

PAAVO YLI-OLLI, Associate professor

ILKKA VIRTANEN, Professor

University of Vaasa

School of Business Studies

Vaasa, Finland

Predictability of stock returns in a thin security market

Empirical evidence with data from Helsinki Stock Exchange 1975—1986

ABSTRACT

Stock market efficiency is a crucial concept when forecasting of future stock returns is discussed. Efficient markets are divided into three levels of market efficiency depending upon the type of information used by investors. In a strong-form of efficient market security prices rapidly reflect all information. A semi-strong form of efficiency respectively holds that no publicly available information can be successfully used in the prediction of stock returns. Finally, under a weak form of efficiency, information of historical stock returns is of no value for predicting the future stock returns. On a thin security market, like in the Helsinki Stock Exchange, many anomalies and deviations from market efficiency have been obtained. It is shown in the paper that the average Finnish stock market returns do not follow the general random walk model. Both the monthly and quarterly stock returns can be forecasted using either univariate time series analysis or multivariate econometric modelling. Also a procedure is presented to generate composite forecasts from univariate time series models and econometric models.

1. INTRODUCTION

Stock market efficiency has been extensively tested in recent years in the U.S. and Europe (see e.g. Fama 1970; Granger and Morgenstern 1970; Jennergren and Korsvold 1974; Korhonen 1977; Hawawini and Michel 1984; Wahlroos and Berglund 1983 and 1986).

Market efficiency is a crucial concept when predictability of stock returns are discussed (see the definition of market efficiency and the problems arising in its empirical testing Fama 1970: 383; Copeland—Weston 1983: 285—287; and Foster 1986: 312—319). The market is efficient in a weak sense if share prices fully reflect the information implied by all prior price movements. The market is efficient in a semi-strong sense if share prices respond instantaneously and without

bias to newly published information. And finally, the market is efficient in a strong sense if share prices fully reflect all relevant information including data not yet publicly available.

As such the weak form is directly opposed to the basic premises of technical analysis or univariate time series analysis (e.g. that presented by Box and Jenkins 1970) where the behaviour of the series is explained by its own variability (tests of the weak form of market efficiency are usually associated with random walk hypotheses of stock returns). Similarly, the semi-strong form of efficiency holds that all publicly available information is of no value in the prediction of future stock returns (see e.g. Hagin 1979: 11—36). In such a market securities will be traded at prices which are close to their true value and investors will be unable to systematically earn above normal profits by using fundamental analysis e.g. multivariate econometric models for estimating future stock returns.

According to most empirical results, market in U.S. and in some European countries is efficient in the semi-strong form or the results as a rule are in support of the weak form of efficiency at least (see e.g. Hawawini and Michel 1984; Jennergren and Korsvold 1974; Wahlroos and Berglund 1983 and 1986). However, we can also find opposite results (for U.S. market see Umstead 1977: 427—441; in Europe the returns of some German and Scandinavian stocks exhibit statistically significant dependence over time, see e.g. Hawawini and Michel 1984: 8—25). On a thin security market — like in Finland — there may exist many anomalies and deviations from market efficiency, as found by Berglund (1986). These anomalies and deviations from market efficiency are the starting point of this research. For if the market were efficient enough it would not be possible to predict stock returns.

In a statistical investigation concerning the predictability of stock prices we can forecast, first the price level (e.g. Hansmann and Zetsche 1985; and Virtanen and Yli-Olli 1987), second price changes (e.g. Granger and Morgenstern 1970: 58—59) and third total returns where current dividend yield is added to the price changes (e.g. Umstead 1977).

In Finland, in the Helsinki Stock Exchange we have in public use the so called Unitas and KOP stock market indices published by two Finnish commercial banks (the Unitas index by SYP and the KOP index by KOP, respectively). KOP and Unitas indices do not include the dividend component of market returns. In practice the dividend component of market returns in Finland has been extremely stable during the period to be examined and variation in total returns from month to month has almost entirely been due to price fluctuations. The correlation coefficient between the theoretically correct total return index (counted by Berglund, Wahlroos and Grandell 1983) and the Unitas index was .991. The respectively correlation between the total return index and the KOP index was .978 (see Berglund, Wahlroos and Grandell 1983: 39). In this research stock returns computed from the Unitas index during the eleven-year period from January 1975 to March 1986

are used as a surrogate of the total returns. The models will be estimated using both monthly and quarterly data.

The purposes of this study are:

1. To analyze whether stock returns are predictable in a thin security market like the Helsinki Stock Exchange.
2. To compare the forecasting results based on univariate time series analysis and econometric models with each other and to develop composite forecasting models and examine the possible forecasting improvement of those models to the time series models and econometric models.
3. To compare the forecasting results of stock market prices and stock returns with each other.

