FORECASTING PRICES OF PAPER PRODUCTS: FOCUSING ON THE RELATION BETWEEN AUTOCORRELATION STRUCTURE AND ECONOMIC THEORY

RUNAR BRÄNNLUND, KARL-GUSTAF LÖFGREN AND SARA SJÖSTEDT[*]

ABSTRACT

In this paper we introduce a relatively recent forecasting technique, which here is used to forecast prices of Swedish forest products. The technique is, like vector autoregressive models (VAR), based on the idea that time series of prices and quantities from different sectors of the forest industry typically co-vary over time. A second aim is to test whether the Swedish forest industry act as a price-taker on the world market. If they do we show that current prices should be bad predictors of future quantities, and vice versa, which in turn implies that multivariate techniques do not necessarily yield better forecasts than univariate techniques. The results show that we can not reject the "small open economy" hypothesis, i.e. that Swedish producers are price takers. This result is founded on the econometric result that the inclusion of quantities do not significantly improve the forecasts of prices.

Keywords: forecasting technique, Swedish forest products, VAR-models.

~

INTRODUCTION

The forest industry is one of the branches of industry which shows a typical business cycle, although it is not clear how this cycle is related (lead or lag) to the general

[*] Runar Brännlund, Department of Economics, Umeå University, SE-901 87 Umeå, Sweden.

Karl-Gustaf Löfgren, Department of Economics, Umeå University, SE-901 87 Umeå, Sweden.

Sara Sjöstedt, Department of Mathematical Statistics, Umeå University, SE-901 87 Umeå, Sweden

The authors acknowledge comments and input of ideas from Professor Jan G de Gooijer, Faculty of Economics and Econometrics, University of Amsterdam, Assistant Professor Per Johansson, Department of Economics, Umeå University, and two anonymous referees.

business conditions. Both prices and quantities are rather volatile, and it is quite a challenge to forecast them, without fully understanding the underlying market structures. It is obviously very likely that the performance of a forecasting technique, and the underlying economic structure are closely related. This is, however, very seldom brought up explicitly.

The purpose of this paper is twofold. Firstly, we introduce a relatively recent forecasting technique, which here is used to forecast future prices of Swedish forest products. The technique is, like vector autoregressive models (VAR), based on the idea that time series of prices and quantities from different sectors of the forest industry are mutually correlated (co-vary) over time. It has the additional virtue of trying to divide the information inherent in the time series into ”important” and ”unimportant”, or alternatively, structural and non-structural. The former is used for forecasting purposes, while the latter is discarded. The signals, or information pieces, are aggregated into a combined signal to form ”important information” by weighting them in a manner such that the autocorrelation over time of the combined signal is maximized, subject to othogonality restrictions.

It should be intuitively clear that if the time series of the different sectors do not co-vary over time, there is no reason to believe that a multivariate forescasting technique based on this idea is superior to a univariate forecasting method, like ARIMA. It turns out that ARIMA forecasts frequently match, or even outperform, those of our multivariate technique, and this fact must contain important information about the underlying economic structure which generated the time series. We show that a small open economy assumption — the typical firm is a price taker in world markets — implies that current prices should be important predictors of future prices, and that quantities are likely to be good predictors of future quantities. However, current prices are bad predictors of future quantities, and vice versa. The second purpose of the paper is to formally test these conjectures.

Since business cycle phenomena typically are recurrent fluctuations of output and prices about trend and co-movements among other aggregate time series, it would be noteworthy if the joint signals, which constitute the maximum correlation over time, the so called maximum autocorrelation factors (MAFs), did not contain information about these co-movements, and, hence, are useful for forecasting purposes. This idea is, of course, too simple to be completely new. The idea on how to extract the important information for forecasting purposes is closely related to a paper by Box & Tiao (1977), where a canonical analysis of multiple time series is conducted, and where the components of the transformed autoregressive process are ordered from least to most predictable. Our forecasting idea is closely related to the contents of Sims (1981) and Reinsel (1983), where autoregressive index models are discussed.

A more general analysis of the statistical theory behind the MAFs can be found in Switzer & Green (1984), Switzer (1985), Conradsen *et al.* (1986), Löfgren, Ranneby & Sjöstedt (1993), and Sjöstedt (1993). The technique was applied by Löfgren, Ranneby & Sjöstedt (1993) in a relatively succesful attempt to forecast Swedish GDP components.

There are many previous studies dealing with demand for paper products, and they are typically founded in economic theory. In Sweden Åberg (1968) produced the first thorough econometric analysis of paper and paper board demand in Western Europe. He based his specification on Nerlove-Houthakker consumption functions[1]. Buongiorno (1978) used a pooling approach to estimate price elasticities in the world demand for paper and paperboard, and Baudin & Lundberg (1984) produced new FAO-forecasts for wood products. Uutela (1987) contains the demand analysis for the global forest sector model. Very few of them, however, deals with forecasts.

Below we start by introducing very briefly the theoretical background and the forecasting instrument used. Next we introduce the data set and our forecasting results, which are compared to a univariate forecasting technique ARIMA applied to the same time series. We then move to an inference section and use a moving blocks bootstrapping technique to generate standard errors for the coefficients of the most strongly correlated MAFs and for the

[1] See Nerlove (1958), Houthakker (1965).

autocorrelations for the MAFs. This rough and very indirect test of the small open economy assumption turns out to be indecisive, since most coefficients in the MAFs have quite large variances. However, it shows which MAFs have a significant autocorrelation over time, and therefore should be used to forecast the time series. To seriously test the small open economy assumption, we move to a VAR(2) model and test whether quantities Granger-cause the price series, and if prices Granger-cause the quantity series. This test essentially turns out in favour of a small open economy description of the market conditions in the Swedish forest sector. The paper concludes with a summing up of the results.

