

Forecasting and forecast-combining of quarterly earnings-per-share via genetic programming

Arturo Rodríguez.
Facultad de Economía y Negocios
Universidad de Chile
arodriguez@unegocios.cl

Joaquín Trigueros
Capula Global

Abstract

In this study we examine different methodologies to estimate earnings. More specifically, we evaluate the viability of Genetic Programming as both a forecasting model estimator and a forecast-combining methodology. When we compare the performance of traditional mechanical forecasting (ARIMA) models and models developed using Genetic Programming we observe that Genetic Programming can be used to create time-series models for quarterly earnings as accurate as the traditional linear models. Genetic Programming can also effectively combine forecasts. However, Genetic Programming's forecast combinations are sometimes unable to improve on Value Line. Moreover, simple averaging of forecasts results in better predictive accuracy than Genetic Programming-combining of forecasts. Hence, as implemented in this study, Genetic Programming is not superior to traditional methodologies in either forecasting or forecast combining of quarterly earnings.

Keywords: Genetic Programming, Earnings forecasts, Combining Forecast.

Resumen

En este estudio examinamos distintas metodologías para estimar beneficios. En concreto, analizamos la viabilidad de la programación genética como modelo de estimación y también como método para combinar distintos modelos de predicción. En nuestro estudio encontramos que la programación genética es al menos tan eficaz como los métodos tradicionales de estimación de series de tiempo. Sin embargo las predicciones de nuestros algoritmos genéticos no predicen mejor los beneficios que el *Value Line*. Además, los modelos de predicción combinados utilizando programación genética no son superiores a un simple promedio de los estimados de los modelos. De manera, que en nuestro estudio la programación genética no resulta superior a las metodologías tradicionales.

Palabras claves: Programación genética, predicción de beneficios, combinación de predicciones.

1. Introduction

There is often a need in accounting research for a proxy of the markets' expectation of earnings. Quarterly earnings forecasts from both analysts and mechanical models have been traditionally used as such proxies. Mechanical models include ARIMA models introduced by Foster (1977), Watts (1975)-Griffin (1977), Brown and Rozeff (1979) and individually identified models. Lee and Chen (1990) introduce models that account for structural change in a firm's quarterly earnings series. Mechanical models, although easy to implement, have unfortunately never been able to consistently beat analysts' forecasts, of which Value Line (VL) is an example. However, Lee and Chen (1990) show that both Value Line and ARIMA models have marginal earnings forecasting power. Thus, a

combination of them might prove better than each individual forecast. Indeed, Newbold, Zumwalt and Kannan (1987) show that when for a group of utilities the Value Line forecast is combined with the Brown-Rozeff forecast, a forecast error smaller than either's is obtained. Newbold *et al.* use simple regression to combine their forecasts. Lately, with the advent of new nonlinear computing paradigms, other researchers have turned to these nonlinear methods for forecast combining. Donaldson and Kamstra (1996) use Artificial Neural Networks (ANNs) to combine forecasts of stock market volatility for several countries. They show that in that context, forecasts combined using ANNs are superior to those combined using traditional linear procedures. Hibon and Evgeniou (2005) by using simulations show that the accuracy of forecast combinations is significantly better and less variable than that of individual forecasts. Also, Terui and Van Dijk (2002) show that combined forecast performed particularly well for macroeconomic time series with time varying coefficients. Finally, Yang (2004) provides the theoretical basis for the combined forecast superiority over their individual components. In this study we evaluate the viability of Genetic Programming, a nonlinear methodology, as (1) a forecasting model estimator and (2) a forecast-combining methodology. When we compare the performance of traditional mechanical forecasting (ARIMA) models and models developed using Genetic Programming we observe that Genetic Programming can be used to create time-series models for quarterly earnings as accurate as the traditional linear models. Genetic Programming can also effectively combine forecasts. However, unlike Donaldson and Kamstra and Newbold *et al.*, Genetic Programming's forecast combinations are sometimes unable to improve on Value Line, their most accurate component. Moreover, simple averaging of forecasts results in better predictive accuracy than Genetic Programming-combining of forecasts. Hence, as implemented in this study, Genetic Programming is not superior to traditional methodologies in either forecasting or forecast combining of quarterly earnings. In

section 2 we give a brief summary of quarterly earnings forecasting and forecast combining research. In section 3 we present the data used in the study. Section 4 presents our methodology. Section 5 presents results. Section 6 concludes.