2. CONSTRUCTION OF FORECASTING MODELS

The empirical variable in our analysis used to measure the market return in Helsinki Stock Exchange consists of the logarithmic differences of the Uunitas stock market index. There are three main reasons for the use of the logarithmic transformation before taking the first differences of the index. The first one is the empirical fact that there have been considerable changes in the value of the index, which tend to invalidate the assumption of a constant relationship between the absolute values of variables. The second, technical, reason is that, when using logarithms, the efficiency of the estimates is increased because heteroscedasticity in regression analysis is reduced (see e.g. Driehuis 1972: 11—12). In time-series analysis, stationarity in variance can be achieved, respectively (Makridakis, Wheelwright and McGee 1983: 439). And third, the use of logarithmic transformation also appears reasonable from an interpretative point of view since the first difference of the logarithms will closely approximate the percentage change for the period.

The models will be estimated using both monthly and quarterly data. The monthly data contain more detailed and precise information and they are, therefore, expected to produce more accurate models. However, the quarterly models will also be estimated in order to control and confirm the results based on monthly data, because some independent variables in our econometric analysis are observed only quarterly. The monthly values of those variables are interpolated. Further, the quarterly models generate one-step forecasts for three months, instead of forecasts for one month in monthly models.

The data to be analyzed consist of the values of the Uunitas index in years 1975—86. The values in the years 1975—84 (119 monthly values or 39 quarterly values in the logarithmic first-difference form) are used for estimating and testing purposes, the rest of the data (from January 1985 to March 1986 in monthly

data or from I/1985 to I/1986 in quarterly data) will be saved for measuring the forecasting accuracy of the derived models. The development of the logarithmic transformed Unitas index is presented, both in the level and in the first-difference form, in Figures 1—4 in the Appendix.

2.1. Univariate time-series models

General AutoRegressive Integrated Moving Average (ARIMA) models developed by Box and Jenkins (1970) have been extensively applied also by accounting and finance researchers to investigate for instance the behaviour of reported accounting numbers and stock market prices.

ARIMA models express the current value of a series (y_t) as a function of its past values as well as current and past values of a noise or error series (e_t):

$$(1) \quad \phi^p(B)(1-B)^d y_t = \theta^q(B)e_t.$$

In (1) B is the backward shift or lag operator such that

$$(2) \quad B^k y_t = y_{t-k}, \quad k = 1, 2, \dots,$$

$(1-B)^d$ is the differencing operator of order d , $\phi^p(B)$ is the autoregressive polynomial in B of order p , $\theta^q(B)$ is the moving average polynomial in B of order q , and e_t is the noise series, that is assumed to be independent and normally and identically distributed over time. Model (1) is also called ARIMA (p, d, q)-model.

Model (1) is for non-seasonal processes. If we have to add seasonal components to the model, we get an ARIMA (p, d, q)(P, D, Q) S -model, which in terms of the polynomials becomes

$$(3) \quad \phi^p(B)\Phi^P(B^S)(1-B)^d(1-B^S)^D y_t \\ = \theta^q(B)\Theta^Q(B^S)e_t.$$

In (3) S is the number of periods per season, Φ^P and Θ^Q , the seasonal autoregressive and seasonal moving average polynomials, respectively, as well as the seasonal differencing operator $(1-B^S)^D$ are now all expressed in powers of B^S .

In the recent paper by the authors (Virtanen and Yli-Olli 1987) the univariate Box—Jenkins methodology was applied to model both monthly and quarterly stock prices on the Helsinki Stock Exchange. The empirical variable in that study was the logarithm of the Unitas stock market index. In this study we deal with the first differences of the same logarithmed Unitas index. In terms of ARIMA methodology this means that the models become identical except the order of differencing (parameter d in (1) and (3)), which in the case of stock returns becomes one less than in the case of the logarithmed index itself.

The paper mentioned above contains a detailed description of the derivation of the different ARIMA models for the logarithm of the Unitas index. Therefore, in this paper only the main results in different phases of the model development

APPENDIX.

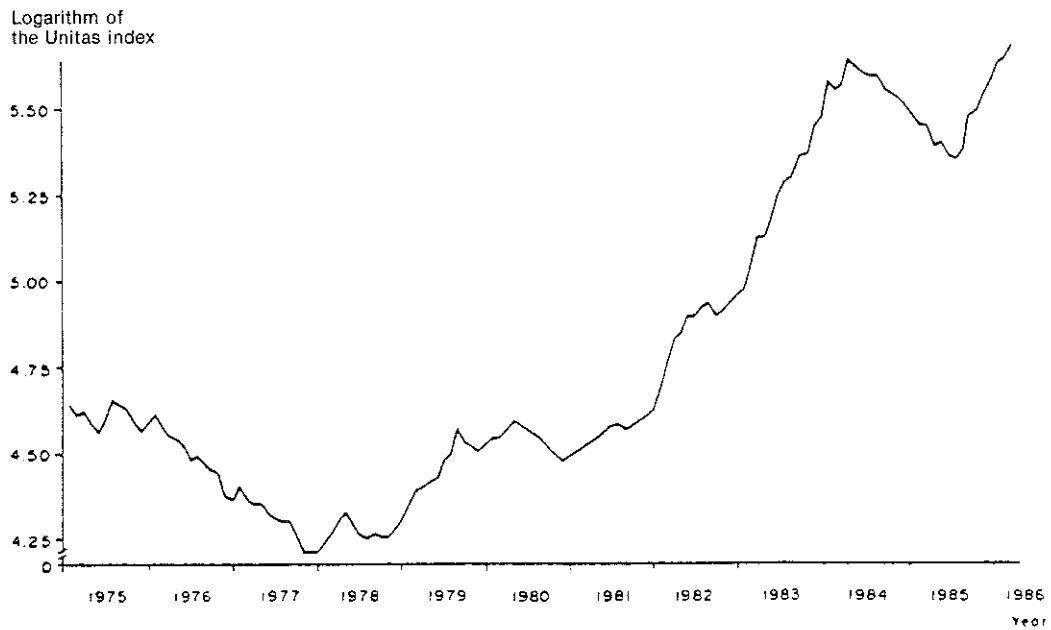


Figure 1. The logarithms of the Uritas stock market index (monthly data).

