ΣΠΟΥΔΑΙ / SPOUDAI

ΕΤΟΣ 1993 ΑΠΡΙΛΙΟΣ - ΙΟΥΝΙΟΣ ΤΟΜΟΣ 43 ΤΕΥΧ. 2

YEAR 1993 APRIL - JUNE VOL. 43 No 2

THE USE OF ERROR COMPONENTS MODELS IN BUSINESS FINANCE. A REVIEW ARTICLE AND AN APPLICATION

By

G. Karathanassis
Athens University of Economics and Business

N. Philippas
University of Piraeus

Abstract

This study applies and tests several stock valuation models of companies whose shares are traded in the Athens Stock Exchange. The relevant equations are estimated for the five major sectors of the Athens Stock Exchange (Banks, Textiles, Foods, Buildings, Commercials) using a specification which combines cross sectional and time series data. This is the Error Components Model.

In view of the results obtained the most important variables across sectors appear to be dividends followed by retained earnings. The contribution of the remainder of the independent variables has been mixed.(JEL G10)

1. Introduction

In the past few years business theorists and practitioners have at their disposal such data that allow them to apply efficient advanced techniques for the estimation of relevant parameters. One such approach is the combination of cross section and time series data. The significance of this approach and its potential use by various specialists in the broad area of Business Finance is what prompted us to write this article. Our objective is to provide a critical appreciation of the approach so that potential users will familiarize themselves with its advantages and limitations.

The article is divided into four sections. In the second section we deal the fundamentals of pooling cross-section and time series data. In the third, in an effort to apply the theory, we test a share valuation model with data from the Athens Stock Exchange. The fourth and final section is concerned with conclusions.

2. Pooling of Cross-Section and Time Series Data

The estimation of functions that combine time series and cross-section data is an occurence common enough in empirical business research. Usually we observe a number of companies, households, individuals etc. over a number of years. The combination of time series and cross section data offers researchers a significant number of degrees of freedom which allows them to overcome the constraints of the assumptions of the classical least squares regression model. Perhaps the most serious underlying assumption is that both the slope coefficient and the intercept are fixed and identical from observation to observation. This assumption is violated since individuals are likely to differ in their response to some economic or other stimuli[1].

The common approach is the introduction of unobservable cross-section and time effects. The introduction of all these variables allows one to capture all those important individual or time effects which affect the dependent variable but which cannot be measured explicitly.

Furthermore the introduction of these effects helps to reduce the degree of autocorrelation, heteroscedasticity and multicollinearity[2].

Algebrically the relationship may be written as follows:

$$Y_{it} = \alpha + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \qquad i = 1, 2, \ldots, N \qquad (1)$$

$$t = 1, 2, \ldots, T \qquad (2)$$

ςηερε $\varepsilon_{it} = \mu_i + \lambda_t + w_{it}$

- Y_{it} is the dependent variable
- X_{kit} is the kth nonstochastic explanatory variable
- μ_i is the unobservable cross sectional effect, which is invariant over time but differs among cross-sections,
- λ_t is the unobservable time effect, which is constant among cross-sections but differs over time,
- w_{it} is the unobservable remainder effect, which differs both across time and among cross sections.

The parameters of the above specification can be estimated making a number of different assumptions regarding the nature of the stochastic term ε_{it}. These assumptions, used widely in applied research are:

Assumption 1. The terms μ_i and λ_t are unknown constants while the term w_{it} is a random variable.

Assumption 2. All the terms mentioned above are random variables.

The first assumption leads to the dummy variable model[3] while the second leads to the error components model[4].

With the dummy variable model the resulting estimates will be unbiased and consistent, but will not be the most efficient in comparison with other estimating techniques. Another disadvantage of this is the use of a significant number of degrees of freedom. Further, application of this approach eliminates a large amount of the variation among both the explained and the explanatory variables when the variation between cross-sections and between time periods is large[5].

