Journal of Official Statistics
Vol. 7. No. 4, 1991, pp. 437–455

A Multivariate Time Series Analysis of Fertility, Adult Mortality, Nuptiality, and Real Wages in Sweden 1751–1850: A Comparison of Two Different Approaches

Mats Hagnell[1]

Abstract: The relationships between fertility, adult mortality, nuptiality, and real wages in Sweden from 1751 to 1850 are studied using multivariate time series analysis methods. This approach allows an empirical determination of the relationships between the four time series, a considerable advantage where existing theory provides insufficient guidance.

Both the vector autoregressive moving average (VARMA) model and the vector autoregressive (VAR) model are used and the two approaches are compared with regard to parsimony, interpretation, and post-sample forecasting performance. We also compare the VARMA and the VAR approaches with regard to their sensitivity to outliers.

Key words: Multivariate time series model; VARMA model; VAR model; historical demography; outliers.

1. Introduction

In this article we study the relationships between fertility, adult mortality, nuptiality, and real wages in Sweden from 1751 to 1850, using annual data. We explore the short-term interaction between these four variables without making any a priori assumptions. To accomplish this goal we use multivariate time series analysis which allows an empirical determination of the relationships between the four time series, a considerable advantage where existing theory provides insufficient guidance.

The short-run relationship between demographic and economic time series in Sweden during the eighteenth and nineteenth centuries has been studied previously in a number of papers from the departments of Economic History and Statistics at the University of Lund (Andersson and Hagnell 1989; Bengtsson and Broström 1986; Bengtsson and Ohlsson 1985; Bring 1987; Hagnell and Salomonsson 1989a; Larsen 1986). In these papers, however, only one-sided causation is considered, i.e., one dependent variable is explained by one or several explanatory variables. The time series models used in these papers are distributed lag-models or transfer function models.

[1] Department of Statistics, University of Lund, Box 7008, S-220 07 Lund, Sweden.

Acknowledgements: The author would like to thank Sven Berg and Tommy Bengtsson for their helpful comments and Tommy Bengtsson for providing the data. Financial support is acknowledged from HSFR, The Swedish Council for Research in the Humanities and Social Science, and NFR, the Swedish Natural Science Research Council.

When we study the relationships between demographic and economic time series, in most cases we have to take into consideration the possibility that one variable, which explains another variable can itself be explained by this other variable. For example, an increase of the adult mortality is probably followed by lower fertility, as fewer adult women are available to give birth to children. On the other hand, an increase in fertility may lead to an increase in adult mortality due to deaths from childbed fever.

Thus to study more complex relationships between demographic and economic time series we have to use time series models which allow two-way causation or feedback, i.e., models which are non-recursive multiple equation systems.

One approach to the analysis of social science time series is to model data as a theoretically specified structural system. This approach, known as econometric modeling in economics and structural equation modeling in the social sciences, requires a good a priori theory. Unfortunately, in economic demography we rarely have the kind of precise knowledge which is required in structural equation modeling. Even if we assume that we know which variables influence a given variable, we do not know at which lags the influences occur. If, however, adequately long time series are available, we can instead use multiple time series analysis, which allows an empirical determination of the pattern of lag structures between series.

There are two recent main developments in this direction. Sims (1980) has proposed an autoregressive (AR) model for vector-valued or multiple time series. One advantage of the vector AR (VAR) method is that model parameters can be estimated by ordinary least squares algorithms. A disadvantage is that VAR models cannot parsimoniously represent moving average (MA) processes. Since MA processes are commonly encountered in social science data, this is a serious shortcoming. An example of an application of VAR modeling in historical demography is given by Eckstein, Schultz, and Wolpin (1985).

A more promising approach has been developed by Tiao and Box (1981), who have proposed an iterative approach for building multivariate or vector-valued autoregressive moving average (ARMA) models. A disadvantage of vector ARMA (VARMA) models, compared to VAR models at least, is that model parameters are not easily estimated. Nowadays, however, several software packages are available with algorithms for VARMA models. With the estimation problem no longer burdensome, the VARMA model promises to become a powerful tool for the analysis of social science time series data. An example of an application of VARMA modeling in historical demography is given by McCleary and McDowall (1984).

We use two different approaches to multivariate time series analysis of the relationships between our four time series in order to make a comparison between these approaches. First we consider VARMA modeling and then the modeling of the four time series as a VAR system.

In demography, an empirical comparison between different time series models, univariate, transfer function and VARMA models was made by Carter and Lee (1986) and in macroeconomics, an empirical comparison between VAR and VARMA models was made by Fackler and Krieger (1986). In both these cases the main purpose was to compare the forecasting ability of the different models and the post-sample forecasting performance was used as a criterion. The results in Fackler and Krieger (1986) suggested that VARMA models may out-perform VAR models by parsimoniously representing the underlying process, thus improving the efficiency of the parameter estimates. Here, since we use historical data, we mainly make comparisons between our different approaches with regard to substantive interpretation and parsimonious representation. However, a less detailed comparison of the two approaches with regard to their post-sample forecasting performance is also made.

A problem which is common in many applications of time series models is the influence of outliers. Here we have some large outliers in the data, especially for the adult mortality. Our view is that the outliers in our historical time series are probably caused by unknown exogenous events or interventions which disturb the normal relationships between the variables. When, for example, we study how adult mortality depends on real wages, and adult mortality has some large peaks that may be due to epidemics, it seems questionable to include these peaks in the analysis of how adult mortality depends on real wages.

In order to study the influence of outliers we try, for both the VARMA and the VAR approach, to identify the outliers, adjust their values to more normal values, and then perform an analysis on the data with the adjusted values. This analysis can then be compared with the analysis on the original values. Thus we are able to compare the two approaches' sensitivities to outliers.

This paper is organized as follows. Multivariate time series models are surveyed in Section 2 and the data are described in Section 3. We identify and estimate a VARMA model for our data and discuss the casual implications of the estimated model in Section 4. A VAR model is estimated and its implications are discussed in Section 5. A comparison of the VARMA and the VAR model with regard to their post-sample forecasting performance is made in Section 6 while the influence of outliers for the two approaches is treated in Section 7. Our main findings are briefly summarized and some concluding remarks are made in Section 8.

2. Multivariate Time Series Models

In this section we give a short survey of multivariate time series or vector ARMA models and of vector ARMA modeling following the approach given in Tiao and Box (1981). Multivariate ARMA models are also considered in Granger and Newbold (1986).

A vector ARMA model with p autoregressive and q moving average terms or a VARMA (p, q) model may be written as

$$[\mathbf{B}^0 - \boldsymbol{\phi}_1\mathbf{B} - \ldots - \boldsymbol{\phi}_p\mathbf{B}^p]\mathbf{Z}_t = \boldsymbol{\delta} + [\mathbf{B}^0 - \boldsymbol{\Theta}_1\mathbf{B} - \ldots - \boldsymbol{\Theta}_q\mathbf{B}^q]\mathbf{a}_t \quad (2.1)$$

where $\mathbf{Z}_t$ is a column vector of k stationary time series; $\mathbf{a}_t$ is a column vector of k white noise processes; $\boldsymbol{\delta}$ is a column vector of k constants; $\boldsymbol{\phi}_i$ is the ith $k \times k$ AR matrix, $i = 1, \ldots, p$; $\boldsymbol{\Theta}_j$ is the jth $k \times k$ MA matrix, $j = 1, \ldots, q$; $\mathbf{B}$ is a $k \times k$ diagonalized backshift matrix, i.e., for $k = 2$,

$$\mathbf{B}^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ldots, \mathbf{B}^p = \begin{bmatrix} B^p & 0 \\ 0 & B^p \end{bmatrix}$$

and B is the usual backshift operator, i.e., $Bz_t = z_{t-1}$, $B^2 z_t = z_{t-2}$, etc.

