Volume Title: Yields on Corporate Debt Directly Placed

Volume Author/Editor: Avery B. Cohan

Volume Publisher: NBER

Volume ISBN: 0-87014-472-3

Volume URL: http://www.nber.org/books/coha67-1

Publication Date: 1967

Chapter Title: Yields on Industrials

Chapter Author: Avery B. Cohan

Chapter URL: http://www.nber.org/chapters/c1512

Chapter pages in book: (p. 38 - 82)

3

YIELDS ON INDUSTRIALS

This chapter, the end product of which consists of two sets of series on yields on industrial direct placements, has three parts:

The first part describes the preliminary tests made of the variables checked in column 1 of Table 13. This part responds to the following question: How much of the variation in yields on direct placements, time held constant, do those variables together explain?

The second part describes the procedures used to identify the *significant* variables—those variables, that is, which can be said to be the determinants of yields on industrials, time held constant, and to rank them in order of importance. This part responds to the question: Which of the variables are statistically significant and, of these, which are really important and which are of slight importance, time held constant?

The last part of the chapter presents the various series on yields on industrials and describes the procedures used to construct them.

Summary of Procedure

Much of this chapter is technical and will probably not be of interest to the general reader. But for him, if he feels venturesome, and for others perhaps as well, a brief guide to the procedures used and their meaning may be helpful.

In general, two types of series have been constructed [1]—"cross-classified" series and "computed" series. The cross-classified series hold constant through time just two characteristics (variables), albeit two of the most important—size of issuer (as measured by

[1] In addition to simple quarterly averages, of the actual yields in the sample.

total pro-forma capital) and times pro-forma interest earned (coverage). These series represent the changing cost of various classes of direct placements, each of which is roughly homogeneous, with respect to these two variables, through time.

The computed series, on the other hand, hold *all* the relevant variables of an issue constant over time, and therefore represent the changing cost of various classes of direct placements, each of which is, for practical purposes, perfectly homogeneous through time. These series, that is, hold not only size of issuer and coverage constant but also eight other variables which, the analysis suggests, are capable of having some effect on yield. Each computed series is, thus, analogous to a cost of living index based on a rigidly fixed basket of commodities. Each measures the changing cost through time of a direct placement of specified characteristics.

In order to construct the computed series, all the variables capable of influencing yield had to be identified; in order to construct the cross-classified series, the most important variables had to be selected from among all the relevant variables.

The first step toward the accomplishment of these two objectives was to subject the eighteen substantive variables checked in column 1 of Table 13 to a five-step statistical procedure. Each variable was tested to determine whether it contributed anything to our understanding of yield, all other variables held constant. For this purpose, multiple regression techniques were used. Multiple regression is the statistical counterpart of the economists' "other things being equal." It enabled us to observe, first, whether some given variable was capable of affecting yield, all other variables held constant, and, if so, to measure that effect on a percentage scale.

Table 14, column 1, lists the substantive variables which survived the foregoing sifting process. Column 2 gives, for each such variable, the number which measures the size of the effect of the given variable on yield. Each of these numbers (regression coefficients) can be interpreted as an estimate of the percentage effect on yield when the variable with which it is associated is varied by 1 per cent. The sign of the regression coefficient indicates the direction

TABLE 14

Industrials: Significant Variables, Their Regression Coefficients and Percentage Impact on Yield

	Variable	Regression Coefficient	Percentage Impact on Yield[a]
X_2	Total pro-forma capital	-.0475	9
X_3	Average term (years)	-.0228	1
X_{4r}	Pro-forma interest	+.0683	12
X_5	Type of security	-.0232	2
X_6	Industrial class	-.0178	1
X_7	Years nonrefundable	+.0010	1
X_8	Size of issue	-.0195	3
X_{12}	Earnings before interest and taxes	-.0322	6
X_{13}	Maturity (years)	-.0212	1
X_{15}	Dollars of long-term debt per dollar of total capital	-.0349	2

[a]See text for explanation.

of effect. Thus, on the average over the sample, when EBIT (X_{12}), was 1 per cent higher, yield was lower by 322/10,000 of 1 per cent. When maturity (X_{13}) was longer by 1 per cent, yield was lower by 212/10,000 of 1 per cent—and so forth for each of the other regression coefficients.

The regression coefficients given in Table 14 are averages over the whole eleven-year period, 1951–61. In addition to these eleven-year averages, separate regression coefficients were obtained for each of the forty-four calendar quarters during the period. These quarterly coefficients are given in Appendix D.

In the next step, the total impact of each variable was assessed. Total impact is a combined measure which takes account of two things—the variability of each variable and the importance of that variability as measured by the appropriate regression coefficient. The regression coefficients alone tell us how much yield will vary with each 1 per cent change in each variable. But some variables vary little and others much. Thus, for example, size and earnings vary from a few hundred thousand dollars to several hundred million dollars, or by a multiple of perhaps a thousand. Maturities, on the other hand, range from eight or nine to twenty-five years, or by a multiple of about three.

Column 3 of Table 14 provides a combined estimate of total impact for each variable and indicates that, taking both variability and the importance of that variability into account, X_2, X_{4r}, and X_{12} have the most impact on yield.

