PSYCHOMETRIKA—VOL. 62, NO. 2, 215–236
JUNE 1997

FITTING ARMA TIME SERIES BY STRUCTURAL EQUATION MODELS

STEF VAN BUUREN

DEPARTMENT OF STATISTICS
TNO PREVENTION AND HEALTH, LEIDEN

This paper outlines how the stationary ARMA(p, q) model can be specified as a structural equations model. Maximum likelihood estimates for the parameters in the ARMA model can be obtained by software for fitting structural equation models. For pure AR and mixed ARMA models, these estimates are approximately unbiased, while the efficiency is as good as those of specialized recursive estimators. The reported standard errors are generally found to be valid. Depending on sample size, estimates for pure MA models are biased 5–10% and considerably less efficient. Some assets of the method are that ARMA model parameters can be estimated when only autocovariances are known, that model constraints can be incorporated, and that the fit between observed and modelled covariances can be tested by statistical methods. The method is applied to problems that involve the evaluation of pregnancy as a function of perceived bodily changes, the effect of policy interventions in crime prevention, and the influence of weather conditions on absence from work.

Key words: lagged variables, Box-Jenkins model, covariance structures, PROC CALIS, intervention analysis, autocorrelation.

1. Introduction

The structural equation model is a general and flexible scheme for specifying linear relations among observed and unobserved variables. The model subsumes many common linear models as a special case. Other names that refer to the same class of models are dynamic simultaneous equations models, reticular action models (RAM), latent variable models, path models, and covariance structure models. Software for fitting structural models include LISREL (Jöreskog & Sörbom, 1985), EQS (Bentler, 1989), COSAN (Fraser & McDonald, 1988) and SAS® PROC CALIS (SAS Institute, 1990).

This paper investigates how the stationary ARMA model of Box and Jenkins (1976) can be formulated and fitted as a structural equations model. The ARMA model provides a parsimonious and elegant way to describe univariate time series. The model has been successfully applied to time series in many branches of science. Writing an ARMA model as a structural model is useful because this opens up the possibility to extend the univariate ARMA model to, for example, ARMA models that incorporate contemporaneous relations, or to constrained ARMA models. A convenient feature of the approach is that solutions can be fitted by widely available software for cross-sectional data. In addition, a practical advantage is that access to the raw data is not needed because parameters can be estimated as long as the right autocovariances are known. Proper understanding of ARMA modeling may aid in generalizing covariance structure models to the case of correlated observations.

Box and Jenkins have shown how the ARMA model produces particular patterns in the covariances among lagged variables. The purpose of this paper is to describe the lagged

This work is dedicated to the memory of my father, Gerrit van Buuren. I thank Erik Meijer for his comments on an earlier version of this paper, and Peter Smulders who contributed data and background knowledge in the weather example. I am grateful to the editor, the associate editor and several anonymous reviewers for their insightful suggestions and thorough review of the manuscript.

Requests for reprints should be sent to Stef van Buuren, Dept. of Statistics, TNO Prevention and Health, P.O. Box 2215, 2301 CE Leiden, THE NETHERLANDS. Email: S.vanBuuren@pg.tno.nl.

0033-3123/97/0600-94180$00.75/0

covariance matrix by a structural model. An early attempt to do this is Cattell's P-technique (Cattell, 1963). Molenaar (1985) improved upon P-technique by considering a latent factor model that can be fitted with LISREL. Other work in this area has been done by Hagnell (1992), who analyzes a multiple economic series by four different lagged variable LISREL models, and Molenaar, de Gooijer, and Schmitz (1992), who formulate a nonstationary dynamic factor model with time dependent factor structure as a LISREL problem. The paper by Wood and Brown (1994) summarizes this line of research and contains in-depth simulations of some of these techniques.

Time series models can also be integrated into structural models via the state space model. See for example the paper by Jones (1991) for a comprehensive overview of the relations between these approaches. More specifically, MacCallum and Ashby (1986) and Oud, van den Bercken, and Essers (1990) have shown how the state space model can be written as a LISREL model. Since ARMA models can be written as state space models, they are indirectly also LISREL models. A disadvantage of this parametrization is that it leads to large, sparse parameter matrices that grow with the number of time points.

Guttman (1954) was the first to realize that the covariance matrix of many psychological tests taken in time forms a simplex, where measures closer in time correlate higher than measures more distant in time. Anderson (1960) pointed out the relation between Guttman's simplex and the autoregressive model. More recent work on correlation patterns include the papers by Mukherjee and Maita (1988), who show how many popular dynamic models in psychology can be translated into structured covariance matrices, and Browne (1992), who extends simplex theory to negative correlations.

For panel data, covariance modelling of time dependent data by autoregressive models is well established (e.g., Cook & Campbell, 1979; Jöreskog, 1978, 1979; McArdle & Aber, 1990). Many modeling issues and path models for panel data are similar to those found in time series analysis, so this literature contains a wealth of useful knowledge. However, the covariance matrix arising from the panel design differs in a fundamental way from the lagged covariance matrix. The covariance matrix of panel data quantifies the variation between sample elements. In contrast, the covariance matrix of time series represents the variation between time points, possibly sampled from one single subject or process only.

The present paper relies on the extensive use of lagged variables, both observed and latent, to capture time dependent aspects in the data. Viewed in this way, time series analysis is just another form of multivariate analysis. A similar approach was taken in van Buuren (1990, 1992). No systematic accounts seems to exist in which the specification of the univariate Box-Jenkins model as a structural equation model is the primary objective. This paper is intended to fill this gap. It concentrates on the univariate case and establishes links with Box-Jenkins theory. On a conceptual level, the use of error variables as latent components leads to very compact models. The ARMA model can be portrayed as a path diagram. Furthermore, the quality of the estimates will be investigated. The method is illustrated by applying it to real time series.

2. Method

Let v_t denote a realization of a time series v at time t. The backshift operator B is defined such that $v_{t-1} = Bv_t$ and that $v_{t-j} = B^j v_t$. If B^j is applied to the observations of all time points $t = 1, \ldots, T$, the resulting scores can be stored as the j-th lagged variable. The covariance matrix of lagged variables is called the lagged covariance matrix. ARMA models possess known theoretical covariance structures. Using lagged variables, software for structural equations can be used to fit such patterns to time series.

Let f_t denote values generated by an independent white noise process with variance σ^2. The ARMA(p, q) model is defined by

$$v_t = \phi_1 v_{t-1} + \cdots + \phi_p v_{t-p} + f_t - \theta_1 f_{t-1} - \cdots - \theta_q f_{t-q}, \tag{1}$$

where the ϕ's and θ's are unknown parameters. This paper is restricted to stationary and invertible ARMA(p, q) processes where all roots of the characteristic equations $1 - \phi_1 B - \cdots - \phi_p B^p = 0$ and $1 - \theta_1 B - \cdots - \theta_q B^q = 0$ lie outside the unit circle.

Suppose that η is an $m \times 1$ random vector of dependent variables and that ξ is an $n \times 1$ random vector of independent variables, then the Bentler-Weeks structural model is defined by (Bentler & Weeks, 1980)

$$\eta = \beta_0 \eta + \gamma \xi, \tag{2}$$

where β_0 ($m \times m$) and γ ($m \times n$) are parameter matrices. Let Φ denote the covariance matrix of ξ, which contains elements that are either free, constrained or fixed. Let G be a known matrix that selects observed variables from η and ξ. The covariance structure of the observed variables under model (2) can be represented as

$$\Sigma = G(I - B_0)^{-1}\Gamma\Phi\Gamma'(I - B_0)'^{-1}G', \tag{3}$$

where B_0 and Γ are specified as

$$B_0 = \begin{pmatrix} \beta_0 & 0 \\ 0 & 0 \end{pmatrix}, \Gamma = \begin{pmatrix} \gamma \\ 0 \end{pmatrix}.$$

See Bentler and Weeks (1980) for exact definitions of these matrices and their precise roles in the structural equations model (2).

Let the $T \times 1$ vector v_0 contain the time series to be analyzed, and that v_p is the p-th lagged variable of v_0. The ARMA(p, 0) can be expressed as a single equation in the Bentler-Weeks model as

$$v_0 = \gamma\xi = (\phi_1, \ldots, \phi_p, 1)\begin{pmatrix} v_1 \\ \vdots \\ v_p \\ f_0 \end{pmatrix}. \tag{4}$$

The variance-covariance matrix of ξ in the model is

$$\Phi = \begin{pmatrix} \tau_0 & \tau_1 & \cdots & \tau_{p-1} & 0 \\ \tau_1 & \tau_0 & \cdots & \tau_{p-2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tau_{p-1} & \tau_{p-2} & \cdots & \tau_0 & 0 \\ 0 & 0 & \cdots & 0 & \sigma^2 \end{pmatrix}, \tag{5}$$

where τ_i denotes the i-th order autocovariance in the ARMA(p, 0) model and where σ^2 is the variance of the noise process f_0. The τ_i-parameters are not free and depend on γ. This relation is nonlinear and cannot be easily incorporated into the structural model. Therefore τ_i is estimated directly from the raw data. One of the reviewers remarked that this might cause loss of efficiency in analyzing real data as opposed to simulated data. For $p = 1$, by substituting (4) and (5) into (3) yields $\tau_0 = \sigma^2/(1 - \phi_1^2)$, which is identical to the known theoretical variance of the AR(1)-model. Because Box and Jenkins express variances of higher order AR-models in terms of infinite MA-models, the variance structure of (3) for $p > 1$ cannot be directly related to their results.

