Division of Research
Graduate School of Business Administration
The University of Michigan

August, 1977

ANALYSIS OF STOCK REPURCHASES
WITH A RANDOM COEFFICIENT REGRESSION MODEL

Working Paper No. 149

by

Roger L. Wright, Terry Dielman, Timothy J. Nantell

FOR DISCUSSION PURPOSES ONLY

ABSTRACT

Developments in data base management require extension of the scope and capabilities of interactive statistical packages and analysis techniques if the more complex and rich data structures are to be used to full advantage. Perhaps the most fundamental and pervasive type of data structure is comprised of multivariate time series of observations for each of a cross-section of individual entities. A theoretical foundation for statistical analysis of data like these has been provided by the random coefficient regression techniques developed by Swamy [7,8,9]. This paper will briefly describe Swamy's method and illustrate its use in analyzing a 168-month equity rate of return series of each of 40 firms that repurchased their stock once during the 1961 to 1974 period. (A somewhat related analysis has been provided by Rosenberg [4,6].)

The General Random Coefficient Regression Model

Suppose that $R(i,t)$ denotes the equity rate of return on stock i for month t ($i=1,\ldots,N$; $t=1,\ldots,T$), that is,

$$R(i,t)=[P(i,t)-P(i,t-1)+D(i,t)]/P(i,t-1) \qquad (1)$$

where $P(i,t)$ is the price of stock i at the end of month t and $D(i,t)$ is the dividend paid to stock i during month t. Under the Random Coefficient Regression (RCR) Model, we assume that $R(i,t)$ can be written as a linear function of K explanatory variables $X(1;i,t),\ldots,X(K;i,t)$:

$$R(i,t)=b(i,1)X(1;i,t)+\ldots+b(i,K)X(K;i,t)+e(i,t). \qquad (2)$$

Each regression coefficient is allowed to vary between stocks (as well as between variables) and is viewed as a random variable. In particular the vector of regression coefficients $[b(i,1),\ldots,b(i,K)]$ ($i=1,\ldots,N$) is assumed to have been independently drawn from the multivariate normal distribution with expected value vector $[b(1),\ldots,b(K)]$ and unknown variance-covariance matrix V. The residual $e(i,t)$ is assumed to be normally distributed with mean zero and variance $v(i)$, independent of all other residuals and all regression coefficients.

Several features of this model should be emphasized. The regression coefficients for each stock are assumed to be invariant over time, but varying from stock to stock. By assuming that the regression coefficients come from a common distribution, they can be analyzed coherently as a set using efficient statistical techniques. However, if desired, the RCR Model can accomodate an assumption that some of the regression coefficients are fixed from stock to stock.

The Market Model

The familiar market model can be easily formulated in terms of the RCR Model. Let R(m,t) denote the rate of return of the market for month t. We can assume

$$R(i,t)=b(i,0)+b(i,1)R(m,t)+e(i,t). \tag{3}$$

Here b(i,1) is the "beta coefficient" of stock i, representing the systematic relationship between R(i,t) and R(m,t). Under our RCR formulation, these beta coefficients are assumed to be fixed for each stock over the time period of observation, but normally distributed between stocks.

Estimation and Hypothesis Testing Under the RCR Model

The statistical estimation procedure developed by Swamy for the RCR Model is a generalization of the familiar weighted regression technique usually used for heteroscedastic regression models. The first step is to calculate separate regression equations for each stock. The population mean b(k) of the random regression coefficient b(i,k) is estimated as a weighted average of the separately estimated regression coefficients for each stock. The variance-covariance matrix V describing the distribution of the regression coefficients is estimated from the variances among the separately estimated coefficients between stocks, with allowance made for variation arising from the residuals. These estimates have desirable large-sample statistical properties, but little can be said yet about their behavior with small samples.

Under the RCR Model one can test a number of types of hypotheses of which two are most important:

(1) Ho: b(1)=0,...,b(k)=0 $(k \leq K)$, and

(2) Ho: b(i,1),...,b(i,k) $(k \leq K)$ are all fixed (i.e., nonrandom).

Like the estimation procedures, these tests require large samples.

Interactive statistical analysis is usually characterized by an interplay of several techniques:

(1) Specification and estimation of linear regression models involving various carefully chosen explanatory variables

(2) Selecting between alternative models using appropriately formulated hypothesis tests (along with other considerations)

(3) Graphical examination of regression models and their underlying data

The estimation techniques and hypothesis tests that are available for the RCR Model provide the essential tools for interactive analysis of a cross-section of time series.

