Structural Econometric Modeling and Time Series Analysis

F. C. Palm

Economische Faculteit
Vrije Universiteit
Postbus 7161
1007 MC Amsterdam

ABSTRACT

We discuss the Structural Econometric Modeling and Time Series Analysis (SEMTSA) approach put forward by Zellner and Palm, which provides a synthesis of econometric and time series methods in modeling economic time series. The approach aims at giving guidance for checking the data admissibility of the dynamic specification of a model in its various forms, in particular the transfer function form and the final equation form. We review the SEMTSA approach, discuss recent developments, and briefly compare the SEMTSA with other methodologies for econometric modeling. Finally some remarks are made about problems that remain to be solved.

1. INTRODUCTION

Since the publication in 1970 of the influential book by Box and Jenkins [4], univariate ARIMA time-series models have been applied on a large scale in forecasting. Initially, these models were seen as naive ad hoc alternatives for existing econometric models, which are often large systems of nonlinear dynamic equations. Comparisons of the forecast performance of univariate time-series models and multiequation macroeconomic models were undertaken and surprisingly led to the conclusion that in short- and medium-term prediction, the univariate ARIMA schemes frequently outperformed the causal econometric models.

This finding raised a number of methodological questions. What is the relationship between ARIMA models and causal econometric models? What is the role of time-series methods in econometric model building? These and other related issues were addressed by Zellner and Palm [35, 36] and have since been on the agenda of many research projects (see e.g. [30]).

In this paper, we discuss the "structural econometric modeling and time-series analysis" (SEMTSA) approach put forward by Zellner and Palm
[35, 36], hereafter referred to as ZP, which aims at a synthesis of econometric and time-series methods in modeling economic time series. In Section 2, we review the SEMTSA approach. We also discuss developments which have taken place in this area since the beginning of the seventies and briefly compare the SEMTSA with other methodologies for econometric modeling. Finally, Section 3 contains some concluding remarks about problems that remain to be solved.

2. THE SEMTSA APPROACH

To outline the SEMTSA approach, we consider the following linear dynamic simultaneous-equation model (SEM):

\[ H_{11}(L) y_t + H_{12}(L) x_t = F_{11}(L) \varepsilon_{1t}, \]

where \( y_t \) and \( x_t \) are vectors of endogenous and exogenous variables; \( \varepsilon_{1t} \) is a vector of disturbances; and \( H_{11}(L) = \sum_{i=0}^{r} H_{11i} L^i \), \( H_{12}(L) = \sum_{i=0}^{k} H_{12i} L^i \), and \( F_{11}(L) = \sum_{i=0}^{q} F_{11i} L^i \), whose elements are finite polynomials in the lag operator \( L \) of (maximum) degree \( p \), \( r \), and \( q \) respectively. These matrices do not have a common nonunimodular left factor. The matrix \( H_{11}(L) \) is nonsingular, i.e., \( \text{det} H_{11}(z) \neq 0 \), where \( z \) is a complex variable. Further, we assume that \( \varepsilon_{1t} \) is serially independent and normally distributed with mean zero and identity covariance matrix. The last assumption is a matter of normalization, as \( F_{11}(0) \) is not necessarily an identity matrix. We assume strict exogeneity of \( x_t \) with respect to \( \varepsilon_{1t} \), i.e., \( x_t \) and \( \varepsilon_{1t} \) are independent for all \( t \) and \( t' \). Although strict exogeneity is not always needed, it has the advantage of preserving the independence of the \( x_t \)'s and the disturbances under linear transformations of the model. For a detailed discussion of the properties of ARMAX systems we refer the reader to [5].

2.1. Different Forms of the Model

The system (2.1) with \( H_{11}(0) \) a nonsingular matrix with unknown coefficients is usually called the structural form of the simultaneous-equation model. The parameters of the structural form often have an interpretation in terms of economic theory and are (implicitly) assumed to be invariant to changes in the exogenous variables \( x_t \). Without additional prior restrictions, the structural parameters are not identified. The reduced form of (2.1) corresponds to the solution of (2.1) for the vector \( y_t \), obtained by premultiplying (2.1) by \( H_{11}^{-1}(0) \).