The special case when we have only autoregressive terms in the VARMA model (2.1) is a vector autoregressive model with p terms, a VAR (p) model

$$\mathbf{Z}_t = \boldsymbol{\delta} + \boldsymbol{\phi}_1\mathbf{Z}_{t-1} + \ldots + \boldsymbol{\phi}_p\mathbf{Z}_{t-p} + \mathbf{a}_t \quad (2.2)$$

The VARMA model (2.1) can be written in shorter notation as

$$\boldsymbol{\phi}_p(\mathbf{B})\mathbf{Z}_t = \boldsymbol{\delta} + \boldsymbol{\Theta}_q(\mathbf{B})\mathbf{a}_t \quad (2.3)$$

where $\phi_p(\mathbf{B})$, the autoregressive part, is a matrix polynomial of degree p in $\mathbf{B}$ and $\Theta_q(\mathbf{B})$, the moving average part, is a matrix polynomial of degree q in $\mathbf{B}$.

In order to have a better appreciation of the VARMA models, we present a simple case, which is a bivariate model, $k = 2$, with $p = 1$ and $q = 1$ or a VARMA (1, 1) model

$$(\mathbf{I} - \phi\mathbf{B})\mathbf{Z}_t = \delta + (\mathbf{I} - \Theta\mathbf{B})\mathbf{a}_t \qquad (2.4)$$

which can also be written as

$$Z_{1t} = \delta_1 + \phi_{11}Z_{1,t-1} + \phi_{12}Z_{2,t-1} + a_{1t} - \Theta_{11}a_{1,t-1} - \Theta_{12}a_{2,t-1}$$

$$Z_{2t} = \delta_2 + \phi_{21}Z_{1,t-1} + \phi_{22}Z_{2,t-1} + a_{2t} - \Theta_{21}a_{1,t-1} - \Theta_{22}a_{2,t-1}.$$

In the model (2.4), diagonal elements of ϕ (ϕ_{11}, ϕ_{22}) and Θ (Θ_{11}, Θ_{22}) represent AR and MA structures within each series. Off-diagonal elements are causal effects between pairs of series and pairs of shocks. The parameter Θ_{12}, for example, is the causal effect of $a_{2,t-1}$ on Z_{1t}, while Θ_{21} is the causal effect of $a_{1,t-1}$ on Z_{2t}.

In general, model (2.4) represents two-way causation or feedback between Z_{1t} and Z_{2t}. However, if it is unlikely that Z_2 "causes" Z_1, then we expect Θ_{12} and ϕ_{12} to be zero. Equation (2.4) then forms a recursive model and the two equations in (2.4) reduce to a univariate ARMA (1, 1) model for Z_{1t} and a model that may be rewritten as a transfer function model with Z_{2t} as the dependent and Z_{1t} as the independent variable.

The class of vector ARMA models (2.1) is extensive and may contain models with a large number of parameters. Given a vector time series $\mathbf{Z}_t$ of finite length $t = 1, \ldots, n$, the aim is to find a model which contains as few parameters as possible and at the same time adequately represents the dynamic and stochastic relationships in the data at hand. Extending the basic ideas in Box and Jenkins (1976) for the univariate series, Tiao and Box (1981) proposed an iterative approach for building multivariate ARMA models consisting of three main phases: (i) tentative model identification, (ii) estimation, and (iii) diagnostic checking.

The number of ARMA parameters to be estimated may be sucessively reduced by constraining some parameters to zero. Parameters that could not be distinguished from zero during the estimation process are here constrained to zero to obtain parsimonious models.

Two valuable tools in the tentative identification stage are the auto and cross correlation matrices (CCM) and partial autoregression (PAR) matrices. Sample cross correlation matrices (CCM) are given by $\mathbf{R}(k) = \{r_{ij}(k)\}$, where $r_{ij}(k)$ is the sample cross correlation of lag k between series component i and series component j, i.e., the correlation between $Z_{i,t}$ and $Z_{j,t-k}$.

If $\mathbf{Z}_t$ follows an MA(q) model, i.e., an VARMA (0, q) model, then the true CCM's are all zero matrices for lags greater than q. That is, the cross correlation matrices "cut off" after q lags, much like the autocorrelations of a univariate MA(q) model.

If $\mathbf{Z}_t$ follows a vector AR(p) model, a VAR(p) model, then the CCM's have a die-out pattern. In order to better characterize autoregressive models, Tiao and Box (1981) introduced the partial autoregression (PAR) matrices. The PAR matrices are the multivariate counterpart to the partial autocorrelation function (PACF) in univariate ARMA modeling. If $\mathbf{Z}_t$ follows a vector AR(p) model, then the true PAR matrices are all zero matrices for lags greater than p. That is, the partial autoregression matrices "cut off" after p lags, much like the partial autocorrelations for a univariate AR(p) model.

To help determine a tentative order of an autoregressive model, we can also use the likelihood ratio statistic corresponding to testing the null hypothesis that the partial autoregression matrix of lag 1 is a zero matrix against the alternative that it is not a zero matrix. This statistic is, for the null hypothesis, asymptotically distributed as a chi-square statistic.

To estimate time series models, computer programs are necessary. While programs for univariate ARIMA models and transfer functions models are available in many general-purpose statistical packages, programs for multivariate time series are available only in some of the packages especially developed for time series analysis. We have used the SCA package, described in Liu and Hudak (1986), which contains a module for multivariate time series analysis.

3. The Data

The data used for the estimation are annual time series for all of Sweden during the period 1751–1850 for the crude birth rate, the death rate for ages 20–50 years, the crude marriage rate, and an index of real wages. We will use the following notation for the four variables:

F = fertility, measured by the crude birth rate,
M = adult mortality, the death rate for ages 20–50 years,
N = nuptiality, measured by the crude marriage rate, and
R = the real wages index.

We use the ages 20-50 in calculating the mortality, here called the adult mortality, since we know from previous studies (Andersson and Hagnell 1989; Bengtsson and Ohlsson 1985) that the mortality differs substantially between this age group and the age group over 50 years old.

The real wages index is calculated as the ratio between the wages for agricultural labour and an index for the living cost. The later index depends mainly of grain prices, especially the price of rye. These prices, as well as the wages, were decided in the autumn and held fixed until the next autumn. Here real wages are mainly a measure of the standard of living the following year. The real wages index was developed at the Department of Economic History, University of Lund, (Bengtsson and Jörberg 1975: Jörberg 1972) and applies to all of Sweden.