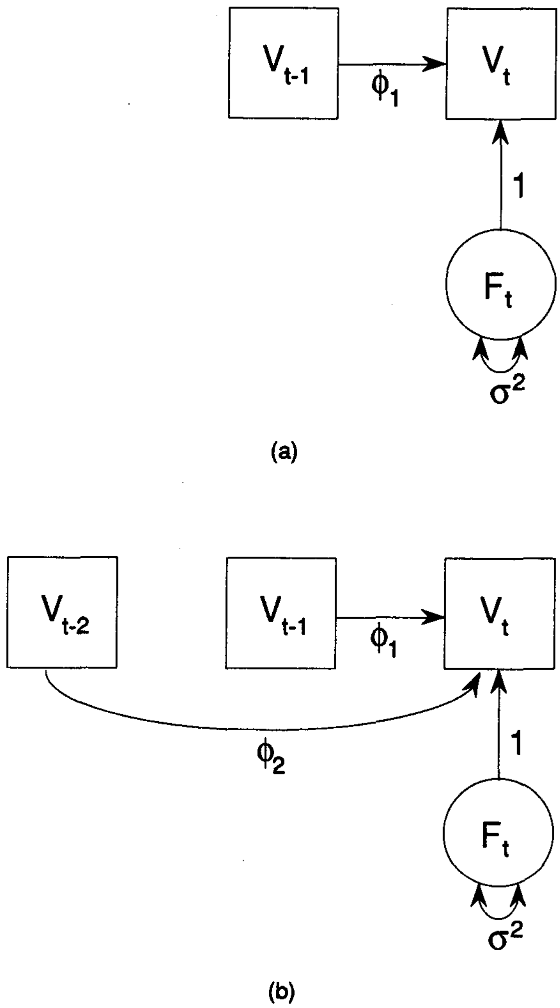

FIGURE 1.
Path diagrams of the ARMA(1, 0) model (a) and the ARMA(2, 0) model (b).

Figure 1 depicts the path diagrams for $p = 1$ and $p = 2$ in RAM nomography (see McArdle & Aber, 1990, pp. 165–167). Slings (two-head arrows) associated with the variances and covariances of manifest exogenous variables are not drawn. These (co)variances are automatically fixed to their sample values and consequently not very interesting. Since the corresponding slings divert attention from the core model, they are omitted. One could consider drawing an arrow in the AR(2) model between v_{t-2} and v_{t-1}, restrict the corresponding path coefficient $\phi_1(v_{t-2}, v_{t-1})$ to $\phi_1(v_{t-1}, v_t)$, and add an error term to v_{t-1}. Apart from introducing extra complexities into the model, this will lead to the wrong answer because $\phi_1(v_{t-2}, v_{t-1})$ is estimated without taking the influence of the second predictor, in this case v_{t-3}, into account. The equality constraint will somehow pool $\phi_1(v_{t-2}, v_{t-1})$ and $\phi_1(v_{t-1}, v_t)$, which produces bias in ϕ_1.

The pure moving averages ARMA(0, q) model can be specified as a structural equation model as follows:

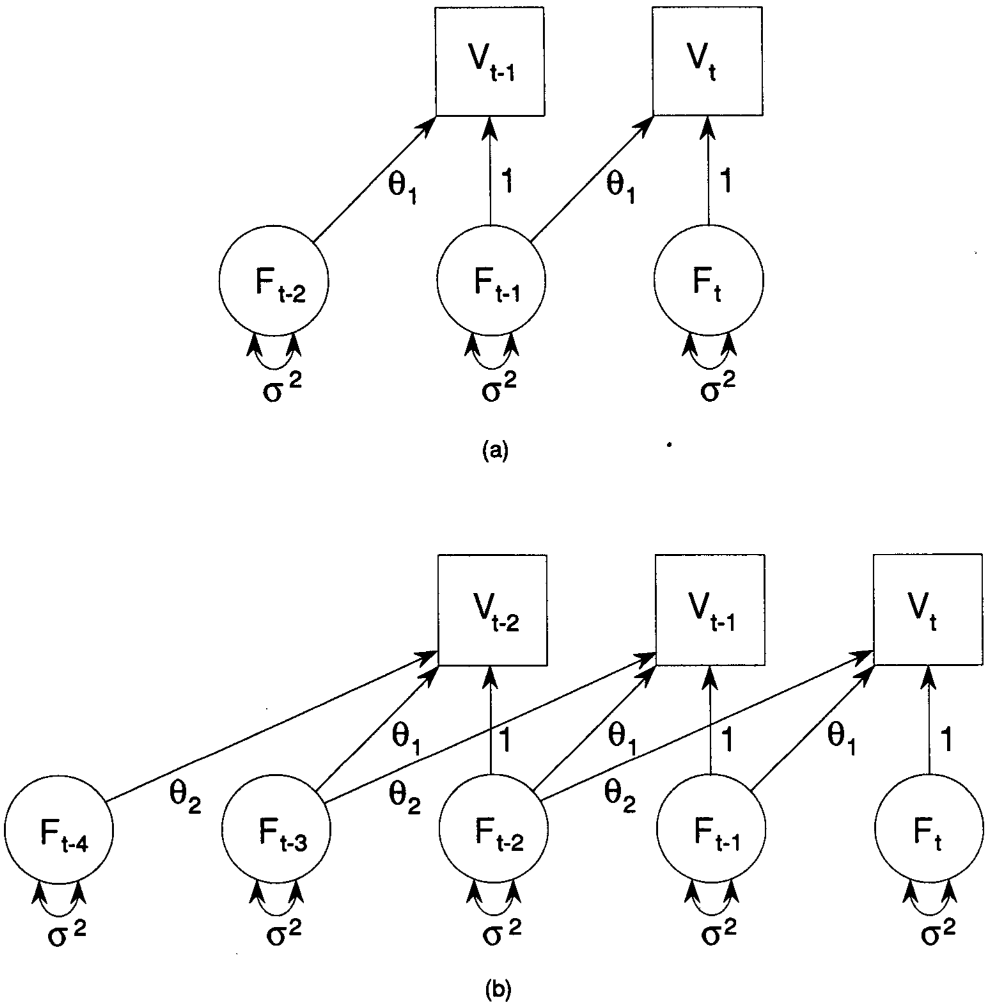

FIGURE 2.
Path diagrams of the ARMA(0, 1) model (a) and the ARMA(0, 2) model (b).

$$\begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_q \end{pmatrix} = \gamma\xi = \begin{pmatrix} 1 & -\theta_1 & \cdots & -\theta_q & & & & \\ & 1 & -\theta_1 & \cdots & -\theta_q & & & \\ & & \ddots & \ddots & & \ddots & & \\ & & & 1 & -\theta_1 & \cdots & -\theta_q \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{2q} \end{pmatrix}, \qquad (6)$$

a form that also appears in dynamic factor analysis (Molenaar, 1985). The variance-covariance matrix of ξ is equal to $\Phi = \mathrm{diag}(\sigma^2, \ldots, \sigma^2)$ of order $2q \times 2q$. The effect of requiring orthogonality of the f's in this model is that these become serially uncorrelated processes (Geweke & Singleton, 1981). The variances of $f_0, \ldots, f_{2q}$ are unknown, but since they represent different lags of the same random process f_0, their estimates are constrained to be equal.

Figure 2 contains path diagrams for $q = 1$ and $q = 2$. Note that the same basic MA(q)-model is repeated $q + 1$ times, thus yielding a chained equation structure. Because it is impossible to estimate all parameters from the basic model only, lagging the entire model is necessary to increase the degrees of freedom. It is known that the ARMA(0, q) process only up to q lags of v_0 have nonzero autocovariances, so properly estimating its parameters requires that $q + 1$ lagged models are to be chained. The theoretical variance of v_0 by (3) is equal to $\tau_0 = \sigma^2(1 + \theta_1^2 + \cdots + \theta_q^2)$, which is identical to the result given in Box and Jenkins.