The set of transfer-function (TF) equations associated with (2.1) is obtained through premultiplication of (2.1) by the adjoint matrix \( H_{11}^*(L) \) of
$H_{11}(L)$:

$$[H_{11}(L)]y_t = -H_{11}^*(L)H_{12}(L)x_t + H_{11}^*(L)u_{1t}, \quad (2.2)$$

where $|H_{11}(L)|$ is the determinant of $H_{11}(L)$, a scalar polynomial in $L$ of finite degree ($\leq mp$) and $u_{1t} = F_{11}(L)e_{1t}$. As pointed out by ZP [35], the autoregressive polynomials of the TFs in (2.2) are identical provided $H_{11}(L)$ has no special structure such as that of a diagonal, block diagonal, triangular, or block triangular matrix. Further, the equations in (2.2) form a system of seemingly unrelated dynamic regression equations, each including one endogenous variable, all exogenous variables, and a moving-average disturbance. Each element of the vector $H_{11}^*(L)u_{1t}$, say the $i$th one, can be represented as a moving average $\theta_i(L)v_{it}$ of order smaller or equal to $q^* = (m - 1)p + q$, with $v_{it}$ a white noise of mean zero and variance $\sigma_v^2$ (see also [12] on this point).

Strictly speaking, the TFs are given by the matrix polynomials $-H_{11}^*(L)H_{12}(L)$ and $H_{11}^*(L)F_{11}(L)$ respectively. The inverse of $H_{11}(L)$ exists, and the TFs are one-sided polynomials in $L$, as $\det H_{11}(z) \neq 0$ for $|z| \leq 1$ (causality assumption).

Under the additional assumption that the exogenous variables $x_t$ are generated by a multivariate ARMA model

$$H_{22}(L) \quad x_t = u_{2t}, \quad (2.3)$$

where $H_{22}(L) = \sum_{l=0}^k H_{22l}L^l$, $u_{2t} = F_{22}(L)e_{2t}$, $F_{22}(L) = \sum_{l=0}^k F_{22l}L^l$, and $e_{2t}$ are normally distributed and serially independent with mean zero and identity covariance matrix, the final equations for $x_t$ are obtained by premultiplying (2.3) by the adjoint matrix $H_{22}^*(L)$ of $H_{22}(L)$:

$$[H_{22}(L)]x_t = H_{22}^*(L)u_{2t}. \quad (2.4)$$

The determinant of $H_{11}(L)$ is such that $|H_{22}(z)| \neq 0$, $|z| \leq 1$. After substitution of (2.4) into (2.2), we get

$$[H_{22}(L)]|H_{11}(L)]y_t = -H_{11}^*(L)H_{12}(L)H_{22}^*(L)u_{2t}$$

$$+ |H_{22}(L)|H_{11}^*(L)u_{1t}, \quad (2.5)$$

which is called the set of final equations (FEs) for the endogenous variables.
Notice that the disturbance term of each equation in (2.5) can be represented as a moving average of order smaller than or equal to \( \max((m-1)p+r+(k-1)s+f; ks+(m-1)p+q) \) in one variable. Also, except when \( H_{11}(L) \) and/or \( H_{22}(L) \) are matrices with triangular or diagonal structure or when other coincidental situations (e.g. common factors on both sides of an equation) occur, all endogenous variables will have the same autoregressive polynomial in the FEs.