Plots of the four time series are shown in Figures 3.1–3.4. The adult mortality and the real wages index seem stationary but with some large outliers for the adult death rate. Although there is some indication of a decreasing trend for fertility and nuptiality, we do not consider it necessary to difference the four series to achieve stationarity. For vector time series, even if some components exhibit nonstationary behaviour, linear combinations of the elements of $\mathbf{Z}_t$ may often be stationary and simultaneous differencing of all series can lead to unnecessary complications in VARMA model fitting.

4. Identification and Estimation of a VARMA Model

We begin the identification of a VARMA model for the four variables by studying the cross correlations. A condensed summary of the pattern of cross correlations matrices $\mathbf{R}(k)$ for the first 10 lags is provided in Table 4.1 in terms of the plus, minus and dot symbols. The indicator symbol "+" is assigned when an element of $\mathbf{R}(k)$ is greater than two times its estimated standard error, the symbol "−" for values less than minus two standard errors, and "." for values in between. Since there is no abrupt cut-off after some lag but rather a die-out pattern, the VARMA model cannot be a low order MA process but must be an AR process or a mixed model.

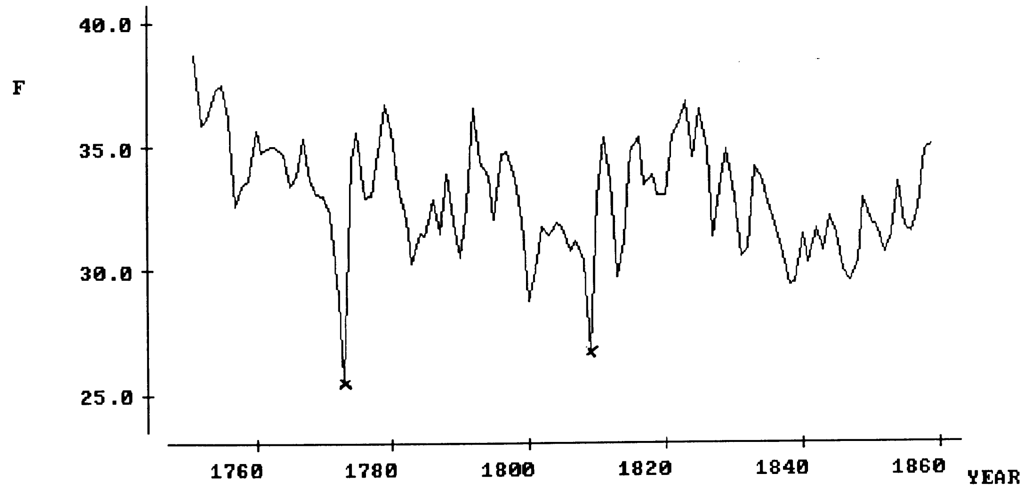

Fig. 3.1. Fertility 1751–1859

The partial autoregression matrices **P**(k) in summary symbols and the likelihood ratio statistic up to the sixth lag are given in Table 4.2. It is clear from Table 4.2 that little improvement occurs after lag 2. For lags greater than 2, nearly all of the elements of the partial autoregression matrix are small compared with their estimated standard errors, and the likelihood ratio statistic, which is approximately distributed as chi-square with 16 degress of freedom, fails to show significant improvement.

A tentative identification is an AR(2) or an ARMA(1, 1) model. In order to choose between these two alternatives we fit an AR(1) model and study the pattern of the residual cross correlation matrices. Table 4.3 gives the pattern of the cross correlations of the residuals after the AR(1) fit. The cut-off after lag 1 indicates that the residuals follow an MA(1) process. Consequently we choose

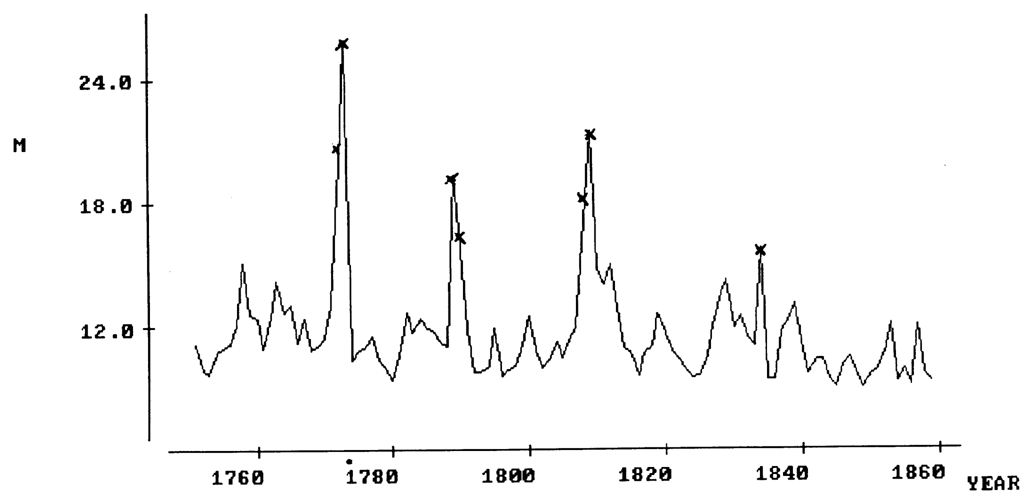

Fig. 3.2. Adult mortality 1751–1859

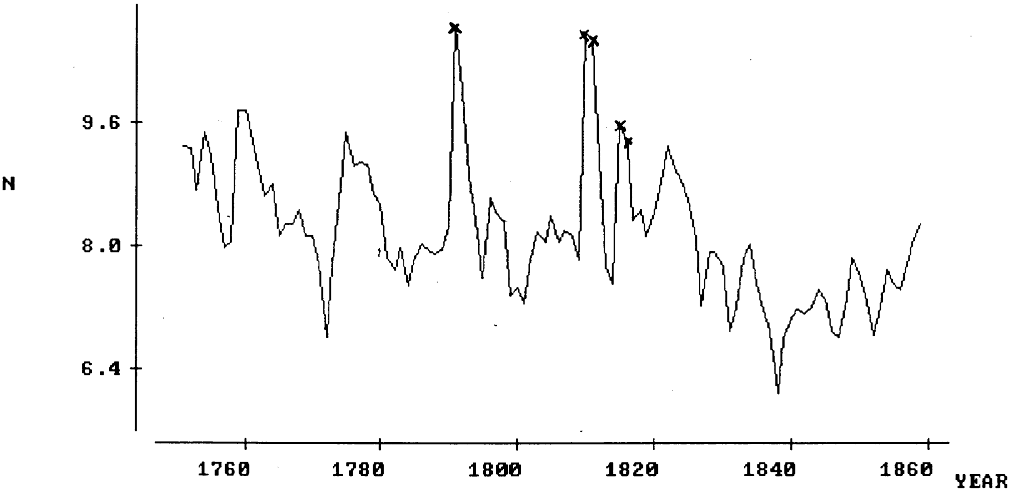

Fig. 3.3. Nuptiality 1751–1859

an ARMA(1, 1) model as the best VARMA model for our four original variables. The AR(2) alternative will be considered more in detail in Section 6 as it is the best fitting VAR model.