Provided that $q < p$, the general ARMA(p, q) model with p, $q > 0$ can be written as a special case in the Bentler-Weeks framework as $\eta = \beta_0\eta + \gamma\xi$ where

$$\eta = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_q \end{pmatrix}, \quad \xi = \begin{pmatrix} v_{q+1} \\ \vdots \\ v_{p+q} \\ f_0 \\ \vdots \\ f_{2q} \end{pmatrix}, \quad \beta_0 = \begin{pmatrix} 0 & \phi_1 & \cdots & \phi_{q-1} & \phi_q \\ & 0 & \phi_1 & \cdots & \phi_{q-1} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & \phi_1 \\ & & & & 0 \end{pmatrix} \tag{7}$$

$$\gamma = \begin{pmatrix} \phi_{q+1} & \cdots & \phi_p & 0 & \cdots & 0 & 1 & -\theta_1 & \cdots & -\theta_q & & & & \\ \phi_q & \cdots & \cdots & \phi_p & 0 & \cdots & 0 & 1 & -\theta_1 & \cdots & -\theta_q & & & \\ \vdots & \ddots & & & \ddots & \ddots & & \ddots & \ddots & \ddots & & \ddots & & \\ \phi_1 & \cdots & \phi_q & \cdots & \cdots & \phi_p & 0 & \cdots & 0 & 1 & -\theta_1 & \cdots & -\theta_q & \end{pmatrix}.$$

Note that in addition to v_0, variables $v_1, \ldots, v_q$ are endogenous variables, so they should be arranged in the η-component of the Bentler-Weeks model. Writing the model for the case $q \geq p$ is quite simple, but it differs only marginally from (7) in the division of v's over η and ξ which is of little further theoretical interest so this equation is not given.

The upper-left block of Φ contains p process autocovariances, while the lower-right block holds the variances of the $2q + 1$ latent shock factors $f_0, \ldots, f_{2q}$. The matrix is equal to

$$\Phi = \begin{pmatrix} \tau_0 & \cdots & \tau_{p-1} & & & & \nu_{1,1} & \cdots & \nu_{q,1} \\ \vdots & & \vdots & & & & & \ddots & \vdots \\ \tau_{p-1} & \cdots & \tau_0 & & & & & & \nu_{q,s} \\ & & & \sigma^2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \sigma^2 & & & \\ \nu_{1,1} & & & & & & \sigma^2 & & \\ \vdots & \ddots & & & & & & \ddots & \\ \nu_{q,1} & \cdots & \nu_{q,s} & & & & & & \sigma^2 \end{pmatrix}. \tag{8}$$

In some models, v's and the f's are correlated by definition, so their covariances should be included into the model. The parameters are represented by $\nu_{1,1}, \ldots, \nu_{q,s}$, with $s = \min(p, q)$. The precise inclusion rule is as follows: Suppose that j is bounded by $q + 1 \leq j \leq 2q$ and that i is bounded by $q + 1 \leq i \leq \min(p + q, j)$, then $\sigma(v_i, f_j)$, the covariance between v_i and f_j, should be included as a free parameter. For example, in an ARMA(1, 1) or an ARMA(2, 1) model $\nu_{1,1} = \sigma(v_2, f_2)$ must be free, in an ARMA(1, 2) model $\nu_{1,1} = \sigma(v_3, f_3)$ and $\nu(2, 1) = \sigma(v_3, f_4)$ are free, and in an ARMA(2, 2) model $\nu_{1,1} = \sigma(v_3, f_3)$, $\nu_{2,1} = \sigma(v_3, f_4)$ and $\nu_{2,2} = \sigma(v_4, f_4)$ are free parameters. It is also possible that the ν-block is empty.

Figure 3a contains the ARMA(1, 1) path model. Note the inclusion of the bidirec-

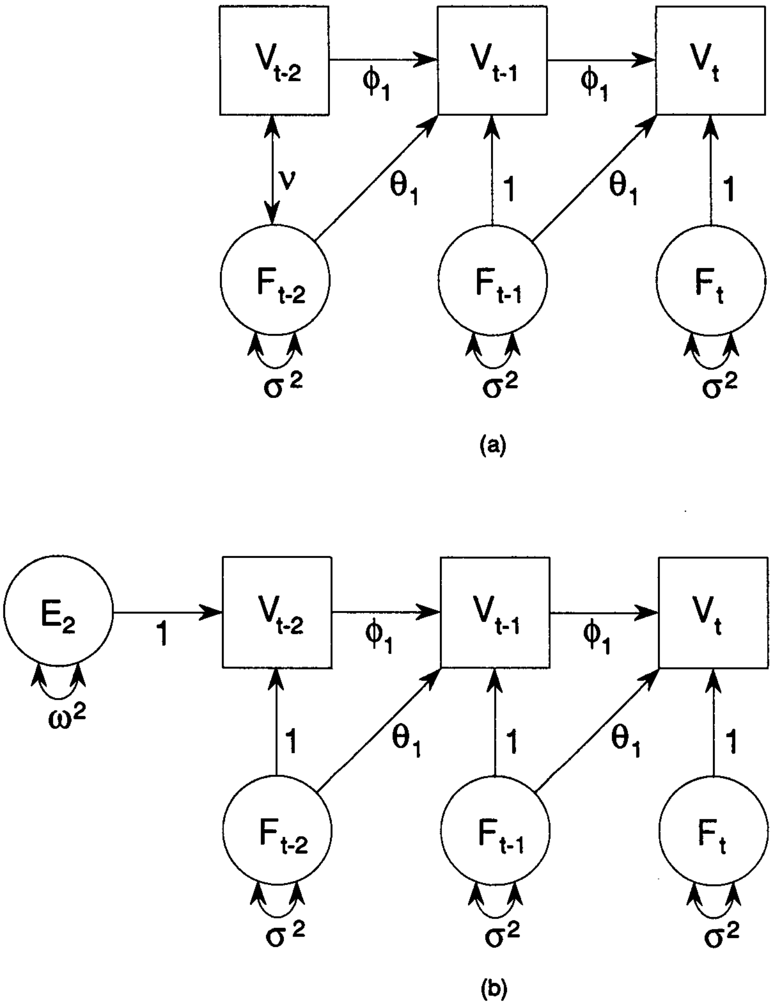

FIGURE 3.
Path diagrams of the ARMA(1, 1) model. Covariance method (a), Noise method (b).

tional arrow to account for $\sigma(v_2, f_2)$. One of the reviewers suggested to repeat the pattern of moving average coefficients and add an error term to each v_i that has less than $q + 1$ arrows pointing at it. This has the advantage of giving extra degrees of freedom. This is called the 'noise method', while model (8) is identified as the "covariance method". Figure 3b illustrates the idea.

In their general forms, (7) and (8) may appear complicated. However, in most popular cases the actual specifications in terms of program code are quite compact. The appendix contains PROC CALIS code for fitting all ARMA models with $p + q \leq 2$. Note that this code relies on the fact that PROC CALIS automatically sets the predicted (co)variances of the exogenous manifest variables equal to the sample quantities and adjusts the degrees of freedom accordingly. Translation of the code into other software packages typically requires the addition of explicit modelling statements with respect to these variances.

3. Simulation

If the ARMA model can written as a structural model it can also be fitted to time series by LISREL, EQS, COSAN, PROC CALIS or other general purpose structural equations software. It is interesting to study how well the resulting estimates compare to those derived by specialized recursive estimation methods that have been developed for ARMA models. This section contains the results of a small simulation study investigating a variety of models.

The SAS/IML® function ARMASIM was used to generate 100 simulated time series of $T = 500$ time points, each for a known ARMA(p, q) model. Parameters are estimated by PROC CALIS (SAS, 1990) and by SPSS® ARIMA (SPSS, 1990), the latter of which implements Mélard's method (Mélard, 1984), a fast and specialized recursive estimation method for ML estimation of ARMA parameters. This analysis was repeated under 10 different ARMA models. The difference between the replication mean and the a priori value is the approximate bias of the estimator. The standard deviation around the mean is a measure of efficiency. PROC CALIS estimates were obtained by the method of maximum likelihood. After some experimentation the Levenberg-Marquart optimization method turned out to be best suited for this problem.

Table 1 contains the results of the simulation. Ten different ARMA models are investigated: two AR(1), two AR(2), two MA(1), two MA(2), and two mixed ARMA(1, 1) models. The order of the models is identified in the columns labeled p and q respectively. A column labeled "PAR" identifies the unknown parameters under each model, a column labeled "POP" contains a priori values. Next, the average ($\bar{x}$) and its standard deviation ($\sigma(\bar{x})$) are given for both CALIS and ARIMA.

Table 1 shows that for pure autoregressive and for mixed models there are not many differences in terms of bias and efficiency between both methods, that is, both CALIS and ARIMA yield virtually unbiased estimates of comparable precision. For pure MA models, ARIMA is unbiased and precise, but the CALIS estimates are slightly off-target and deviate up to 5% from the true value. This is especially visible in the MA(2) models. As outlined in section 2, two structural representations can be used for the ARMA(1, 1)-model. The table reports the results for the noise method. The covariance method produces almost identical estimates, but minimization by the noise method generally proceeds somewhat smoother and is more reliable.

Especially for pure MA-models, convergence is sometimes problematic. The worst case is the model with $\theta_1 = 1.20$ and $\theta_2 = -0.80$. In this case, mild and severe convergence problems show up in about 10 to 20 percent of the runs. Typical symptoms are boundary estimates (mild), negative variance estimates (severe) and huge standard errors (severe). MA-models appear to be rather difficult to estimate anyway. Even Mélard's algorithm goes astray quite a few times in both MA(1) and MA(2) problems. In practice, one could try alleviate such problems by specifying different starting values or by selecting alternative minimization procedures, but it will not cure the patient. Here is clearly room for improvement.