FORECASTING USING THE MINIMUM / MAXIMUM AUTOCORRELATION FACTORS

In a multivariate time series setting we consider linear transformations which maximize the temporal autocorrelation functions. The MAFs for stationary time series can be derived as follows (similar derivations are found e.g. in Switzer, 1985, and Conradsen *et al.*, 1986).

Let $z_t' = [z_{1t},\ldots,z_{pt}]$ be a *stationary* p-dimensional discrete time series. For $\gamma \in R^p$ and for each fixed (integer valued time lag) Δ we can maximize (or minimize) corr $[\gamma' z_t, \gamma' z_{t+\Delta}]$ over γ in subspaces of R^p. In particular, given Δ, we can find vectors $\gamma_1,\ldots,\gamma_p$ where for each i, γ_i is chosen so that the correlation between $\gamma' z_t$ and $\gamma' z_{t+\Delta}$ is maximized under the restriction corr $[\gamma_i' z_t, \gamma_j' z_t] = 0$ for all $j < i$. The linear combinations $\gamma_i' z_t$, $i = 1,\ldots,p$ are the min / max autocorrelation factors. Let Var $[z_t] = \Sigma_0$ and Var $[z_{t+\Delta} - z_t] = \Sigma_\Delta$.

We can now write

$$\begin{aligned}\mathrm{Var}\left[\gamma' z_{t+\Delta} - \gamma' z_t\right] &= \gamma' \Sigma_\Delta \gamma \\ &= 2\gamma' \Sigma_0 \gamma - 2\,\mathrm{cov}\left[\gamma' z_t, \gamma' z_{t+\Delta}\right]. \end{aligned} \tag{1}$$

This implies that

$$\mathrm{corr}\left[\gamma' z_t, \gamma' z_{t+\Delta}\right] = 1 - \frac{1}{2}\frac{\gamma' \Sigma_\Delta \gamma}{\gamma' \Sigma_0 \gamma}. \tag{2}$$

Maximizing (2) is equivalent to minimizing

$$R(\gamma) = \frac{\gamma' \Sigma_\Delta \gamma}{\gamma' \Sigma_0 \gamma}.$$

However, $R(\gamma)$ is at its minimum if γ is chosen equal to the eigenvector corresponding to the smallest eigenvalue of Σ_Δ with respect to Σ_0 (see e.g. Rao, 1973, p. 74).

Let $\lambda_1 \leq \ldots \leq \lambda_p$ be the eigenvalues, and $\gamma_1, \ldots, \gamma_p$ the corresponding eigenvectors of Σ_Δ with respect to Σ_0, i.e. λ_i and γ_i satisfy:

$$\Sigma_\Delta \gamma_i = \lambda_i \Sigma_0 \gamma_i \quad i = 1, \ldots p. \tag{3}$$

Now put $mf_{it} = \gamma_i' z_t \quad i = 1, \ldots p$. Then $mf_t = [mf_{1t}, \ldots, mf_{pt}]$ are called the MAFs.

The MAFs have the following properties:

(*i*) $$\operatorname{corr}\left[mf_{it}, mf_{jt}\right] = 0 \quad i \neq j,$$

(*ii*) $$\operatorname{corr}\left[mf_{it}, mf_{i,t+\Delta}\right] = 1 - \frac{1}{2}\lambda_i \quad \text{for all } i,$$

(*iii*) $$\operatorname{corr}\left[mf_{1t}, mf_{1,t+\Delta}\right] = \sup_{\gamma} \operatorname{corr}\left[\gamma' z_t, \gamma' z_{t+\Delta}\right],$$
$$\operatorname{corr}\left[mf_{pt}, mf_{p,t+\Delta}\right] = \inf_{\gamma} \operatorname{corr}\left[\gamma' z_t, \gamma' z_{t+\Delta}\right],$$
$$\operatorname{corr}\left[mf_{it}, mf_{i,t+\Delta}\right] = \sup_{\gamma \in M_i} \operatorname{corr}\left[\gamma' z_t, \gamma' z_{t+\Delta}\right],$$
$$\text{where } M_i = \left\{\gamma \,\middle|\, \operatorname{corr}\left[\gamma' z_t, mf_{jt}\right] = 0 \quad j < i\right\},$$

(iv) mf_t = are invariant with respect to linear transformations of the original time series.

For a proof of (*i*) to (*iv*), see e.g. Löfgren, Ranneby & Sjöstedt (1993) or Sjöstedt (1993).

It is worth noting that although similar to principal component analysis, there are a few important differences.

One of them is that MAF maximizes the autocorrelation over time whereas principal components maximize the variance of linear combinations of the time series. Another difference is (*iv*) above, i.e. the MAFs are invariant with respect to linear transformations of the original time series.

The forecasting instrument is obtained by an OLS-estimation of a linear model:

$$\hat{z}_{t+\delta} = \hat{\alpha}_0 + \sum_{i=1}^{k} \hat{\alpha}_i m_{it} \quad k \le p, \tag{4}$$

where $m_{1t},\ldots,m_{kt}$ are the k most strongly correlated MAFs. Now $(m_{1t},\ldots,m_{pt})$ corresponds to the MAFs in descending absolute correlation order, so that negative signs are ignored. Note that δ (the forecast horizon) can be different from Δ (the horizon for the maximization procedure).