The VARMA(1, 1) model was fitted using the exact maximum likelihood method. More precisely, we use the model

$$(\mathbf{I} - \boldsymbol{\phi}\mathbf{B})\mathbf{Z}_t = \delta + (\mathbf{I} - \boldsymbol{\Theta}\mathbf{B})\mathbf{a}_t. \quad (4.1)$$

The parameter estimates for the fit of the unrestricted model (4.1) are shown in Table 4.4, where the figures within parentheses are the absolute t-values for the parameter estimates.

The fit of the unrestricted model (4.1) requires estimation of four constants, 16 AR terms and 16 MA terms, in all 36 parameters. A more parsimonious model is obtained by setting to zero those coefficients whose estimates are small compared to their standard errors, i.e., have small absolute

Fig. 3.4. Real wages 1751–1859

Table 4.1. Pattern of sample cross correlations matrices, **R***(k), k = 1, . . . , 10, for the four variables*

Lag =	1				2				3				4				5			
	F	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*
F	+	−	+	+	+	.	+	+	+	.	+	.	+	.	+	.	+	.	+	.
M	−	+	.	−	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	+	+	+	.	.	+	+	−	.	+	+	−	.	.	+	−	.	.	+	−
R	+	.	.	+	+	.	.	+	+	.	.	+	.	.	.	+	.	.	.	+

Lag =	6				7				8				9				10			
	F	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*
F	.	.	+	.	.	+	+	.	.	.	.	.	.	.	.	.	.	.	.	.
M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	.	.	+	−	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
R	.	.	.	+	.	.	.	+	.	.	.	+	.	.	.	+	.	.	.	.

t-values. This was done for model (4.1) in several sucessive steps. The parameter estimates for the fit of the final restricted model (4.1) are shown in Table 4.5 where all estimated parameters are significant. In this restricted model we have to estimate only 6 AR terms and 3 MA terms, a considerable simplification in comparison with the unrestricted model where we had to estimate 16 terms of each kind.

The final restricted model (4.1) fits the data almost as well as the unrestricted model. The difference in -2*(log likelihood) between the restricted and unrestricted model is 30.9, which gives a *p*-value greater than 10% for a chi-square distribution with degrees of freedom equal to the number of parameters set equal to zero, i.e., $16 + 16 - 6 - 3 = 23$.

As the final step in the model building, we performed a diagnostic check by calculating the cross correlation matrices of the residuals. Table 4.6 shows the pattern of residual cross correlations for the final restricted VARMA(1, 1) model. It suggests that the restricted model provides an adequate representation of the data. The significant entries in the residual cross correlation matrix for lag 5 return in every fitted model and are possibly due to some sort of five-year cycle. This five-year relationship is also seen clearly in the univariate model for the central marriage rate.

We did a further diagnostic check by overfitting. A VARMA(2, 1) unrestricted model was fitted and then sucessively reduced in several steps by constraining insignificant parameters to zero. The final restricted model was exactly the same as the final restricted VARMA(1, 1) model. The extension of the VARMA(1,1) model with a VAR

Table 4.2. Pattern of partial autoregression matrices, P*(k), and likelihood ratio statistic* $X^2(k)$*, k = 1, . . . , 6, for the four variables*

Lag =	1				2				3				4				5				6			
	F	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*
F	.	.	+	+	.	.	−	−	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
M	+	+	−	−	.	.	.	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	−	+	+	.	.	.	.	−	.	.	.	.	.	.	.	.	.	.	+	.	.	.	.	.
R	.	.	.	+	+	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$X^2(k)$	205.42				34.57				21.32				15.97				16.78				9.94			

Table 4.3. Pattern of sample cross correlations matrices for the residuals after the AR(1) fit

Lag =	1				2				3				4				5				6			
	F	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*
F	+	.	.	.	.	.	.	.	.	.	.	−	.	.	.	.	.	.	+	.	.	.	.	.
M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	+	.	.	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.	.	.	.
R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

(2) term did not give a significantly better fit to the data.

The final restricted fitted VARMA(1, 1) model (4.1), with estimated parameters given in Table 4.5, explains the relationships among these four time series as the four-equation system

$$F_t = 14.3 + 1.35N_{t-1} + 0.12R_{t-1} + 0.34a_{F,t-1} + a_{Ft} \quad (4.2)$$

$$M_t = 5.85 + 0.51M_{t-1} - 0.14a_{R,t-1} + a_{Mt} \quad (4.3)$$

$$N_t = 1.38 + 0.09M_{t-1} + 0.70N_{t-1} + 0.04a_{R,t-1} + a_{Nt} \quad (4.4)$$

$$R_t = 14.1 + 0.78R_{t-1} + a_{Rt}. \quad (4.5)$$

Of these four equations, (4.2) explains the fertility in Sweden from 1751 to 1850 in terms of the nuptiality and the real wages, both one year earlier, and also of the fertility in preceding years. The effects are positive: increasing nuptiality and real wages result in higher fertility. The influence from real wages on fertility is easily interpreted. However, the coefficient of 1.35 representing influence from nuptiality on fertility is too large to represent only a direct causal effect. This large effect from nuptiality on fertility has been observed previously (Carlsson 1970; Lee 1975, 1981) and is attributed to some unknown common exogenous factor. However, the size of the coefficient is reduced when outliers are taken into consideration. This will be discussed in more detail in Section 7.

From (4.3) adult mortality is seen to depend on the random shock for real wages the preceding year and on adult mortality the preceding year. As one expects, increased real wages are associated with decreased mortality.

Furthermore, (4.4) explains the nuptiality in terms of adult mortality and real wages,

Table 4.4. Estimation results for the unrestricted model (4.1)

δ	φ				Θ			
14.15	−.07	.10	1.52	.11	−.29	.18	−.02	−.01
(2.7)	(.3)	(.7)	(3.3)	(2.9)	(1.4)	(1.3)	(.1)	(.3)
4.64	.34	.50	−.64	−.07	−.18	−.06	.36	.12
(.7)	(1.3)	(2.9)	(1.1)	(1.6)	(.6)	(.3)	(.5)	(2.1)
3.36	−.16	.11	.99	.01	−.17	.05	.20	.02
(1.9)	(2.0)	(2.2)	(6.0)	(.8)	(1.8)	(.9)	(1.4)	(1.3)
−20.44	.54	1.42	−.80	.89	1.03	1.53	−.37	.14
(1.3)	(.7)	(3.0)	(.5)	(7.4)	(1.1)	(2.8)	(.2)	(.9)

Table 4.5. Estimation results for the final restricted model (4.1)

δ	Φ				Θ			
14.25 (8.0)	.	.	1.35 (7.8)	.12 (6.4)	−.34 (4.5)	.	.	.
5.85 (6.3)	.	.51 (6.7)	.	.	.	.	.	.14 (4.0)
1.38 (2.5)	.	.09 (5.2)	.70 (11.8)	.	.	.	.	−.04 (4.2)
14.10 (3.9)	.	.	.	.78 (14.2)	.	.	.	.

both one year earlier, and also of nuptiality the preceding year. The effects are positive, so increasing mortality and real wages result in higher nuptiality. This is in accordance with what we would expect as more deaths result in more remarriages of the newly widowed and perhaps in more marriage opportunities due to inheritance.