A series length of 500 may give an idea of the asymptotic behavior of the estimator, but does not often occur in practice. Therefore the simulations were repeated with a more realistic value of $T = 50$. The results are given in Table 2. Of course, standard errors increase, but the overall pattern is similar to that in Table 1. So, for $T = 50$ estimates are on the average quite reasonable, though not as good as Mélard's method.

The means of the parameter estimates do not tell whether one could reasonably expect both methods to give identical result for a given set of data. The last column in Table 2 contains the correlations between CALIS and ARIMA estimates, as computed from the same data. The results for pure AR models are encouraging. Correlations vary between 0.77 and 0.99. However, correlations for the pure MA models are practically zero. This means that, though the estimates are on the average correct, the values for a particular data set may be very different. Thus, CALIS estimates for pure MA models are unreliable. For mixed models, the situation is much more favorable.

TABLE 1

Simulation results (100 replications) of fitting 10 ARMA(p,q) models by SAS PROC CALIS and by SPSS[x] ARIMA (T=500). For each method, averaged estimates and standard error of this average are given.

				CALIS		ARIMA	
p	q	PAR	POP	$\bar{x}$	$\sigma(\bar{x})$	$\bar{x}$	$\sigma(\bar{x})$
1	0	ϕ_1	.80	.80	.03	.80	.03
		σ^2	1.00	1.00	.06	1.00	.06
		ϕ_1	-.80	-.80	.02	-.79	.02
		σ^2	1.00	1.00	.06	1.00	.06
2	0	ϕ_1	.80	.80	.04	.80	.04
		ϕ_2	-.20	-.20	.05	-.20	.05
		σ^2	1.00	1.00	.06	1.00	.06
		ϕ_1	1.20	1.19	.02	1.19	.02
		ϕ_2	-.80	-.79	.02	-.79	.02
		σ^2	1.00	1.01	.07	.99	.06
0	1	θ_1	.80	.82	.15	.80	.03
		σ^2	1.00	.98	.14	1.00	.07
		θ_1	-.80	-.80	.16	-.80	.03
		σ^2	1.00	.99	.15	1.00	.07
0	2	θ_1	.80	.83	.10	.80	.05
		θ_2	-.20	-.20	.10	-.21	.05
		σ^2	1.00	.96	.09	1.00	.06
		θ_1	1.20	1.24	.14	1.20	.03
		θ_2	-.80	-.76	.19	-.80	.03
		σ^2	1.00	.97	.14	1.00	.06
1	1	ϕ_1	.80	.79	.04	.79	.04
		θ_1	.20	.19	.07	.19	.07
		σ^2	1.00	1.00	.07	1.00	.07
		ϕ_1	.80	.79	.03	.79	.03
		θ_1	-.20	-.20	.05	-.20	.05
		σ^2	1.00	.99	.07	.99	.07

TABLE 2

Simulation results (100 replications) of fitting 10 ARMA(p,q) models by SAS PROC CALIS and by SPSS^x ARIMA (T=50). For each method, averaged estimates, se of this average, and averaged estimates of the standard errors are given ('na' means 'not available'). The column labeled 'r' gives the correlation between CALIS and ARIMA estimates.

				CALIS			ARIMA			
p	q	PAR	POP	$\bar{x}$	$\sigma(\bar{x})$	$\overline{\sigma(x)}$	$\bar{x}$	$\sigma(\bar{x})$	$\overline{\sigma(x)}$	r
1	0	ϕ_1	.80	.76	.08	.10	.79	.07	.09	.77
		σ^2	1.00	1.03	.22	.20	1.02	.22	na	.97
		ϕ_1	-.80	-.77	.09	.09	-.76	.09	.09	.97
		σ^2	1.00	1.03	.21	.20	1.02	.21	na	.97
2	0	ϕ_1	.80	.79	.13	.14	.80	.12	.14	.97
		ϕ_2	-.20	-.23	.13	.14	-.22	.13	.14	.98
		σ^2	1.00	.99	.22	.20	1.02	.22	na	.99
		ϕ_1	1.20	1.17	.09	.09	1.19	.08	.09	.87
		ϕ_2	-.80	-.77	.09	.09	-.78	.09	.09	.89
		σ^2	1.00	1.10	.22	.27	1.02	.22	na	.82
0	1	θ_1	.80	.73	.24	.32	.81	.10	.11	.00
		σ^2	1.00	1.07	.30	.33	1.02	.22	na	.70
		θ_1	-.80	-.71	.25	.28	-.83	.09	.09	-.07
		σ^2	1.00	1.05	.30	.30	1.02	.21	na	.69
0	2	θ_1	.80	.87	.21	.25	.80	.13	.14	.66
		θ_2	-.20	-.22	.22	.29	-.22	.18	.14	.58
		σ^2	1.00	.89	.22	.24	1.02	.22	na	.81
		θ_1	1.20	1.34	.26	9.11	1.21	.11	.11	.21
		θ_2	-.80	-.79	.26	12.94	-.83	.12	.11	.16
		σ^2	1.00	.85	.23	12.37	1.00	.22	na	.59
1	1	ϕ_1	.80	.70	.18	.16	.75	.15	.13	.78
		θ_1	.20	.12	.21	.23	.15	.22	.20	.85
		σ^2	1.00	.99	.21	.14	1.02	.22	na	.97
		ϕ_1	.80	.71	.14	.10	.77	.10	.10	.68
		θ_1	-.20	-.29	.21	.23	-.24	.18	.16	.62
		σ^2	1.00	1.00	.23	.17	1.03	.22	na	.89

Time series observations are not independent. An important question therefore is whether the standard errors that are reported by CALIS can be used for significance testing. The answer is yes. Two extra columns, labeled $\overline{\sigma(x)}$, in Table 2 contain the standard errors as reported by CALIS and ARIMA, averaged over the replications. Actually, the median instead of the mean is reported since this measure is more robust to occasional outliers. The reported standard errors are OK if $\overline{\sigma(x)}$ is equal to $\sigma(\bar{x})$. If $\overline{\sigma(x)} < \sigma(\bar{x})$ the reported standard errors underestimate the true standard error, and therefore give p-values that are too small. This happens only slightly in the ARMA(1,1) model, for both CALIS and ARIMA, and sometimes for the σ^2 estimates in other models. The latter parameters are hardly ever tested though. Apart from the troublesome MA(2) problem, estimates are generally quite close, and if they deviate then $\overline{\sigma(x)} > \sigma(\bar{x})$. Thus if CALIS standard errors are used for significance testing or for constructing confidence intervals, results are reasonably accurate. Also, if bias occurs it is likely to be into the conservative direction.

In summary, CALIS estimates are approximately asymptotically unbiased but in some cases more efficient estimators exist. Standard errors reported by CALIS can be used for parameter testing or for constructing confidence intervals. Convergence problems may show up, especially in pure MA models. CALIS and ARIMA estimates can be quite dissimilar for pure MA models. Therefore, CALIS should not be used for fitting these models. Except for pure MA models, the result lend credit to the notion that fitting ARMA models within the Bentler-Weeks framework produces results that will not be too far off the mark.

4. Examples

Bivariate Autoregressive Model with Contemporaneous Relations

This example illustrates (a) how structural modeling helps to formulate multivariate ARMA models that incorporate both simultaneous and lagged relations, and (b) how constraints can be used to test for Granger causality.

Schmitz (1990) has written a fine introduction into many aspects of multivariate time series analysis. To illustrate his methods, Schmitz analyzes the relationship between two time series from the field of developmental psychology, the evaluation of pregnancy (p) and the rating of bodily changes (b), by a multivariate ARMA model. The basic data consist of two daily measurements on 82 points of time and were obtained by reading off the scores from Schmitz's figure. The correlation between p and b is 0.63.

Using the auto-correlation function (ACF) and the partial autocorrelation function (PACF), Schmitz identified the bivariate AR(1) model

$$\begin{pmatrix} p \\ b \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} p_{t-1} \\ b_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_p \\ \varepsilon_b \end{pmatrix}, \tag{9}$$

where the ϕ's are unknown parameters, and where ε_p and ε_b are white noise processes with variances σ_p^2 and σ_b^2 and with covariance σ_{pb}. A column in Table 3 labeled "A1" contains the PROC CALIS estimates for this model. The estimates differ from those given in Schmitz (1990) as Schmitz reports standardized coefficients while Table 3 gives the raw estimates.

Since ϕ_{21} is only just significantly different from zero at $\alpha = 0.05$, the relation between p_{t-1} and b_t is weak. If the relation actually exist, then it is an instance of Granger causality (Granger, 1969). Loosely speaking, a predictor is Granger causal if it precedes the dependent variable in time and if it has an independent contribution to the prediction. Granger causality can be formally tested by restricting the corresponding coefficient(s) to zero, and

TABLE 3

Maximum likelihood estimates for six models applied to a bivariate time series from Schmitz (1990) (standard errors are in parentheses).