With these findings in hand, the cross-classified series were constructed. The variables X_{4r} and X_{12} were combined into a single variable—coverage—and a trade-off factor was found between it and X_2. Two issues with different coverage ratios and sold by companies of different size were considered to represent approximately equivalent "quality" if the two variables, adjusted for the trade-off between them, for the first issue, were equal to the same two variables, similarly adjusted, for the second issue. This, of course, is just a roundabout way of saying that an issue sold by a large company with a low coverage ratio may be equivalent in quality to an issue sold by a small company with a high coverage ratio. The technical problem, which is discussed in detail below, lay in estimating the trade-off factor, in the above sense, between size of issuer and coverage. The factor actually used was tested on public offerings and found to conform fairly well to that implicit in agency ratings (see Appendix C). After the trade-off factor was ascertained, three classes were established and the actual observations deposited in them. Average yields within each class were computed quarterly. (The resulting series are given in Table 28 and in Chart 7.)

A *composite* cross-classified series, quarterly, was then obtained by averaging the three series across each quarter. This series is given in Table 28 and Chart 7, where it is compared with yields on FHA mortgages and long-term governments. It is compared with yields on Moody's new issues (Baa and A) in Chart 9.

Computed series were then obtained as follows: first, average values over the whole period for each variable for *each* of the above three classes were computed.[2] In other words, an average value for X_2 (total pro-forma capital) was obtained for class I by averaging over X_2 for all issues falling in class I over the entire eleven-year period. The same was done for each of the other nine variables. The same procedure was then followed for classes II and III (see Table 29). These values were then inserted into each of the forty-four quarterly regression equations described above, and yields calculated quarterly for each class. The three computed series will be found in Table 30 and in Chart 11, where they are compared with their cross-classified counterparts.

Finally, a computed composite series was obtained by using, for each variable, its average value over all observations for the year 1956. For example, the value used for X_2 was an average over X_2 for all the direct placements sold in the calendar year 1956, and so forth for each other variable. These values are given in Table 31.[3] The computed composite series itself is given in Table 32 and Chart 12 where it is compared with its cross-classified counterpart.

The rest of this chapter describes procedure in detail.

The Preliminary Tests

As indicated in Chapter 2, discussion with life insurance company financial managers had suggested that most of the variation in yield, time held constant, could be explained by some combination of the eighteen substantive variables checked in column 1 of Table 13.

In order to form a judgment regarding the shape of the function relating the foregoing variables to yield, scatter diagrams were

[2] That is, the three classes used to construct the cross-classified series.

[3] They are virtually identical to the average values for each variable over the entire period.

drawn for six quarters fairly well distributed through the period.[4] These quarters were chosen because, within each, the level of yields on outstandings had been reasonably stable and, therefore, each such quarter could be regarded as a reasonably close approximation to a strict cross section of time.[5]

The scatter diagrams showed yield on each variable for each of these six quarters. Six diagrams were thus obtained for yield on total pro-forma capitalization, yield on average term, yield on times pro-forma interest earned, and so forth. Examination of these diagrams strongly suggested that yield varied proportionately with the following thirteen variables:[6] total pro-forma capitalization, average term, times pro-forma interest earned, size of issue, the coefficient of variation of the ratio of EBIT to sales, EBIT, maturity, the ratio of pro-forma working capital to pro-forma long-term debt, ratio of pro-forma long-term debt to pro-forma total capitalization, average size of issue, the coefficient of variation of EBIT, the ratio of EBIT to sales, and sales itself. Logarithms were therefore used for these thirteen variables. Natural numbers were indicated for the remaining five variables.[7] The preliminary hypothesis then was of the form: $\text{Log}\, Y = a + b_1 \,\text{Log}\, X_1 + b_2 X_2 + \ldots$, where Y is yield; X_1 represents the thirteen variables for which logarithms were used, and X_2 represents the five variables for which natural numbers were used.

Six regressions were then run on this function—one for each of the six quarters for which the scatter diagrams had been drawn. The results of these six regressions are given in Table 15. All showed high R^2's and highly significant F's.[8] For these six quarters,

[4] First quarter of 1951, second quarter of 1952, fourth quarter of 1954, fourth quarter of 1955, third quarter of 1956, and first quarter of 1961.

[5] This procedure assumed, of course, that if the level of yields on outstandings were reasonably stable the level of yields on direct placements would have been reasonably stable also.

[6] That is, when the given variable varied by some specified (constant) percentage, yield would also vary by some specified (constant) percentage.

[7] Logarithms could not, of course, be used for the trend variables inasmuch as, in many cases, trend was negative.

[8] The symbol R^2 refers to the proportion of the variation in yield explained by the independent variables; F is a test of the significance of the amount of

TABLE 15

Industrials: Fourteen Regressions, Yield on Eighteen Variables, R^2, F, Probability of F, Degrees of Freedom, Selected Quarters, 1951–61

Quarter and Year	R^2	F	$P_{F \leq}$	Degrees of Freedom
1/1951[a]	.834	6.69	.01	24
2/1951	.852	7.05	.01	22
1/1952[a]	.810	5.22	.01	22
2/1952	.838	4.60	.01	16
1/1954	.919	10.78	.01	17
2/1954	.792	4.64	.01	22
4/1954[a]	.902	7.66	.01	15
2/1955	.849	4.70	.01	15
4/1955[a]	.850	4.39	.01	14
2/1956	.738	2.66	.05	17
3/1956[a]	.696	2.54	.05	20
4/1956	.683	2.16	.05	18
1/1961[a]	.857	5.34	.01	16
2/1961	.727	2.95	.05	20

[a] Six original regressions.

then, the preliminary hypothesis was, in fact, explaining a high percentage of variation in yield.