Finally, from (4.5) real wages depend only on real wages the preceding year. In this context it is natural to interpret the real wages index as an independent or exogenous variable which does not depend on any of the other three variables.

Another way to illustrate the properties of the estimated system is to observe the system's responses to random shocks, called impulse responses (Sims 1980). If the system is stable, the impulse responses will dampen, i.e., the system will return to its steady state. However, as pointed out earlier (Eckstein et al. 1985), a simple interpretation of these impulse responses critically depends upon the assumption that the random shocks for different variables are uncorrelated. Here, we find relatively large correlations between the shock in F and the shocks for the other variables (−0.41, 0.52 and 0.37 for M, N, and R, respectively) while the remaining correlations are smaller. Ignoring the correlations between shocks, we briefly describe the impulse responses for the system.

Due to the simple structure of our estimated system, the impulse responses follow a relatively simple pattern. A shock in fertility does not give any responses in the other variables while a shock in nuptiality results only in exponentially decreasing impulse responses in fertility. The impulse responses for shocks in adult mortality and real wages are depicted in Figures 4.1 and 4.2, respectively.

From Figure 4.1 we find that a shock in adult mortality gives impulse responses in both fertility and nuptiality that first increase and then smoothly decrease towards zero. The overall positive responses for fertility may seem puzzling at first but are

Table 4.6. Pattern of sample cross correlations matrices for the residuals of the final restricted VARMA(1,1) model

Lag =	1				2				3				4				5				6			
	F	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*
F	.	.	.	.	.	.	.	.	.	.	.	−	.	.	.	.	.	.	+	+	.	.	.	.
M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.	.	.	.
R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

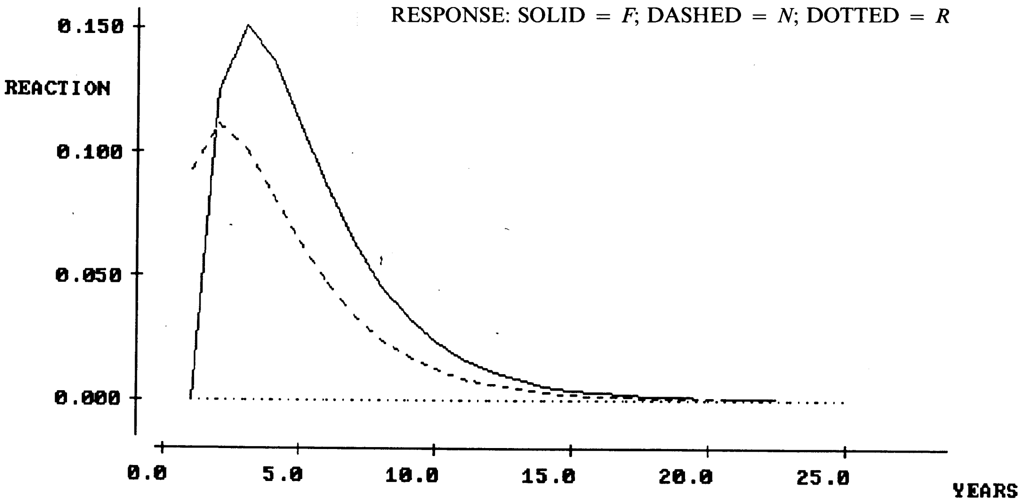

Fig. 4.1. Impulse response to M shock

explained by the fact that fertility in our system only depends indirectly on adult mortality through nuptiality. Since we use yearly data, we cannot catch the first anticipated negative response in our model. In Figure 4.2 we observe that the response in fertility to a shock in real wages increases at first and then decreases smoothly. The impulse responses in mortality increase rapidly towards zero while the impulse responses in nuptiality decrease rapidly towards zero.

The results of our estimated system are generally simpler than those found by others. For example, Eckstein et al. (1985) used Swedish data on fertility, non-infant mortality (i.e., mortality for persons one year and older), real wages and some other variables during the period 1750–1869.

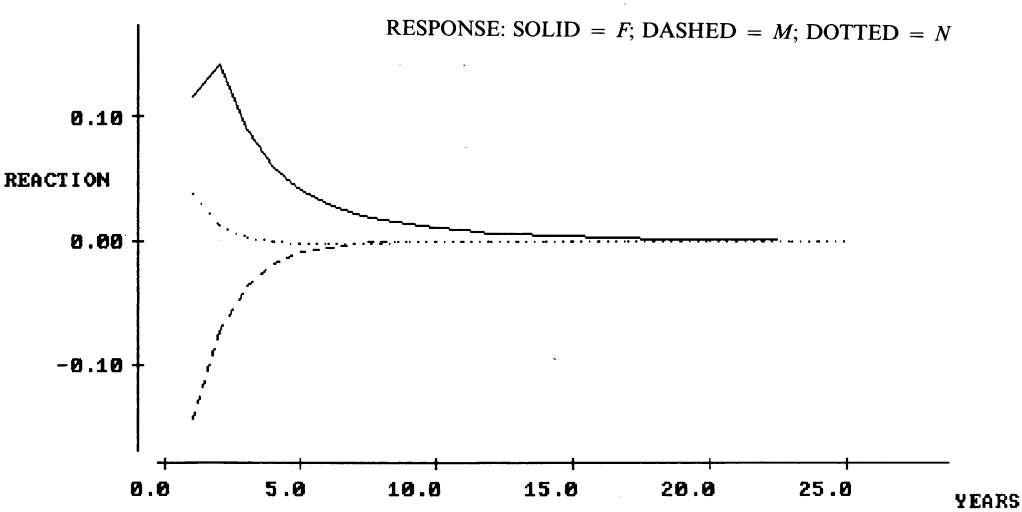

Fig. 4.2. Impulse response to R shock

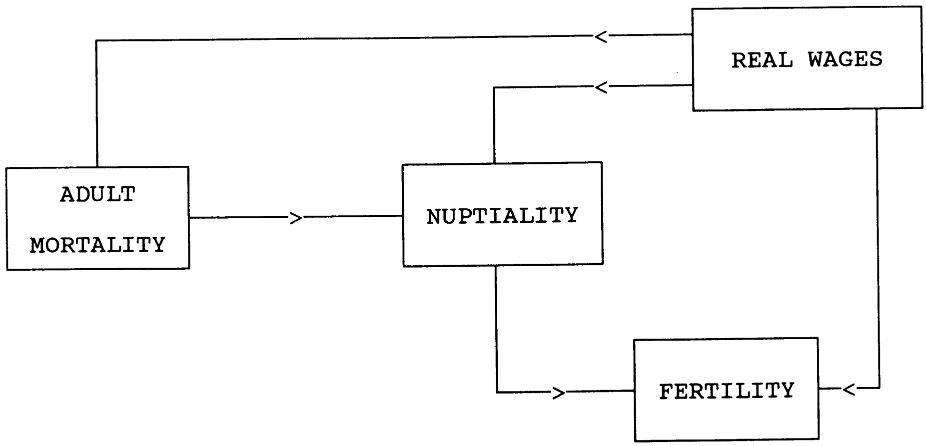

Fig. 4.3. Relationships between the four variables according to the VARMA model

Their corresponding impulse responses show much more complex and cyclic behaviour which they are sometimes able to interpret. One reason for their complicated pattern of impulse responses is their use of a more complicated model, an unrestricted VAR(4) model. The other reason is their analysis of non-infant mortality where they put together age groups with very different reactions to the other variables. We, on the other hand, are analysing mortality for the relatively homogeneous age group 20–50 years and get a simpler pattern of impulse responses.