Furthermore it is extremely difficult to attach a sound economic meaning to the dummy variables. Finally this approach is especially sensitive to possible errors in variables[6].

The above problems may be overcome using a specification that treats the μ_i and λ_t as random variables.

With this approach the relationship may be written as:

$$Y_{it} = \alpha + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \qquad i= 1, 2, ..., N \qquad (3)$$

$$t= 1, 2, ..., T$$

where

$$\varepsilon_{it} = \mu_i + \lambda_t + w_{it} \qquad (4)$$

In equation 5 the total random effect ε_{it} consists of three random effects, the first accounting for firm effects, the second for time effects and the third is an overall cross-section and time series effect.

The estimation of the coefficients involves the use of a modified Aitken procedure consisting of two stages[7]. In the first stage the estimates of the variable of the error components may be obtained using least squares with dummy variables, while in the second one could use the generalised least squares estimator:

$$\hat{B} = (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} Y \qquad (5)$$

2.1. Specification Tests

Are the variance components significant?

To test whether the firm and time effects are statistically significant Breusch and Pagan[8] have developed a test the properties of which have been tested by others[9]. Under the Breusch and Pagan test the null hypothesis is expressed as follows: $H_0 = \sigma^2_\mu = \sigma^2_\mu = 0$. This means that there are no firm and time effects and the alternative hypothesis is $\sigma^2_\mu, \sigma^2_\lambda \neq 0$. Under the null hypothesis the variance-covariance matrix of the stochastic error term is $\Omega = \sigma^2_w I_{NT}$ and the simple least squares method gives a maximum likelihood estimate for the variance $\sigma^2_w = \frac{\hat{u}\,\hat{u}'}{NT}$ where $\hat{u}$ stands for the least squares residuals.

Breusch and Pagan propose a Langrange-multiplier statistic the properties of which have been analysed by Moran (1971) and Chant (1974) who have shown that under the null hypothesis it is distributed as X^2. If the value of the LM statistic is lower than its theoretical value $X^2_{(2)}$ for a given level of statistical significance we then accept the null hypothesis that $\sigma^2_\mu = \sigma^2_\mu = 0$. Conversely if the value of LM is greater than its theoretical value $X^2_{(2)}$ we reject the null hypothesis that $\sigma^2_\mu = \sigma^2_\mu \neq 0$ accepting thus the error components model.

2.2. Monte Carlo Experiments

Baltagi[10] via Monte Carlo experiments examined among other things the finite sample properties of the GLS estimators, the performance of the Breusch and Pagan tests and the frequency of negative variance estimates. His conclusions were:

a) The Breusch and Pagan test performed satisfactorily save those cases where one of the variance components was zero or very small but significantly different from zero.

b) Provided the variances of the error components do not tend to zero use of the estimated generalised least squares estimator is preferred to the simple least square and/or covariance estimator.

c) If the number of cross-sections N is larger than the time periods T it is better to view the time series and cross section effects as being random and adopt the error components specification.

d) If N and T are small, around 10, it is difficult to decide which estimator is more efficient (GLS or covariance).

e) When the estimated variances of the error components are negative, it is proposed that the negative variance be replaced with zero[11].

3. An Application Using Data from the Athens Stock Exchange

3.1. The Economic Model

According to valuation theory[12] the two main sources of value for shares (P) are dividends (D) and retained earnings (RE) per share. Retained earnings may represent new future growth in earnings per share. Once the assumption of certainty is removed we must take account of the influence of risk on share values. We should consider both business (CV) and financial risk (OE). The business risk variable reflects our inability to obtain perfect estimates of profits. Financial risk on the other hand represents the additional variability injected into earnings per share as a result of using debt capital. A number of previous researchers[13] have used size (S) as an independent variable in a various cost of capital models. Size has been used mainly as a proxy for risk. The discussion so far suggests the following relationship:

$$P = F\ (D,\ RE,\ CV,\ OE,\ S). \qquad (6)$$

Our approach specifies a number of variables as being important determinants of share prices. What we have is a testable hypothesis and only by resorting to testing will we be able to say anything concrete about importance of the dependent variables affecting share prices.

