	0.00	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06
-0.01	0.16	0.00	0.31	0.04	0.59	0.13	0.15	-0.23	-0.16	-0.21	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39
0.09	-0.01	0.16	0.00	0.31	0.04	0.59	0.13	0.15	-0.23	-0.16	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45
-0.01	0.09	-0.01	0.16	0.00	0.31	0.04	0.59	0.13	0.15	-0.23	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39
0.05	-0.01	0.09	-0.01	0.16	0.00	0.31	0.04	0.59	0.13	0.15	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16
-0.01	0.05	-0.01	0.09	-0.01	0.16	0.00	0.31	0.04	0.59	0.13	0.20	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51
0.02	-0.01	0.05	-0.01	0.09	-0.01	0.16	0.00	0.31	0.04	0.59	0.07	0.20	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00

Correlation matrix for the transformed variables when $\rho_{a_1 a_2} \approx .87$.

Some measures of multicollinearity for the two theoretical correlation matrices with $\rho_{a_1 a_2} \approx .87$.

Original variables				Transformed variables			
i	λ_i	%	Σ%	i	λ_i	%	Σ%
1	5.187	23.6	23.6	1	3.905	17.7	17.7
2	4.855	22.1	45.6	2	3.439	15.6	33.4
3	4.355	19.8	65.4	3	3.354	15.2	48.6
4	2.705	12.3	77.7	4	2.602	11.8	60.5
5	1.227	5.6	83.3	5	2.008	9.1	69.6
6	0.844	3.8	87.1	6	1.604	7.3	76.9
7	0.575	2.6	89.8	7	1.086	4.9	81.8
8	0.401	1.8	91.6	8	0.816	3.7	85.5
9	0.324	1.5	93.1	9	0.671	3.0	88.6
10	0.273	1.2	94.3	10	0.543	2.5	91.0
11	0.239	1.1	95.4	11	0.503	2.3	93.3
12	0.219	1.0	96.4	12	0.475	2.2	95.5
13	0.213	1.0	97.3	13	0.299	1.4	96.8
14	0.200	0.9	98.3	14	0.147	0.7	97.5
15	0.190	0.9	99.1	15	0.136	0.6	98.1
16	0.096	0.4	99.6	16	0.127	0.6	98.7
17	0.043	0.2	99.8	17	0.075	0.3	99.0
18	0.021	0.1	99.8	18	0.070	0.3	99.4
19	0.013	0.1	99.9	19	0.043	0.2	99.6
20	0.009	0.0	99.9	20	0.040	0.2	99.7
21	0.007	0.0	100.0	21	0.029	0.1	99.9
22	0.006	0.0	100.0	22	0.029	0.1	100.0
	22.000	100.0			22.000	100.0	

Sum of variance inflation factors:

$\Sigma\ 1/\lambda_i = 642.09$ $\qquad$ $\Sigma\ 1/\lambda_i = 181.66$

The spectral condition number:

$\lambda_1/\lambda_{22} = 915.50$ $\qquad$ $\lambda_1/\lambda_{22} = 135.66$

Determinant of the correlation matrix:

$|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 9.42407 \cdot 10^{-16}$ $\qquad$ $|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 1.70391 \cdot 10^{-10}$

Haitovskys test variable:

$\chi^2_H(231) = 0.000$ $\qquad$ $\chi^2_H(231) = 0.000$

1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.04	0.03	0.66	0.46	0.32	0.23	0.16	0.11	0.08	0.05	0.04	0.03	0.02
0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.04	0.48	0.66	0.46	0.32	0.23	0.16	0.11	0.08	0.05	0.04	0.03
0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	0.06	0.10	0.48	0.66	0.46	0.32	0.23	0.16	0.11	0.08	0.05	0.04
0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	0.08	-0.23	0.10	0.48	0.66	0.46	0.32	0.23	0.16	0.11	0.08	0.05
0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	0.12	-0.36	-0.23	0.10	0.48	0.66	0.46	0.32	0.23	0.16	0.11	0.08
0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	0.17	-0.28	-0.36	-0.23	0.10	0.48	0.66	0.46	0.32	0.23	0.16	0.11
0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.24	-0.08	-0.28	-0.36	-0.23	0.10	0.48	0.66	0.46	0.32	0.23	0.16
0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.34	0.11	-0.08	-0.28	-0.36	-0.23	0.10	0.48	0.66	0.46	0.32	0.23
0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.49	0.20	0.11	-0.08	-0.28	-0.36	-0.23	0.10	0.48	0.66	0.46	0.32
0.04	0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.70	0.17	0.20	0.11	-0.08	-0.28	-0.36	-0.23	0.10	0.48	0.66	0.46
0.03	0.04	0.06	0.08	0.12	0.17	0.24	0.34	0.49	0.70	1.00	0.06	0.17	0.20	0.11	-0.08	-0.28	-0.36	-0.23	0.10	0.48	0.66
0.66	0.48	0.10	-0.23	-0.36	-0.28	-0.08	0.11	0.20	0.17	0.06	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30	0.25	0.08
0.46	0.66	0.48	0.10	-0.23	-0.36	-0.28	-0.08	0.11	0.20	0.17	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30	0.25
0.32	0.46	0.66	0.48	0.10	-0.23	-0.36	-0.28	-0.08	0.11	0.20	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17	0.30
0.23	0.32	0.46	0.66	0.48	0.10	-0.23	-0.36	-0.28	-0.08	0.11	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11	0.17
0.16	0.23	0.32	0.46	0.66	0.48	0.10	-0.23	-0.36	-0.28	-0.08	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42	-0.11
0.11	0.16	0.23	0.32	0.46	0.66	0.48	0.10	-0.23	-0.36	-0.28	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55	-0.42
0.08	0.11	0.16	0.23	0.32	0.46	0.66	0.48	0.10	-0.23	-0.36	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36	-0.55
0.05	0.08	0.11	0.16	0.23	0.32	0.46	0.66	0.48	0.10	-0.23	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14	-0.36
0.04	0.05	0.08	0.11	0.16	0.23	0.32	0.46	0.66	0.48	0.10	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71	0.14
0.03	0.04	0.05	0.08	0.11	0.16	0.23	0.32	0.46	0.66	0.48	0.25	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00	0.71
0.02	0.03	0.04	0.05	0.08	0.11	0.16	0.23	0.32	0.46	0.66	0.08	0.25	0.30	0.17	-0.11	-0.42	-0.55	-0.36	0.14	0.71	1.00

Correlation matrix for the original variables when $\rho_{a_1 a_2} \approx .98$.

1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	-0.02	0.04	0.67	0.04	0.35	0.01	0.18	-0.01	0.10	-0.01	0.05	-0.01	0.03
-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	-0.02	0.15	0.67	0.04	0.35	0.01	0.18	-0.01	0.10	-0.01	0.05	-0.01
0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	0.08	0.17	0.15	0.67	0.04	0.35	0.01	0.18	-0.01	0.10	-0.01	0.05
-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.04	-0.26	0.17	0.15	0.67	0.04	0.35	0.01	0.18	-0.01	0.10	-0.01
0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	0.15	-0.18	-0.26	0.17	0.15	0.67	0.04	0.35	0.01	0.18	-0.01	0.10
-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.06	-0.24	-0.18	-0.26	0.17	0.15	0.67	0.04	0.35	0.01	0.18	-0.01
0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.28	-0.01	-0.24	-0.18	-0.26	0.17	0.15	0.67	0.04	0.35	0.01	0.18
-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	-0.08	0.05	-0.01	-0.24	-0.18	-0.26	0.17	0.15	0.67	0.04	0.35	0.01
0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.53	0.16	0.05	-0.01	-0.24	-0.18	-0.26	0.17	0.15	0.67	0.04	0.35
-0.02	0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	-0.11	0.09	0.16	0.05	-0.01	-0.24	-0.18	-0.26	0.17	0.15	0.67	0.04
0.04	-0.02	0.08	-0.04	0.15	-0.06	0.28	-0.08	0.53	-0.11	1.00	0.05	0.09	0.16	0.05	-0.01	-0.24	-0.18	-0.26	0.17	0.15	0.67
0.67	0.15	0.17	-0.26	-0.18	-0.24	-0.01	0.05	0.16	0.09	0.05	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28	0.20	0.07
0.04	0.67	0.15	0.17	-0.26	-0.18	-0.24	-0.01	0.05	0.16	0.09	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28	0.20
0.35	0.04	0.67	0.15	0.17	-0.26	-0.18	-0.24	-0.01	0.05	0.16	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15	0.28
0.01	0.35	0.04	0.67	0.15	0.17	-0.26	-0.18	-0.24	-0.01	0.05	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06	0.15
0.18	0.01	0.35	0.04	0.67	0.15	0.17	-0.26	-0.18	-0.24	-0.01	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39	-0.06
-0.01	0.18	0.01	0.35	0.04	0.67	0.15	0.17	-0.26	-0.18	-0.24	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45	-0.39
0.10	-0.01	0.18	0.01	0.35	0.04	0.67	0.15	0.17	-0.26	-0.18	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39	-0.45
-0.01	0.10	-0.01	0.18	0.01	0.35	0.04	0.67	0.15	0.17	-0.26	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16	-0.39
0.05	-0.01	0.10	-0.01	0.18	0.01	0.35	0.04	0.67	0.15	0.17	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51	0.16
-0.01	0.05	-0.01	0.10	-0.01	0.18	0.01	0.35	0.04	0.67	0.15	0.20	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00	0.51
0.03	-0.01	0.05	-0.01	0.10	-0.01	0.18	0.01	0.35	0.04	0.67	0.07	0.20	0.28	0.15	-0.06	-0.39	-0.45	-0.39	0.16	0.51	1.00