The forecasting idea behind equation (4) can be presented as follows. Let

$$m_{t+\delta}^{k} = \left[m_{1t+\delta},\ldots,m_{kt+\delta}\right]' = \Gamma_k' z_{t+\delta} \quad k \le p, \tag{5}$$

denote the k most strongly correlated MAFs, where m_{t+d}^{k} is a $k\times 1$ vector and Γ_k^t is a $k\times p$ matrix. If $\Delta = \delta$ we have, due to the strong autoecorrelation, that:

$$m_{t+\delta}^{k} \approx a + Bm_t^k \tag{6}$$

where B is a $k \times k$ matrix.

Moreover:

$$z_{t+\delta} = (\Gamma_k \Gamma_k')^{-1} \Gamma_k m_{t+\delta}^{k} \tag{7}$$

and hence

$$z_{t+\delta} = (\Gamma_k \Gamma_k')^{-1} \Gamma_k m_{t+\delta}^{k} \approx \tilde{a} + \tilde{B} m_t^k \tag{8}$$

where $\tilde{a}$ and $\tilde{B}$ are estimated by OLS. Equation (8) tells us that $z_{t+\delta}$ is approximately a linear function of the most strongly correlated MAFs at time t. The choice of k is not obvious, but k should be less than p since the minimum autocorrelation factors contain noise, which may make the forecast less precise[2].

To estimate the MAFs the time series generally need to be made stationary.[3] Depending on the underlying model different methods can be used. The forecasting instrument will in general consist of two steps (parts). One part is taking care of the forecast of the stationary part, and one taking care of the rest.

In this paper we will check for unit roots and transform the data into stationary time series by taking differences. For a time series with a unit root we have:

$$z_t = \beta + z_{t-1} + e_t \tag{9}$$

where e_t is stationary. Differencing z_t by lag one gives:

$$w_t = z_t - z_{t-1} = \beta + e_t \tag{10}$$

which since it is stationary can be forecasted by MAFs. With $\delta = \Delta$ we obtain the δ steps ahead forecast from the regression equation:

$$\hat{w}_{t+\delta} = \hat{\alpha}_0 + \sum_{i=1}^{k} \hat{\alpha}_i m_{it} \quad k \le p, \tag{11}$$

where $m_{1t},\ldots, m_{kt}$ are the k most strongly correlated MAFs. The one step ahead forecasting equation for z_{t+1} is $\hat{z}_{t+1}=z_t+\hat{w}_{t+1}$. For $\delta > 1$ we have to forecast $z_{t+\delta-1}$, and the final step is a repeated application of the equation:

$$\hat{z}_{t+\delta} = \hat{z}_{t+\delta-1} + \hat{w}_{t+\delta} \tag{12}$$

This is essentially the technique, which will be used below in our MAF approach to forecast the world market prices of paper products in the Swedish forest industry.

One can, of course, ask how a VAR-model:

$$\hat{w}_{t+\delta} = \hat{\beta}_0 + \sum_{i=1}^{p} \hat{\beta}_i w_{it} \tag{13}$$

[2] One can easily show that a forecasting instrument where all the MAFs are used is equivalent to a VAR model where all the variables are independent variables with lag one. For the intuition, see the end of this section.

[3] For results on MAFs for non-stationary time series, see Sjöstedt (1993, ch. 7).

differs from Equation (11). It can be shown that the model in Equation (13) is equivalent to using all p MAFs in equation (11). The reason is the following: The vectors $(\alpha_0,\ldots,\alpha_p)$ and $(\beta_0,\ldots,\beta_p)$ are estimated by ordinary least squares and the MAFs are linearly independent mappings of w_t . Therefore it is true that:

$$\underset{\beta}{\mathrm{Min}} \sum \left[w_{t+\delta} - \beta_0 - \sum_{i=1}^{p} \beta_i w_{it} \right]^2 = \underset{\alpha}{\mathrm{Min}} \sum \left[w_{t+\delta} - \alpha_0 - \sum_{i=1}^{p} \alpha_i m_{it}(w_t) \right]^2 . \qquad (14)$$

Our idea is hence to sharpen the VAR-instrument by eliminating the noise accumulated in the lowest correlated MAFs.

EMPIRICAL APPLICATION

Data

The data we use are monthly time series data that are collected from Swedish official statistics. The paper products we are considering in this study are Newsprint, Kraftpaper, Sackpaper, and Coated paper, and the time period spans from January 1978 to January 1994. We have at least three reasons as to why we are considering these specific paper products. The first reason is that these products belong to a group of paper products for which official data exists and are reasonable compatible over a longer time span. The problem with comparability over time is due to the fact that the classification nomenclature was changed in 1988. A second reason for choosing these products is that they belong to a group of products which are very important for the Swedish paper industry. A third reason is that the chosen assortments differ with respect to the phase in their product life cycle. The rate of growth of Sackpaper exports have declined considerably during the time period studied, and since the beginning of the eighties the growth rate is even negative. Newsprint and Kraftpaper also seem to be in a mature stage of their life cycle, however their growth rate seems not to have reached the stagnating stage yet as sackpaper. Coated paper, on the other