2. Quarterly earnings forecasting and forecast-combining research

Research on the time-series properties of quarterly earnings dates back to the 1970's. The core models for quarterly earnings processes are those proposed in Foster (1977), Watts (1975) - Griffin (1977) and Brown and Rozeff (1979). In Box-Jenkins notation, where ARIMA models are described in the form $(p,d,q) \times (P,D,Q)$, the Foster model is $(1,0,0) \times (0,1,0)$ plus a constant, the Watts-Griffin model $(0,1,1) \times (0,1,1)$, and the Brown-Rozeff model $(1,0,0) \times (0,1,1)$. Evidence has been presented in favor of all three models. Their forecasting accuracy, though not as good as Value Line's, remained unbeaten until Lee and Chen (1990) introduced models incorporating structural change information. However, the Lee and Chen models were not able to make better forecasts than Value Line.

Research in forecast combining dates back to the late 1960's with the publication of Reid (1968,1969) and Bates and Granger (1969). Forecast combining research is extensive. It repeatedly shows that combinations of forecasts are more accurate than individual forecasts. Most combination techniques are linear and involve the averaging of individual forecasts. The simplest type of averaging is equally-weighted averaging. It is the usual benchmark against which other methods are evaluated. Many other weighting schemes have also been proposed. To mention a few, Doyle and Fenwick (1976), proposed historical weightings. Bunn and Kappos (1982) assigned weights according to a function of the historical records of the most accurate forecast in all previous periods. Gupta and Wilton (1987) introduced the odds-matrix method of computing

weights. Holmen (1987) used the constrained linear combination method to combine forecasts of short-term earnings. Russell and Adam (1987) proposed weightings based on actual forecast errors. Non-averaging combining techniques include constrained multiple objective linear programming, unconstrained multiple objective linear programming, and artificial neural networks. Of particular interest to us is Newbold *et al.*, (1987), who use OLS to combine pairs of forecasts for a set of utilities' quarterly earnings series. All forecast pairs include a Value Line forecast. The other member of the pair is one of several mechanical models. Among the models considered are the Watts-Griffin and Brown-Rozeff ARIMA models. Newbold *et al.*, find that combining Brown-Rozeff with Value Line gives the most accurate forecasts. The combined Value Line-Brown-Rozeff forecasts are more accurate than either individual forecast. Also of interest to us is the Donaldson and Kamstra (1996) study. Donaldson and Kamstra argue that non-linear combining methodologies may sometimes fare better than linear techniques. They successfully use artificial neural networks to combine stock market volatility forecasts. The ANNs' forecast combinations are more accurate than those obtained by several averaging methods, including equal-weighted averaging and OLS. These results motivated us to consider combining Value Line forecasts with those of different ARIMA models. However, like Donaldson and Kamstra, we chose a non-linear combining technique. Specifically, we chose to evaluate Genetic Programming as both a forecasting and a forecast-combining methodology.

3. Data

We used all the utilities from the Lee and Chen (1990) study which satisfied the following two requirements:

- i) A Value Line forecast was available for our required time periods.

- ii) Earnings per share before extraordinary items (COMPUSTAT item # 19) was available for our required time periods.

Earnings per share (EPS) values were available for our firms from the first quarter of 1963 (1963Q1) onwards. Value Line forecasts were available from 1968Q3 onwards. We used all available EPS data for estimation of the ARIMA models under consideration but truncated the data sets used to estimate the Genetic Programming models to 1968Q3 onwards to match the availability of Value Line data. Table 1 presents the firms used in our study.