The arrow scheme in Figure 4.3 summarizes our findings concerning the relationships between the four variables.

As is seen from Figure 4.3 the VARMA model (equations (4.2) to (4.5)) is a recursive equation system. By rearranging the variables we get parameter matrices $\boldsymbol{\phi}$ and $\boldsymbol{\Theta}$ which are upper triangular, i.e., the equation system is recursive. Since we have no feedback in the system we could have estimated equations (4.2) to (4.4) by a transfer function model for each equation seperately and equation (4.5) by a univariate ARMA model. This way of estimating equations (4.2) to (4.5) would be inefficient, however, since simultaneous estimation is more efficient. Furthermore, it was not evident that we had a recursive system until we already had estimated the VARMA model.

5. Identification and Estimation of a VAR Model

In Section 4, where we identified a VARMA model for our data, we saw that a VAR(2) model was not inconsistent with the pattern of cross correlation matrices and partial autoregression matrices. Consequently we choose a VAR(2) model as the best VAR model for our four original variables. The VAR(2) model was fitted using the exact maximum likelihood method. More precisely, we use the model

$$(\mathbf{I} - \boldsymbol{\phi}_1\mathbf{B} - \boldsymbol{\phi}_2\mathbf{B})\mathbf{Z}_t = \boldsymbol{\delta} + \mathbf{a}_t. \quad (5.1)$$

The parameter estimates for the fit of the unrestricted model (5.1) are shown in Table 5.1, where the figures within parentheses are the absolute t-values for the parameter estimates.

The fit of the unrestricted model (5.1), like the unrestricted model (4.1), requires estimation of four constants and 32 AR terms, in all 36 parameters. A more parsimonious model is obtained by setting to zero those AR coefficients whose estimates are small

Table 5.1. Estimation results for the unrestricted model (5.1)

δ	ϕ_1				ϕ_2			
10.67	.12	−.11	1.54	.15	.27	.13	−.76	−.10
(3.4)	(.9)	(1.5)	(5.2)	(5.4)	(2.4)	(1.6)	(2.5)	(3.4)
6.22	.52	.63	−.93	−.18	−.28	−.08	.45	.11
(1.3)	(2.7)	(5.8)	(2.1)	(4.4)	(1.7)	(.7)	(1.0)	(2.4)
2.08	−.01	.05	.71	.04	.02	.05	−.07	−.04
(1.8)	(.2)	(2.0)	(6.5)	(3.6)	(.4)	(1.6)	(.6)	(4.0)
−1.54	−.54	.01	.19	.84	1.32	.61	−1.69	−.14
(.1)	(1.1)	(.1)	(.2)	(8.2)	(3.0)	(2.0)	(1.4)	(1.2)

compared to their standard errors. This was done for model (5.1) in several successive steps. The parameter estimates for the fit of the final restricted model (5.1) are shown in Table 5.2 where all estimated parameters are significant. In the final restricted model we have to estimate only 12 AR terms compared with the unrestricted model where we had to estimate 32 AR terms.

The final restricted model (5.1) fits the data almost as well as the unrestricted model. The difference in −2*(log likelihood) between the restricted and unrestricted model is 32.6, which gives a p-value slightly less than 5% for a chi-square distribution with degrees of freedom equal to the number of parameters set to zero, i.e., $32 - 12 = 20$.

As the final step in the model building, a diagnostic check was performed by calculating the cross correlation matrices of the residuals. Table 5.3 shows the pattern of residual cross correlations for the final restricted VAR(2) model. It is very similar to Table 4.6 where the pattern of residual cross correlations is given for the final restricted VARMA(1, 1) model (4.1). However, the only difference, the negative entry in Table 5.3 for lag 1, indicates that the restricted VAR(2) model is slightly inferior to the restricted VARMA(1, 1) model.

We now turn to the causal implications of the estimated VAR(2) model. The final restricted fitted VAR(2) model (5.1), with estimated parameters given in Table 5.2, explains the relationships among the four variables in terms of the four-equation system

$$F_t = 13.9 + 1.53N_{t-1} + 0.15R_{t-1} - 0.05R_{t-2} + a_{Ft} \quad (5.2)$$

$$M_t = 3.30 + 0.59F_{t-1} + 0.54M_{t-1} - 1.06N_{t-1} - 0.13R_{t-1} + a_{Mt} \quad (5.3)$$

$$N_t = 1.82 + 0.09M_{t-1} + 0.70N_{t-1} + 0.04R_{t-1} - 0.05R_{t-2} + a_{Nt} \quad (5.4)$$

$$R_t = 16.9 + 0.74R_{t-1} + a_{Rt}. \quad (5.5)$$

Equations (5.2) to (5.5) for the VAR(2) model are essentially similar to equations (4.2) to (4.5) in Section 4 for the VARMA(1, 1) model. Especially, we observe that the estimates of the corresponding parameters are approximately equal. The main difference is that in the VAR model we have two significant relationships which are not present in the VARMA model. These two relationships are: that past fertility influences adult mortality, and that past nuptiality influences adult mortality. That adult mortality depends on past fertility may partly be due to deaths from childbed fever. However, the coefficient of 0.59 is too

Table 5.2. Estimation results for the final restricted model (5.1)

δ	Φ_1				Φ_2			
13.91 (8.3)	.	.	1.53 (9.9)	.15 (5.9)	.	.	.	−.05 (2.2)
3.30 (.8)	.59 (3.9)	.54 (6.1)	−1.06 (3.2)	−.13 (4.4)	.	.	.	.
1.82 (2.5)	.	.09 (4.8)	.70 (12.1)	.04 (4.3)	.	.	.	−.05 (5.1)
16.85 (3.8)	.	.	.	.74 (10.9)	.	.	.	.

large for this to be the only explanation. There must also be some sort of spurious relationship, or it could be an effect of outliers. This is discussed further in Section 7.

It is also hard to explain why adult mortality seems to depend on past nuptiality. A possible explanation may be that this relationship in some way expresses the fact that married people have lower death rates than unmarried.

We summarize our findings concerning the relationships between the four variables in the arrow scheme in Figure 5.1.

For the VAR model, in contrast to the VARMA model, we thus get a significant relationship which is not easily explained. Another difference between the VAR model and the VARMA model is that since the VARMA(1, 1) model has only 9 significant parameters besides the constant terms, it is more parsimonious than the VAR(2) model with its 12 significant parameters.