3.2. Sources of Data and Definition of Variables

In this study we report results for five (5) sectors. Banks (11), Textiles (11), Food and Spirits (8), Building and Materials (7) and Commercials (8).

All data were extracted from the Athens Stock Exchange Year Book. The companies for which we could obtain data were observed from 1984 to 1989 inclusive. The companies were thus observed for six years continuously. For the construction of a number of variables, namely growth and risk we needed continuous data for the period from 1980 to 1989. All variables are expressed in real terms using the implicit GNP deflator.

Regarding the functional form of the regression equations since there is not a priori reasoning for the choice of a specific functional form we experimented with both linear and logarithmic formulations[14]. Our results suggested that as a first approximation the relationships were linear in absolute terms. The relationship stated above was tested in a stepwise fashion. The statistical findings are shown in Appendix 2.

Our objective is to determine the extent to which changes in share prices across time and companies could be explained by those variables advanced by theory as being important determinants of value.

Specifically we tested the following relationship:

$$P = F\ (\overset{+}{AD},\ \overset{+}{RE},\ \overset{-}{CV},\ \overset{?}{OE},\ \overset{+}{SIZE}) \qquad (7)$$

Where P is price per share. It is the arithmetic average of the monthly average prices.

AD is average dividends per share. Ex-ante we do not really know whether investors use current dividends or some normalised value of dividends in order to have an estimate of expected dividends. We have used both current dividends per share and a five year average of dividends. An average of dividends in conjuction with pooling our time series and cross-section data may purge the transitory effects of this variable which may still exist in any single cross-section. According to valuation theory we expect a positive relationship between price and dividends.

RE is the variable representing retained earnings per share. We expect a positive relationship between share price and retained earnings.

CV is the variable representing business risk. Specifically $CV = S_u/\overline{E}$, where Su is the standard error of the regression of earnings per share before the deduction of interest and taxes against time. To be more precise: $E_t = a + b(t) + u_t$.

$$t = t\text{-}4, \ldots, t$$

$$\overline{E} = \sum_{t-4}^{t} E_t / 5$$

We expect a negative relationship between share price and business risk.

OE is the additional variability injected into the earnings per share as a result of using fixed interest debt capital. It is represented by the ratio of own to debt capital. It is extremely difficult a priori to decide the sign of the relationship between share prices and financial risk. Theoretically the issue has not been resolved and it still remains an empirical problem.

Size is the ratio of total assets to the total number of shares. The theoretical justification for inclusion of this variable is that a big company may be better diversified than a smaller one and be, thus less risky. We expect a positive relationship between share price and size.

3.3. Empirical Results

According to the theoretical relationships advanced by valuation theory, we expected both dividends and retained earnings to be positively related to share prices. Our empirical findings are in accordance with the theoretical relationships of section 3. Thus our theoretical predictions are empirically validated since both components of return exert a positive and significant influence on share prices. Given the sign, magnitude and statistical significance of the dividends coefficient we can say that dividends exert a greater influence on share prices than retained earnings. The importance of dividends vis-a-vis retained earnings may indicate investors preferences for certain dividends over uncertain future earnings expected from the reinvestment of retained earnings. Alternatively, this finding may reveal that dividends act as a vehicle in conveying information about future earnings. As regards the importance of the other variables no precise conclusions can be reached as their contribution to the dependent variable has not really been worth discussing. This is probably due to the fact that Greek investors do not as yet have the required information correctly to appraise shares. They look upon dividends as being the only variable capable of conveying important information about the future prospects of companies. We should also state that by application of the Breusch and Pagan test we found that the time effects were insignificant while the firm effects were statistically significant. The results are shown in appendix 1.