Table 1
Firms in the study

	NAME	INDIVIDUAL ARIMA PROCESS
1	Carolina Power & Light	(0,0,1)(0,1,1) ₄
2	Central Hudson Gas & Electric	(0,0,1)(0,1,1) ₄
3	Cincinnati Gas & Electric	(0,0,1)(0,1,1) ₄
4	Cleveland Electric Illumination	(0,0,1)(0,1,0) ₄
5	Dayton Power & Light Inc.	(0,0,1)(0,1,1) ₄
6	Detroit Edison Co.	(0,0,4)(0,0,0) ₄
7	Idaho Power Co	(4,0,0)(0,0,0) ₄
8	Interstate Power Co	(0,0,0)(0,1,1) ₄
9	Long island Lighting	(0,0,4)(0,1,0) ₄
10	New York state Electric & Gas	(0,0,0)(1,0,0) ₄
11	Niagara Mohawk Power	(2,0,0)(1,0,0) ₄
12	Northern states Power/MN	(0,0,1)(0,1,1) ₄
13	Pacific Gas & Electric	(0,0,0)(1,0,0) ₄
14	Potomac Electric Power	(0,0,0)(0,1,1) ₄
15	San Diego Gas & Electric	(4,0,0)(0,0,0) ₄
16	Union Electric co	(0,0,1)(0,1,1) ₄
17	Washington Water Power	(0,0,1)(0,1,1) ₄

The utility firms used in the study are a subset of the firms used in Lee and Chen (1990) for which we found all required information. When using an individually identified ARIMA model, we used the one identified by Lee and Chen in their study.

4. Methodology

We assessed the forecasting performance of ten different models. These are presented in Table 2. Model 1 is the Value Line forecast. Models 2-5 are traditional ARIMA models. Model 6 is “Genetic Programming ARIMA”. Its dependent variables are lagged values of EPS and lagged errors. Model 7 (VIBFG) combines the forecasts of Value Line, and the individually identified, Brown & Rozeff, Foster, and Watts-Griffin models. Model 8 is an enhanced version of model 7. Lagged values of EPS and lagged errors are added to the set of model 7 independent variables. Model 9 combines the forecasts of model X (Genetic Programming ARIMA) with the traditional forecasts. Model 10 is an enhanced model 9 which again includes lagged EPS values and lagged errors. The operations we allowed in the models constructed by the Genetic Programming algorithm included not only the traditional arithmetic operators of the linear ARIMA models but also three averaging operators. These are presented in Table 3.

Each model we estimated was in reality a collection of models estimated with data determined by a rolling window. In pre-specified model estimation (ARIMA) only the model parameters change as the window rolls forward in time. In Genetic Programming estimation, the entire functional form changes. We made the first Genetic Programming estimation using the quarterly earnings series for the period 1968Q3 to 1980Q4. We then collected 1, 2 and 3 period forecasts for 1981Q1 - 1981Q3, the associated validation period. We next shifted the estimation/validation window over by one quarter and repeated this process. Table 4 presents beginning and ending validation/testing periods. For each period, the dependent variable was the quarterly EPS value for the period. Depending on the model, the independent variables were lagged EPS values, lagged model errors, the one-period-ahead Value Line forecast, and values given by the different ARIMA models used in the study.

Table 2
Forecasting Models

	MODEL NAME	MODEL COMPONENTS	DESCRIPTION	REFERENCE
1	VL	Multiple	Black Box	Value Line
2	INDIV	Lagged X_{i,a_i}	ARIMA	Lee and Chen (1990)
3	BR	Lagged X_{i,a_i}	ARIMA	Brown and Rozeff (1979)
4	FOSTER	Lagged X_i	ARIMA	Foster (1977)
5	GW	Lagged X_{i,a_i}	ARIMA	Watts (1975) Griffin(1977)
6	X	Lagged X_{i,a_i}	Pure GP ARIMA	This study
7	VIBFG	VL, INDIV, BR,FOSTER, G	GP Forecast Combination	This study
8	VIBFGX	VL,INDIV,BR, FOSTER,G, lagged X_{i,a_i}	Enhanced GP Forecast Combination	This study
9	GPVIBFG	Model 5, VL, INDIV, BR, FOSTER, G	GP Forecast Combination	This study
10	GPVIBFGX	Model 5, VL, INDIV, BR, FOSTER, G, lagged X_{i,a_i}	Enhanced GP Forecast Combination	This study