Correlation matrix for the transformed variables when $\rho_{a_1 a_2} \approx .98$.

Some measures of multicollinearity for the two theoretical correlation matrices with $\rho_{a_1 a_2} \approx .98$.

Original variables				Transformed variables			
i	λ_i	%	Σ%	i	λ_i	%	Σ%
1	5.367	24.4	24.4	1	4.008	18.2	18.2
2	5.047	22.9	47.3	2	3.547	16.1	34.3
3	4.545	20.7	68.0	3	3.464	15.7	50.1
4	2.799	12.7	80.7	4	2.721	12.4	62.5
5	1.239	5.6	86.3	5	2.104	9.6	72.0
6	0.757	3.4	89.8	6	1.660	7.5	79.6
7	0.569	2.6	92.4	7	1.131	5.1	84.7
8	0.402	1.8	94.2	8	0.844	3.8	88.5
9	0.309	1.4	95.6	9	0.689	3.1	91.7
10	0.255	1.2	96.8	10	0.544	2.5	94.1
11	0.224	1.0	97.8	11	0.516	2.3	96.5
12	0.206	0.9	98.7	12	0.430	2.0	98.4
13	0.194	0.9	99.6	13	0.244	1.1	99.6
14	0.031	0.1	99.7	14	0.021	0.1	99.6
15	0.029	0.1	99.9	15	0.019	0.1	99.7
16	0.014	0.1	99.9	16	0.018	0.1	99.8
17	0.006	0.0	100.0	17	0.011	0.0	99.9
18	0.003	0.0	100.0	18	0.010	0.0	99.9
19	0.002	0.0	100.0	19	0.006	0.0	99.9
20	0.001	0.0	100.0	20	0.006	0.0	100.0
21	0.001	0.0	100.0	21	0.004	0.0	100.0
22	0.001	0.0	100.0	22	0.004	0.0	100.0
	22.000	100.0			22.000	100.0	

Sum of variance inflation factors:

$\Sigma\ 1/\lambda_i = 4310.49$ $\qquad$ $\Sigma\ 1/\lambda_i = 1198.88$

The spectral condition number:

$\lambda_1/\lambda_{22} = 6611.98$ $\qquad$ $\lambda_1/\lambda_{22} = 994.06$

Determinant of the correlation matrix:

$|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 2.17814 \cdot 10^{-23}$ $\qquad$ $|X'X| = \lambda_1 \cdot \ldots \cdot \lambda_{22} = 3.93816 \cdot 10^{-18}$

Haitovskys test variable:

$\chi^2_H(231) = 0.000$ $\qquad$ $\chi^2_H(231) = 0.000$