Eight additional regressions were then run in order to see how the hypothesis would behave under circumstances which seemed to be somewhat less favorable—i.e., in quarters in which some movement of yield on outstandings had, in fact, occurred. The results of these additional eight regressions are also given in Table 15. The R^2's ranged from .68 for the fourth quarter of 1956 to .92

variation being explained by the independent variables; P_F is the probability, on the basis of chance alone, of obtaining an F as high as that actually obtained.

for the first quarter of 1954. Of the fourteen R^2's taken together, two were .90 or better, nine were .80 or better and all but two were .72 or better.

These results were taken to mean that the preliminary hypothesis was explaining a satisfactory percentage of variation in yield (especially because time could not be held absolutely constant) and that therefore no substantial purpose was to be served by adding variables or experimenting with other forms of function.[9]

The Significant Variables

This part of the chapter is concerned with the question: Which of the foregoing variables had a measurable effect on yield more or less consistently through the period?

The procedure followed was, first, to choose twenty-two cross sections for each of the twenty-two half years, 1951–61. Twenty-two was decided upon rather than eleven or forty-four, or some other number, because twenty-two represented the "optimum" between degrees of freedom, on the one hand, and number of cross sections on the other.[10] Had more than twenty-two cross sections been chosen (the most likely larger number was forty-four, one for each quarter), degrees of freedom would have been negative in some periods. On the other hand, had fewer than twenty-two been chosen (the most likely smaller number was eleven, one for each year), degrees of freedom would have been larger and the cross section wider than seemed necessary.

Simple correlations were then obtained among all the variables, both dependent and independent (Y on X_1 through X_{19}) for *each* cross section. Thus, twenty-two simple correlation coefficients were obtained for Y on X_1, twenty-two for Y on X_2, and so forth. Twenty-two simple correlation coefficients were also obtained for

[9] Residuals were, for the most part, approximately normally distributed although a somewhat better fit would probably have been obtained in four of the fourteen quarters had the squared logarithms of some variables been used. These four quarters were 1/1951, 2/1952, 1/1954, 2/1954.

[10] In the judgment of the author.

TABLE 16

Industrials: Weighted Average Correlation Coefficients, Y on Each X and Each X on Each Other X, 1951–61

	Y	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}
Y	1.00																		
X_2	-.59 (0)	1.00																	
X_3	-.47 (0)	+.61 (22)	1.00																
X_4	-.33 (1)	+.13 (16)	+.05 (13)	1.00															
X_5	-.22 (4)	-.03 (11)	+.03 (13)	+.06 (13)	1.00														
X_6	-.07 (7)	+.04 (11)	+.07 (12)	-.04 (7)	-.09 (5)	1.00													
X_7	-.14 (4)	-.41 (22)	+.40 (22)	-.01 (10)	.01 (14)	-.05 (6)	1.00												
X_8	-.54 (0)	+.81 (22)	+.63 (22)	-.06 (8)	+.14 (16)	-.04 (6)	+.28 (21)	1.00											
X_9	-.07 (5)	-.04 (7)	-.03 (9)	+.14 (18)	-.02 (8)	+.05 (11)	-.07 (6)	-.03 (9)	1.00										
X_{10}	-.04 (9)	-.01 (12)	-.02 (10)	+.11 (18)	-.03 (8)	-.03 (9)	-.04 (8)	-.04 (10)	+.84 (21)	1.00									
X_{11}	+.24 (22)	-.25 (0)	-.21 (1)	-.23 (1)	-.01 (13)	-.13 (3)	-.11 (5)	-.18 (2)	-.04 (11)	-.09 (11)	1.00								
X_{12}	-.60 (0)	+.96 (22)	+.63 (22)	+.36 (22)	-.01 (13)	-.01 (8)	+.74 (21)	+.74 (22)	+.01 (10)	+.01 (12)	-.31 (0)	1.00							
X_{13}	-.46 (0)	+.61 (22)	+.91 (22)	+.07 (15)	-.05 (7)	+.07 (14)	+.35 (20)	+.59 (22)	.00 (11)	+.00 (12)	-.21 (0)	+.60 (22)	1.00						
X_{14}	-.08 (7)	+.04 (9)	-.11 (6)	+.52 (22)	+.13 (17)	-.01 (9)	-.07 (7)	-.22 (3)	+.10 (13)	+.03 (13)	+.03 (13)	+.01 (11)	-.10 (5)	1.00					
X_{15}	+.20 (20)	-.11 (5)	+.05 (18)	-.70 (0)	-.02 (10)	-.01 (11)	+.07 (16)	+.14 (19)	-.14 (4)	-.06 (6)	+.02 (13)	-.13 (4)	+.02 (14)	-.67 (0)	1.00				
X_{16}	-.56 (0)	+.80 (22)	+.65 (22)	-.05 (10)	+.16 (18)	-.01 (10)	+.29 (20)	+.99 (22)	-.07 (7)	-.03 (11)	-.17 (2)	+.76 (22)	+.58 (22)	-.21 (3)	+.13 (3)	1.00			
X_{17}	+.18 (19)	-.19 (1)	-.20 (1)	-.16 (4)	-.01 (9)	-.14 (6)	-.07 (4)	-.11 (4)	-.03 (10)	-.03 (10)	+.82 (22)	-.25 (1)	-.20 (0)	.15 (11)	+.00 (15)	-.11 (4)	1.00		
X_{18}	-.22 (2)	+.28 (22)	+.18 (21)	+.27 (20)	-.01 (8)	-.17 (6)	+.11 (17)	+.30 (22)	+.01 (14)	+.12 (14)	-.21 (4)	+.38 (22)	+.18 (21)	-.18 (3)	+.00 (10)	+.29 (22)	-.05 (7)	1.00	
X_{19}	-.55 (0)	+.91 (22)	+.58 (22)	+.26 (21)	-.01 (11)	+.09 (13)	+.28 (21)	+.69 (22)	.00 (11)	-.00 (12)	-.25 (0)	+.91 (22)	+.56 (22)	+.09 (14)	-.14 (5)	+.69 (22)	-.25 (0)	-.03 (11)	1.00