Traditional models evaluated in the study in the study were financial analyst forecasts (Value Line - VL) and several time series models (Individually identified models – INDIV, Brown and Rozeff – BR, Foster – FOSTER, and Griffin-Watts – GW). Our contribution was to evaluate a GP ARIMA model and several GP forecast-combining and “enhanced” forecast-combining models.

Table 3
Formula Operations

OPERATION	DESCRIPTION
+	Addition
-	Subtraction
×	Multiplication
%	Protected Division : $x\%y = x/y$ if $y \neq 0$, 1 otherwise
avg2	$\text{avg2}(x,y) = \frac{x+y}{2}$
gavg2	$\text{gavg2}(x,y) = \sqrt{ x \times y }$
gavg3	$\text{gavg3}(x,y,z) = \sqrt[3]{x \times y \times z}$

Basic set of operations available for the construction of Genetic Programming models.

Table 4
ARIMA Training/Validation/Test Dates

	TRAINING PERIOD	VALIDATION PERIOD	TESTING PERIOD
First Validation Estimation	1963 Q1-1980Q4	1981Q1-1981Q3	NA
Last Validation Estimation	1964Q4-1982Q3	1982Q4	NA
First Testing Estimation	1965Q1-1982Q4	NA	1983Q1-1983Q3
Last Testing Estimation	1966Q4-1984Q3	NA	1984Q4-1985Q2

This table describes the time spans of the different data sets used in developing and testing our formulas. Training data sets were truncated to begin in the third quarter of 1968 to match the availability of Value Line data.

The Genetic Programming algorithm can be run under different conditions or parameters. We did not know what parameters were best for evolving good forecasting models. Therefore, we evolved formulas using an in-sample period using a variety of parameters. We then chose as “best” parameters those that resulted in the smallest forecasting errors in a validation period. Finally, we rolled the time windows forward and used these parameters to estimate models and produce forecasts in a test period. We varied four algorithm parameters: the method of selecting “parent” formulas for reproduction, the population size, the number of generations for which the formulas were evolved, and the weightings assigned to the estimation errors. These parameters appear in Table 5.

We considered forecasting horizons separately when choosing parameters. For each forecasting horizon (one, two or three periods ahead) we chose the best training parameters in one of two ways:

- i) We averaged the validation-period errors for all firms and then picked the parameters that led to the smallest average error. We call this the GLOB method.
- ii) For each firm we picked the firm-specific parameters that resulted in the minimum errors in the validation period. We call this the INDIV method.

Whenever forecasting for a specific horizon in the out-of-sample (test) period we would the estimate Genetic Programming model using the best parameters for the horizon.

Table 5
GP Training Parameter Variations

1. Selection Method	1. Best 2. Fitness	Method of selecting "parent" models for potential combination by the crossover operator.
2. Population Size	1. 100 2. 500 3. 1000	Size of model population.
3. Generations	1. 2 2. 5 3. 10	Number of populations successively evolved using the selection-crossover operators.
4. Error weights	1. Equal 2. Front	Method of weighting the errors. Frontal weighting gives higher weights to errors made in more recent forecasts.

We developed formulas using different parameters and chose the parameters that resulted in best forecasting performance in a validation period. We then used these parameters to develop new formulas and tested their forecasting power in an out-of-sample period.