The equations in (2.5) form a system of seemingly unrelated ARMA equations for the endogenous variables \( y_t \). Cross-equation correlations only appear through the cross-correlations of the disturbances. These univariate ARMA models are derived from the joint process for \( z_t = (y'_t, x'_t)' \), and therefore they are not "naive" and "ad hoc" alternative specifications, a point already emphasized by ZP [35]. The subscript \( t \) distinguishes the random vector \( z_t \) from the complex variable \( z \). In a statistical sense, the set of FEs can be interpreted as the set of marginal processes, given the past of the corresponding variable only.

It is worthwhile to notice that in (2.1), \( y_t \) is not necessarily stationary. The vector \( y_t \) is stationary conditionally on the present and past of \( x_t \). When specifying (2.3), \( x_t \) is usually assumed to be stationary. Alternatively, if the \( d \)th difference of \( x_t \) is stationary, we can take \( H_{22}(L) = H_{22}(L)(1-L)^d \) with \( \det H_{22}(z) \neq 0 \) for \( |z| < 1 \).

When \( z_t = (y'_t, x'_t)' \) is weakly stationary with mean zero, using the generalized version of Wold's decomposition theorem (see [13, p. 158]), we can represent the process \( z_t \) as

\[
\begin{align*}
    z_t &= \begin{pmatrix} C(L) & e_t \\ \end{pmatrix},
    \quad (m+k) \times 1
\end{align*}
\]

where \( C(L) \) is an infinite polynomial matrix operator in \( L \), holomorphic within the unit circle, with \( i, j \)th element given by

\[
    c_{ij}(L) = \sum_{l=0}^{\infty} c_{ijl} L^l,
\]

and \( e_t \) is a vector white noise with mean zero and identity covariance matrix.

A system like (2.1) and (2.3) arises when \( z_t \) has an ARMA representation, i.e. when

\[
    C(L) = H^{-1}(L) F(L)
\]
with $H(L)$ and $F(L)$ block-triangular and block-diagonal respectively:

$$
H(L) = \begin{bmatrix}
H_{11}(L) & H_{12}(L) \\
0 & H_{22}(L)
\end{bmatrix}
\quad \text{and} \quad
F(L) = \begin{bmatrix}
F_{11}(L) & 0 \\
0 & F_{22}(L)
\end{bmatrix}.
$$

(2.9)

Also, the parameters of the ARMA representation $H(L)z_t = F(L)\varepsilon_t$ do not necessarily have an interpretation in terms of economic theory. In fact, $z_t$ may be generated by some nonlinear structural model and still admit an ARMA representation. The normality assumption is then probably inappropriate.

To illustrate this point, we take the following example of a nonlinear model:

$$
y_t = \beta x_t^2 + \varepsilon_{1t}, \quad \varepsilon_{1t} \sim N(0, \sigma_1^2),
$$

(2.10a)

$$
x_t = (1 - \theta L)\varepsilon_{2t}, \quad \varepsilon_{2t} \sim N(0, \sigma_2^2),
$$

(2.10b)

with $\varepsilon_{1t}$ and $\varepsilon_{2t}$ independent for all $t$ and $t'$. Conditionally on $x_t$, $y_t$ is independently normally distributed. Equation (2.10a) could have an economic interpretation. Marginally with respect to $x_t$, the process of $y_t$ is MA(2). For a discussion of nonlinear transformations of ARMA processes, we refer to [7]. The joint process for $y_t$ and $x_t$ has an MA(2) representation with nonzero mean.

A few more comments on the model (2.1) and (2.3) are in order. By imposing the structure (2.9) on $H(L)$ and $F(L)$, there is unidirectional Wiener-Granger causality from $x_t$ to $y_t$ only. The assumptions that $F_{12}(L) = 0$ and that the contemporaneous correlations of the disturbances in (2.1) and (2.3) are zero, are not required for unidirectional causality. As stated above, $x_t$ is strictly exogenous with respect to $\varepsilon_{1t}$. If the parameters of (2.1) and (2.3) are variation free, $x_t$ is also weakly exogenous for the parameters in (2.1) and strongly exogenous with respect to $\varepsilon_{1t}$. For more details on exogeneity concepts, we refer the reader to [8].