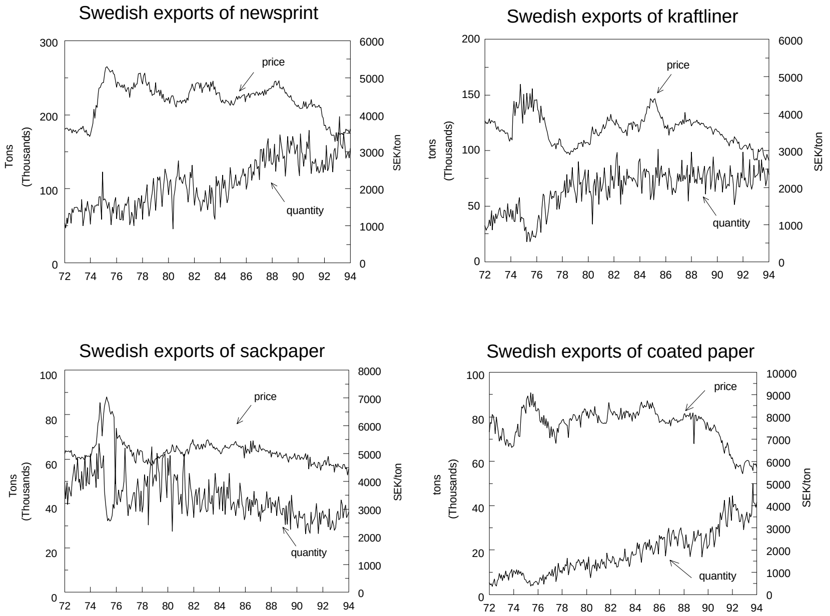

FIGURE 1. EXPORTS OF PAPER FROM SWEDEN, QUANTITY AND PRICES. MONTHLY DATA 1978 –1994.

hand, seems to be a product in the beginning of the life cycle, since the exports is steadily increasing.

The quantity data is measured as total exports in tons, and the price for each assortment is calculated as the export value in SEK, divided by quantity. This implicit price per ton does then include exchange rate changes since paper prices are in US dollars. This nominal price has then to be deflated by a producer price index for the manufacturing sector in Sweden. A visual description of the data is presented in Figure 1.

Estimating the MAFs

Having the data we can now proceed with the multivariate forecasting method previously described. Let $z_t = (Nq_t, Np_t, Kq_t, Kp_t, Sq_t, Sp_t, Cq_t, Cp_t)$ correspond to the eight paper price and quantity series used. Augmented Dickey-Fuller tests (Fuller 1976, Dickey & Fuller 1979) on z_t could not reject a unit root, and after a first order difference they showed a stationary behaviour. The MAFs were estimated on the first order differenced time series, $w_t = \nabla z_t = z_t - z_{t-1}$.

Table 1. MAF Weights for $\Delta = 1$ Derived from the First Order Differenced Price and Quantity Series.

Series	Corr(1)	∇Nq_t	∇Np_t	∇Kq_t	∇Kp_t	∇Sq_t	∇Sp_t	∇Cq_t	∇Cp_t
mf_{16t}	−0.796	0.0005	−0.397	0.004	0.194	−0.001	0.528	−0.003	−0.070
mf_{6t}	0.084	0.001	−0.411	0.014	0.497	−0.044	−0.070	−0.006	−0.141
mf_{1t}	0.412	−0.0003	−0.254	−0.001	−0.651	−0.003	−0.156	0.009	−0.060
Series	**Corr(1)**	∇Nq_{t-1}	∇Np_{t-1}	∇Kq_{t-1}	∇Kp_{t-1}	∇Sq_{t-1}	∇Sp_{t-1}	∇Cq_{t-1}	∇Cp_{t-1}
mf_{16t}	−0.796	0.0002	0.194	−0.005	−0.198	−0.003	−0.643	0.009	0.179
mf_{6t}	0.084	0.001	−0.681	0.014	0.272	−0.042	0.040	−0.013	−0.128
mf_{1t}	0.412	−0.0002	−0.173	−0.001	−0.640	−0.003	−0.201	0.009	−0.060

They were constructed as linear combinations of both and to catch the dynamics in the price series, i.e.

$$mf_t = \Gamma' \begin{pmatrix} w_t \\ w_{t-1} \end{pmatrix} \tag{15}$$

where Γ is a 16×16 matrix. The MAF-weights and autocorrelations for three selected MAFs are presented[4] in Table 1. The MAFs were estimated on the time period 78:1 up to 94:1 where no structural changes were assumed. The autocorrelations in the MAFs were maximized for time lag $\Delta = 1$.

For example, from Table 1 the sixteenth MAF with strongest (negative) autocorrelation is

$$\begin{aligned} mf_{16t} &= 0.0005\nabla Nq_t - 0.397\nabla Np_t + 0.004\nabla Kq_t + \\ &0.194\nabla Kp_t - 0.001\nabla Sq_t + 0.528\nabla Sp_t - 0.003\nabla Cq_t - \\ &0.070\nabla Cp_t + 0.0002\nabla Nq_{t-1} + 0.194\nabla Np_{t-1} - 0.005\nabla Kq_{t-1} \\ &-0.198\nabla Kp_{t-1} - 0.003\nabla Sq_{t-1} - 0.643\nabla Sp_{t-1} + 0.009\nabla Cq_{t-1} \\ &+0.179\nabla Cp_{t-1} \end{aligned} \tag{16}$$

with corr $(mf_{16t}, mf_{16t+1}) = -0.796$. Note that in mf_{16} the weights for the same time series at the two different time points are of the same size but with opposite signs. This turns out to be true for all negatively autocorrelated MAFs.