We performed the study twice. The first time our Genetic Programming estimations involved minimizing sum of squared errors. The second time we minimized the sum of absolute errors. Sum of squared errors is the criterion traditionally minimized in most statistical estimations. A sample model estimated by the algorithm is shown in Equation 1.

$$X_t = \text{gavg3}(X_{t-2}, X_{t-1}, \text{gavg3}(X_{t-4}, 0.21444, X_{t-1})) \quad (1)$$

5. Results

Table 6 presents the average out-of-sample percentage errors made by all models. Value Line has the smallest errors of the non-combining models. We can also see that the performance of Genetic Programming ARIMA (model X) is comparable but not superior to that of the linear ARIMA models. Some Genetic Programming

combination methods also manage to improve on the component forecasts for certain forecasting horizons. However, a simple arithmetic average of the traditional methods always dominates the Genetic Programming combination forecasts.

We also considered the possibility that Genetic Programming might work for some types of time series better than for others. More specifically, we divided our firms into high and low earnings-volatility groups. Table 6 also presents the forecast errors the models made for both groups of firms. One can see that the performance of Genetic Programming relative to the other methodologies does not seem to be a function of volatility.

Like Lee and Chen we also looked at the marginal forecasting power of the mechanical forecasting models versus Value Line by estimating the following regression

$$Z_t = \alpha + \beta_{\text{model}} f_{\text{model}} + \beta_{\text{vl}} f_{\text{vl}} + \varepsilon_i \quad (2)$$

where Z_t is actual earnings, f_{model} is the forecast of a model, and f_{vl} is the Value Line forecast. The values of the estimated coefficients β_{model} and β_{vl} represent the information content of the Value Line and model forecasts. Table 7 presents the results of the regression for different models. Except for the INDIV Genetic Programming two-quarters-ahead models, all Genetic Programming models seem to contain some information. However, the amount of information is no larger than that of the traditional models and is markedly less than that of Value Line.

Table 6
Average Percentage Forecasting Error in Out-of-Sample Periods

Panel A: All Firms

MODEL	GP PARAMETER CHOICE	QUARTERS AHEAD		
		1	2	3
VL	N/A	13.1	14.1	14.8
INDIV	N/A	15.7	16.7	18.1
BR	N/A	15.0	16.6	17.1
FOSTER	N/A	14.6	13.8	14.0
GW	N/A	14.3	14.6	14.5
AVERAGE OF VL & ARIMA MODELS	N/A	13.0	13.5	14.4
Objective Function: Sum of Absolute Errors				
X	GLOB	15.2	15.9	16.3
X	INDIV	16.8	16.8	15.9
VIBFG	GLOB	13.9	13.7	15.4
VIBFG	INDIV	13.6	14.3	14.8
VIBFGX	GLOB	13.3	14.0	14.7
VIBFGX	INDIV	13.2	14.3	15.4
GPVIBFG	GLOB	13.3	13.6	14.9
GPVIBFG	INDIV	13.5	14.0	15.2
GPVIBFGX	GLOB	13.5	14.2	15.4
GPVIBFGX	INDIV	13.2	14.5	14.3
Objective Function: Sum of Squared Errors				
X	GLOB	16.2	16.6	16.2
X	INDIV	15.7	17.2	16.9
VIBFG	GLOB	13.1	14.6	16.2
VIBFG	INDIV	13.6	14.7	15.7
VIBFGX	GLOB	13.4	14.0	15.6
VIBFGX	INDIV	13.9	13.6	15.3
GPVIBFG	GLOB	13.3	14.7	16.5
GPVIBFG	INDIV	13.4	14.4	16.0
GPVIBFGX	GLOB	13.4	14.3	15.5
GPVIBFGX	INDIV	13.2	13.8	15.4

(Continued)