The SEMTSA approach aims at giving guidance for checking the data-admissibility of the dynamics of a model which is assumed to be a reasonable approximation of the data generation process (DGP). The approach is complementary to other model evaluation procedures and tests for disturbance autocorrelation, structural stability, functional form, etc. A model should be data coherent in its different forms. As the TF equations are dynamic regression equations, the lag length and the parameter value for the individual equations in (2.2) can be determined equation by equation from the data. The empirical results can be subsequently confronted with the TF specification derived from an initial model in the form of (2.1).
Any incompatibility between the results of the empirical analysis of the individual transfer functions and those derived from the tested structural form is an indication of a misspecification in one or both forms of the model and can be used to reformulate the model. Examples of how to respecify the model when an incompatibility is detected are given by ZP [35, 36]. The transfer functions can also be used to study the dynamic properties of a model. The roots of the characteristic equation associated with (2.1) are obtained by solving \( |H_{1i}(z^{-1})| = 0 \). These roots can be calculated from the estimated autoregressive polynomials of the transfer functions. They determine the time path of the expectations of \( y_t \) given the exogenous variables.

Similarly to what has been said about the transfer functions, any incompatibility between the results of the empirical analysis of the individual final equations (e.g. along the lines proposed by Box and Jenkins [4]) and those for the structural model is an indication of a misspecification of the system of final equations (2.5) and/or of the finally accepted structural form of the model. The role of the empirical analysis of the final equation form for the formulation of a simultaneous-equation model has been discussed and illustrated by ZP [35].

2.2. An Example

To illustrate these points, we take a simple three-equation macroeconomic model for consumption expenditures \( c_t \), aggregate income \( y_t \), and "autonomous" expenditures \( a_t \):

\[
\begin{align*}
\phi(L)c_t &= \alpha + \beta(L)y_t + \epsilon_{1t}, \\
y_t &= c_t + a_t, \\
\Delta a_t &= \gamma + \epsilon_{2t},
\end{align*}
\]

with \( \epsilon_{it} \sim N(0, \sigma^2) \), \( i=1,2 \), \( \epsilon_{1t} \) and \( \epsilon_{2t} \) independent for all \( t \) and \( t' \); \( \Delta = 1 - L \); \( \phi(L) = (1 - \phi_1 L - \cdots - \phi_p L^p) \) with \( \phi(z) \neq 0 \) for \( |z| < 1 \); and \( \beta(L) = \beta_0 + \beta_1 L + \cdots + \beta_p L^p \). Assume that there is specification uncertainty concerning the order of \( \phi(L) \) and \( \beta(L) \). The TF equations for \( c_t \) and \( y_t \) are

\[
\begin{align*}
\delta(L)c_t &= \beta(L)a_t + \alpha + \epsilon_{1t}, \\
\delta(L)y_t &= \phi(L)a_t + \alpha + \epsilon_{1t},
\end{align*}
\]

with \( \delta(L) = \phi(L) - \beta(L) \).

The polynomials \( \phi(L) \) and \( \beta(L) \) could be determined straightforwardly from the estimates of each individual TF equation. Notice that each equation in (2.12) is a dynamic regression model with the same white-noise disturbance
and a nonstationary regressor $a_t$. One can apply a consistent order determination criterion for ARMAX models to each equation separately to estimate $p$, $r$, $a$, and the $\phi_i$’s and $\beta_j$’s consistently. The computational advantage of such a procedure is quite obvious in the present example. Although a single-equation procedure is not fully efficient, it is computationally very attractive and it may yield strongly consistent estimates of the integer-valued parameters $p$ and $r$. Alternatively, an initial model with say $\phi(L) = 1 - \phi L$ and $\beta(L) = \beta$ might be rejected by the information in the data. Also, if the initial model is rejected, the results of the TF analysis usually yield an indication of how the model has to be respecified to become data admissible.