4. Conclusions

In many areas of Business Finance the nature of data is such that only combination of time and cross-section data can be used to test various business relationships. We now possess a sufficient number of statistical tools and criteria to be able to use the appropriate specification for combined cross-section and time series data. Chow's (1960) test should used for data homogeneity. Hausman's criterion is important in deciding whether one should apply an error components or a covariance specification. Breusch and Pagan'n test is valuable for knowing if the cross-sectional and time effects are statistically significant.

All methods for estimating the variances of the error components, with the possible exception of that proposed by Wallace and Hussain (1969), yield identical estimates for the regression estimates. The phenomenon of negative variance may be due to:

a) bad model specification or

b) the fact the true variance tend to zero but is statistically different from zero.

In general, when N>T the generalised least squares estimator seems to be more appropriate than the covariance estimator. [Arora (1973), Taylor (1980), Baltagi (1981)]. Only when N and T are small, around 10, we could not decide which estimator is more efficient. [Swamy and Arora (1972), Taylor (1980), Baltagi (1981)].

Appendix 1

Test for Statistical Significance of Firm Effects
Using the Criterio of Breusch and Pagan

Sector	Value X^2	Critical Value $X^2_{0.05,1}$
BANKS	28.5	3.841
TEXTILES	39.5	3.841
FOOD-SPIRITS-MILLS	5.5	3.841
BUILDING MATERIALS	6.2	3.841
COMMERCIALS	43.1	3.841

Appendix 2

Table 1

BANKS

Estimation Method: Generelised Least Squares

SPECIFICATION VARIABLES	I	II	III
C	-792,0 (-2,9)*	-573,4 (-1,6)	-803,8 (-2,0)
AD	18,4 (8,1)	17,9 (8,1)	19,3 (8,2)
RE	2,9 (2,1)	3,5 (2,4)	3,2 (2,2)
OE	3.,634,5 (1,5)	2.842,2 (1,0)	3.713,0 (1,4)
Size	—	-0,005 (-0,4)	-0,003 (-0,2)
CV	199,6 (1,4)	—	205,8 (1,4)
$\bar{R}^2$	0,725	0,72	0,73

* Numbers in brackets are t values

Table 2

TEXTILES

Estimation Method: Generalised Least Squares

SPECIFICATION VARIABLES	I	II	III
C	36,3 (1,8)*	12,6 (0,9)	23,5 (1,5)
AD	8,1 (12,6)	7,0 (11,9)	6,8 (11,7)
RE	0,5 (1,6)	0,6 (2,1)	0,5 (1,9)
OE	-3,4 (-0,4)	-5,9 (-0,8)	-1,7 (-0,2)
Size	—	0,018 (5,8)	0,017 (5,3)
CV	-27,2 (-2,8)	—	-15,7 (-1,9)
$\bar{R}^2$	0,91	0,92	0,93

* Numbers in brackets are t values

Table 3

FOOD - SPIRITS - MILLS

Estimation Method: Generalised Least Squares

SPECIFICATION VARIABLES	I	II	III
C	130,6 (4,5)*	95,6 (4,0)	131,3 (4,5)
AD	4,0 (3,0)	6,3 (4,3)	4,6 (2,8)
RE	0,3 (0,5)	0,6 (1,1)	0,4 (0,7)
OE	-92,1 (-3,5)	-87,5 (-3,2)	-95,0 (-0,6)
Size	—	-0.018 (-1.1)	-0.009 (-0.6)
CV	-38,0 (-2,2)	—	-35,2 (-2,0)
$\bar{R}^2$	0,69	0,66	0,69

* Numbers in brackets are t values

Table 4

BUILDING MATERIALS

Estimation Method: Generalised Least Squares

SPECIFICATION / VARIABLES	I	II	III
C	13,1 (0,3)*	27,2 (0,4)	29,0 (0,4)
AD	5,0 (11,4)	6,1 (12,1)	6,1 (11,4)
RP	6,8 (5,0)	8,7 (5,5)	8,7 (5,8)
OE	2,7 (0,3)	9,8 (0,9)	10,0 (0,9)
Size	—	-0,09 (-1,4)	-0,09 (-1,4)
CV	-15,3 (-0,7)	—	-2,7 (-0,1)
$\bar{R}^2$	0,89	0,90	0,90