Panel B: Low-Volatility Firms

MODEL	GP PARAMETER CHOICE	QUARTERS AHEAD		
		1	2	3
VL	N/A	10.1	12.2	12.5
INDIV	N/A	15.6	17.5	20.3
BR	N/A	14.0	15.9	16.9
FOSTER	N/A	12.0	11.4	10.9
GW	N/A	11.8	12.6	12.6
AVERAGE OF VL & ARIMA MODELS	N/A	10.5	11.7	12.7
Objective Function: Sum of Absolute Errors				
X	GLOB	14.8	14.5	15.0
X	INDIV	15.1	14.6	14.0
VIBFG	GLOB	10.7	12.0	14.4
VIBFG	INDIV	10.4	12.2	12.7
VIBFGX	GLOB	10.8	12.4	12.4
VIBFGX	INDIV	10.6	12.7	13.2
GPVIBFG	GLOB	10.9	11.9	13.3
GPVIBFG	INDIV	10.8	11.8	13.8
GPVIBFGX	GLOB	11.3	12.8	13.0
GPVIBFGX	INDIV	10.4	13.0	12.1
Objective Function: Sum of Squared Errors				
X	GLOB	15.8	16.3	15.4
X	INDIV	14.3	17.4	16.6
VIBFG	GLOB	11.1	13.7	15.0
VIBFG	INDIV	11.4	14.2	15.1
VIBFGX	GLOB	11.5	12.4	13.9
VIBFGX	INDIV	11.7	12.0	13.0
GPVIBFG	GLOB	11.1	13.5	16.0
GPVIBFG	INDIV	10.7	12.6	14.8
GPVIBFGX	GLOB	11.2	12.9	14.5
GPVIBFGX	INDIV	11.7	12.2	13.7

(Continued)

Panel C: High-Volatility Firms

MODEL	GP PARAMETER CHOICE	QUARTERS AHEAD		
		1	2	3
VL	N/A	15.4	15.6	16.6
INDIV	N/A	15.7	16.0	16.4
BR	N/A	15.8	17.0	17.3
FOSTER	N/A	16.7	15.7	16.4
GW	N/A	16.3	16.2	16.0
AVERAGE OF VL & ARIMA MODELS	N/A	15.0	14.9	15.8
Objective Function: Sum of Absolute Errors				
X	GLOB	15.5	17.0	17.2
X	INDIV	18.1	18.5	17.3
VIBFG	GLOB	16.3	15.0	16.2
VIBFG	INDIV	16.2	15.9	16.5
VIBFGX	GLOB	15.2	15.3	16.5
VIBFGX	INDIV	15.2	15.6	17.1
GPVIBFG	GLOB	15.2	15.0	16.1
GPVIBFG	INDIV	15.5	15.7	16.2
GPVIBFGX	GLOB	15.3	15.4	17.2
GPVIBFGX	INDIV	15.4	15.6	16.0
Objective Function: Sum of Squared Errors				
X	GLOB	16.4	16.9	16.8
X	INDIV	16.8	17.0	17.1
VIBFG	GLOB	14.7	15.4	17.2
VIBFG	INDIV	15.2	15.2	16.3
VIBFGX	GLOB	14.8	15.4	16.9
VIBFGX	INDIV	15.6	14.9	17.0
GPVIBFG	GLOB	15.0	15.7	16.8
GPVIBFG	INDIV	15.5	15.8	17.0
GPVIBFGX	GLOB	15.2	15.4	16.2
GPVIBFGX	INDIV	14.3	15.0	16.7

We first searched for training parameters that resulted in formulas with the smallest forecast errors in a validation period. We then used these parameters to train formulas for the test periods. Parameters were chosen in one of two ways: (1) GLOB – choose those training parameters that result in the smallest average error over *all* firms in the validation period or (2) INDIV: train individual models for each firm using a firm's own best parameters in the validation period

Table 7
Marginal Forecasting Power of Model X (GP ARIMA)
and ARIMA vis-à-vis Value Line