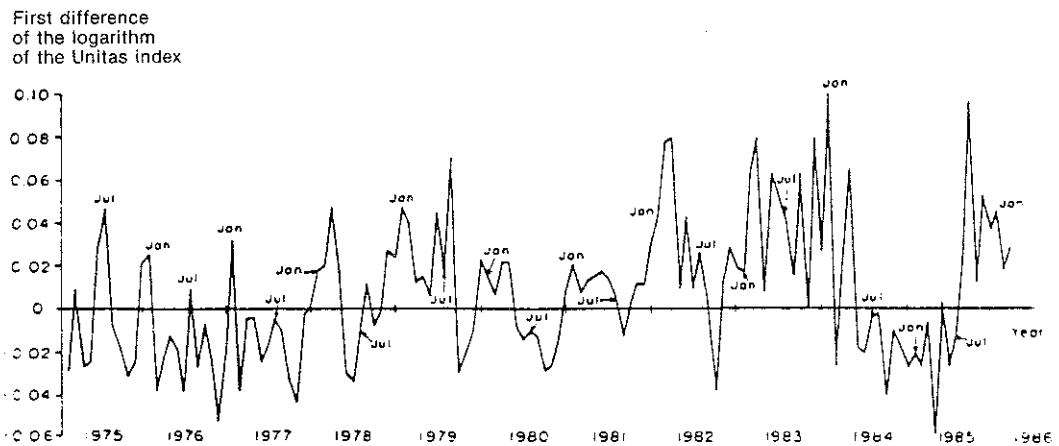


Figure 2. The first differences of the logarithms of the Uritas index (monthly data).

will be presented. The results have been adjusted to fit with the current response variable.

For monthly returns, two different tentative models are obtained: a non-seasonal ARIMA (0, 1, 1)-model, or model



Figure 3. The logarithms of the Uritas stock market index (quarterly data).

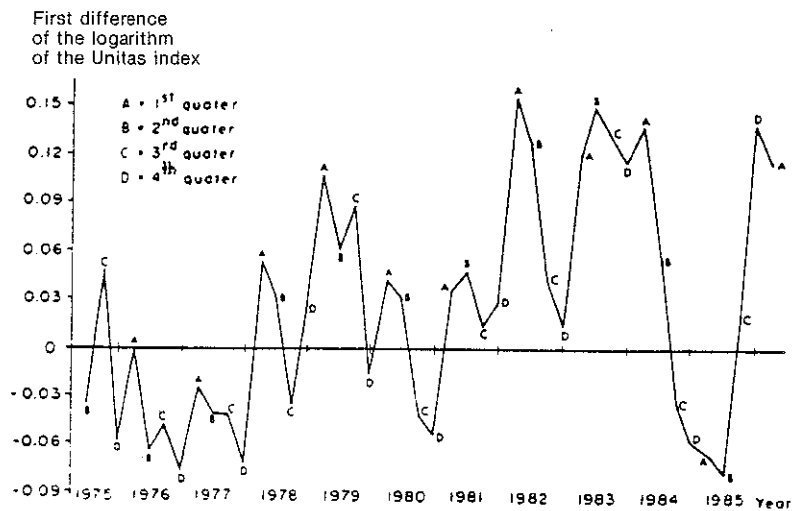


Figure 4. The first differences of the logarithms of the Uritas index (quarterly data).

$$(4) \quad (1-B)y_t = (1-\theta B)e_t$$

and a seasonal ARIMA $(0, 1, 1)(0, 1, 1)^{12}$ -model, i.e.

$$(5) \quad (1-B)(1-B^{12})y_t = (1-\theta B)(1-\Theta B^{12})e_t.$$

Both models (4) and (5) contain a differencing of the return series y_t . This is due to a clear, although not very strong trend in the return series (cf. for example Figure 2 in the Appendix). The differenced return series, which is now stationary both in the mean and in the variance, indicates a clear MA(1)-process. This component is included in both of the models (4) and (5). A seasonal component (after taking the one-year seasonal differences of the series) may also be added into the model. This is motivated by the moderate clear seasonality pattern in the data: the return from stocks is high at the beginning of each year and low in early summer and late autumn (cf. Figure 2 in the Appendix). A closer analysis of the data suggests that the seasonal component should be of MA(1)-type. Model (5) contains this seasonal MA(1)-component.

The behaviour of the quarterly return is very similar to that of the monthly return. The main difference is that the seasonal variation becomes much more clear-cut, due to reduced random variation in the more aggregated and thus smoothed raw data, cf. Figure 4 in the Appendix.