Similarly, the FE$s for c_t$ and $y_t$ associated with (2.11) are

$$\delta(L) \Delta c_t = \alpha' + \beta(L) e_{zt} + \Delta \epsilon_{tt},$$

$$\delta(L) \Delta y_t = \alpha'' + \phi(L) e_{zt} + \Delta \epsilon_{tt}$$

with $\alpha' = \beta(0) \gamma$, $\alpha'' = \phi(0) \gamma$. The order of the polynomials in (2.13) could be determined empirically from the estimated autocorrelation functions for $\Delta c_t$ and $\Delta y_t$ along the lines of the Box-Jenkins [4] approach. Alternatively, a consistent model selection criterion could be used to determine the order of the univariate models for $\Delta c_t$ and $\Delta y_t$ (see e.g. [14]). The third equation in (2.11) can be checked directly in this way, as it is already in FE form. If the FE$s associated with the initial model are not in agreement with the information in the data, the model has to be reformulated such that it becomes data admissible.

Also the degree of nonstationarity of the variables in the model can be investigated. In the model (2.11), all three variables are integrated of order one, which we denote as $I(1)$. This again is a testable hypothesis. It could be checked using a test for unit roots (see e.g. [22, 6]).

The analysis of the final equations as a means for checking the dynamics of a simultaneous-equation model has been pursued by Anderson et al. [3], Palm [24], Prothero and Wallis [25], Trivedi [29], Wallis [30], Wallis [31] for modeling seasonality, and Zellner and Palm [36] among others. Zellner [34] discusses some of the statistical problems associated with the SEMTSA approach that require further research.

When the implications of a linear structural model are in agreement with the results of the empirical analysis of the transfer functions and final equations, the model can be used to predict the postsample period. If postsample data are available, the predictive performance of the structural form can be compared with that of the transfer functions and/or the final equations. If it predicts less well than the transfer functions or the final equations, there are good reasons for believing that the structural model is
misspecified. When the structural model is nonlinear, forecasts from its ARMA representation may be more accurate in a mean-square-error sense than some forecasts from the nonlinear structural form (see [33] for an example). When all three forms predict badly, the model either is misspecified or has undergone a structural change during the postsample period.

The procedures outlined here ought to be considered as a guideline for modeling systems of dynamic equations. Although an analysis under full information is applicable to small and medium-size models, in practice, because of specification uncertainty and the dimension of the model, the analysis will often have to be pursued under limited information, as discussed above for single TF and final equations.

2.3. The Appropriateness of the Linear Dynamic Model

It is quite important to realize that economic time series frequently cover a long period over which part and perhaps most of the structure of the model has been subject to change. One might then be inclined to drop the constant-parameter model and adopt a variable-coefficient framework instead. For instance, Sims [26] estimated three- and six-variable vector autoregressions in which the parameters vary according to a random walk. It is however questionable whether a model with drifting coefficients is a sensible alternative, or at any rate the only one, to constant-parameter models. Experience from empirical research shows that many economic relationships are remarkably stable over time. For instance, a stable relation between aggregate consumption expenditures on nondurables and aggregate disposable income has been found for several countries (see e.g. [15]) for a rather long period after World War II. However, unconditionally on income, the process for consumption expenditures seems to have undergone structural changes. Such a finding has been a major argument in econometrics for adopting conditional models instead of closed or marginal models. The system (2.1) is a conditional model for $y_t$ given $x_t$. If the parameters of (2.1) are invariant, the TF parameters are constant too, so that TF equation analysis is perfectly valid when $x_t$ is superexogenous for the parameters in (2.1) [i.e., $x_t$ is weakly exogenous for the parameters in (2.1), and the conditional model is structurally invariant for a change in the distribution of $x_t$].

In applications, however, it will frequently be necessary to specify the process of $x_t$ as the following examples show:

1. To forecast $y_t$ in the future, forecasts for $x_t$ are needed. The FEIs for the exogenous variables (2.4) or the system (2.3) can be used to generate future values for the exogenous variables.