Parameter	A1	A2	B1	B2	C1	C2
β_b	-	-	.47(.11)	.50(.10)	-	-
ϕ_{11}	.43(.11)	.35(.10)	.35(.10)	.38(.09)	.43(.11)	.60(.09)
ϕ_{12}	.30(.12)	.36(.12)	.08(.13)	-	.30(.13)	-
β_p	-	-	-	-	.37(.09)	.38(.08)
ϕ_{11}	.18(.10)	-	.18(.10)	-	.02(.10)	-
ϕ_{12}	.48(.11)	.61(.09)	.48(.11)	.61(.09)	.37(.11)	.38(.09)
σ_p^2	.13(.02)	.13(.02)	.11(.02)	.11(.02)	.13(.02)	.14(.02)
σ_b^2	.10(.02)	.11(.02)	.10(.02)	.11(.02)	.09(.01)	.09(.01)
σ_{pb}	.04(.01)	.05(.01)	-	-	-	-
ρ_p^2	.41	.41	.51	.51	.41	.36
ρ_b^2	.40	.37	.40	.38	.50	.50
DW_p	2.21	2.21	2.15	2.22	2.21	2.32
DW_b	2.07	2.07	2.07	2.07	2.00	2.02
χ^2	0	3.15	0	3.52	0	5.62
d.f	0	1	0	2	0	2
p	-	.08	-	.17	-	.06

test the constrained against the unconstrained model. Model A2 is a constrained version of A1 in which $\phi_{21} = 0$. The difference in χ^2-statistic, which summarizes the difference between the observed and predicted lagged covariance matrices, is $\chi^2_\Delta = 3.15$ with d.f. = 1 ($p = 0.08$). Models A1 and A2 are not significantly different at $\alpha = 0.05$, implying that the contribution of p_{t-1} in predicting b_t is nonsignificant so the relation between p_{t-1} to b_t is not Granger causal. Figure 4a gives the path diagram of Model A2.

Note that the Durbin-Watson (DW) statistic for A1 and A2 is close to 2, indicating that the residuals have an insignificant first-order autocorrelation. Note also that models A1 and A2 allow for correlated residuals. In fact, the residual correlation $r(\varepsilon_p, \varepsilon_b) = \sigma_{pb}/\sigma_p\sigma_b = 0.42$ is substantial. So, while the structural part of model A2 explains the major time dependent relationships in the data, there is still a lot of contemporaneous information left unaccounted for in the residuals.

It is attractive to develop a model that not only explains serial dependency, but also contemporaneous relationships, if any. To do so, an additional contemporaneous model is fitted, just like in a conventional covariance structure model. Model B1 extends (9) by including b_t as a predictor for p_t so that $p_t = \beta_b b_t + \phi_{11}p_{t-1} + \phi_{12}b_{t-1} + \varepsilon_p$ and by restricting $\sigma_{pb} = 0$. Since ϕ_{12} and ϕ_{21} are no longer significant, a second, constrained solution (B2) was computed in which these coefficients were set to zero. Figure 4b contains the corresponding path diagram. There are some remarkable differences between A2 and

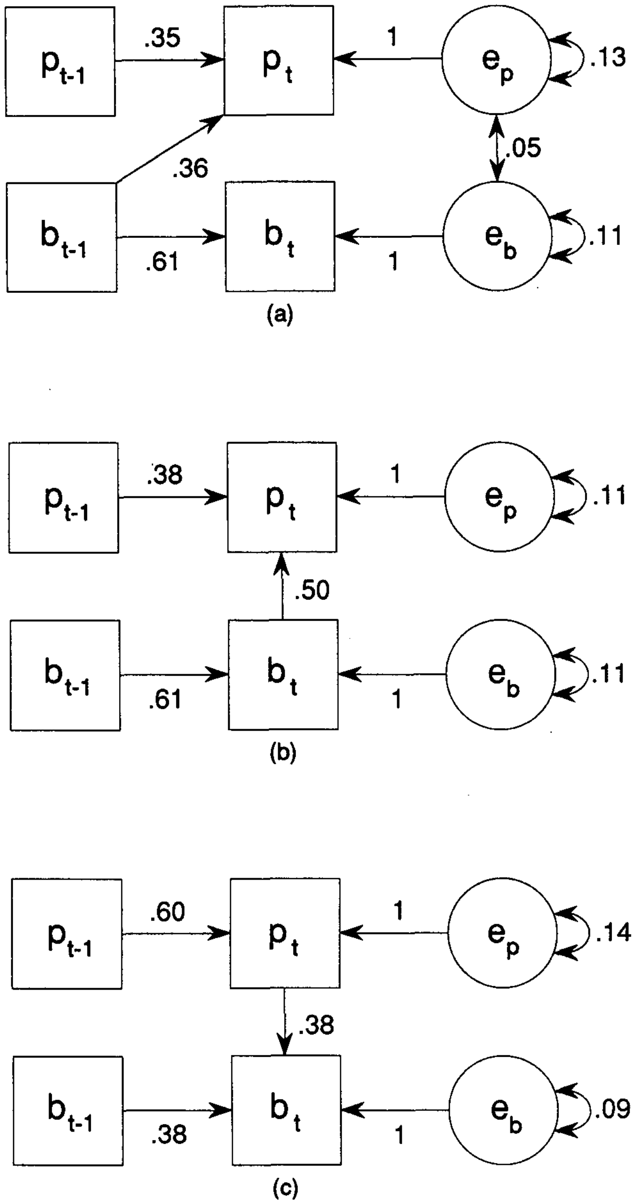

FIGURE 4.
Path diagrams for three different models applied to a bivariate time series: evaluation of pregnancy (p) and rating of bodily changes (b) (Source: Schmitz, 1990). Model (a) is the multivariate ARMA model identified by Schmitz. Models (b) and (c) allow for contemporaneous relations between p and b.

B2. The most important one is that model A2 only explains time dependencies, while model B2 takes account of all time- and cross-relationships. Also, B2 fits the data better, uses less parameters, and has larger proportions of explained variance (compare the squared multiple correlations ρ_p^2 and ρ_b^2).

Models C1 and C2 "reverse the arrow" between p_t and b_t. Models C2 also fits the observed cross-autocovariance matrix, but not as well as B2. The Durbin-Watson statistic for this model increases to 2.32, which falls into the inconclusive region of the Savin-White tables, so C2 may not adequately deal with time dependencies.

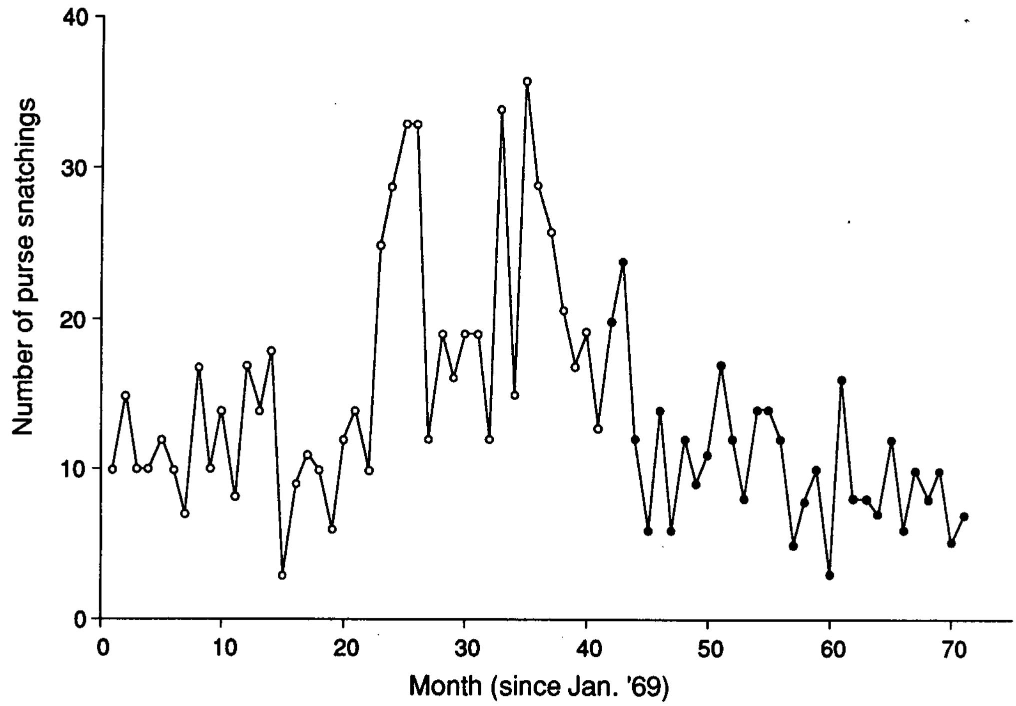

FIGURE 5.
Hyde Park Purse Snatchings Series. The prevention plan starts at time point 42.

Summarizing, PROC CALIS was used to mix cross-sectional and time series models. Also, constraints were applied where this was necessary or useful. Based on the data alone, the preferred model is B2.