Note: Numbers in parentheses refer to number of positive relationships between the pairs of variables.

each pair of X variables (X_2 on X_3, X_2 on X_4, and so forth). These simple relationships are summarized in matrix form in Table 16.[11] The figure −.59 in the first column is, for example, the weighted average of the twenty-two simple correlations of Y on X_2; −.47 is the weighted average of twenty-two simple correlations of Y on X_3, and so forth. Under the column labeled X_2, +.61 is the weighted average of the twenty-two simple correlations of X_2 and X_3, and so forth for each and every pair of X variables. The number in parenthesis under each such average correlation is the number of times the relationship between the pair of variables in question was positive. Thus, the correlation between X_2 and X_8 was positive in each and every cross section, whereas the correlation between Y and X_8 was positive in none, i.e., it was consistently negative.

A summary of this matrix follows.

1. Of the eighteen (substantive) variables, the following nine were consistently correlated with yield (See Table 17): X_{12}, X_2, X_{16}, X_{19}, X_8, X_3, X_{13}, X_4, and X_{11}. Of these nine variables, five were "size" variables and two were "duration" variables.

2. Five variables were slightly but nevertheless more or less consistently correlated with yield: X_{18}, X_5, X_{15}, X_{17}, and X_7.

3. Four variables appeared not to be correlated at all with yield: X_{14}, X_6, X_9, and X_{10}.

In addition, the matrix indicates that, as one would expect, substantial intercorrelations existed among the "independent" variables and especially within certain broad classes of variables. Tables 18 and 19 summarize these "grouped" intercorrelations.

1. The *size* variables (X_2, X_8, X_{12}, X_{16}, X_{19}) were all highly intercorrelated and all were highly correlated with yield.

2. The *duration* variables (X_3, X_{13}) were highly intercorrelated and both were highly correlated with yield. In addition, *both* duration variables were highly correlated with *all* the size variables.

3. The *financial security variables* really fall into two subgroups: X_4, X_{14}, X_{15} on the one hand, and X_{11} and X_{17} on the other. The two subgroups are only slightly intercorrelated.

[11] Quarter of year (X_1) is not included in Table 16.

TABLE 17

Industrial: Weighted Average, Correlation of Yield with Each Independent Variable, and Number of Plus Signs, 1951–61

	Variables[a]	Correlation with Yield	Number of Plus Signs
X_{12}	EBIT	-.60	0
X_2	Total capitalization	-.59	0
X_{16}	Average size	-.56	0
X_{19}	Sales	-.55	0
X_8	Size of issue	-.54	0
X_3	Average term	-.47	0
X_{13}	Maturity	-.46	0
X_4	Times pro-forma interest earned	-.33	1
X_{11}	Coefficient of variation: ratio EBIT to sales	+.24	22
X_{18}	Ratio EBIT to sales	-.22	2
X_5	Type of security	-.22	4
X_{15}	Ratio pro-forma long-term debt to pro-forma total capitalization	+.20	20
X_{17}	Coefficient of variation: EBIT	+.18	19
X_7	Years nonrefundable	-.14	4
X_{14}	Ratio of working capital to pro-forma long-term debt	-.08 -.08	7
X_6	Industrial classification	-.07	7
X_9	Relative trend: EBIT	-.07	5
X_{10}	Relative trend: ratio of EBIT to sales	-.04	9

[a] Quarter of year (X_1) not included.

TABLE 18

Industrials: Weighted Average Correlations Among Various Size and Duration Variables, 1951–61

		Size Variables					Duration Variables	
	Y	X_2	X_8	X_{12}	X_{16}	X_{19}	X_3	X_{13}
X	1.00							
X_2	-.59 (0)	1.00						
X_8	-.54 (0)	+.81 (22)	1.00					
X_{12}	-.60 (0)	+.96 (22)	+.74 (22)	1.00				
X_{16}	-.56 (0)	+.80 (22)	+.99 (22)	+.76 (22)	1.00			
X_{19}	-.55 (0)	+.91 (22)	+.69 (22)	+.91 (22)	+.69 (22)	1.00		
X_3	-.47 (0)	+.61 (22)	+.63 (22)	+.63 (22)	+.65 (22)	+.58 (22)	1.00	
X_{13}	-.46 (0)	+.61 (22)	+.59 (22)	+.60 (22)	+.58 (22)	+.56 (22)	+.91 (22)	1.00