[4] A straightforward algorithm for the estimation of the MAFs is found e.g. in Switzer (1985) or Sjöstedt (1993).

For positively autocorrelated MAFs the weights also tend to be of the same size but with common signs for the same time series. Although the coefficients corresponding to the quantity series are smaller than those of the price series, there is yet no way to tell whether they are unimportant or important for forecasting purposes, since we do not know their sampling distributions.

Here standard errors are estimated using the moving blocks bootstrap technique.[5] The reason for this non-standard bootstrapping technique is that the time series is not i.i.d.; the forecasting technique is based on the presumption that observations are autocorrelated. This is a resampling technique that constructs replicates of the time series. From the original time series all possible blocks of a certain length (= 8 here) are constructed. Blocks are then randomly chosen with replacement and linked together until the new constructed series has the same length as the original.[6]

One thousand replicate multiple time series were constructed, resampled from the time period 78:1 up to 94:1. For each of the replicates MAFs were estimated with autocorrelations maximized for time lag $\Delta = 1$. Based on them standard errors were estimated. Table 2 shows the results for the autocorrelations of all the MAFs for time lag $\Delta = 1$.

The standard errors are of the same size. Most of the autocorrelations are stable and significantly different from zero. The MAF weights of the sixteenth MAF (mf_{16}) on the other hand, have barely any significant weights, see Table 3. Note that the larger weights in front of the price series have large variances and the smaller weights in front of the quantity series have small variances. Since the weights are so unstable no conclusions can be drawn about their respective relative importance for forecasting purposes. The weights of the other MAFs are also unstable and no conclusions can be drawn from them either. The exercise shows, however, which autocorrelations are sig-

5 See e.g. Politis & Romano (1992), Carlstein (1986), Künsch (1989).

6 The moving blocks bootstrap method works for eigenvalue problems since both eigenvalues and eigenvectors are smooth functions of the covariance matrices, see Politis & Romano (1992).

TABLE 2. AUTOCORRELATIONS FOR TIME LAG $\Delta = 1$ FOR ALL MAFS.

MAF No.	Autocorrelation*	Mean	Bias	Standard Error
1	0.412	0.440	0.027	0.046
2	0.337	0.351	0.014	0.039
3	0.251	0.273	0.022	0.033
4	0.206	0.209	0.002	0.033
5	0.148	0.142	−0.006	0.032
6	0.084	0.078	−0.006	0.035
7	0.054	0.005	−0.049	0.039
8	−0.049	−0.091	−0.042	0.044
9	−0.469	−0.351	0.118	0.049
10	−0.558	−0.429	0.129	0.037
11	−0.578	−0.493	0.085	0.032
12	−0.619	−0.549	0.070	0.031
13	−0.659	−0.603	0.056	0.029
14	−0.728	−0.661	0.066	0.030
15	−0.771	−0.717	0.054	0.027
16	−0.796	−0.769	0.027	0.027

* The autocorrelation at time lag 1 for MAFs estimated from the original time series. The mean, bias and standard error is obtained from 1000 bootstrap replicates.

nificantly different from zero. Most of them are, and this information is comforting for the MAF forecasting technique.

The Forecasting Equation

The forecasts are based only on twelve previous years. For example, when prices 93:1 were predicted the models were estimated on the time period 81:1 up to 92:12. This was to avoid errors due to potential structural changes. When the MAFs had been estimated (based on the previous twelve years), the next step was to estimate the forecasting Equations (17)

$$w_{t+\delta} = \alpha_0 + \sum_{i=1}^{k} \alpha_i m_{it} \quad k \le p, \tag{17}$$

for the differenced price series w_t, with $\delta = 1$ and 6 months. The equations were estimated by OLS for k = 1,...,16 MAFs.

TABLE 3. MAF- WEIGHTS FOR THE SIXTEENTH MAF.

WEIGHTS*	MEAN	BIAS	STANDARD ERROR
0.000	0.000	−0.000	0.002
−0.397	−0.314	0.082	0.350
0.004	0.001	−0.003	0.004
0.194	0.095	−0.099	0.241
−0.001	0.001	0.001	0.008
0.528	0.295	−0.233	0.316
−0.003	−0.003	0.000	0.008
−0.070	−0.035	0.036	0.120
0.000	0.000	0.000	0.002
0.194	0.238	0.045	0.406
−0.005	−0.002	0.003	0.004
−0.198	−0.079	0.119	0.253
−0.003	−0.002	0.000	0.006
−0.643	−0.374	0.268	0.236
0.009	0.005	−0.004	0.007
0.176	0.079	−0.097	0.124

* The MAF-weights estimated from the original time series. The mean, bias and standard error is obtained from 1000 bootstrap replicates.