The process behind the quarterly return series resembles the monthly return process with the exception that there is now no sign of any non-seasonal MA(1)-component in it. Further, a seasonal four-quarter MA(1)-component must be included in the model. We thus have a tentative ARIMA $(0, 1, 0)(0, 1, 1)^4$ -model, or model

$$(6) \quad (1-B)(1-B^4)Y_t = (1-\Theta B^4)E_t,$$

where we have used capital letters Y and E referring to the quarterly return and error term, respectively.

The parameters of the tentative models (4)—(6) were estimated by a non-linear Gauss—Marquardt algorithm using the backcasting method (Box and Jenkins 1970; Dixon 1983). The results from the estimation are presented in Table 1.

Table 1. Summary of estimation and test statistics associated with the ARIMA-models.

Model	Parameter	St. error	t-value	D.o.f.
(0, 1, 1)	$\hat{\theta} = 0.759$	0.0614	12.4	117
(0, 1, 1)(0, 1, 1) ¹²	$\hat{\theta} = 0.753$	0.0627	12.0	104
	$\hat{\Theta} = 0.840$	0.0359	23.4	104
(0, 1, 0)				
(0, 1, 1) ⁴	$\hat{\Theta} = 0.833$	0.0556	15.0	33

We see that the estimates are stable and consistent. The first-order non-seasonal MA-parameter θ gets in both of the monthly models a value near to 0.75. The seasonal MA(1)-parameter Θ , both in the monthly model (5) and in the quarterly model (6), is of the size 0.84. We can further see that in all the models the true

values of the parameters fall with extremely high probability into the interval $(0, 1)$, as desired.

The diagnostic checking and testing of the models was based on studying the residuals — to see if any pattern was remained unaccounted for — and on studying the sampling statistics — to see if the models could be simplified (cf. Makridakis, Wheelwright and McGee 1983: 446).

The analysis showed, first, that all the models presented in Table 1 have produced residuals which are essentially random: none of the individual autocorrelations or partial autocorrelations was significantly different from zero, and the Ljung-Box statistics (Dixon 1983: 690) revealed no evidence of inadequacy for any of the models. Second, the results of Table 1 show that all the t -values associated with the estimates of the parameters are highly significant, the inclusion of the corresponding components being therefore motivated. To get an idea how well the different models fit with the data, we have computed some residual statistics commonly used for this purpose. Table 2 presents the measures RSS (residual sum of squares), RMS (residual mean square) and RRMS (square root of RMS) for the different models.

Table 2. The residual statistics for the estimated ARIMA-models.

Model	RSS	D.o.f.	RMS	RRMS
$(0, 1, 1)$	0.0942	117	0.000805	0.0284
$(0, 1, 1)(0, 1, 1)^{12}$	0.0761	104	0.000732	0.0271
$(0, 1, 0)(0, 1, 1)^4$	0.0613	33	0.001859	0.0431

The weak form of market efficiency hypothesizes a random walk model for the stock returns. The preceding results show, however, that the Finnish stock returns, measured with the percentage change of the general price level, do not follow the general random walk model. At least one moving average component (in addition to differencing) must be included in the ARIMA-models.

2.2. *Econometric models*

2.2.1. *The selection of explanatory variables*

The results of univariate time-series models showed that the average stock returns in the thin Finnish security market did not follow the general random walk model. As such the result seems to indicate a deviation from the weak form of market efficiency. However, we have used aggregated data and we have omitted

transaction costs in the analysis. The final tests of the weak form of efficiency should be carried out by using the daily stock returns of individual firms. Useful statistical techniques for this problem are e.g. serial correlation tests, run tests or filter strategies (see e.g. Fama and Blume 1966; and Jennergren 1975).

In this section we develop a forecasting model based on multiple regression analysis, in order to predict the monthly and quarterly stock returns. Our econometric model includes six groups of explanatory variables which a priori can be supposed to affect the development of stock returns. In our recent paper (Virtanen and Yli-Olli 1987) we can in principle find the same groups of explanatory variables. However, those variables were in the level form because the endogenous variable was stock market price index.

The first explanatory variable is according to the results of univariate time series models the endogenous variable itself lagged one period (coefficient a priori positive). This variable does not totally exhaust the effect of past historical development itself. However, composite stock return forecasting models to be presented in the following section include this full past history.

The second explanatory variable is the change in the future cash-flow of the firms. As a surrogate of the aggregated future cash-flow of the firms we use the anticipated order stock next period as compared to now in Finnish industry («decreases» answers; the regression coefficient is a priori negative; see in detail Teräsvirta 1984; 3—4 and 21). In an efficient market stock returns change significantly only in response to unanticipated changes in prospects for future cash-flow.

The third exogenous variable is the change in the price of money, that is the change in the returns of bonds or bank deposits. The regression coefficient of this variable is a priori negative.

The fourth explanatory variable is the change in the supply of money. According to Sprinkel's forecasting framework, changes in the stock of money and stock returns are leading indicators for an economy. The lead time for monetary supply is longer than the lead time for stock returns (see Bicksler 1972: 229—230). In that case the changes in monetary supply can be used to predict the general stock returns. According to our hypothesis the rise in money supply will — *ceteris paribus* — rise the stock returns.