Intervention Analysis

A common goal in time series analysis is to determine whether some external event influences the level or shape of a series. Typical examples come from $N = 1$ research designs in clinical psychology. Here the question is whether a therapeutic treatment has an effect on the client. The main worry in intervention analysis is serial dependency of the series. Serial dependency invalidates the t-test for testing the difference between the pre-intervention mean and the post-intervention mean. As a solution, one may fit an ARMA model to the series, derive the residual, and apply the t-test to it. The classic reference for such procedures in the behavioral sciences is Glass, Willson, and Gottman (1975).

The Hyde Park Purse Snatching series (MacCleary & Hay, 1980) consists of 71 counts of purse snatchings in Hyde Park, Chicago during the period of January 1969 to September 1973. The series is plotted in Figure 5. At time point 42 Operation Whistlestop started, a community crime prevention program. Amongst others, the project involved distributing whistles to citizens which could be used to alarm the police. Figure 5 shows that after the intervention point the number of purse snatchings decreases, so the intervention seems to have a desirable effect. Indeed, the t-test is statistically significant with $p < 0.001$. The significance of this five-star result is debatable however, since it is hard to maintain that the observations are independent. The first 10 autocorrelation are .50, .54, .37, .30, .27, .16, .25,

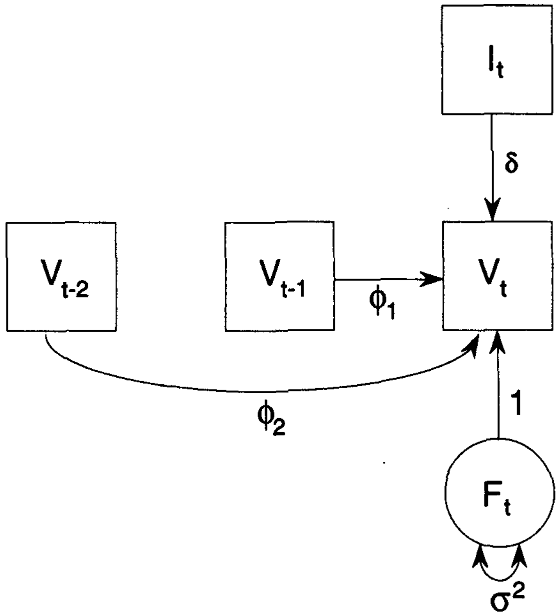

FIGURE 6.
Path diagram of an intervention model of abrupt, constant change under an ARMA(2, 0) model.

.19, .19 and .26, which is typical for an autoregressive process. In fact, MacCleary and Hay identified an ARMA(2, 0) model for the series.

PROC CALIS was used to estimate an intervention effect of abrupt, constant change under three autoregressive noise models with respectively $p = 0, 1, 2$. This was done by dummy coding the intervention variable I and estimate its effect on the mean of the series. The associated intervention parameter δ can then be tested for significance. The path diagram for the model $p = 2$ is given in Figure 6. Table 4 contains the parameter estimates with their standard errors and the values of the Durbin-Watson and the Ljung-Box-Pierce (LBP) statistics.

The first row of Table 4 is the analysis without the noise model. After the intervention about six purse snatchings less per month are counted than before. However, both DW and LBP-statistics indicate that the data contain considerable autocorrelation, so the estimated standard error for δ is suspect. When the first and second lags of the series are included, the LBP-statistic drops dramatically, and the DW-statistic approaches the neutral value of 2. But simultaneously δ declines. So the more lags are included, the less effect we see. In this example, controlling for autocorrelation by the ARMA(2,0) model annihilates the intervention effect. Using other estimation methods, similar results were found by MacCleary and Hay (1980) and van Buuren (1990).

Optimal Prediction Lags By Sliding

Smulders and van Deursen (1995) study the influence of the weather on absenteeism from work. Thirteen daily weather measurements during the year 1990 were obtained from the Royal Netherlands Meteorological Institute (KNMI). Two measures of absence are used: absence incidence, which is equal to the percentage of new spells on a given day, and absence prevalence, which is the total absence percentage for a given day. Absence data were obtained from personnel records of ten organizations that were located within a distance of 40 kilometers from the weather station.

TABLE 4

Intervention estimates (δ) and ϕ-estimates under three noise models for the Hyde Park Purse Snatching series (T=71) (DW = Durbin-Watson statistic, LBP = Ljung-Box-Pierce statistic (25 df)).

Noise Model	δ(se)	ϕ_1 (se)	ϕ_2 (se)	DW	p	LBP	p
None	-6.2(1.69)			1.22	<.01	85.2	.00
ARMA(1,0)	-3.7(1.69)	.40(.11)		2.30	<.10	35.5	.08
ARMA(2,0)	-2.7(1.60)	.26(.11)	.36(.11)	2.00	>.10	24.1	.55

Using linear regression with lagged predictors, Smulders and van Deursen were able to show slight effects of weather on both prevalence and incidence. The analysis below uses correction terms like weekly patterns and holiday periods that are modeled as an integral part of the regression model, which enhances the efficiency of the analysis. A structural time series model is fitted that expresses today's prevalence as a function of yesterday's prevalence and today's incidence. This model prewhitens the absence data so that any cross-lag effects are not determined by the autocorrelation in the individual absence series. A latent weather variable is constructed as a linear combination of observed weather indicators. The effect of weather on absence prevalence can be either direct or being routed via the incidence. To test for the optimal lag, the latent weather variable is 'slided along' the absence model, that is, tried at different lags. By comparing the estimates at different lags, this procedure suggests the optimal delay at which the weather is effective.

Absence incidence turns out to be related to (a) day of the week (coded as six dummy variables using Sunday as the reference), (b) whether or not it is a holiday period, and (c) yesterday's incidence. Absence prevalence largely depends on (i) day of the week, (ii) yesterday's prevalence, and (iii) today's incidence. Using linear regression, no noticeable autocorrelation remained in either of the residuals (the DW-statistics are close to 2). For example, the first-order autocorrelation of the residual of prevalence drops from 0.94 (if only a constant term were included) to 0.02 for the autoregressive model given above. Both regression models are combined into one path model, which now contains two coupled autoregressive models, each of order one. The modification indices for this model suggested that a separate correction for Saturday was unnecessary, so Saturdays and Sundays will be further treated as equivalent. Figure 7 contains the resulting diagram.

The model of Figure 7 fits the lagged-autocovariance model very well (e.g. Goodness of Fit Index (Adjusted) = 0.96, Bentler & Bonnett's NFI = 0.99, Bollen's $\Delta^2 = 0.99$).

As a next step, weather measures are added to the absence model. Smulders and van Deursen used factor analysis to identify four primary weather indicators: mean daily temperature, mean wind velocity, hours of precipitation and percentage of sunshine. For each indicator, six lagged variables were constructed. The total window size of the analyses is arbitrarily chosen to cover one week.

CALIS was applied to make a optimal linear combination of all four indicators at every lag d for $d = 0, \ldots, 6$. Figure 8 contains that the path diagram of this analysis. The diagram represents a range of path models, each of which corresponds to a different delay. To avoid clutter, errors e_1 and e_2 are deleted from the picture. The most interesting part

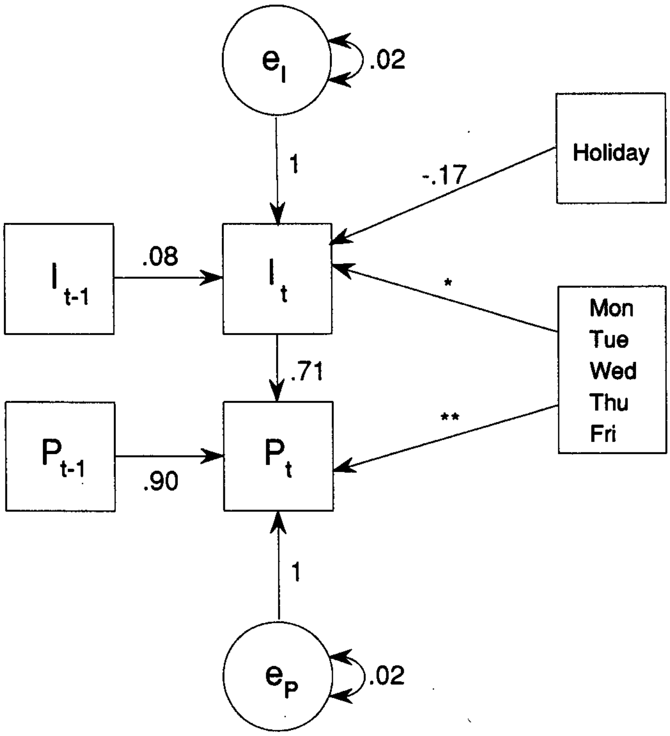

FIGURE 7.
Model for incidence (I) and prevalence (P) of employee absenteeism, including a 'day of week' correction (* = 1.06, 0.50, 0.49, 0.46, 0.33; ** = −1.12, −0.28, −0.24, −0.23, −0.11 for Mon, Tue, Wed, Thu, Fri respectively).

of the model is to the influence of weather on absence as measures by α_d and β_d. According to the model, the effect of weather on prevalence can be either direct or through the incidence.