2. To simulate the model (2.1), assumptions on the path or on the stochastic process of $x_t$ are required.
(3) The FEbs can be used to form expectations of the current or future values of the exogenous variables in models with rational expectations (see e.g. [32]).

(4) If the exogenous variables \( x_t \) are measurements with error, it may be appropriate to specify their process jointly with that of the latent variables (see [18, 2] for system identification from noisy data, and [9] for dynamic factor analysis).

(5) Model encompassing (see e.g. [16]) often requires marginalization with respect to conditioning variables that are not included in all models under consideration.

There are important instances in which \( x_t \) has to be modeled anyway although it is weakly exogenous with respect to some parameters. The analysis of the FEbs for the exogenous variables may lead to the detection of structural changes in the process (2.3). Therefore, this analysis is recommended whenever the exogenous variables have to be modeled. If the model for a superexogenous \( x_t \) is stable, the FEbs for the endogenous variables (2.5) are stable too over the sample period.

Quite obviously, some types of exogenous variables (for instance a deterministic variable) cannot be modeled by an ARIMA process. They cannot be eliminated and have to be carried along in the analysis of the FEbs for the endogenous variables.

Finally, we would like to point out that univariate ARIMA processes might also be fairly stable over time. For example, according to the efficient-market hypothesis, stock returns follow approximately a random walk, a result which has been found to hold for many stock-market series with remarkable parameter stability over long periods. Also, the models for many highly aggregated economic time series have been found to be approximately a random walk. This finding might result from the fact that aggregation flattens the spectrum of the differenced series, which is ARMA with possibly high-order lags but small coefficients that cannot be detected from the data. Structural changes in the process of some components of the aggregate are therefore almost negligible at the aggregate level.

It should be clear from this discussion that the analysis of the dynamic properties of economic time series can yield insights which are useful for the construction of the model. The final model should be in agreement with the properties of the single series, a point also stressed by Granger [10].

In the past, several authors have questioned the need and appropriateness of moving-average disturbances in a linear dynamic SEM such as (2.1) and (2.3). The arguments against the use of MA disturbances are:

(1) the identification and estimation of the model are simpler when the disturbances are serially uncorrelated, and
(2) the interpretation of the MA parameters in terms of economic behavior
(i.e. in terms of conditional reactions of economic agents given the environ-
ment) is not trivial.

These arguments are sensible. In particular, it is possible that economic
agents base their decisions on a limited number of variables, and not on the
complete past of these variables as is sometimes assumed in infinite-
distributed-lag models. Then, at the microeconomic level, conditional models
with infinite lags (as implied e.g. by an MA disturbance) might not be
required. However, econometric models are usually transformations of some
(micro)economic model with unobserved or latent variables. The observed
variables are often aggregates over time, commodities, and/or agents, possi-
bly measured with errors, and subject to seasonal variation. The transforma-
tion of the theoretical model into a model for observables is formally equiva-

tent to marginalization with respect to the unobservables and frequently
leads to models with MA disturbances. An example which illustrates this point
is a univariate ARMA(p, q) model for the aggregate of two independent
AR(p_1) processes where p \leq p_1 + p_2 and q \leq \max\{ p_1, p_2 \} (see [12]). A
second example is the (additive) unobserved-component ARIMA model (see
e.g. [23]), where the trend, seasonal, and irregular components are ARI. The
marginal process for the series obtained by substituting the models for the
components is a restricted ARIMA scheme.

Therefore, there are good reasons for considering econometric models with
MA disturbances such as e.g. in (2.1). Of course, an invertible MA model can
be approximated by a high-order AR model. However, besides the approxima-
tion error, the AR model will suffer from a lack of parsimony and from
problems with the interpretation of the parameters.