Illustration of RCR Analysis Using the Market Model

176 firms repurchased their stock between 1961 and 1974. The data that we use here describe those 40 of these firms that repurchased their stock exactly once during this period. These repurchase offers were mostly large tender offers, but a few were smaller open market offers. The data provide the monthly equity rate of return R(i,t) for each of the 40 stocks over a 168-month period selected, so that month 108 consistently represents the announcement date of the repurchase offer. In addition, we have the market rate of return corresponding to the observations of each stock.

Model 1

The first model that we will analyze was formulated to evaluate the behavior of the rate of return around the announcement date of the repurchase offer. The model will incorporate any shift in the rate of return during the quarter preceeding the announcement date and also in the month of announcement. We will accomplish this by adding two dummy variables to equation (3), denoted Z(-1;i,t) and Z(0;i,t). Z(-1;i,t) is a dummy variable which is equal to one for each of the three months preceeding the announcement date, and equal to zero otherwise. Similarly, Z(0;i,t) is a dummy variable which is one for the actual month of announcement, and zero otherwise. (In this and subsequent models, all variables will be measured as deviations from their within-stock means so that the constant terms of the regression equations will be dropped.) With this notation Model 1 is:

$$R(i,t)=b(i,1)R(m,t)+b(i,2)Z(-1;i,t)+b(i,3)Z(0;i,t)+e(i,t). \quad (4)$$

The RCR estimates for Model 1 summarized in Table 1 show that the rate of return of these stocks during the announcement month averaged 15.2 percent above the level expected from the market rate of return. Furthermore, although there was considerable deviation (11.9 percent) from stock to stock, the mean of 15.2 percent was very significantly different from zero. (The degrees of freedom for the T-statistic is 39.) Further RCR analysis shows that we can reject the hypothesis that the coefficients b(i,3) are identical for all firms. This means that the estimate of the population standard deviation of these coefficients (.119) is significantly different from zero.

Table 1 also suggests that the rates of return are significantly below normal during the quarter preceeding announcement. However, an RCR test indicates that there is no significant randomness in the coefficients b(i,2) and, when the RCR analysis is repeated with b(i,2) assumed to be nonrandom, then the coefficients b(i,2) are not significantly different from zero. Apparently RCR analysis requires a very careful specification of the assumed model.

Table 1

RCR Estimates for Model 1

Coefficient	Mean	Standard Error	T Statistic	Population Standard Deviation
b(i,1)	1.29	.09	17.3	.41
b(i,2)	-.016	.005	-3.4	**
b(i,3)	.152	.026	5.8	.119

** Indicates a negative estimate.

Model 2

A further analysis of repurchase effects involving future quarters was conducted also using the dummy variable approach. These coefficients for two, three, and four quarters ahead resulted in test statistics indicating no randomness as well as a mean effect of zero, thus implying no effect on rate of return.

Model 3

To examine the effect of the stock repurchase on the systematic risk of the firms we considered a third model:

$$R(i,t)=b(i,1)R(m,t)+b(i,2)Z(0;i,t)+b(i,3)S(t)+e(i,t).$$

This model was estimated specifying all three coefficients as random. The results are showen in Table 2 below. The variable S(t) was created to measure a possible shift or jump in the risk (i.e, in the beta coefficient) of the firms. For each month before the repurchase announcement the variable

S(t) is set equal to 0; for months after the announcement S(t) is set equal to R(m,t). When the regressions are run this results in the coefficient of S(t) capturing any sudden change in the firms' beta in months following the repurchase.

Table 2

RCR Estimates for Model 3

Coefficient	Mean	Standard Error	T Statistic	Population Standard Deviation
b(i,1)	1.29	.067	19.4	.141
b(i,2)	.13	.024	5.5	.014
b(i,3)	-.159	.043	-3.7	.004

The Table 2 results show that b(i,3) is definitely significant and indicates a downward shift in the risk of the firms involved. The betas drop an average of .159 while the variation of this drop between firms is very small, .004.

Model 4

A fourth model was constructed to determine whether the change in systematic risk of the firms appeared only as a sudden shift of the beta coefficients or whether there was a trend discernible. This question might prove particularly interesting because of work done involving the modeling of the behavior of the betas over time. Rosenberg [4] and Rosenberg and McKibben [6] have published recent articles reporting on research in this area. Our approach also investigates this type of behavior, although we include no implicit random time component.

The model developed for this purpose is:

$$R(i,t)=b(i,1)R(m,t)+b(i,2)Z(0;i,t)+b(i,3)T(t)+b(i,4)S(t)+e(i,t)$$

where T(t) has been constructed as follows to capture any trend in the betas:

$$T(t)=rR(m,t) \quad t=1,2,.,.,.,168.$$

T(t) is thus the market rate of return in each time period times the number of the period. If there is a trend present in the beta coefficient we will have taken account of this by the introduction of the variable T(t) and its resulting coefficient, b(i,3). To determine the presence of trend we merely test the significance of b(i,3).

