BEBR

FACULTY WORKING
PAPER NO. 1292

Direct versus Indirect ARIMA Forecasts of Defined Variables: Some Further Evidence Based on Corporate Accounting Data

Peter A. Silhan

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1292

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

October 1986

Direct Versus Indirect Arima Forecasts of Defined Variables:
Some Further Evidence Based On Corporate Accounting Data

Peter A. Silhan, Associate Professor
Department of Accountancy

Direct Versus Indirect ARIMA Forecasts of Defined Variables:
Some Further Evidence Based on Corporate Accounting Data

ABSTRACT

In a recent study, Kang (1986) provided some empirical evidence indicating that a number of defined economic variables could be predicted more accurately with ARIMA models based indirectly on their components than directly on the defined variables, themselves. The study presented here uses corporate accounting data data to provide some large-sample evidence on such comparisons. It shows for corporate profit margins that the more parsimonious direct forecasts were not outperformed by the less parsimonious indirect forecasts.

Direct Versus Indirect ARIMA Forecasts of Defined Variables: Some Further Evidence Based on Corporate Accounting Data

1. INTRODUCTION

A "defined" variable is a variable constructed from two or more other variables. Real GNP, for example, is defined as nominal GNP divided by the GNP deflator. In a recent study, Kang (1986) provided some empirical evidence indicating that a number of defined economic variables (real GNP, real interest rate, and M1 velocity) could be predicted more accurately with ARIMA models based <u>indirectly</u> on the underlying components than with corresponding ARIMA forecasts based <u>directly</u> on the defined variables, themselves. He also found, however, that for at least one defined variable (M1 multiplier) the "direct" forecasts were not outperformed by the "indirect" forecasts. Furthermore, because the above results were based, in effect, on samples of one for each of only four macroeconomic variables, he suggested that additional research would be needed to learn more about ARIMA forecasts of a variety of defined variables.

In the current study, corporate accounting data are used instead of macroeconomic data. This makes it possible to use a large sample for comparing direct versus indirect ARIMA forecasts. Net profit margin, which is calculated by dividing net income by net sales, is used as the variable of interest.

Net profit margin is one of the most widely used variables in the business community. In a recent behavioral experiment, this was evidenced when financial analysts requested this ratio more often than any other ratio when they were asked to assess corporate profitability (Biggs, 1984). Business periodicals, such as Business Week, routinely publish net profit margins by company by industry on a quarterly basis.

2. TIME SERIES AGGREGATION

Much theoretical and empirical research has focused on the issue of time series aggregation. In particular, the choice between direct and indirect forecasts of a single aggregated variable has been of considerable interest.

In the statistics literature, Engel (1984) provides a unified approach to time series aggregation. He identifies three basic types of aggregation (sums, products, and intertemporal aggregations) and discusses several necessary and sufficient conditions which determine a variety of aggregation functions. Because intertemporal aggregations and sums are not consistent with the way that defined variables are constructed, it appears that defined variables, such as the ones examined by Kang (1986), can be viewed as multiplicative aggregations that are products of their components.

Analytically, Wecker (1978) and Dossou-Gbete, Ettinger, and De Falguerolles (1980) focus attention on the products of dependent and independent ARMA processes, respectively. They set forth the necessary and sufficient conditions under which ARMA processes would result from the product of ARMA processes. Wecker (1978) identifies a number of situations in economics where ARMA processes arise as the product of two independent AR processes. One situation, for example, is where corporate sales, Z, are derived as the product of national demand, X, and market share, Y.

However, theoretical analyses, such as the ones cited above, cannot address the issue of predictive ability. This is because in practice the underlying processes must be identified and estimated. Therefore in this area predictive ability has always been an empirical issue. Wei and Abraham (1981, p. 1343) point out that "there is no guarantee that the forecast based on a component series is . . . more efficient than the forecast from a single univariate aggregate series." Lutkepohl (1984, p. 213) notes that "if the underlying processes are not known and have to be specified and/ or estimated on the basis of the available sample information, the resulting MSE's of the univariate [direct forecasts] may be smaller than the MSE's of the [direct] forecasts."

In the accounting literature, potential gains in predictive ability due to disaggregation have been discussed with respect to a number of accounting disclosure policies, such as interim reporting (e.g., Cogger, 1981; Hopwood, McKeown, and Newbold, 1982) and segment reporting (e.g., Ang, 1979; Barnea and Lakonishok, 1980; Hopwood, Newbold, and Silhan, 1982; Silhan, 1982, 1984). Interim reporting and segment reporting, however, involve intertemporal aggregations and sums, respectively. The issue of comparing direct versus indirect forecasts of various defined variables, such as net profit margin, has not received such attention.

Kang (1986) suggests that aggregation research, such as the research cited above, can be extended to defined variables because indirect forecasts of defined variables are constructed in the same way as indirect forecasts of aggregated variables. In essence, he views aggregation as a special case of definition. Therefore, forecasting defined variables appears to be more fundamental than forecasting aggregated variables.