As we have already pointed out above, the system (2.1) and (2.3) is
formulated in the levels of the variables. A class of models that link economic
equilibrium conditions with time-series model specifications are the error-cor-
rection models. Equilibrium conditions in economics often state that a pair of
variables \( y_{1t} \) and \( y_{2t} \) obeys the equation

\[ y_{1t} = y_{2t}, \quad (2.14) \]

or that a linear combination of two or several variables equals zero in
equilibrium:

\[ \alpha'y_t = 0, \quad (2.15) \]

where \( \alpha \) is a vector of constant coefficients. The variable \( w_t = \alpha'y_t \) is the
equilibrium error. If all the elements of $y_t$ are integrated of order 1 [$I(1)$], one might expect $w_t$ to be $I(0)$. The vector $y_t$ is then said to be cointegrated.

Assuming that $H_{11}(L) = 0$, the model (2.1) is specified as an error-correction model

$$
H_{11}(L)(1 - L)y_t = \beta w_{t-1} + F_{11}(L)e_{1t},
$$

(2.16)

where $H_{11}(L) = H_{11}(L)(1 - L) - \beta \alpha' L$ with $H_{11}(z) \neq 0$, $|z| \leq 1$, $H_{11}(0) = I$, and where $\beta$ is an $m$-dimensional vector. When more than one $\alpha$ exists such that $\alpha' y_t$ is $I(0)$, $w_{t-1}$ and $\beta$ will be a vector and a matrix respectively. Using the results of Theorem 1 in [11], the FEs associated with (2.16) can be shown to be ARIMA processes, the order of which can be checked empirically along the lines discussed above. In particular, it is important to determine accurately the order of integration of the individual series $y_{it}$. Notice finally that the system (2.16) can be interpreted as a closed dynamic model with a stationary latent variable $w_{t-1}$ (with $\alpha$ unknown), which is generated by the system itself.

2.4. Order Determination of ARIMA Models

The determination of the lag length is a major problem in the analysis of FE and TF equations. Several new model selection criteria have been proposed recently. It is probably sufficient to mention Akaike’s [1] information criterion (AIC). More important is the result that some of these criteria [e.g. the Bayesian information criterion (BIC)] are strongly consistent for the order of the ARMA models. Strong consistency is a rather strong property. But what are the small-sample properties of order selection criteria and procedures currently in use?

In Figure 1, we give the cumulative frequency with which a given ARMA model has been selected in a simulation study of 200 runs. A discussion about order selection and detailed simulation results are given in [27]. The data are generated by the following two models:

Model 1: $\text{AR}(1)$: $(1 - 0.9L)y_t = \varepsilon_t,$

Model 2: $\text{ARMA}(1,1)$: $(1 - 0.9L)y_t = (1 - 0.4L)\varepsilon_t,$

(2.17)

with $\varepsilon_t \sim \mathcal{N}(0, 1)$ and sample size $T = 50$. In Figure 1, the cumulative frequency is plotted against the distance $\delta$ between the true model and some alternative model. The measure for the distance put forward by Snee [27] is the difference between residual variance ratios of the true model (with
Fig. 1. Cumulative frequency of selecting an ARMA model: (a) upward testing, (b) downward testing.
parameters $\phi$ and $\theta$) and some alternative model with parameters $\hat{\phi}$ and $\hat{\theta}$:

$$
\delta[\theta, \theta], \hat{\theta}] = \frac{E \hat{e}_t^2 - E e_t^2}{E e_t^2}.
$$

(2.18)

The parameters $\hat{\phi}$ and $\hat{\theta}$ are the probability limits of the maximum-likelihood estimates of the parameters of the alternative model given the true model. For (2.17), $\delta = T(E \hat{e}_t^2 - 1)$.