6. Post-sample Forecasting Performances of the Two Approaches

Even though we use historical data, we will make a small comparison of the VARMA and the VAR models with regard to their post-sample forecasting performances. We will forecast the four variables in the period 1851–1859 where no great changes have occurred from the estimation period 1751–1850. Only one-step ahead forecasts for each of these nine years are considered, using the models derived from the estimation period.

The forecasting performance of the restricted VARMA(1, 1) model, equations (4.2) to (4.5), for the four variables is summarized in Table 6.1. Besides providing the expected forecast variance (given that the model is true, i.e., the noise variance in eq. (4.2) to (4.5), respectively), and the mean squared error (MSE), based on the differences between forecasts and the actual observations, the table also includes the relative

Table 5.3. Pattern of sample cross correlations matrices for the residuals of the final restricted VAR(2) model

Lag =	1				2				3				4				5				6			
	F	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*	*F*	*M*	*N*	*R*
F	.	−	.	.	.	.	.	.	.	.	.	−	.	.	.	.	.	.	+	+	.	.	.	.
M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.	.	.	.
R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

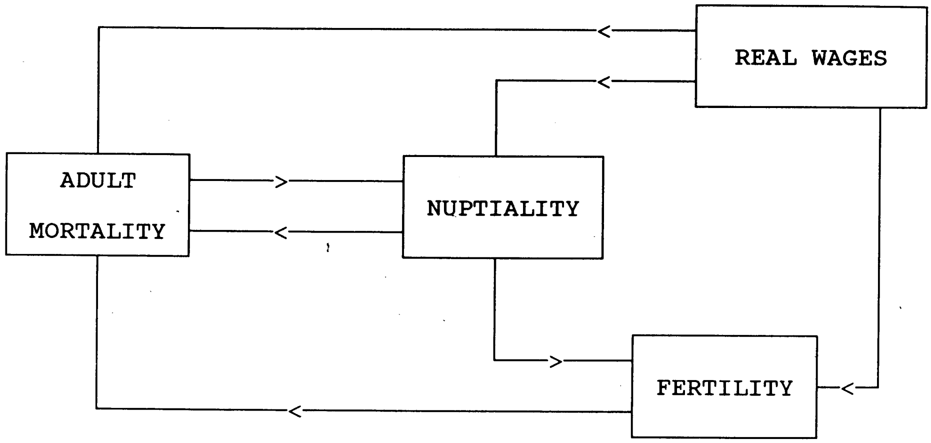

Fig. 5.1. Relationships between the four variables according to the VAR model

MSE's for the unrestricted VARMA(1, 1) model, Table 4.4, and for the restricted VAR(2) model, eq. (5.2) to (5.5).

Table 6.1 shows that the MSE's for all four variables are considerably smaller than expected. This is mainly due to the influence of the outliers in the estimation period which is discussed in more detail in the next section. Here, the expected forecast variance for mortality is 4.44, but when outliers are removed it reduces to 1.10. As expected the unrestricted VARMA model gives worse forecast than the restricted VARMA model. We notice that this is especially true for fertility. This is interesting because from the interpretation of the restricted VARMA model in Section 4, we had reason to believe that the fertility equation was underspecified, i.e., it should have included a mortality term. This was not verified here.

Finally, we find that the VAR model is inferior to the VARMA model even with regard to post-sample forecasting performance.

7. Influence of Outliers on the Two Approaches

Methods for identifying outliers in univariate time series models and transfer function models have been discussed by several authors, e.g., Abraham and Chuang (1989), Chang and Tiao (1983), and Tsay (1986). No method, however, has been proposed to identify outliers in VARMA models. If a VARMA model happens to be recursive then it can be rewritten as a set of transfer function models and one univariate ARIMA

Table 6.1. Forecasting performance of the restricted VARMA model

Variable	Expected forecast variance	MSE	MSE restricted VARMA/MSE unrestricted VARMA	MSE restricted VARMA/MSE restricted VAR
F	2.005	1.172	28%	79%
M	4.444	2.414	90%	80%
N	0.285	0.056	79%	90%
R	30.92	16.81	59%	97%

model. By using this description of a recursive VARMA model, methods for identifying outliers in transfer function models can be used to identify outliers in a recursive VARMA model.

The method for identifying outliers suggested by Chang and Tiao (1983) is implemented in SCA for univariate and transfer function models. Since the VARMA(1, 1) model in equations (4.2)–(4.5) is recursive, it is possible to rewrite the model as three transfer function models and one univariate model. Then we can use SCA to identify the outliers. Details are given in Hagnell and Salomonsson (1989b). Two outliers were found for F_t, seven for M_t, five for N_t, and none for R_t.

After having adjusted the identified outliers with their residuals, we proceeded with a multivariate time series analysis on the adjusted values. Another way to handle the outliers is to incorporate them explicitely in the model building, using intervention components representing the effects of the outliers, see Hagnell (1990) and Hagnell and Salomonsson (1989b). However, these two approaches for dealing with outliers lead to the same conclusions for our data.

The patterns of cross correlations matrices $\mathbf{R}(k)$ and partial autoregression matrices $\mathbf{P}(k)$ for the adjusted values were similar to those in Table 4.1 and Table 4.2, so we identified a VARMA(1, 1) model as the best fitting VARMA model and a VAR(2) model as the best fitting VAR model.

The final restricted VARMA(1, 1) model gives the following four-equation system

$$F_t = 14.2 + 0.30F_{t-1} + 1.07N_{t-1} + 0.13a_{R,t-1} + a_{Ft} \quad (7.1)$$

$$M_t = 4.11 + 0.64M_{t-1} - 0.08a_{R,t-1} + a_{Mt} \quad (7.2)$$

$$N_t = 2.76 + 0.06M_{t-1} + 0.78N_{t-1} + 0.06a_{R,t-1} + a_{Nt} \quad (7.3)$$

$$R_t = 16.8 + 0.74R_{t-1} + a_{Rt}. \quad (7.4)$$

Compared to models (4.2)–(4.5) for the original values, the residual variances are reduced, especially for mortality. In (7.1) we obtain a lower value for the coefficient of the N_{t-1} term. In general we get larger estimates for the diagonal parameters in models (7.2) and (7.3), whereas the off-diagonal parameters are smaller in magnitude, compared to (4.3) and (4.4), except the one for $a_{R,t-1}$ in (7.3).

The final restricted VAR(2) model gives the following four-equation system

$$F_t = 12.1 + 0.31F_{t-1} + 1.05N_{t-1} + 0.13R_{t-1} - 0.10R_{t-2} + a_{Ft} \quad (7.5)$$

$$M_t = 2.38 + 0.17F_{t-2} + 0.60M_{t-1} - 0.05R_{t-1} + a_{Mt} \quad (7.6)$$

$$N_t = 1.85 + 0.06M_{t-1} + 0.77N_{t-1} + 0.04R_{t-1} - 0.05R_{t-2} + a_{Nt} \quad (7.7)$$

$$R_t = 17.1 + 0.74R_{t-1} + a_{Rt}. \quad (7.8)$$

As a rule, the same changes occur here as for the VARMA(1, 1) model, i.e., we obtain larger estimates for the diagonal parameters and smaller ones for the off-diagonal parameters. For mortality, the N_{t-1} term is no longer significant and the explanatory effect from fertility changes from lag 1 to lag 2.