* Numbers in brackets are t values

Table 5

COMMERCIAL COMPANIES

Estimation Method: Generalised Least Squares

SPECIFICATION / VARIABLES	I	II	III
C	37,4 (0,8)*	61,5 (1,2)	35,7 (0,7)
AD	4,7 (4,9)	4,1 (3,1)	4,5 (3,6)
RP	2,4 (2,5)	1,8 (1,8)	2,3 (2,3)
OE	-17,6 (-2,3)	-19,2 (-2,3)	-18,3 (-2,3)
Size	—	0,08 (0,6)	0,03 (0,2)
CV	69,3 (2,1)	—	68,1 (2,0)
$\bar{R}^2$	0,90	0,89	0,90

* Numbers in brackets are t values

Footnotes

1. This assumption has been criticised with particular severity. See Klein (1953), Theil (1957) and Zellner (1962b).

2. To the extent that they affect the dependent variable.

3. The well known covariance analysis, which in the past had been used extensively in the area of production functions. See Mundlak (1963), Hock (1962).

4. For a number of variations of the error components model see, Balestra and Nerlove (1966), Chamberlain and Griliches (1975), Lillard and Willis (1978), Lillard and Weiss (1979), Hausman and Taylor (1981).

5. Maddala (1971).

6. Mundlack (1978) and Hausman (1978)

7. Amemiya (1971).

8. Breusch and Pagan (1980).

9. Moran (1971) and Chant (1974).

10. Baltagi (1981).

11. Wallace and Hussain (1969).

12. Miller and Modigliani (1961).

13. Benishay (1961).

14. See for instance Zarembka's work (1968).

15. The performance of the rest of the variables was mixed. It is interesting to note that there are similarities in the performance of the main sources of value, that is dividends and retained earnings, between our study and other studies. See for instance Fisher (1961) and Friend and Puckett (1964), among many others.

References

Amemiya, T. (1971), "The Estimation of Variance in a Variance Components Model", International Economic Review, pp. 1-13.

Arora, S.S. (1973), "Error Components Regression Models and Their Applications", Annals of Economic and Social Measurement, 2, pp. 451-461.

Balestra, P. and *Nerlone, M.* (1966), "Pooling Cross-Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas", Econometrica, 34, pp. 585-612.

Baltagi, B.G. (1981), "Pooling: An Experimental Study of Alternative Testing and Estimation Prodecures in a Two-Way Error Component Model", Journal of Econometrics, 17, pp. 21-49.

Benishay, H. (1961), Variability in P/E ratios of Equities, American Economic Review.

Breusch, T.S. and *Pagan A.R.* (1980), "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics" Review of Economic Studies 47, pp. 239-253.

Chamberlain, G. and *Grilighes, Z.* (1975), "Unobservables with a Variance-Components Structure: Ability, Schooling and the Economic Success of Brothers", International Economic Review, 16.

Chant, D. (1974), "On Asymptotic Tests of Composite Hypothesis in Nonstandard Conditions", Biometrica, 61, pp. 291-298.

Chow, G.C. (1960), "Tests for Equality between Sets of Coefficients in Two Linear Regressions", Econometrica 28, pp. 591-605.

Fisher, G.R. (1961), "Some Factors Influencing Share Prices", Economic Journal, pp. 124-141.

Friend, I and *Puckett M.,* (1964), "Dividends and Stock Prices", American Economic Review, pp. 656-682.

Fuller, W.A. and *Battese, G.E.* (1973), "Transformations for Estimation of Linear Models with Nested Error Structure", Journal of the American Statistical Association, 68, pp. 626-632.

——, (1974), "Estimation of Linear Models with Crossed - Error Structure", Journal of Econometrics, 2, pp. 67-78.