Defined variables can be (1) developed at every level of aggregation and (2) constructed from diverse components (Kang, 1986, p.82). Therefore, an aggregation problem with respect to a defined variable can be more complex than a similar problem with respect to another type of aggregated variable. This is due, in part, from the fact that the underlying components of a defined variable are more likely to be heterogeneous than those associated with intertemporal aggregations and sums.

The current study compares the predictive ability of direct versus indirect ARIMA forecasts of corporate profit margins. This variable can be viewed as the product of two ARIMA processes (the net income series multiplied by the inverse of the net sales series) or equivalently as the quotient of two series (the net income series divided by the net sales series). The results presented here are based on the second definition.

3. RESEARCH DESIGN

To evaluate the predictive ability of direct versus indirect forecasts of corporate profit margins, ARIMA forecasts (Box and Jenkins, 1970) of margins (for the direct forecasts) and earnings and sales (for the indirect forecasts) were projected into a five-year holdout period (1978-82). Forecast errors were computed for this period.

Metrics

Two metrics, mean absolute error (MAE) and mean error (ME), were used to measure forecasting performance. Notationally, these metrics can be represented as follows:

$$MAE = Avg\ [Abs\ (P - A)]$$

and

$$ME = Avg\ [(P - A)]$$

where Abs is the absolute value operator, P is predicted net profit margin, and A is actual profit margin. These metrics can be viewed as measures of accuracy and bias, respectively.

Data Sample

Every manufacturing and retailing company with a complete sales and earnings history for the 68 consecutive quarters ending with the fourth quarter of 1982 (1966-I to 1982-IV) was screened from the quarterly COMPUSTAT industrial tape. Each company was required to have only one fiscal year and one industry affiliation throughout this 17-year period. In all, 172 firms qualified for inclusion in this sample.

Quarterly net income (Item 8) and quarterly net sales (Item 2) were selected as the COMPUSTAT variables of interest. Profit margins were constructed by dividing Item 8 (net income) by Item 2 (net sales).

ARIMA Models

Firm-specific ARIMA models were individually identified and estimated for margins, earnings, and sales for each of four two-year holdout periods (1978-82). Together, re-identification and re-estimation tend to produce the most accurate univariate ARIMA forecasts (McKeown and Lorek, 1978). An automated search and estimation routine was used to individually identify and estimate each time series (see Hopwood (1980) for a general discussion of these procedures). In all, there were 2,064 models individually identified and estimated (172 firms x 4 two-year periods x 3 variables). In all, there were 11,008 quarterly predictions (t+1 to t+8) for 1978, 1979, 1980, and 1981, that were based on 48, 52, 56, and 60 observations, respectively.

4. EMPIRICAL RESULTS

The results presented here do not support using an indirect approach when forecasting net profit margins. On average, the indirect ARIMA forecasts of corporate profit margins did not outperform the comparable direct ARIMA forecasts.

Accuracy

Table 1 presents the MAE comparisons for the eight horizons measured (t+1 to t+8). It shows that for virtually every horizon projected into the holdout period the direct ARIMA forecasts were not outperformed by the indirect ARIMA forecasts.

Bias

Table 2 presents the ME comparisons across the eight horizons examined. It shows that the direct forecasts generally tended to overpredict corporate margins, while the indirect forecasts tended to underpredict corporate margins.

Directional Agreement

Table 3 provides information about directional agreement. It shows that the direct and indirect forecasts were in agreement with respect to underpredictions ($P < A$) and overpredictions ($P > A$) in approximately 80 percent of the comparisons. In all, there were 5,504 comparisons (8 horizons x 4 two-year periods x 172 firms).

The direct forecasts underpredicted (overpredicted) in 48.5 percent (50.5 percent) of the quarterly predictions, while the indirect forecasts underpredicted (overpredicted) in 56.9 percent (43.1 percent) of the quarterly predictions. Together, these results and the results presented in Table 2 indicate that even though the indirect forecasts underpredicted more often than the direct forecasts, the average bias was not worse overall in an absolute sense. That is, the absolute value of the average bias, which represents the ME relative to zero, was essentially the same for both sets of forecasts (.002 for the direct forecasts and .002 for the indirects).

5. CONCLUDING REMARKS

Comparisons between direct and indirect ARIMA forecasts of corporate profit margins show that at least for this defined variable there is no apparent advantage to be gained by using the underlying components (net income and net sales). This study, which is based on a large sample of COMPUSTAT firms (N = 172), also suggests that regardless of the overall accuracy (MAE) of a given set of forecasts, there may be are other differences which should be examined (ME and directional agreement) before choosing one forecasting approach over the other. Future research should thus consider these differences as well as differences in overall accuracy.

REFERENCES

Ang, J.S. (1979), "Aggregate Versus Component Forecasts of Financial Statement Items," Review of Business and Economic Research, 15, 30-42.