4. The *profitability* variables (X_{11}, X_{12}, X_{18}) were not markedly intercorrelated.

5. The two *growth* variables (X_9 and X_{10}) were moderately intercorrelated.

THE STEPWISE REGRESSIONS

To this point, three things have been done: (1) a hypothesis has been formulated and tested and found reasonably satisfactory; (2) the cross-section periods have been chosen so as to "optimize" number of cross sections and degrees of freedom; (3) simple correlations have been obtained and averaged over the twenty-two cross sections, the dependent variable on each "independent"

TABLE 19

Industrials: Weighted Average Correlations Within Various Classes of Independent Variables

Financial Security

	Y	X_4	X_{14}	X_{15}
Y	1.00			
X_4	-.33 (1)	1.00		
X_{14}	-.08 (7)	+.52 (22)	1.00	
X_{15}	+.20 (20)	-.70 (0)	-.67 (0)	1.00

Variability

	Y	X_{11}	X_{17}
Y	1.00		
X_{11}	+.24 (22)	1.00	
X_{17}	+.18 (19)	+.82 (22)	1.00

Profitability

	Y	X_{12}	X_{18}
Y	1.00		
X_{12}	-.60 (0)	1.00	
X_{18}	-.22 (2)	+.38 (22)	1.00

Growth

	Y	X_9	X_{10}
Y	1.00		
X_9	-.07 (5)	1.00	
X_{10}	-.04 (9)	+.84 (22)	1.00

variable and each independent variable on each of the other independent variables.

These simple correlations have identified those independent variables which are most likely to be important in explaining variations in yield. They have also suggested rather strongly that some of the "independent" variables are highly intercorrelated and that, therefore, not all the variables which seem important may really be so when the others are held constant.

The next step in the procedure responds, therefore, to the following question: Which variables are and which are not really necessary to a satisfactory explanation of variation in yield, time held constant? This question, in turn, has two parts. First, of those variables that are intercorrelated, which are the most important and, of the rest, which in fact contribute something *additional* to our understanding of variation in yield? For example, X_2 and X_{12} are both highly correlated with yield but both are "size" variables and they are highly correlated with each other. Are both necessary? And what about the other size variables? Do they make any contribution to an explanation of variation in yield after the separate effects of X_2 and X_{12} have been taken into account?

Second, what about the variables which seem to be acting independently (e.g., X_7), or which are correlated little, if at all, with the other independent variables? Are they really making a *separate* contribution to our understanding of variation in yield?

In an attempt to answer these questions, the variables were introduced successively, in each cross section, in the order of their subscripts as given in Table 17. For the purpose of deciding which variable should be introduced first, guidance was provided by the simple correlation matrix. In effect, the first substantive variable introduced was a size variable (X_2); the second, a duration variable (X_3); the third, a financial security variable (X_4), and so forth. The second size variable (X_{12}) was not introduced until at least one variable from each other group had been introduced.

Thus, for the first cross section (i.e., the first half of 1951), Y was regressed on X_1 and a regression coefficient for X_1 obtained; Y was then regressed on X_1 and X_2 together, coefficients for both X_1 and X_2 being obtained; X_3 was then introduced, and so on through X_{19}. Therefore, for the first half of 1951, nineteen regression equations were obtained—the first with one variable, the second with two, etc. One set of nineteen such equations was thus obtained for each of the twenty-two half years in 1951–61.

Table 20 gives R^2, F, Probability of F, and degrees of freedom for *the nineteenth equation* in each of these twenty-two cross sections.[12]

[12] Comparison of the R^2's in Table 20 with those given in Table 15 indicate some deterioration, especially in 1956. In general, it seems clear that the basic hypothesis tends to explain less variation in yield when the length of the cross section, in terms of time, is increased. This, of course, is what one would expect,

TABLE 20

Industrials: Twenty-Two Regressions, Yield on Nineteen Variables, R^2, F, Probability of F, Degrees of Freedom, Semiannually, 1951–61

	R^2	F	$P_{F \leqq}$	Degrees of Freedom
1951				
1	.783	12.2	.01	64
2	.876	11.4	.01	31
1952				
1	.769	9.8	.01	56
2	.815	6.2	.01	27
1953				
1	.709	4.7	.01	37
2	.797	5.6	.01	27
1954				
1	.743	8.7	.01	57
2	.805	8.7	.01	40
1955				
1	.707	5.3	.01	42
2	.766	9.3	.01	54
1956				
1	.560	2.7	.01	40
2	.590	4.9	.01	64
1957				
1	.484	1.8	.05	36
2	.843	2.0	.10	7
1958				
1	.691	3.4	.01	29
2	.714	4.3	.01	33
1959				
1	.736	2.5	.05	17
2	.799	4.8	.01	23
1960				
1	.534	1.3		21
2	.663	4.0	.01	39
1961				
1	.721	7.3	.01	54
2	.697	2.9	.05	24

The principal product of these twenty-two stepwise regressions was a succession of "t" statistics as follows: the initial "t" obtained for any given variable was the "t" which resulted when that variable was entered into the regression—i.e., when X_2 was entered into the regression in the first cross section, a "t" was obtained which constituted a test of the significance of X_2 given X_1. One such initial "t" for X_2 was obtained for *each of the twenty-two cross sections,* or, in other words, a total of twenty-two "t's" bearing on the significance of X_2 and X_1. The same procedure was followed for each variable X_3 through X_{19}. Table 21 gives these *initial* "t's" for each variable and each cross section.