Figure 9 contains estimates of α_d and β_d for different values of d. Some of these analyses required as much as 150 maximum likelihood iterations. The sample size ($n = 365 - d$) is sufficiently large to compare these values to standard normal probabilities. It appears that the largest effect of weather on incidence occurs after zero ($\alpha_0 = -0.36$), three ($\alpha_3 = -0.35$) and four days ($\alpha_4 = -0.34$). The mean daily temperature usually was the predominant factor in the latent weather indicator. Since the average incidence for males is about 0.5%, the model implies that, if everything else is held constant, an increase in mean temperature of 10 degrees Celsius lowers the incidence after three days with 0.035%. The 90% confidence intervals that are drawn indicate that this three-day effect is just significant at a P-level of 0.10. Despite this, the evidence is inconclusive with respect to the question which lag influences absence most, especially after correcting for multiple testing. It is fair to conclude that weather does affect the absence incidence after some days, but also that the effect is rather small.

5. Discussion

For pure AR and mixed ARMA models, covariance structure estimates are approximately unbiased and almost as efficient as those provided by a true ARMA algorithm. For pure MA models, parameter estimates are biased up to 5–10%, and considerably less efficient. In the latter case, the correspondence with standard methods is poor. It is wise

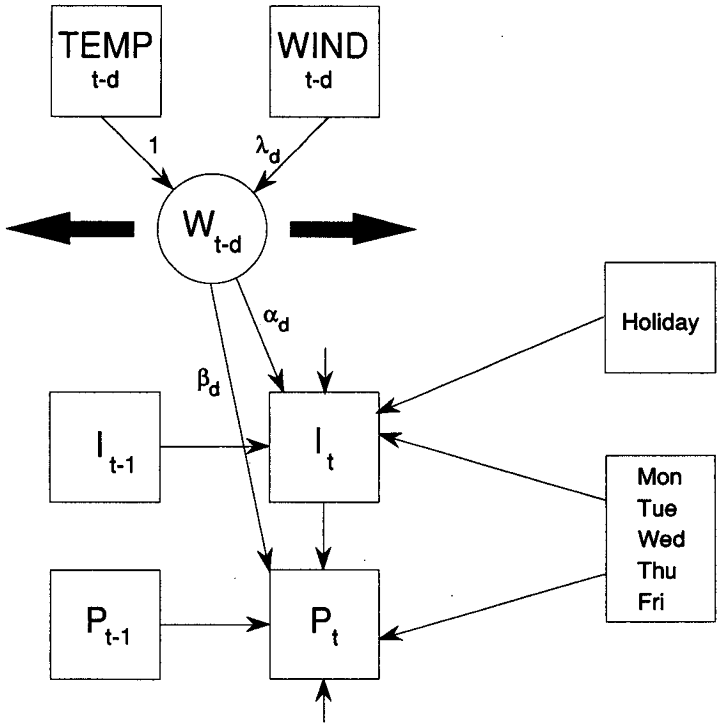

FIGURE 8.
Sliding across time to find the optimal delay between weather conditions and absence. The sliding factor d ranges from 0 (instantaneous weather effect) to 6 days.

not to use CALIS for estimating the parameters of pure MA models. CALIS standard errors are often quite reasonable. Of course, if the primary goal is to fit univariate ARMA models, conventional methods could best be applied. However, the structural model adds substantial flexibility and offers a large choice of models, some of which are quite useful in time series analysis.

One of the critical assumptions in covariance structure analysis is independence of observations. The analogous criterion in time series analysis is that model residuals must conform to white noise. The use of this criterion was stressed by Ghaddar and Tong (1981). Checking for white noise is certainly not common practice however. This state of affairs can, at least partly, be attributed to the fact that no adequate strategy for dealing with serial correlation in covariances structure models has been developed yet. Integrating lagged variables into the structural component provides one solution to the problem of autodependence. Properly chosen lagged variables takes away autocorrelation from the residuals and puts it where it belongs, that is, into the structural model.

Many different modelling strategies exist in time series analysis, some of which are fiercely defended. Because modelling is part of every analysis, it is a great opportunity for debate. Cook and Campbell (1979, chap. 7) distinguish eight different strategies for analyzing concomitancies in time series and discuss their pros and cons. No single method is optimal. A modelling strategy that is often quite useful is to start with a clear research question and a plausible minimal path specification of the process of interest. Preferably, there is an outcome measure, a causal factor of interest, and some covariates that have to be accounted for when assessing the relation between the cause and the effect. Either type of variable could possibly be measured by multiple indicators. For ample data, split the

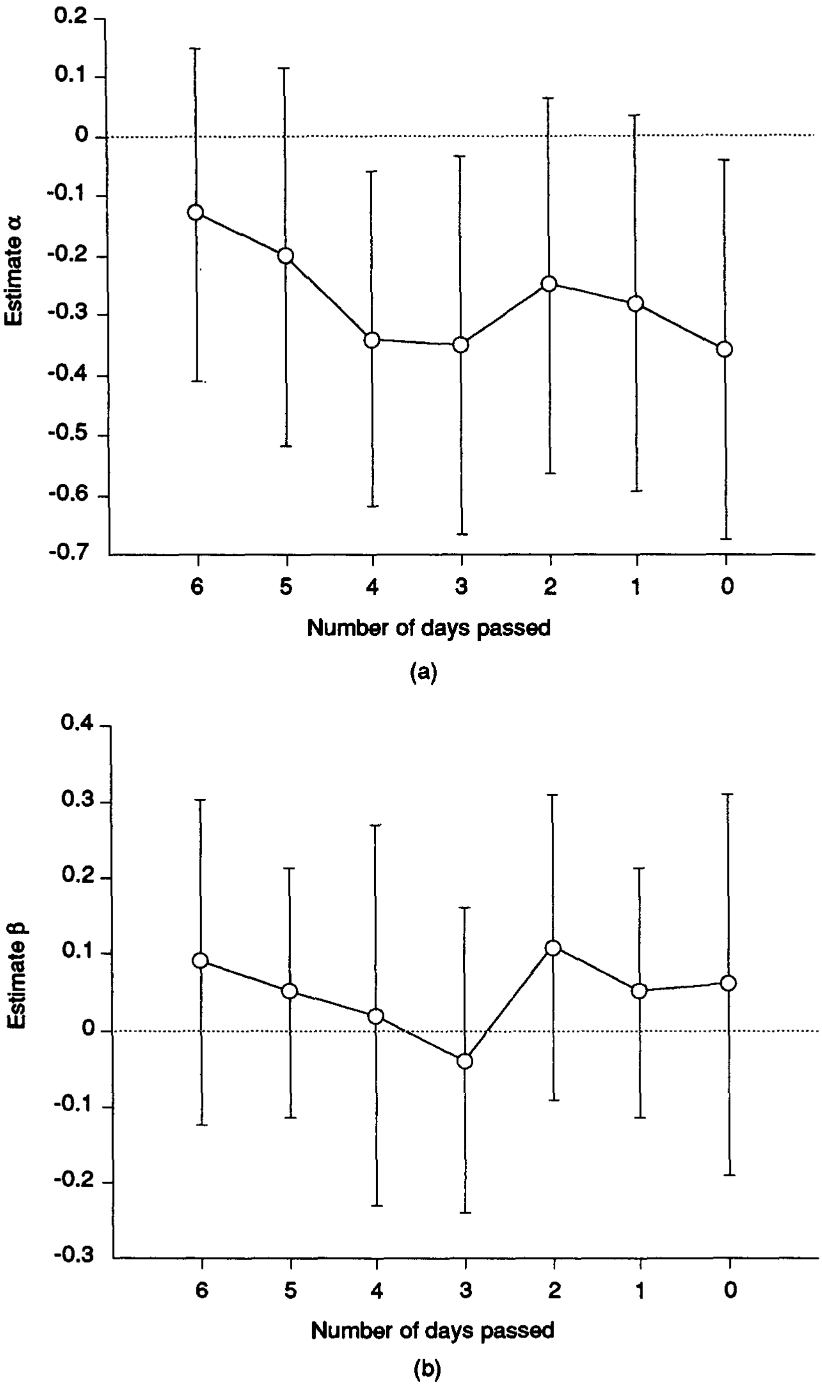

FIGURE 9.
Regression estimates plus 90% confidence interval for the influence of weather on absenteeism incidence (a) and prevalence (b) as a function of the number of the days passed since the weather.

data into two sets, one for fitting and one for validation. Subsequently, fit the minimal model without the effect of interest and test each residual on the lack of autocorrelation. Eliminate trend and seasonality in the series by incorporating polynomial and modulating functions of time as covariates. If an autocorrelated residual is detected, try adding an additional lagged variable for predicting the outcome. This can be either a lag of (a) the outcome variable itself, (b) the residual error variable, and (c) another observed or latent variable. Some authors restrict to options (a) and (b), which leads to prewhitening by ARMA models. Working this way along all autocorrelated residuals, try to capture all serial dependency into the path model. Avoid correlated residuals as far as possible. Introduce the main effect of the study and estimate its magnitude and standard error. Compute how well the model fits the observed covariance matrix as compared to the one without the causal relation. If a validation set is available, fix the free parameters to the model estimates and report how well the model fits the covariance matrix of the validation data.