This model was first examined with all coefficients random. The results in Table 3 indicate that the coefficients b(i,3) and b(i,4) are very possibly fixed rather than random. When tested for randomness the statistic for b(i,3) is not significant. This result plus the negative estimate of the population standard deviation for b(i,4) led us to rerun the model with these two coefficients constrained to be fixed. The results are shown in Table 4. The T-statistic shown in Table 4 indicates that b(i,3) is equal to zero. We therefore conclude that the change in systematic risk for those 40 firms occurs as a shift after the repurchase announcement rather than as a trend.

Table 3

RCR Estimates for Model 4

Coefficient	Mean	Standard Error	T Statistic	Population Standard Deviation
b(i,1)	1.611	.153	10.5	.604
b(i,2)	.145	.025	5.9	.014
b(i,3)	-.003	.001	-2.3	.000
b(i,4)	-.094	.040	-2.3	**

** Indicates a negative estimate.

Table 4

Revised RCR Estimates for Model 4

Coefficient	Mean	Standard Error	T Statistic	Population Standard Deviation
b(i,1)	1.298	.139	9.3	.604
b(i,2)	.130	.024	5.5	.014
b(i,3)	.001	.001	.6	...
b(i,4)	-.233	.069	-3.4	...

Testing Assumptions of the RCR Approach

In the adoption of a model such as Swamy's RCR Model certain assumptions are made. To determine the validity of the assumptions made in the case of our stock-return data several tests have been performed.

For the RCR estimation results presented we assumed no serial correlation of errors within the time series regressions. This assumption was tested using the Durbin-Watson statistic for firms with 100 or fewer observations and the von Neumann statistic for those with more than 100. At the .05 level of significance the tests indicated serial correlation in 6 out of the 40 regressions for each of the models previously described. To examine the effect of the serial correlation on our estimation and inference procedures we estimated the correlation coefficient, transformed the data for those firms with correlated residuals, and then reran the regressions using the transformed data. This method is equivalent to using generalized least squares to obtain the coefficient estimates. The values from these transformed regressions were used in the pooled estimates and in the inference. Little appreciable change was noted from the results previously reported.

In order to use the procedures we have adopted for inference, assumptions were made concerning the normality of both the residuals in the time series regressions and the coefficients considered random in each model. In particular we assumed each residual, e(i,t), to be normally distributed with mean zero and variance v(i); the coefficient vector [b(i,1),...,b(i,K)] was assumed to have been drawn from a multivariate normal distribution with mean vector [b(1),...,b(K)] and variance-covariance matrix V.

We used several approaches to examine the normality of both errors and coefficients. These approaches included:

1. Examination of skewness and kurtosis;

2. Examination of plots of the cumulative distribution functions (CDF) of the errors and coefficients

3. Use of the Lilliefors test for normality. This test makes use of the sample CDF for standardized variables by comparing it to the standard normal CDF to determine whether or not an assumption of normality is justified.

The Lilliefors statistic for the time series residuals indicated normality in about 70 percent of the distributions (i.e., for 28 of the 40 firms). For the random coefficients, results vary widely depending on the model. The following examples illustrate this.

Example 1.

Market Model: R(i,t)=b(i,0)+b(i,1)R(m,t)+e(i,t).

For a given value of N the 95th percentile of the Lilliefors statistic is given by $.886/\sqrt{N}$. Values falling above $.886/\sqrt{N}$ indicate a good possibility that the distribution from which our values were drawn is non-normal. Since we are dealing with 40 firms, the 95th percentile value for our tests of the coefficients is $.886/\sqrt{40} = .14009$. Results for the simple market model are given in Table 5.

Table 5

Results of Lilliefors Test Using Simple Market Model

Coefficient	Lilliefors Statistic
b(i,0)	.13786
b(i,1)	.06973

In the simple model we cannot reject the hypothesis of normality of the coefficients.

Example 2

Model 3: $R(i,t)+b(i,1)R(m,t)+b(i,2)Z(0;i,t)+b(i,3)S(t)+e(i,t)$.

Results of the Lilliefors test are shown in Table 6. Again we cannot reject the hypothesis that the sample coefficients come from normal distributions.

Table 6

Results of Lilliefors Test Using Model 3

Coefficient	Lilliefors Statistic
b(i,1)	.09142
b(i,2)	.11670
b(i,3)	.12350

Example 3

Model 4: $R(i,t)=b(i,1)R(m,t)+b(i,2)Z(0;i,t)+b(i,3)S(t)+b(i,4)T(t)+e(i,t)$.

Results are given in Table 7. In this instance the test shows evidence that the coefficients b(i,1), b(i,2) and b(i,4) may not be normally distributed.