Using the initial "t's," three tests were then run: a $\bar{t}$ test, a sign test, and what might be called a "distribution of t's" test.

1. The $\bar{t}$ test, was designed to ascertain whether some given variable, X_i, added *consistently and significantly* (in the statistical sense) to regression when entered into the regression, i.e., given the preceding variables. For example, if b_1 (the coefficient of variable X_1) showed a $\bar{t}$ greater than 2.00, this would mean that a large proportion of the twenty-two signs were in one direction *and* that a large proportion of the individual t's were high. Inasmuch as each individual t is a test of *statistical* significance, a high $\bar{t}$ constitutes strong evidence that the variable in question was consistently (i.e., over the twenty-two cross sections) statistically significant when entered. Table 21 shows, for example, that, when entered, X_2 produced twenty-one t's equal to or less than -2.00 and a $\bar{t}$ of -5.42; when X_3 was entered, it produced six t's equal to or less than -2.00 and a $\bar{t}$ of -1.01.[13]

2. Variables which show a low $\bar{t}$ may, however, be significant. A low $\bar{t}$ means merely that such variables are not *statistically* significant, i.e., that the amount of variation being explained by

given the fact that the longer the cross section, the greater the expected variation in yield with respect to time. But see below, Table 24.

[13] Strictly, probabilities rather than "t's" should have been distributed and averaged, using Fisher's technique, in order to give due weight to differing degrees of freedom. This refinement did not seem worth the large amount of additional computation which it would have required. See last column of Table 20.

TABLE 21

Industrials: Significance of Each Variable, $X_2 - X_{19}$, When that Variable Was Introduced into Regression, Semiannually, 1951–61[a]

	$^tX_2$	$^tX_3$	$^tX_4$	$^tX_5$	$^tX_6$	$^tX_7$	$^tX_8$	$^tX_9$	$^tX_{10}$
1951									
1	-9.90	-2.19	-3.27	-4.25	-3.12	+ .92	-1.99	-1.74	+1.02
2	-9.54	+1.30	-4.38	-1.55	-1.22	-2.13	+1.30	- .57	+ .48
1952									
1	-6.81	+ .60	-5.00	-4.56	-1.05	+1.85	-2.30	+ .93	+2.35
2	-6.38	+ .19	- .12	-4.53	- .54	- .30	- .12	- .16	- .10
1953									
1	-6.09	-1.25	- .52	-3.00	-3.36	+1.63	+ .72	- .98	- .32
2	-3.75	-3.32	-3.97	-3.51	-1.61	+1.80	-1.89	- .002	+ .99
1954									
1	-6.20	-2.01	-1.83	-2.02	-1.04	+3.26	-2.77	+ .37	+1.02
2	-5.60	- .89	-2.61	-4.54	-2.92	+2.76	-3.30	+ .14	- .02
1955									
1	-7.62	-1.22	-3.38	-1.06	- .84	+1.68	- .65	- .12	+ .64
2	-6.92	-1.62	-4.22	-5.74	-1.43	- .13	- .28	- .81	-2.59
1956									
1	-2.18	-2.60	-1.53	+1.13	+2.31	+ .38	-2.42	-2.65	-1.06
2	-6.46	-1.09	-2.63	-0.06	-1.21	-2.03	-1.75	- .49	+ .47
1957									
1	-1.88	+ .20	-2.93	-2.99	-1.34	+ .43	-1.53	+ .52	+ .37
2	-2.69	+0.15	- .69	-2.11	- .85	- .66	-1.23	+ .26	+ .21
1958									
1	-2.94	-2.23	-2.76	- .89	- .13	+1.15	-2.44	+ .34	- .37
2	-6.00	-1.05	-1.80	-1.73	+ .57	+2.21	+ .54	+ .28	-1.11
1959									
1	-3.51	-1.29	-0.06	+ .47	-1.61	+1.65	- .48	-1.72	+1.04
2	-4.11	-1.83	-3.44	-2.41	- .31	- .39	-2.00	- .42	- .32
1960									
1	-2.52	+ .29	-2.08	- .74	-1.19	+ .20	-1.18	-1.30	+0.92
2	-7.00	-1.04	-1.08	-1.13	- .91	+1.14	+1.53	- .06	+1.01
1961									
1	-5.23	-3.18	-4.95	-1.59	+1.15	+ .04	-1.34	+1.37	+ .12
2	-5.88	+ .49	-1.02	- .64	+ .79	- .71	-1.37	- .38	- .47
$\bar{t}$	-5.42	-1.07	-2.47	-2.09	-0.90	+0.67	-1.14	-0.33	+0.13

(continued)

TABLE 21 (concluded)