We later chose the k that minimizes the prediction error for the price series. The autocorrelations in the MAFs were maximized for the same distance as the forecasting horizon, i.e. $\delta = \Delta$. Note that for $\delta = 1$, when all sixteen MAFs are in the forecasting equation it is equivalent to an ordinary VAR(2) model:

$$w_{t+\delta} = A_0 + A_1 w_t + A_2 w_{t-1}. \tag{18}$$

The one month ahead forecasts were constructed as:

$$\hat{z}_{t+1} = z_t + \hat{w}_{t+1} \tag{19}$$

and, similarly, the six month ahead forecasts as:[7]

$$\hat{z}_{t+6} = z_t + \hat{w}_{t+6}. \tag{20}$$

[7] In Equation (20) $w_t = z_t - z_{t-1}$, whereas in Equation (21) $w_t = z_t - z_{t-6}$.

The MAF forecasting equation we finally chose, i.e. the number of MAFs k in (18), was the one with the lowest Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|\hat{z}_{n+t+\delta} - z_{n+t+\delta}\right|}{z_{n+t+\delta}} \times 100\%. \tag{21}$$

where T is the number of forecasts.

Forecasting using ARIMA

To get an idea of how well the multivariate MAF forecasting technique works, and to determine if quantities are relevant for the price forecasts, we have compared it with univariate ARIMA models (Box & Jenkins (1970)). Since our objective is to forecast prices, the ARIMA technique simply means that we do not consider the quantities, or other prices than lagged own prices, in the model formulation. For each of the four time series our objective is to identify the ARIMA process which fits the data best, or produces the best forecast. Since this is a univariate technique this amounts to take the series one by one. Following Box & Jenkins (1970), the criteria used are visual inspections of autocorrelation functions and partial autocorrelation functions. In addition we used the Ljung-Box (Q) test statistic for higher order serial correlation, as well as Akaike's information criterion.

As already mentioned, a common result for all four price series is that first order differencing is required to obtain stationarity. The results from the identification process on the first order differences are presented in Table 4.

From the results in Table 4 it should be clear that the chosen specifications provides rather good fit to the actual data. In addition, the Q-statistics do not indicate on any problems with serial correlation.[8]

The results in Table 4 are used to obtain one month and six month ahead forecasts (out of sample) for each price series.

[8] We have tried alternative specifications, but the results are almost unaltered.

Table 4. Results from ARIMA Modeling (t-values within Parentheses).

Series	AR1	AR2	Constant	adj.R^2	Q(12)	AIC
News	−0.48 (−6.98)		−7.24 (−9.50)	0.94	11.3	2785
Kraft	−0.22 (−3.19)		−2.08 (−0.36)	0.93	10.52	2793
Sack	−0.63 (−8.27)	−0.15 (−2.06)	−3.70 (−0.74)	0.80	8.4	2880
Coated	−0.48 (−6.83)	−0.25 (−3.52)	−10.0 (−1.13)	0.93	0.74	3089

MAF versus ARIMA

The *out of sample* forecasting performance was measured by MAPE based on 60 forecasts from 89:1 up to 94:1. MAPE for the naive random walk model[9], ARIMA, VAR(2), and MAF-models are displayed in Tables 5 and 6. The MAF-model with the lowest MAPE is the one presented together with its corresponding number of MAFs.

For one month ahead forecasts News and Sack paper prices had lowest MAPE for the ARIMA-models. For Kraft and Coated paper prices the MAF approach worked better. The MAF-forecasting errors never exceeded MAPE for the naive forecasts while the ARIMA forecasting errors did so for both Kraft and Coated paper prices.

Now for the six month ahead forecasts the ARIMA models had lowest MAPE for all price series. This may indicate that the eight time series are not connected to each other. In the next section we will discuss this in more detail.

Inference: “Testing” the Small Open Economy Hypothesis

If quantities do not contribute to the explanation of future prices, this would support a small open economy (quantity adjuster-price taker) description of the Swedish

[9] Naive forecasts are defined as $\hat{z}_{t+\delta} = z_t$.

TABLE 5. MAPE FOR $\delta = 1$ MONTH AHEAD FORECASTS.

	MAPE				
	Naive	ARIMA	VAR	MAF	No.of MAFS
News	1.885	1.714	1.948	1.791	6
Kraft	2.166	2.302	2.248	2.099	3
Sack	2.684	2.250	2.403	2.358	13
Coated	2.109	2.175	2.615	2.087	4

forest sector. To see this, let $q_{it}^{d}(p_t, x_{it})$ and $q_{it}^{s}(p_t, x_{it})$ be respectively the demand and supply from country i in the world market. p_t represents the world market price at time t and x_{it} and y_{it} are country specific exogenous variables, which together determine the position of the aggregate demand and supply curves. The equilibrium price, p_t^*, would be given from the equality between aggregate world demand and supply:

$$\sum_{i=1}^{n} q_{it}^{d}(p_t^*, x_{it}) = \sum_{i=1}^{m} q_{it}^{s}(p_t^*, y_{it}) \tag{22}$$

It is a function of the vectors $x_t = [x_{1t},\ldots,x_{nt}]$ and $y_t = [y_{1t},\ldots,y_{mt}]$ or:

$$p_t^* = p_t^*(x_t, y_t) \tag{23}$$

Country i's sales in the world market (exports), q_{it}^{ns}, under the small open economy assumption would be given by the difference between supply and demand in country i at the world market price or:

TABLE 6. MAPE FOR $\delta = 6$ MONTH AHEAD FORECASTS.