Graybill, F.A. (1961), "An introduction to Linear Statistical Models", McGraw Hill Book Company, New York.

Griliches, Z. and *Hausman J.A.,* (1986), "Errors in Variables in Panel Data", Journal of Econometrics, 31, pp. 93-118.

Hausman, J.A. and *Taylor, W.E.* (1978), "Specification Tests in Econometrics", Econometrica, 46, pp. 1251-1272.

Hausman, J.A. and *Taylor, W.E.* (1981), "Panel Data and Unobservable individual Effects", Econometrica, 49, pp. 1377-1398.

Henderson, C.R. (1953), "Estimation of Variance and Covariance Components", Biometrics, 9, pp. 397-401.

——, (1971), Comment on "The Use of Error Components Models in Combining Cross Section with Time Series Data", Econometrica, 39, pp. 397-401.

Hock, I (1962), "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data", Econometrica, 30, pp. 34-53.

Judge et al. (1985), "The Theory and Practice of Econometrics", Wiley Series, New York.

Karathanassis, G. and *Philippas, N.* (1988), "Estimation of Bank Stock Price Parameters and the Variance Components Model" Applied Economics, 20, pp. 497-507.

Klein, L.R. (1962), "A Textbook of Econometrics", New York: Macmillan.

Kmenta, J. (1971), "Elements of Econometrics", New York: Macmillan.

Kuh, E. (1963), "Capital Stock Growth: A Micro-Econometric Approach", North Holland.

Lillard, L.A. and *Willis R.* (1978), "Dynamic Aspects of Earnings Mobility", Econometrica, 46, pp. 985-1012.

Lillard, L.A. and *Weiss, Y.* (1979), "Components of Variation in Panel Earnings Data: American Scientists 1960-70", Econometrica, 47, pp. 437-454.

Maddala, G.S. (1971), "The Use of Variance Components in Pooling Cross-Section and Time Series Data", Econometrica, 39, pp. 341-358.

Maddala, G.S. and *Mount, T.D.* (1973), "A Comparative Study of Alternative Estimators for Variance Components Models", Journal of the American Statistical Association, 68, pp. 324-328.

Miller, M. and *Modigliani F.* (1961), "Dividend Policy, Growth and the Valuation of Shares", Journal of Business.

Mundlak, Y. (1963), "Estimation of Production and Behavioral Functions from a Combination of Cross-Section and Time-Series Data", in C.F. Christ ed., Measurment in Economics, Stanford University Press.

——, (1978a), "On the Pooling of Time-Series and Cross-Section Data", Econometrica, 46, pp. 69-85.

Nerlove, M. (1971a), "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections", Econometrica, 39, pp. 359-382.

——, (1971b), "A Note on Error Components Models", Econometrica, 39, pp. 383-396.

Rao, C.R. (1970), "Estimation of Heteroscedastic Variance in Linear Models", Journal of the American Statistical Association, 65, pp. 161-172.

——, (1972), "Estimation of Variance and Covariance Components in Linear Models", Journal of the American Statistical Association, 67, pp. 112-115.

Swamy, P.A.V.B. and *Arora S.S.* (1972), "The Exact Finite Sample Properties of the Estimator of Coefficients in the Error Components Regression Models", Econometrica, 40, pp. 253-223.

Taylor, W.E. (1980), "Small Sample Considerations in Estimation From Panel Data", Journal of Econometrics, 14, pp. 203-223.

Theil, H. (1957), "Specification Errors and the Estimation of Economic Relationships", Rev. Int. Statist. Inst., 25, pp. 41-51.

Wallace, T.D. and *Hussain* (1969), "The Use of Error Components Models in Combining Cross Section with Time Series Data", Econometrica, 37, pp. 55-72.

Zarembka, P. (1968), "Functional Form in the Demand for Money", Journal of the American Association, pp. 502-511.

Zellner, (1962b), "Estimation of Cross Section Relations: Analysis of a Common Specification Error", Metro-Economica, 14, pp. 111-117.