The fifth endogenous variable is inflation. According to classical Fisherian hypothesis, common stock of an unlevered firm serves as an effective inflation hedge during anticipated inflation (see Lintner 1975: 270). However, the results of Lintner (1975), Fama and Schwert (1977), Modigliani and Cohn (1979), Feldstein (1980), Kanninen and Kurikka (1984), Pearce and Roley (1985) and Virtanen and Yli-Olli (1987) show that inflation (both anticipated and unanticipated) may have a variety of effects on the real earnings of firms depending on the net monetary position of the firms (see also Sharpe 1985: 251—252) and on the prevailing tax system. However, the latest results of Kanninen and Kurikka (1984) and Virtanen and Yli-Olli (1987) suggest that inflation seems to be good rather

than bad news for the stock market in Finland.

The sixth exogenous variable includes psychological aspects. According to empirical experience we have found that stock returns are parallel in different countries. In addition we have noticed that stock returns in Finland seem to follow stock returns in Sweden. However, during the period to be examined the stock exchanges of Stockholm and Helsinki were isolated from each other and only at the end of the period there were some few firms whose stocks were quoted on both exchanges. The economy of both these countries is very open, export to some extent similar and trade between them on a high level. If stock returns in Finland follow stock returns in Sweden the possible explanation to this is, besides psychological aspects, that the international business cycles will reach Finland lagged compared to Sweden. According to our hypothesis Finnish stock returns follow Swedish stock returns.

2.2.2. *Empirical results*

The least-squares results for econometric models using monthly and quarterly data are presented in Table 3. The results show that the signs of all coefficients — both in the monthly and quarterly equations — are parallel to the hypotheses. The sign of inflation coefficient is opposed to the recent Finnish results (see Kurikka and Kanninen 1984 and Virtanen and Yli-Olli 1987) but parallel to most recent theoretical and empirical results (see e.g. Lintner 1975; Fama and Schwert 1977; Modigliani and Cohn 1979; and Feldstein 1980). In the earlier research (Virtanen and Yli-Olli 1987) the dependent variable was stock prices and the lag-structure between the dependent variable and inflation variable was thirteen months using monthly data and five quarters using quarterly data. In this research, where dependent variable is stock returns, the lag is only three months in the monthly model and one quarter in the quarterly model respectively. The coefficients of all explanatory variables are also significant at least at 5 % level.

The Durbin—Watson statistic shows that autocorrelation is not a problem in the models. Durbin—Watson statistic is biased if lagged endogenous variable appears as an explanatory variable. The bias tends to decrease if there are also exogenous explanatory variables in the model (see Malinvaud 1966: 460—465). However, the absence of any autocorrelation was verified also by Durbin's method, which allows the endogenous variable as an explanatory variable (Durbin 1970: 410—421).

The models in Table 3 include explanatory variables from all the groups — excluding change in the price of money — presented in the preceding section. The own history seems to be an important explanatory variable to stock returns in Finland. In addition, stock returns in Finland follow Swedish stock returns. The lag is four months or two quarters. The anticipated order stock in Finnish industry is a leading indicator to the stock returns. Finally, inflation seems to have a

Table 3. OLS regression summary for the monthly and quarterly econometric models.

Monthly model	Parameter estimates and their t-values (in parenthesis)								100R ²	D—W	RMS
	Constant	Y_{t-1}	$X_{1,t-4}$	$X_{2,t-3}$	$X_{3,t-1}$	$X_{4,t-3}$					
	0.00578 (0.82)	0.261 (3.15)	0.122 (2.42)	-0.00327 (-2.96)	-1.935 (-3.83)	1.201 (2.38)		34.2	2.21	0.000666	
Quarterly model	Constant	Y_{t-1}	$X_{1,t-2}$	$X_{2,t-1}$	$X_{3,t-1}$	$X_{4,t-1}$	100R ²	D—W	RMS		
	-0.0229 (-0.70)	0.353 (3.18)	0.282 (3.43)	-0.00266 (-2.76)	-2.478 (-3.00)	2.554 (3.45)	72.1	1.79	0.00149		

The legend

- Y_t : average stock market returns in Helsinki Stock Exchange
 $X_{1,t}$: average stock market returns in Stockholm Stock Exchange
 $X_{2,t}$: changes in cash-flow expectations in Finnish industry
 $X_{3,t}$: rate of inflation
 $X_{4,t}$: changes in the stock of money
100R² : coefficient of determination (in per cent)
D—W : Durbin—Watson statistics
RMS : residual mean square

negative effect to stock returns on the short term and the rise in the supply of money will rise stock returns in Finland.

Compared to our earlier results (Virtanen and Yli-Olli 1987) the results are now to a certain extent very similar. The econometric models in both the studies include the same explanatory variables. Only the price of money is in this research replaced by the volume of money. In fact, the supply of money already proved to be the alternative explanatory variables to inflation in the case where dependent variable was stock prices. However, we have also differences worthy of notice. The first one is that the own history of the predicted variable was clearly the dominating explanatory variable, especially in monthly but also in quarterly models, when predicting of stock prices was considered. In this study the relative importance of own history is on the same level with the other explanatory variables.