	MAPE			
	Naive	ARIMA	MAF	No.of MAFS
News	4.554	3.871	4.569	4
Kraft	4.757	4.311	4.677	4
Sack	2.759	2.258	2.799	1
Coated	4.795	4.394	4.485	5

$$q_{it}^{ns}\left(p_t^*(x_t, y_t), x_{it}, y_{it}\right) = q_{it}^s(\cdot) - q_{it}^d(\cdot) \qquad (24)$$

Even if p_t^* and p_{t+1}^* are strongly correlated q_{it}^{ns} could be a bad predictor of p_{t+1}^*, since changes over time of the domestic exogenous variables x_{it} and y_{it} would have a stronger impact on the domestic quantity, than their impact on the world market price. Domestic demand and supply conditions would under quantity adjustment add extra noise. On the other hand, q_{it}^{ns} could under these conditions be a good predictor of q_{it+1}^{ns}, since (x_{it}, y_{it}), and (x_{it+1}, y_{it+1}) like p_t and p_{t+1} are likely to be correlated. Note that since prices are measured in domestic currency, the impacts of varying exchange rates are included in the analysis.

We already have a rather strong indication that quantities are unimportant for the price forecasts, since the forecasting performance of the univariate ARIMA-model matches that of the MAF approach.

In order to “refute” or “support” the small open economy assumption, a more formal and straightforward approach suggests itself. To investigate which time series are important in forecasting the price series we can use the Granger-Causality tests (Granger (1969), or e.g. Lütkepohl (1991 p. 93–95). Our tests are based on a VAR(2)-model[10] of estimated on the time period 78:1 up to 94:1. The VAR(2)-model, which is a special case of the MAF approach, where all MAFs are used in the forecasting equation, can be written as

$$\begin{pmatrix} \nabla p_t \\ \nabla q_t \end{pmatrix} = \nu + \begin{pmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{pmatrix} \begin{pmatrix} \nabla p_{t-1} \\ \nabla q_{t-1} \end{pmatrix} + \begin{pmatrix} B_{11} B_{12} \\ B_{21} B_{22} \end{pmatrix} \begin{pmatrix} \nabla p_{t-2} \\ \nabla q_{t-2} \end{pmatrix} \qquad (25)$$

where p_t and q_t are the four price and quantity series respectively, and ν a vector of constants and ε_t white noise. To test if the quantities explain (Granger-causes) the price series we tested

H_0: $A_{12} = B_{12} = 0$, quantities *do not* explain prices. (26)

[10] According to the information criteria of Akaike, Hannan & Quinn and Schwartz, 4, 1 and 0 lags respectively would be the best choices.

The Wald statistic divided by the number of coefficients set to zero in H_0, was used as a test statistic. This statistic, Lf, has an approximate F-distribution (see Lütkepohl (1991, p. 93–95) for details). Testing H_0 it turned out that it could not be rejected (Lf = 1.40, p-value = 0.069). An opposite test to see if the prices explains the quantities was constructed as

$$H_0\text{: } A_{21} = B_{21} = 0, \text{ prices } \textit{do not} \text{ explain quantities.} \quad (27)$$

This hypothesis could not be rejected either (Lf = 1.28, p-value = 0.14). Hence prices and quantities seem to be unconnected to each other. This may explain why the univariate ARIMA model did so well in comparison to the MAF approach. A multivariate approach should be preferred only if there is covariation and connection between the time series. The obvious question to ask now is if the prices are explained solely by lags of itself,

$$H_0\text{: prices are explained only by lags of itself.} \quad (28)$$

This hypothesis was rejected (Lf = 1.38, p-value = 0.034). The conclusion should be to use only the four price series to forecast the price series. The MAF forecasting method was run on the four price series solely. It did only marginally improve the forecasting performance.

With respect to the relevance of the small open economy assumption, our causality tests support it, and so does the fact that a univariate ARIMA model forecasts the price series as well as the multivariate MAF technique. This result is consistent with the conclusions in Horn (1979), who used an indirect structural approach to test the small open economy assumption in the Swedish forest sector. It, however, disagrees with the results in Wiberg (1987), who used a structural mark-up approach based on Appelbaum (1979). He found that Swedish forest products were sold at prices significantly above marginal costs in the German market, which is inconsistent with the behavior of a price taker.

CONCLUSIONS

Although one can, in general, do better than a naive martingale forecast of the prices for some of the key paper products in the Swedish forest industry, it turns out to be rather difficult to do “significantly” better[11]. The multivariate forecasting technique, MAF, we introduced has the trivial, but nonetheless fundamental purpose of splitting the forecasting information in the time series into important and unimportant. The important information is used for forecasting purposes and the unimportant (noise) is discarded. The technique has been found to do quite well in forecasting GNP components over the business cycle, see Löfgren, Ranneby & Sjöstedt (1993), but here it was almost outperformed by the univariate ARIMA technique.

This is interesting for at least one reason. It indicates that quantities and perhaps also prices of other forest products are unimportant for forecasting purposes. As we argue, this would be consistent with a small open economy description of the Swedish forest sector. We run two formal tests of whether quantities matter: one rough test based on bootstrapping to find out whether the quantity weights in the MAFs are significant or not; and one block of Granger causality tests of a special case of the MAF approach (the VAR 2-model). It turns out that quantities are indeed unimportant for the price forecasts, and that prices are unimportant in forecasting the quantity series. We can, however, reject the hypothesis that price is solely explained by lags of itself. This means that we should be able to do better than the ARIMA model by a multivariate approach using all prices. In practice, as a MAF model based on only prices shows, this turns out to be very difficult.