	SUM ABS. ERRORS		SUM SQ. ERRORS		INDIV	BR	FOSTER	GW
	INDIV GP	GLOB GP	INDIV GP	GLOBGP				
One Period Ahead								
β_{model}	0.18	0.32	0.24	0.22	0.27	0.41	0.30	0.30
(t-statistic)	2.00	3.53	2.49	2.50	2.87	4.16	3.00	3.11
β_{vl}	0.83	0.69	0.77	0.78	0.72	0.59	0.69	0.70
(t-statistic)	9.40	7.63	8.06	8.72	7.29	5.89	6.61	7.04
R^2	0.88	0.89	0.88	0.88	0.88	0.89	0.88	0.88
Two Periods Ahead								
β_{model}	0.22	0.27	0.15	0.20	0.21	0.33	0.26	0.22
(t-statistic)	2.68	2.78	1.56	1.95	2.28	2.95	2.28	2.13
β_{vl}	0.80	0.76	0.86	0.82	0.80	0.68	0.75	0.79
(t-statistic)	9.45	8.02	8.93	8.38	8.54	5.93	6.38	7.34
R^2	0.89	0.89	0.88	0.89	0.89	0.89	0.89	0.89
Three Periods Ahead								
β_{model}	0.34	0.41	0.35	0.41	0.24	0.58	0.48	0.47
(t-statistic)	3.78	3.92	3.56	3.94	2.98	4.63	4.02	4.22
β_{vl}	0.68	0.62	0.66	0.62	0.75	0.43	0.53	0.53
(t-statistic)	7.49	6.02	6.63	6.05	8.53	3.37	4.36	4.59
R^2	0.87	0.87	0.86	0.87	0.86	0.87	0.87	0.87

We estimated the following regression: $z_t = \alpha + \beta_{\text{model}}f_{\text{model}} + \beta_{\text{vl}}f_{\text{vl}} + \varepsilon_t$

where z_t = realized earnings, f_{model} = (non-value-line) forecast, f_{vl} = Value Line forecast. The size and significance of the β coefficients are a measure of the information content of the forecasts.

6. Conclusion

The results of the study indicate that, as implemented, the Genetic Programming methodology is capable of producing forecasts and forecast combinations of average quality. However, it has no advantage over more traditional methods. Moreover, the marginal information captured by the Genetic Programming methodology is no larger than that captured by more traditional models.

References

- BROWN, L. and M. ROZEFF (1979), "Univariate time series models of quarterly earnings per share", a proposed model, *Journal of Accounting Research*, Vol. 17, pp. 179-189.
- BUNN, D. W. and E. KAPPOS (1982), "Synthesis or selection of forecasting models", *European journal of Operational Research*, Vol. 9, pp. 173-180.
- DOYLE, P. and I. FENWICK (1976), "Sales forecasting: Using a combination of approaches", *Long Range Planning*, Vol. 9, pp. 54-60.
- FOSTER, G. (1977), "Quarterly accounting data: time series properties and predictive ability results", *The Accounting Review*, Vol. 52, pp. 1-21.
- GRIFFIN P.A. (1997), "The time series behavior of quarterly earnings: preliminary evidence", *Journal of Accounting Research*, Vol. 15, pp. 71-83.
- GUPTA, S. and P.C. WILTON, (1988), "Combination of forecasts: An extension", *Management Science*, Vol. 33, pp. 356-372.
- HIBON, M. and T. EVGENIOU (2005), "To combine or not to combine: selecting among forecasts and their combinations", *International Journal of Forecasting*, Vol. 21, pp. 15-24.
- LEE, C.W.J. and C. CHEN (1990), "Structural changes and the forecasting of quarterly earnings in the utility industry", *Journal of Accounting and Economics*, Vol. 13, pp. 93-122.
- NEWBOLD, P., J. KENTON ZUMWALT, SRINIVASAN KANNAN (1987), "Combining forecasts to improve earnings per share prediction: An examination of electric utilities industry", *International Journal of Forecasting*, Vol. 3, pp. 229-238.
- DONALDSON, R. and M. KAMSTRA (1996), "A new dividend forecasting procedure that rejects bubbles in asset prices", *Review of Financial Studies*, Vol. 9, pp. 333-383.
- RUSSELL, T.D. and E. ADAM (1987), "An empirical evaluation of alternative forecasting Combinations", *Management Science*, Vol. 33, pp. 1267-1276.



TERIU, N. and H. VAN DIJK (2002), "Combined forecasts from linear and nonlinear time series models", *International Journal of Forecasting*, Vol. 18, pp. 421-438.

YANG, Y. (2004), "Combining forecasting procedures: some theoretical results", *Econometric Theory*, Vol. 20, pp. 176-222.