Provided that one stays within the class of ARMA models, it is useful to transfer and adapt the Box-Jenkins identification methodology to covariance structures. It would surely help if the software could produce plots of the autocorrelation function and the partial autocorrelation function, and compute Durbin-Watson or the Ljung-Box-Pierce statistics.

The user should also be aware of the problem of choosing the window size of the covariance matrix. Molenaar (1985) remarks that it is not possible to compare models that are based on different sizes of the covariance matrix. It therefore makes sense to fix the number of included lags to accommodate to the most complex model one would like to try, and chain the basic model accordingly. If $p_{\max}$ and $q_{\max}$ are the orders for the most complex model, the maximum number of lagged covariances can be chosen simply as $p_{\max} + q_{\max}$. Current experience is that this hardly affects the estimates, but more work is needed to verify this in more general cases. One of the reviewers expected that the standard errors are affected.

Some work on minimization algorithms seems to be called for, especially on those for solving pure MA models. One could experiment with different minimization methods within PROC CALIS or use other programs like EQS and LISREL. It may even turn out worthwhile to implement a recursive estimator for solving MA-type of problems, though properly integrating it into an existing package probably requires a major effort. Also, it seems to be sensible to apply a Toeplitz transformation so as to stabilize and symmetrize the input covariance matrix. Wood and Brown (1994) published a SAS® macro that implements such a transform.

The method can be extended in a number of ways, for example by fitting an ARMA model to a latent component instead of to the observed data. Suppose that a latent variable follows an AR(1) process, then the structural model is closely related to a state space model. Alternatively, if the latent variable follows a pure MA-process, a form of dynamic factor analysis is obtained. More generally, one basic ARMA model can be replicated in some way. For example, different time series may refer to different manifestations of the same physical process, so then replication units are distinct measurements of the same factor as in the common factor model. Alternatively, replications could consist of different individual growth paths. Yet another possibility is to partition a single univariate series into equally long periods and use covariance structure modelling to see if some periods deviate from the basic ARMA model. The replication unit is then, for example, a week or a year.

For one reason or another, many proficient users of multivariate techniques view the analysis of time series as something that uses entirely different concepts and methods. This is unfortunate because it inhibits cross-fertilization. A tremendous body of knowledge on multivariate methods has accumulated in the social sciences. This wisdom is practically dormant as time series data is concerned. The present work hopefully contributes to awaken some of this potential and will assist in a wider understanding and a better appreciation of time series in behavioral research.

Appendix

SAS code for estimating ARMA parameters using PROC CALIS. Variable v0 is the time series, v1 is its first lag, v2 is its second lag, and so on, f0 is the first latent factor, f1 the second, and so on.

```
proc calis cov;      /* ARMA(1,0) model */
    lineqs v0 = phi1 v1 + f0;
    std    f0 = sigma2;
    bounds -1 <= phi1 <= 1;

proc calis cov;      /* ARMA(2,0) model */
    lineqs v0 = phi1 v1 + phi2 v2 + f0;
    std    f0 = sigma2;
    bounds -2 <= phi1 <= 2, -1 <= phi2 <= 1;

proc calis cov;      /* ARMA(0,1) model */
    lineqs v0 = f0 + theta1 f1,
           v1 = f1 + theta1 f2;
    std    f0-f2 = sigma2;
    bounds -1 <= theta1 <= 1;

proc calis cov;      /* ARMA(0,2) model */
    lineqs v0 = f0 + theta1 f1 + theta2 f2,
           v1 = f1 + theta1 f2 + theta2 f3,
           v2 = f2 + theta1 f3 + theta2 f4;
    std    f0-f4 = sigma2;
    bounds -2 <= theta1 <= 2, -1 <= theta2 <= 1;

proc calis cov;      /* ARMA(1,1) model, covariance method */
    lineqs v0 = phi1 v1 + f0 + theta1 f1,
           v1 = phi1 v2 + f1 + theta1 f2;
    std    f0-f2 = sigma2;
    cov    v2 f2 = cov22;
    bounds -1 <= theta1 phi1 <= 1;

proc calis cov;      /* ARMA(1,1) model, noise method */
    lineqs v0 = phi1 v1 + f0 + theta1 f1,
           v1 = phi1 v2 + f1 + theta1 f2,
           v2 = f2 + e2;
    std    f0-f2 = sigma2,
           e2 = e2se;
    bounds -1 <= theta1 phi1 <= 1;
```

References

Anderson, T. W. (1960). Some stochastic process models for intelligence test scores. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 205–220). Stanford, CA: Stanford University Press.

Bentler, P. M. (1989). *EQS structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika, 45,* 298–308.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis forecasting and control* (rev. ed.). San Francisco: Holden-Day.

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57*, 469–497.

Cattell, R. B. (1963). The structuring of change by P- and incremental r-technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 163–198). Madison: University of Wisconsin Press.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin.

Fraser, C., & McDonald, R. P. (1988). COSAN: Covariance structure analysis. *Multivariate Behavioral Research, 23*, 263–265.

Geweke, J. F., & Singleton, K. J. (1981). Maximum likelihood "confirmatory" factor analysis of economic time series. *International Economic Review, 22*, 37–54.

Ghaddar, D. K., & Tong, H. (1981). Data transformation and self-exciting treshold autoregression. *Applied Statistics, 30*, 238–248.

Glass, G. V., Willson, V. L., & Gottman, J. M. (1975). *Design and analysis of time-series experiments*. Boulder, CO: Colorado Associated University Press.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*, 424–438.

Guttman, L. (1954). A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348). New York: Columbia University Press.

Hagnell, M. (1992). An application of regression on a latent variable for time series data. In P. G. M. van der Heijden, W. Jansen, B. Francis & G. U. H. Seeber (Eds.), *Statistical modelling* (pp. 121–129). Amsterdam: Elsevier.

Jones, K. (1991). The application of time series methods to moderate span longitudinal data. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change* (pp. 75–87). Washington, DC: American Psychological Association.

Jöreskog, K. G. (1978). An econometric model for multivariate panel data. *Annales del'Institut National de la Statistique et des Etudes Economique, 30*, 355–366.

Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal developmental investigations. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 303–351). Washington, DC: American Psychological Association.

Jöreskog, K. G., & Sörbom, D. (1985). *LISREL-VI program manual*. Chicago: International Educational Services.

MacCallum, R., & Ashby, F. G. (1986). Relations between linear systems theory and covariance structure modeling. *Journal of Mathematical Psychology, 30*, 1–27.

MacCleary, R., & Hay, R. A. (1980). *Applied time series analysis for the social sciences*. Beverly Hills: Sage.

McArdle, J. J., & Aber, M. S. (1990). Patterns of change within latent variable structural equation models. In A. von Eye (Ed.), *Statistical methods in longitudinal research, Vol. I* (pp. 151–224). Boston: Academic Press.

Mélard, G. (1984). A fast algorithm for the exact likelihood of autoregressive-moving average models. *Applied Statistics, 33*, 104–114.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50*, 181–202.

Molenaar, P. C. M., de Gooijer, J. G., & Schmitz, B. (1992). Dynamic factor analysis of nonstationary multivariate time series. *Psychometrika, 57*, 333–349.

Mukherjee, B. N., & Maiti, S. S. (1988). Iterative maximum-likelihood estimation of parameters of the Toeplitz correlation structure. *British Journal of Mathematical and Statistical Psychology, 41*, 63–99.

Oud, J. H., van den Bercken, J. H., & Essers, R. J. (1990). Longitudinal factor score estimation using the Kalman filter. *Applied Psychological Measurement, 14*, 395–418.

SAS Institute. (1990). *SAS/STAT user's guide, Vol. 1*. Cary, NC: Author.

Schmitz, B. (1990). Univariate and multivariate time-series models: The analysis of intraindividual variability and intraindividual relationships. In A. von Eye (Ed.), *Statistical methods in longitudinal research, Vol. II* (pp. 351–386). Boston: Academic Press.

Smulders, P. G. W., & van Deursen, C. G. L. (1995). *The effects of weather conditions on employee absenteeism* (Unpublished report). Leiden: TNO Prevention and Health.

SPSS. (1990). *SPSS/PC+ Trends*. Chicago: Author.

van Buuren, S. (1990). *Optimal scaling of time series*. Leiden: DSWO Press.

van Buuren, S. (1992). Predictable linear combinations from categorical time series. *Statistica Applicata: Italian Journal of Applied Statistics, 4*, 743–751.

Wood, P., & Brown, D. (1994). The study of intraindividual differences by means of dynamic factor models: Rationale, implementation and interpretation. *Psychological Bulletin, 116*, 166–186.

Manuscript received 1/27/94
Final version received 1/12/96