The second difference is a slight change in the lag-structure between the endogenous and some exogenous variables (Swedish stock returns and changes in cash-flow expectations). When stock returns are as the predicted variable the lag is a little longer compared to the model when price level was as the predicted variable.

The third difference appears in the behavior of inflation. The lag-structure between stock market prices and the inflation was quite long: thirteen months when monthly data was used and five quarters when quarterly data was used. The sign of the regression coefficient was positive. The lag-structure between stock returns and the rate of inflation — in this research — is three months when monthly data are used and one quarter when quarterly data are used. The regression coefficient is now negative. One explanation to this difference is that the effect of inflation on the short term seems to be negative and on the long term positive. In addition, the correlation-structure between the variables (among the explanatory variables also) changed when we moved from the level variables into return or difference-variables.

Finally, it seems appropriate to conclude that the results of both the econometric models compared support each other. We can predict both by using univariate time-series analysis and econometric models the monthly and quarterly stock market prices as well as stock returns in the Helsinki Stock Exchange.

2.3. Composite models

When two sets of one-step forecasts are available, it is well known that a linear combination of the two competing forecasts may outperform both of them. In this section we will develop a model for producing a combined forecast for the market return. In deriving the composite model from an ARIMA-model and an econometric model we follow the general guidelines presented by Granger and

Ramanathan (1984) and in financial analysis recently applied by Guerard and Beidleman (1986) and Virtanen and Yli-Olli (1987).

The estimation of the composite model is based on the predicted or fitted values of the two component models during the estimation period 1975—84. The estimation of the composite model started with the ordinary least squares method in the general form: using unconstrained weights for the component models and including a constant term. In all cases the value of the constant term became, however, near zero and was far from being statistically significant. Therefore, the final estimation was carried out applying OLS with unconstrained weights without any constant term. The results of the estimation are presented in Table 4 (monthly data) and Table 5 (quarterly data).

Table 4. Estimated composite models for monthly data.

Model	Weights for components and their t-values			RMS
	ARIMA (0, 1, 1)	ARIMA (0, 1, 1) (0, 1, 1) ¹²	Econometric	
Component models				
ARIMA (0, 1, 1)	1.000	—	—	0.000805
ARIMA (0, 1, 1)(0, 1, 1) ¹²	—	1.000	—	0.000732
Econometric	—	—	1.000	0.000666
Composite models				
Model I	0.266 (1.85)	—	0.835 (5.64)	0.000624
Model II	—	0.427 (3.34)	0.600 (3.70)	0.000603

Table 5. Estimated composite model for quarterly data.

Model	Weights for components and their t-values		RMS
	ARIMA (0, 1, 0) (0, 1, 1) ⁴	Econometric	
ARIMA(0, 1, 0) (0, 1, 1) ⁴	1.000	—	0.00186
Econometric model	—	1.000	0.00149
Composite model	0.357 (2.95)	0.655 (4.57)	0.00100

The results show that in all cases combining an ARIMA-model with the econometric model reduces the residual as compared with those of the component models. The reduction is largest in the composite quarterly model and slightest in that case where the monthly non-seasonal ARIMA (0, 1, 1) and the monthly econometric model are combined. The results indicate that the main contribution with which an ARIMA-component can improve the prediction ability of an econometric model is in the seasonal pattern included.

The results obtained mainly coincide with those found when the time-series and econometric models were combined for developing a composite model for the general stock price level in Finland (Virtanen and Yli-Olli 1987). Also in that study the improvement in the residual was higher when a seasonal time-series model was combined with an econometric model, and the improvement was highest when quarterly prices were considered. Similarly, the unrestricted OLS produced non-significant constant terms. One slight difference may, however, be noticed. The weights for the components add in price models nearly to one, whereas their sum in return models shows some deviation from unity.

3. FORECAST RESULTS

The objective of this section is to compute, using the different forecasting models derived above, forecasts for stock returns from January 1985 to March 1986 (monthly returns) or from the first quarter of 1985 to the first quarter of 1986 (quarterly returns), and to compare forecast accuracy between the different models. In addition, comparisons are made between these results and those obtained when forecasts for the price level were generated.

The development of the *Unitas* index was quite interesting during the forecasting period. The decrease in the share prices, which began in the middle of 1984 continued until summer 1985 after which the prices began to rise again. This development produced mainly negative returns in the beginning of the period and positive ones at the end of the period (cf. Figures 1—4 in the Appendix).

The forecasts were computed as one-step forecasts for the market return. The forecasts were computed using all the models summarized in Tables 4 and 5, both the individual time-series and econometric models and the composite models.

For evaluating the forecasting ability of the models, a variety of more common accuracy measures will be used. That is because different measures give weight to different aspects in the error series. The following standard measures were computed (for definition of the accuracy measures see e.g. Makridakis, Wheelwright and McGee 1983: 43—54 and Flores 1986): mean error (ME), mean absolute error (MAE), mean of squared errors (MSE, analogical to residual mean square, RMS, used in estimation and testing of the models), and root mean of squared error (RMSE). The usual relative measures, e.g. mean percentage error

(MPE) and mean absolute percentage error (MAPE), are not relevant in this connection, because the predicted variable is of the size near zero on the average.