Notice that the distance (2.18) is nonnegative. It is zero when the true model is nested in the alternative model. For instance, the distance between model 1 and MA($q$) models, $0 \leq q \leq 3$, is 213, 66.6, 30, 6.76 respectively for $T = 50$, and zero for all models with a first-order AR coefficient. For model 2, the distance to MA($q$) schemes, $0 \leq q \leq 3$, is 65.8, 26.9, 14.7, 9.1 respectively. The distance to AR($p$) models, $0 \leq p \leq 3$, is 65.8, 4.5, 0.67, 0.11. The measure $\delta$ in Figure 1 is a generalized distance in the sense that a negative distance value has been assigned to models which are more general than the DGP, with one unit of distance for each abundant parameter.

Four selection procedures are considered. From the set of ARMA($p, q$) schemes with $0 \leq p \leq 3$ and $0 \leq q \leq 3$, we select the model

1. with the smallest value for Akaike's [1] criterion light solid curves (−),
2. with the smallest value for Schwarz's [28] criterion (−−).
3. using sequential upward (downward) testing based on a likelihood-ratio criterion with a nominal size of 5% at each step (−).
4. using sequential upward (downward) testing based on the Ljung-Box [21] residual autocorrelation statistic (with 12 residual autocorrelations) and a nominal size of 5% at each step (−−−).

The upward procedure starts with a white-noise model ($p = q = 0$) and tests step by step against a model with one extra parameter. First an extra AR parameter is introduced. If it is significant, an extra MA parameter is included. Downward testing proceeds in the opposite direction, starting with a model with $p = q = 3$. The details are given in [27].

As expected, Akaike's [1] criterion leads often to overestimation of the order of the model. Schwarz's [28] criterion and the upward and downward procedures more frequently select the correct model. Most importantly, the frequency with which a model is selected strongly depends on its distance to the true model. All procedures often get in close to the true model, but fail to select it precisely. To put it differently, in finite samples, the data can fairly well discriminate between the true model and models for which the residual variance or the one-step-ahead prediction error variance is much larger. In a way, this means that the data are more informative on the TF between $y_t$ and
past $\epsilon_i$'s than on the integer parameters $p$ and $q$, as the measure in (2.18) can be expressed as $\delta = T \sum_{h}^\infty \hat{\Psi}_h$, with $\hat{\Psi}_h$ the coefficient of $L^h$ in $\hat{\Psi}(L) = \phi^{-1}(L) \hat{\theta}(L) \hat{\phi}(L) \phi^{-1}(L)$.

This finding implies that the results from the empirical analysis of FEAs ought to be used to eliminate the structural models for which the associated FEAs are clearly rejected by the data. (It is much more difficult to accurately determine the degrees of the lag polynomials from the data.) In this way, we can narrow down the set of structural models to that for which the dynamic properties of the individual series are data admissible. Similar conclusions can be expected for TF equations which are dynamic regression equations. The simulation results obtained by Kiviet [19, 20] for model specification and misspecification tests point into this direction. The FE and TF analyses will lead us to reject a number of models on these grounds, but the set of data-admissible models can hardly be expected to be a singleton.

3. CONCLUSIONS

SEMTSA gives guidelines on how to take account of the time-series properties of the data when formulating econometric models. The aim is to improve the models by making them data admissible in terms of the properties of the set of marginal processes associated with a multivariate system. The approach is a complement rather than a substitute for other modeling methodologies (such as put forward e.g. by Hendry and Richard [17]). For instance, model encompassing can and ought to be applied to single FEAs and TF equations.

Further progress is expected from the theory of model reduction. Conditioning on $x_i$ is valid for (2.1) when $x_i$ is weakly exogenous for the parameters in (2.1) (assumed to be of interest). Alternatively, marginalization with respect to all $y_i$'s but one in (2.1) is legitimate and yields the TF form, but it usually implies a loss of information. Finally, there is a need for more research on the finite-sample properties of order estimation procedures for dynamic models.

REFERENCES

11. C. W. J. Granger, Co-integrated variables and error-correcting models, Report 83-13, Univ. of California, San Diego, 1983.