	${}^{t}X_{11}$	${}^{t}X_{12}$	${}^{t}X_{13}$	${}^{t}X_{14}$	${}^{t}X_{15}$	${}^{t}X_{16}$	${}^{t}X_{17}$	${}^{t}X_{18}$	${}^{t}X_{19}$
1951									
1	- .19	+ .51	+1.00	- .60	- .37	- .53	- .61	+ .10	+ .18
2	- .86	+2.45	+ .32	- .29	-1.63	- .06	-1.14	+ .08	+1.75
1952									
1	+ .62	+ .68	+ .09	+1.91	-1.11	+ .69	+ .81	-1.04	+ .99
2	-1.76	+2.25	- .55	+1.31	- .26	- .91	-1.65	+ .52	-1.96
1953									
1	- .98	+ .73	-1.55	+1.29	- .25	+ .03	+1.53	- .36	+ .64
2	- .24	+1.05	- .90	- .82	-1.09	+1.46	- .42	- .04	+ .51
1954									
1	-1.35	+1.85	+ .93	-1.43	- .61	-2.03	+2.55	-1.03	-2.60
2	+2.23	+ .46	-1.49	-1.62	- .66	+ .57	+ .52	+ .10	- .21
1955									
1	+ .66	+2.36	- .09	- .35	- .86	- .62	+1.36	+ .34	-1.85
2	+1.07	- .67	- .16	+ .58	+ .74	- .77	- .07	+ .48	+1.04
1956									
1	-1.23	+1.57	+2.04	-1.14	+ .89	- .26	+ .55	+ .25	- .10
2	+ .48	- .04	+ .76	- .22	-1.33	+1.17	+ .33	-1.91	+ .69
1957									
1	- .02	+ .05	+ .32	+1.42	-1.94	+ .40	- .84	- .84	+ .45
2	+1.58	- .20	+1.96	-1.01	+ .28	- .03	+ .44	-2.79	+ .44
1958									
1	+ .96	-3.22	- .53	+ .83	-2.34	- .26	+ .26	- .41	- .79
2	+ .25	- .93	- .48	+ .13	+1.66	+ .47	+1.07	+1.24	+2.52
1959									
1	+ .30	- .72	-1.74	+1.20	- .01	+1.74	-1.86	-1.28	- .42
2	+1.54	- .19	- .60	+2.37	- .02	- .55	+ .86	-1.55	- .14
1960									
1	+ .56	+ .91	- .04	-1.87	+ .53	+ .52	+1.40	+ .01	-1.29
2	- .72	+ .63	-1.57	+ .91	- .25	- .37	+ .67	-1.39	-2.38
1961									
1	+ .85	+1.55	-1.47	+ .97	+ .38	+ .42	-2.17	-2.54	-1.38
2	+1.11	+1.27	-1.79	+ .85	+ .92	- .61	+ .99	+ .86	+ .52
$\bar{t}$	+0.22	+0.56	-0.34	+0.20	-0.33	-0.02	+0.18	-0.51	-0.15

[a]The numbers given in this table are t's $\left(= \frac{b_i}{\sigma_{b_i}}\right)$.

such variables is relatively small. In addition, any variable, whether or not it bears any real relation to the dependent variable, will add something to the sum of squares for regression, simply because of sampling error. The second test then, *the sign test,* is aimed at identifying those variables which, although not statistically significant, are consistently adding small amounts to regression. This test is based on the presumption that if a variable is merely a random number (i.e., if it bears no real relation to the dependent variable) its twenty-two initial "t's" should show roughly equal numbers of plus and minus signs. If its "t's" show an unexpectedly large number of plus or minus signs, a presumption would be created that it (the underlying variable) was *not* a random number, but, rather, bore some consistent relation to the dependent variable. Table 22 indicates that b_6 and b_8 (in addition to b_2, b_4, and b_5) showed strong significance by this test, and b_3, b_7, b_{12}, b_{13}, and b_{15}, marginal significance.

3. It is altogether conceivable, however, that some variable which was not important in the early part of the period became so later, or vice versa; or that some variable, X_j, occasionally acts as a proxy for some other variable, X_i, which latter is important most of the time but not always, or less important at some times than at others. If some given variable, not significant by either of the above two tests, showed an unusually large number of very high or very low t values, that variable was also presumed to be significant when entered. X_{12}, for example, showed a low $\bar{t}$ (+0.56) and fifteen plus signs but, on the other hand, three t's larger than +2.00 (Table 23).

Any variable which failed to show significance by at least one of the above three tests was eliminated from further consideration; it was assumed to be nothing more than a random number. This does not mean, of course, that such variables were, *in fact,* random numbers (although they may have been so); it means merely *that given the preceding variables,* such variables were behaving as *if they were* random numbers. Thus, given X_2, which showed

TABLE 22

Industrials: Number of Plus and Minus Signs Obtained on Partial Regression Coefficients and Binomial Probability of Obtaining at Least Larger Number if Actual Probability Is .50

Coefficient	No. of Plus Signs	No. of Minus Signs	$P_B \lesseqgtr$
b_2	0	22	.000
b_3	7	15	.067
b_4	0	22	.000
b_5	2	20	.000
b_6	4	18	.002
b_7	15	7	.067
b_8	4	18	.002
b_9	8	14	.143
b_{10}	13	9	.262
b_{11}	13	9	.262
b_{12}	15	7	.067
b_{13}	7	15	.067
b_{14}	12	10	.416
b_{15}	7	15	.067
b_{16}	10	12	.416
b_{17}	13	9	.262
b_{18}	10	12	.416
b_{19}	11	11	.584

high significance when entered, X_{19}, also a size variable, consistently added nothing to regression. This does not mean, of course, that had X_{19} been entered *first* it would have added nothing to regression. Given the high degree of correlation between X_2 and X_{19}, the latter would almost surely have shown high significance, had it been entered first.