It may seem too pretentious to claim that our tests support a small open economy description of the Swedish forest sector. To be able to reject such a hypothesis one would ideally want a structural model based on fundamentals. However, as Sims (1980) has pointed out, structural models often imply identification problems, which may blur the real issues involved. There is a trade off between the sharpness of the hypothesis tested, and the possibility to identify the structural equations. This is, however, not the only problem. To be able to use monthly data to identify a structural model also requires an explicit modeling of

[11] Whether this has something to do with the idea that “all prices are martingales”, see Alchian (1974) and Samuelson (1971) we do not know.

behavior outside equilibrium. A choice of an essentially ”atheoretical” approach may be to go too far in the other direction, but we claim nevertheless that our “test results” are too clearcut to contain no information about the structure of the markets under consideration. Moreover, they certainly show that the market structure may have important implications for the choice of forecasting method.

REFERENCES

Alchian, A., 1974. Information Martingales and Prices. *The Swedish Journal of Economics*, 76, 3–11.

Appelbaum, E. (1979) Testing Price Taking Behavior. *Journal of Econometrics* , 9, 283–294.

Baudin, A. & Lundberg, L. (1984) Econometric Model for Demand for Mechanical Wood Products. Working Report, *FAO Programme Outlook Studies for Supply and Demand of Forest Products* (FAO, Rome).

Box, G. E. P. & Jenkins, G. M., 1970. *Time Series Analysis Forecasting and Control* (San Fransisco: Holden Day).

Box, G. E. P. & Tiao, G. C., 1977. A Canonical Analysis of Multiple Time Series. *Biometrika*, 64, 355–365.

Buongiorno, J., 1978. Income and Price Elasticities in the World Demand for Paper and Paperboard. *Forest Science,* 24, 231-236.

Carlstein, E., 1986. The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence. *Annual Statistics*, 14, 1171–1179.

Conradsen, K., Nielsen, B. & Thyrstedt, T., 1986. A Comparison of Min / Max Autocorrelation Factor Analysis and Ordinary Factor Analysis, *IMSOR, Technical University of Denmark* (mimeographed).

Dickey, D. A., & Fuller, W. A., 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of American Statistical Association*, 74, 427–431.

Efron, B. & Tibshirani, R. J., 1993. *An Introduction to the Bootstrap* (New York: Chapman & Hall).

Fuller, W. A., 1976. *Introduction to Statistical Time Series* (New York: John Wiley & Sons).

Granger, C. W. J., 1969. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37, 424–438.

Horn, H., 1979. Skogsindustrins kostnader, relativpriser och marknadsandelar, en pilotstudie. Research Paper 6136, *EFI*, Stockholm.

Houthakker, H. S., 1965. New Evidence on Demand Elasticities. *Econometrica*, 33, 277–288.

Künsch, H., 1989. The Jackknife and the Bootstrap for General Stationary Observations, *Annual Statistics* 17, 1217–1241.

Lütkepohl, H., 1991. *Introduction to Multiple Time Series Analysis* (Berlin: Springer-Verlag).

Löfgren, K. -G., Ranneby, B. & Sjöstedt, S., 1993. Forecasting the Business Cycle without Using Minimum Autocorrelation Factors. *Journal of Forecasting*, 12, 481–498.

Nerlove, M., 1988. Distributed Lags and the Estimation of Long-run Supply and Demand Elasticities: Theoretical Considerations. *Journal of Farm Economics*, 40, 861–880.

Politis, D. N. & Romano, J. P., 1992. A Nonparametric Resampling Procedure for Multivariate Confidence Regions in Time Series Analysis. *Computing Science and Statistics*, proceedings from the 22:nd symposium on the interface (East Lancing MI: Springer Verlag), pp. 98–103.

Rao, C. R., 1973. *Linear Statistical Inference and its Applications (* John-Wiley and Sons).

Reinsel, G. C., 1983. Some Results on Multivariate Autoregressive Index Models. *Biometrica*, 70, 145–156.

Samuelson, P. A., 1971. Stochastic Speculative Price. *Proceedings of the National Academy of Sciences*, February 1971.

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

Sims, C. A., 1981. An Autoregressive Index Model for the U.S., 1948–1975. In *Large-Scale Macro-Econometric Models*, edited by J. Kmenta & J. B. Ramsey (Amsterdam: North Holland).

Sjöstedt, S., 1993. Forecasting Multiple Time Series using Minimum / Maximum Autocorrelation Factors. *Department of Mathematical Statistics*, University of Umeå.

Switzer, P., 1985. Min / Max Autocorrelation Factors for Multivariate Spatial Imagery. In *Computer Science and Statistics: The Interface*, edited by L. Billard, (Amsterdam: Elsevier Science Publishers).

Switzer, P. & Green, A. A., 1984. Min / Max Autocorrelation Factors for Multivariate Spatial Imagery, Technical Report No. 6, *Department of Statistics*, Stanford University.

Uutela, E., 1987. Demand for Paper and Board: Estimation of Parameters for Global Models, . In M. Kallio *et al. The Global Forest Sector: An Analytical Perspective* (New York: John Wiley & Sons).

Wiberg, A., 1987. Svensk Massa- och Pappersindustris Marknadsställning, *Inst. för Skogsekonomi*, Rapport 71, Sveriges lantbruksuniversitet, Umeå.

Åberg, C. J., 1968. *The Demand for Paper and Paperboard in Western Europe 1950–1962* (Stockholm: Almqvist & Wicksell).