The results of the forecasting procedure are summarized in Tables 6 and 7, for monthly and quarterly returns, respectively.

Table 6. Forecasting accuracy summary (monthly returns).

Model	Accuracy measure			
	ME	MAE	MSE	RMSE
ARIMA (0, 1, 1)	0.01152	0.02366	0.00120	0.03462
ARIMA (0, 1, 1) (0, 1, 1) ¹²	0.00840	0.02576	0.00157	0.03964
Econometric model	-0.00352	0.02127	0.00096	0.03093
Composite model I	-0.00104	0.02060	0.00094	0.03076
Composite model II	0.00176	0.02138	0.00110	0.03312

Table 7. Forecasting accuracy summary (quarterly returns).

Model	Accuracy measure			
	ME	MAE	MSE	RMSE
ARIMA (0, 1, 0) (0, 1, 1) ⁴	0.02230	0.08516	0.00908	0.09528
Econometric model	-0.00207	0.05644	0.00382	0.06180
Composite model	0.00632	0.06556	0.00482	0.06974

When we look at the statistics associated with the monthly forecasts, we see that all the models perform quite well. None of the mean error measures are statistically significant at 5 % level, all the models have produced unbiased forecasts. The error variance measures have not increased alarmingly, either. The increase in the forecast error variance, as compared with the error variance during the estimation period, is not statistically significant at 5 % level for any of the models. All the models are thus statistically acceptable.

When we compare the numerical values of the accuracy measures in Table 6 we see that both the time-series models are clearly outperformed by the econometric model and by the two composite models. Composite model I, which has the econometric model and the non-seasonal ARIMA (0, 1, 1) as its components, and the pure econometric model seem to have produced the lowest error statistics in general. The results indicate a certain change in the seasonal pattern of the return series during the forecasting period as compared with the estimation period. The seasonal component in the ARIMA-model has not performed as well as one could expect on the basis of the estimation results.

As expected, forecasting accuracy of the quarterly models (Table 7) is not as high as that of monthly models. This is natural since useful information is lost due to aggregation in the modelling. The one-step forecasts are also computed further (for three months) ahead. The results can be regarded, however, as fully acceptable. All the forecasts are unbiased (mean errors are not statistically significant at 5 % level). And the error variances have not increased significantly (at 5 % level, composite model at 1 % level) from those in the estimation period.

Out of the three different models in Table 7, the econometric model is clearly the best one. The seasonal ARIMA-model has moderately higher error statistics than the econometric model and even the combination of the ARIMA-model and the econometric model is unable to beat its accuracy. The results seem to confirm a change in the seasonal pattern already obtained in connection with the monthly data.

Finally, the forecasting results may also be considered in comparison with the corresponding forecasting results for the stock prices (Virtanen and Yli-Olli 1987). The two sets of forecasts and their error series are comparable, because all the forecasts have been computed as one-step forecasts using true values for the past values of both the predicted variable and the explanatory variables (in ARIMA-models, the error series are in fact identical for the logarithmed Unitas indices and for its first differences).

As the main conclusion from the comparison we can state that, generally speaking, the difference-formed models have produced more accurate forecasts than the level-formed models. For time-series models the two approaches result in identical error series, in econometric models the forecasting of price changes or returns performs clearly better than the forecasting of prices (the relative improvement in the accuracy measures is of the size of 20 % on the average), and in composite models there is a slight difference in favour of the difference-formed models (a relative accuracy improvement with about 10 % in the monthly models and accuracy on the same level in the quarterly models).

4. SUMMARY

The purposes of this study were, first, to analyze whether stock returns are predictable in a thin security market like the Helsinki Stock Exchange. Second, to compare the forecasting results based on univariate time series analysis and econometric models with each other and to develop composite forecasting models and examine a possible forecasting improvement of those models to the time series models and econometric models. Third, to compare the forecasting results of stock market prices and stock returns with each other.

The theoretical framework of this study has derived its origin from the concept of market efficiency. Under the weak form of efficiency information of past stock

returns is of no value for predicting future returns. Thus the weak form of efficiency is directly opposed to the basic premisses of univariate time-series analysis to forecast stock returns. Similarly, the semi-strong form of efficient market hypothesis is opposed to the basic premisses of econometric analysis to forecast stock returns. Many anomalies and deviations from market efficiency even in the weak form serve a possibility to forecast Finnish stock returns.

The empirical results showed that the Finnish stock returns did not follow the general random walk model. At least one moving average component must be included in the ARIMA-models. We found also both theoretically and empirically satisfactory econometric models to predict stock returns. After that we developed a procedure to generate composite forecasts from univariate time series models and econometric models.

Finally, the results of this study supported the earlier results by the authors, where the dependent variable was stock market prices. However, we had also differences worthy of notice. First, the lag-structures between endogenous and exogenous variables differed to some extent from each other. Second, we had some changes in the effects of some exogenous variables on the endogenous variable. Third, the models in return form produced more accurate forecasts than the models in price level form. This was especially the case, when the results of the econometric models were compared.