Barnea, A. and J. Lakonishok (1980), "An Analysis of the Usefulness of Disaggregated Accounting Data for Forecasts of Corporate Performance," Decision Sciences, 11, 17-26.

Biggs, S.F. (1984), "Financial Analysts' Information Search in the Assessment of Corporate Earnings Power," Accounting, Organizations and Society, 9, 313-323.

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

Cogger, K.O. (1981), "A Time-Series Analytic Approach to Aggregation Issues in Accounting Data," Journal of Accounting Research, 19, 285-298.

Dossou-Gbete, S., P. Ettinger, and A. De Falguerolles (1980), "Identification and Preliminary Estimation of ARMA Type Processes: Some Comments," In Time Series, edited by O.D. Anderson, Amsterdam: North-Holland.

Engel, E.M.R.A. (1984), "A Unified Approach to the Study of Sums, Products, Time-Aggregation and Other Functions of ARMA Processes," Journal of Time Series Analysis, 5, 159-171.

Hopwood, W.S. (1980), "On the Automation of the Box-Jenkins Modeling Procedures: An Algorithm with an Empirical Test," Journal of Accounting Research, 18, 289-296.

Hopwood, W.S., J.C., McKeown, and P. Newbold (1982), "The Additional Information Content of Quarterly Earnings Reports: Intertemporal Disaggregation," Journal of Accounting Research, 20, 343-349.

Hopwood, W.S., P. Newbold, and P.A. Silhan (1982), "The Potential for Gains in Predictive Ability Through Disaggregation: Segmented Annual Earnings," Journal of Accounting Research, 20, 724-732.

Kang, H. (1986), "Univariate ARIMA Forecasts of Defined Variables," Journal of Business & Economic Statistics, 4, 81-86.

Lutkepohl, H. (1984), "Forecasting Contemporaneously Aggregated Vector ARMA Series, " Journal of Business and Economic Statistics, 2, 201-214.

McKeown, J.C. and K.S. Lorek (1978), "A Comparative Analysis of the Predictive Ability of Adaptive Forecasting, Re-estimation, and Re-identification Using Box-Jenkins Time Series Analysis on Quarterly Earnings Data," Decision Sciences, 9, 658-672.

Silhan, P.A. (1982), "Simulated Mergers of Existent Autonomous Firms: A New Approach to Segmentation Research," Journal of Accounting Reseaech, 20, 255-262.

Silhan, P.A. (1984), "Company Size and the Issue of Quarterly Segment Reporting," Journal of Accounting and Public Policy, 3, 185-197.

Wecker, W.E. (1978), "A Note on the Time Series Which is a Product of Two Stationary Time Series," Stoc. Proc. Appl.,8, 153-157.

Wei, W.W.S. and B. Abraham (1981), "Forecasting Contemporal Time Series Aggregates," Communications in Statistics, Part A--Theory and Methods 10, 1335-1344.

Table 1

Comparative Accuracy of Direct Versus Indirect Forecasts of Corporate Profit Margins

	Mean Absolute Error				
	Direct		Indirect		Mean
Horizon	Mean	(SD)	Mean	(SD)	Difference
t+1	.013	(.022)	.014	(.026)	-.001
t+2	.014	(.023)	.015	(.024)	-.001
t+3	.016	(.024)	.017	(.024)	-.001
t+4	.021	(.034)	.023	(.044)	-.002
t+5	.021	(.028)	.021	(.029)	.000
t+6	.019	(.025)	.020	(.029)	-.001
t+7	.021	(.024)	.023	(.031)	-.002
t+8	.032	(.047)	.035	(.079)	-.003
Average	.020	(.028)	.021	(.036)	-.001

Table 2

Comparative Bias of Direct Versus Indirect Forecasts of Corporate Profit Margins

	Mean Error			
	Direct		Indirect	
Horizon	Mean	(SD)	Mean	(SD)
t+1	-.001	(.020)	-.003	(.012)
t+2	-.002	(.016)	-.005	(.016)
t+3	.000	(.016)	-.003	(.013)
t+4	.003	(.017)	-.002	(.022)
t+5	.002	(.024)	-.003	(.028)
t+6	.001	(.022)	-.004	(.025)
t+7	.003	(.022)	-.001	(.023)
t+8	.012	(.035)	.002	(.068)
Average	.002	(.022)	-.002	(.027)

Table 3

Directional Agreement Between Direct and Indirect Forecasts of Corporate Profit Margins

	Direct		Indirect		
	P<A	P>A	P<A	P>A	Agreement
t+1	313	375	362	326	80.1%
t+2	377	311	424	264	80.7%
t+3	349	339	412	274	81.8%
t+4	356	332	412	276	80.8%
t+5	297	391	361	327	77.9%
t+6	335	353	381	307	81.4%
t+7	316	372	400	288	77.6%
t+8	329	359	385	303	80.5%
Average	334	354	392	296	80.1%
% Total	48%	52%	57%	43%	