By adjusting for the outliers, we find that the VAR model no longer contains the effect of past nuptiality in the mortality equation which was difficult to interpret. The effect of past fertility in the mortality equation, which is missing in the VARMA model, is still significant but is reduced in size and changes from lag 1 to lag 2.

Both the VARMA and the VAR approaches are sensitive to outliers in the sense that we obtain higher estimates for the diagonal parameters and lower ones for the off-diagonal parameters when the outliers are adjusted. We would probably obtain this effect from removing the outliers for any reasonable model for the four series. For the VAR approach, however, in addition, a dubious dependence vanishes and another uncertain dependence is weakened. So here the VAR approach seems more sensitive to outliers than the VARMA approach.

8. Summary and Concluding Remarks

In this article we study the relationships between fertility, mortality for ages 20–50 years, nuptiality, and real wages in Sweden during the period 1751–1850. We use multivariate time series analysis methods, which allow an empirical determination of the relationships between the four time series, a considerable advantage where existing theory provides insufficient guidance. Two different approaches of multivariate time series analysis, the VARMA and the VAR approach, are used in order to make an empirical comparison between these two approaches.

Using vector ARMA modeling techniques, we identify a VARMA(1, 1) model for the four series. After first having estimated an unrestricted model, we obtain a final restricted model by setting insignificant parameters to zero. This restricted model gives a simple interpretation of the relationships between the four variables. Fertility is influenced by past real wages and past nuptiality. Adult mortality depends only on past real wages. Furthermore, nuptiality is influenced by both past real wages and past adult mortality. The real wages index is found to be exogenous, i.e., not dependent on any of the other three variables. The lag structure is also simple. Only one-year lags in the variables or their random shocks are present.

The simplest vector AR model, which gives an adequate fit to the data, is a VAR(2) model. The final restricted VAR(2) model gives, on the whole, the same interpretation as the final restricted VARMA(1, 1) model. However, the VAR(2) model implies that adult mortality is influenced by past nuptiality which is not easily interpreted. Nor does the VAR model give such a parsimonious model as the VARMA model.

In a comparison of the post-sample performances of the two approaches for nine one-step ahead forecasts, the VAR model was inferior to the VARMA model.

We also compare the two approaches with regard to their sensitivity to outliers. An analysis with adjusted values, where the outliers are removed, shows that the VAR approach is more sensitive to outliers than the VARMA approach. With adjusted values adult mortality is not any more influenced by past nuptiality in the VAR(2) model.

On the whole, the VARMA modeling approach gives a parsimonious and easily interpreted model of the relationships between the four variables. The interpretation of the model remains the same even if we adjust for outliers.

9. References

Abraham, B. and Chuang, A. (1989). Outlier Detection and Time Series Modeling. Technometrics, 31, 241-248.

Andersson, G. and Hagnell, M. (1989). Use of Intervention Analysis for Dealing with Outliers in Some Time Series Models for Economic-Historical Demographic Data. Research Report 1989: 4, Department of Statistics, University of Lund, Lund.

Bengtsson, T. and Broström, G. (1986). A Comparison of Different Methods of Analyzing Cycles in Population Economy. Meddelande från Ekonomisk-historiska institutionen Nr 46, Department of Economic History, University of Lund, Lund.

Bengtsson, T. and Jörberg, L. (1975). Market Integration in Sweden During the 18th and 19th Centuries. Spectral Analysis of Grain Prices. Economy and History, XVIII: 2, 93–106.

Bengtsson, T. and Ohlsson, R. (1985). The Standard of Living and Mortality Response in Different Ages. European Journal of Population, 1, 309–326.

Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis, Forecasting and Control. San Francisco: Holden-Day.

Bring, J. (1987). An Application of Transfer Function Models to the Relationship Between Death-Rates and Real Wages in Sweden 1751–1850. Research Report 1987: 1, Department of Statistics, University of Lund, Lund.

Carlsson, G. (1970). Nineteenth-Century Fertility Oscillations. Population Studies, 24, 413–422.

Carter, L.R. and Lee, R.D. (1986). Joint Forecasts of U.S. Martial Fertility, Nuptiality, Births, and Marriages Using Time Series Models. Journal of the American Statistical Association, 81, 902–911.

Chang, I. and Tiao, G.C. (1983). Estimation of Time Series Parameters in the Presence of Outliers. Technical Report 8, Statistical Research Center, University of Chicago, Chicago.

Eckstein, Z., Schultz T. P., and Wolpin, K.I. (1985). Short-Run Fluctuations in Fertility and Mortality in Pre-Industrial Sweden. European Economic Review, 26, 295–317.

Fackler, J.S. and Krieger, S.C. (1986). An Application of Vector Time Series Techniques to Macroeconomic Forecasting. Journal of Business and Economic Statistics, 4, 71–80.

Granger, C.W.J. and Newbold, P. (1986). Forecasting Economic Time Series. Orlando: Academic Press.

Hagnell, M. (1990). Intervention Analysis in Multivariate Time Series Models: An Application Using a Package for VARMA Models. Short communications, COMPSTAT 1990, Dubrovnik.

Hagnell, M. and Salomonsson, A. (1989a). Death Rates and Real Wages: An Analysis of Granger Causality with Post Sample Data and Different Forecast Horizons. Research Report 1989: 1, Department of Statistics, University of Lund, Lund.

Hagnell, M. and Salomonsson, A. (1989b). Influence of Outliers and Use of Intervention Analysis in Multivariate Time Series Models: An Application to Economic-Historical Data. Research Report 1989: 9, Department of Statistics, University of Lund, Lund.

Jörberg, L. (1972). A History of Prices in Sweden 1732–1914. Lund: Glerup.

Larsen, U. (1986). A Multivariate ARIMA Analysis of the Determinants of Short Term Fluctuations in Nuptiality in Sweden from 1751 to 1913. Research Report 1987: 1, Department of Statistics, University of Lund, Lund.

Lee, R.D. (1975). Natural Fertility, Population Cycles and the Spectral Analysis of Births and Marriages. Journal of the American Statistical Association, 70, 295–304.

Lee, R.D. (1981). Short Run Fluctuations in Vital Rates, Prices and Weather. In E.A. Wrigley and R. Schofield (eds.), Population Trends and Early Modern England, Cambridge, MA: Harvard University Press.

Liu, L.M. and Hudak, G.B. (1986). The SCA Statistical System: Reference Manual for Forecasting and Time Series

Analysis. De Kalb, IL: Scientific Computing Associates.

McCleary, R. and McDowall, D. (1984). A Time Series Approach to Causal Modeling: Swedish Population Growth, 1750–1849. Political Methodology, 10, 357–375.

Sims, C.A. (1980). Macroeconomics and Reality. Econometrica, 48, 1–48.

Tiao, G.C. and Box, G.E.P. (1981). Modeling Multiple Time Series with Applications. Journal of the American Statistical Association, 76, 802–816.

Tsay, R.S. (1986). Time Series Model Specification in the Presence of Outliers. Journal of the American Statistical Association, 81, 132–141.

Received July 1990
Revised November 1991