Table 7

Results of Lilliefors Test Using Model 4

Coefficient	Lilliefors Statistic
b(i,1)	.16015
b(i,2)	.17190
b(i,3)	.06221
b(i,4)	.16483

In each of the three examples presented the values for skewness and kurtosis and the CDF plots support the results of the Lilliefors test.

Conclusion

The random coefficient regression model as presented by Swamy provides a useful framework for viewing financial data such as the stock repurchase data in this paper. Pooling the cross-sectional and time series observations allows an overall view of what happens to the risk and return of the firms involved rather than providing a view of the firms individually. Normality assumptions for inference in the model are not satisfied in all cases, however. This fact leads us to several questions which might be posed for future research.

1. What effect does the presence of outliers have on the violation of normality assumptions? Consider the model R(i,t)=b(i,1)R(m,t)+b(i,2) Z(0;i,t)+b(i,3)S(t)+b(i,4)T(t)+e(i,t). As stated previously the coefficients b(i,1), b(i,2) and b(i,4) do not appear to be normally distributed. The histogram of each coefficient follows.

Histogram for b(i,1):

Midpoint	Count	Each X=1
-1.1158	2	XX
.4034	16	XXXXXXXXXXXXXXXX
1.9225	19	XXXXXXXXXXXXXXXXXXX
3.4417	2	XX
4.9609	0	
6.4801	0	
7.9993	1	X (Firm No. 8 of 40)

Histogram for b(i,2):

Midpoint	Count	Each X=1
-.13867	2	XX
-.01016	9	XXXXXXXXX
.11835	19	XXXXXXXXXXXXXXXXXXX
.24686	5	XXXXX
.37537	3	XXX
.50388	1	X (Firm No. 8 of 40)
.63239	1	X (Firm No. 16 of 40)

Histogram for b(i,4):

Midpoint	Count	Each X=1
-.06402	1	X
-.04765	0	
-.03128	0	
-.01490	10	XXXXXXXXXX
.00147	22	XXXXXXXXXXXXXXXXXXXXXX
.01784	6	XXXXXX
.03422	1	X

We might ask whether removal of the firm(s) with extreme coefficient values would result in distributions which were more nearly normal. (Transformations such as the Box-Cox transformations might also be useful in producing distributions appropriate for inference.)

As shown in the histograms, Firm Number 8 appears to be an "outlier" in terms of the relationship of its regression coefficients to those of the other firms. It should, perhaps, be removed before using the RCR model.

A further study of the characteristics of firms such as Number 8 could prove interesting and useful and might lead us to a better understanding of what factors affect systematic risk and the return. Such study could contribute especially to seeing what would cause a firm's risk to be extremely high compared to other firms'.

2. A second question which future research might explore is: What effects do violations of the assumptions have on our inference procedures? How robust are the RCR model estimates?

3. A final assumption we might question is that of no correlation of regression residuals across firms or no contemporaneous correlation. If correlation across firms were a problem it could be corrected in a fashion similar to that of serial correlation.

References

[1] Fama, E.F.; Fisher, L.; Jensen, M; and Roll, R. "The Adjustment of Stock Prices to New Information." International Economic Review, Feb. 1969, pp. 1-21.

[2] Fiege, Edgar, and Swamy, P.A.V.B. "A Random Coefficient Model of the Demand for Liquid Assets." Journal of Money, Credit and Banking, May 1974. pp. 241-252.

[3] Johnston, J. Econometric Methods. New York: McGraw-Hill Book Co., 1972, pp. 249-54.

[4] Rosenberg, B. "The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression." Annals of Economic and Social Measurement 24, (1973): 399-428.

[5] ___________. "A Survey of Stochastic Parameter Regression." Annals of Economic and Social Measurement 24 (1973): 381-97.

[6] Rosenberg, Barr, and McKibben, Walt. "The Prediction of Systematic Risk in Common Stocks." Journal of Financial and Quantitative Analysis, Mar. 1973, pp. 317-33.

[7] Swamy, P.A.V.B. "Efficient Inference in a Random Coefficient Regression Model." Econometrica, Mar. 1970, pp. 311-23.

[8] ___________. "Linear Models with Random Coefficients." Frontiers in Econometrics, ed. by Paul Zarembka. New York: Academic Press, Inc., 1974, Chap. 5, pp. 143-68.

[9] ___________. Statistical Inference in Random Coefficient Regression Models. Berlin: Springer-Verlag, 1971.

[10] Wright, R., Nantell, T., and Dielman, T. "Interactive Statistical Analysis of a Cross-Section of Multivariate Time Series with an Application in Financial Analysis." Midwest AIDS Proceedings, Cleveland, Ohio 1977.

[11] Elementary Statistics Using MIDAS. Dec. 1976 Statistical Research Laboratory, The University of Michigan. pp. 253-277.