REFERENCES

- Berglund, T. (1986). Anomalies in stock returns on a thin security market. Publications of the Swedish School of Economics and Business Administration No. 37.
- Berglund, T., Wahlroos, B. and Grandell, L. (1983). The KOP and the UNITAS indexes for the Helsinki Stock Exchange in the light of a new value weighted index. *Finnish Journal of Business Economics* 32: 1, 30—41 (in Swedish).
- Bicksler, J.L. (1972). A cross-spectral analysis of the lead-lag structure of money supply-stock prices. In Bicksler, J.L. (ed.): *Studies in Business, Technology, and Economics*. Lexington: Heath Lexington Books.
- Box, G.E.P. and Jenkins, G.M. (1970). *Time-Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.
- Copeland, T.E. and Weston, J.F. (1983). *Financial Theory and Corporate Policy*. Addison—Wesley Publishing Company.
- Dixon, W.J., ed. (1983). *BMDP Statistical Software, 1983 Printing with Additions*. Berkeley: University of California Press.
- Driehuis, W. (1972). *Fluctuations and Growth in a Near Full Employment Economy*. Rotterdam: Rotterdam University Press.
- Durbin, J. (1970). Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica* 38, 410—421.
- Fama, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance* 25: May, 383—417.
- Fama, E.F. and Blume, M. (1966). Filter rules and stock market trading profits. *The Journal of Finance*, May, 383—417.
- Fama, E.F. and Schwert, G.W. (1977). Asset returns and inflation. *Journal of Financial Economics* 5: 2, 115—146.

- Feldstein, M. (1980). Inflation, tax rules and the stock market. *Journal of Monetary Economics* 6, 309—331.
- Flores, B.E. (1986). A pragmatic view of accuracy measurement in forecasting. *Omega* 14: 2, 93—98.
- Foster, G. (1986). *Financial Statement Analysis*. Prentice-Hall International Editions.
- Granger, C.W.J. and Morgenstern, O. (1970). *Predictability of Stock Market Prices*. Lexington: Heath Lexington Books.
- Granger, C.W.J. and Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting* 3, 197—204.
- Guerard, J.B. Jr. and Beidleman, C.R. (1986). A new look at forecasting annual corporate earnings in the U.S.A. *European Journal of Operational Research* 23, 288—293.
- Hagin, R.L. (1979). *The Dow Jones—Irwin Guide to Modern Portfolio Theory*. Homewood: Dow Jones—Irwin.
- Hansmann, K.—W. and Zetsche, W. (1985). Forecasting for portfolio selection with a microcomputer. The Institute of Industrial Research of the University of Armed Forces Hamburg, Discussion Papers No. 7.
- Hawawani, G.A. and Michel, P.A., eds. (1984). *European Equity Markets; Risk, Return and Efficiency*. New York: Garland Publishing, Inc.
- Jennergren, L.P. (1975). Filter tests of Swedish share prices. In Elton, E. and M. Gruber (eds.), *International Capital Markets*. North-Holland Publishing Company.
- Jennergren, L.P. and Korsvold, P.E. (1974). Price formation in the Norwegian and Swedish stock markets. Some random walk tests. *Swedish Journal of Economics* 2, 171—185.
- Kanniainen, V. and Kurikka, V. (1984). On the effects of inflation in the stock market: empirical evidence with Finnish data 1968—1981. *Journal of Business Finance & Accounting* 11: 2, 139—150.
- Korhonen, A. (1977). Stock prices, information and the efficiency of the Finnish stock market: empirical tests. *Acta Academiae Oeconomicae Helsingiensis, Series A: 23*.
- Lintner, J. (1975). Inflation and security returns. *The Journal of Finance* 30: 2, 259—280.
- Makridakis, S., Wheelwright, S.C. and McGee, V. (1983). *Forecasting: Methods and Applications, Second Edition*. New York: John Wiley & Sons.
- Malinvaud, E. (1966). *Statistical Methods in Econometrics*. Amsterdam: North Holland.
- Modigliani, F. and Cohn, R.A. (1979). Inflation, rational valuation and the market. *Financial Analysts Journal* 1979: 2.
- Pearce, D.K. and Roley, V.V. (1985). Stock prices and economic news. *The Journal of Business* 58: 49—67.
- Sharpe, W.F. (1985). *Investments, Third Edition*. Englewood Cliffs: Prentice-Hall, Inc.
- Teräsvirta, T. (1984). Forecasting the output of the Finnish metal industry using business survey data. The Research Institute of the Finnish Economy. Discussion Papers No. 157.
- Umstead, D.A. (1977). Forecasting stock market prices. *The Journal of Finance* 32: 2, 427—441.
- Wahlroos, B. and Berglund, T. (1983). The January effect on a small stock market: lumpy information and tax-loss selling. CMSEMS. Discussion Papers No 579. Northwestern University.
- » — (1986). Anomalies and equilibrium returns in a small stock market. *Journal of Business Research* 14, 423—440.
- Virtanen, I. and Yli-Olli, P. (1987). Forecasting Stock Market Prices in a Thin Security Market. *Omega* 15: 2, 145—155.