TABLE 23

Industrials: $\bar{t}$'s and Distribution of t's When Entered, Partial Regression Coefficients on X_2–X_{19}

Coefficient	$\bar{t}$	Per Cent $\geqq +2.00$	Per Cent $\leqq -2.00$
b_2	-5.42	--	95.5
b_3	-1.01	--	27.3
b_4	-2.47	--	59.1
b_5	-2.09	--	50.0
b_6	-0.90	4.5	13.6
b_7	+0.67	13.6	9.1
b_8	-1.14	--	22.7
b_9	-0.33	--	4.5
b_{10}	+0.13	4.5	4.5
b_{11}	+0.22	4.5	--
b_{12}	+0.56	13.6	4.5
b_{13}	-0.34	--	4.5
b_{14}	+0.20	4.5	--
b_{15}	-0.33	--	4.5
b_{16}	-0.02	--	4.5
b_{17}	+0.18	4.5	4.5
b_{18}	-0.51	--	9.1
b_{19}	-0.15	4.5	9.1

In summary: (1) by the $\bar{t}$ test, three variables showed clear significance—X_2, X_4, and X_5 (Table 21); (2) by the sign test, seven additional variables seemed significant or on the borderline of being so—X_3, X_6, X_7, X_8, X_{12}, X_{13}, and X_{15} (Table 22); (3) by the distribution of t's test, no *additional* variables appeared to be even marginally significant (Table 23).

In order to catch any coefficient which might have failed to show significance by the sign test simply because it was trending from above zero to below zero, or vice versa, weighted trends were fitted to those coefficients which had not otherwise shown significance—b_9, b_{10}, b_{11}, b_{14}, b_{16}, b_{17}, b_{18}, and b_{19}.[14] None showed trend.

Last, each of the coefficients which had shown no significance by the above tests when entered was examined in the light of *subsequent* variables. None showed significance as variables were added.

The next question was: Are all the variables which are presumed to be significant *when entered* really necessary? For example, three size variables, X_2, X_8, and X_{12}, showed significance when entered, but of course X_8 and X_{12} were not in the regression when X_2 was entered and X_{12} was not in the regression when X_8 was entered. Possibly, therefore, X_2 is not necessary given X_8; and X_8, in its turn, may not be necessary given X_{12}. This means that the significance of X_2 must be examined in the light of X_8 and X_{12}, and the significance of X_8, in the light of X_{12}.

RERUNS ON SIGNIFICANT VARIABLES

In response to the above question, the regressions were rerun quarterly on the ten significant variables plus month of quarter. Month of quarter was included as a variable in order to hold time constant to some extent within each quarter.[15] Given the fact that both the ratio of EBIT to total interest (X_4) and EBIT itself (X_{12}) showed significance when entered, X_4 was redefined so as to avoid including the same variable twice and thus dividing its effect between two coefficients. The redefined variable X_{4r} became simply total pro-forma interest.

The results of these forty-four regressions are given in Tables 24 and 25. On the whole, the results given in Table 24 are satisfactory and represent some improvement over those of the twenty-two

[14] Each coefficient was weighted inversely as to its variance.
[15] All variables were introduced simultaneously and *not* stepwise.

TABLE 24

Industrials: Forty-Four Regressions, Yield on Eleven Variables, R^2, F, Probability of F, Degrees of Freedom, Quarterly, 1951–61

Year and Quarter	R^2	F	$P_{F \lesseqgtr}$	Degree of Freedom
1951				
1	.801	11.3	.01	31
2	.785	9.6	.01	29
3	.951	22.6	.01	13
4	.910	15.6	.01	17
1952				
1	.766	8.9	.01	30
2	.806	9.1	.01	24
3	.869	9.0	.01	15
4	.872	5.6	.01	9
1953				
1	.759	4.0	.01	14
2	.814	8.0	.01	20
3	.904	6.0	.05	7
4	.759	4.6	.01	16
1954				
1	.677	4.6	.01	24
2	.704	6.5	.01	30
3	.870	8.5	.01	14
4	.792	8.3	.01	24
1955				
1	.766	4.7	.01	16
2	.755	6.4	.01	23
3	.889	16.1	.01	22
4	.643	4.9	.01	30
1956				
1	.671	2.2	a	12
2	.643	3.9	.01	24
3	.594	4.8	.01	36
4	.626	3.8	.01	25

(continued)

TABLE 24 (concluded)

Year and Quarter	R^2	F	$P_{F \gtreqless}$	Degree of Freedom
1957				
1	.542	1.8	a	17
2	.588	1.9	a	15
3	.986	6.7	.01	1
4	.796	1.1	a	3
1958				
1	.893	6.1	.01	8
2	.757	5.1	.01	18
3	.789	6.1	.01	18
4	.884	7.6	.01	11
1959				
1	.774	0.9	a	3
2	.566	1.4	a	12
3	.895	9.3	.01	12
4	.778	2.9	a	9
1960				
1	.856	1.1	a	2
2	.424	1.1	a	16
3	.704	2.6	.05	12
4	.701	4.9	.01	23
1961				
1	.803	8.9	.01	24
2	.540	2.9	.05	27
3	.654	2.4	a	14
4	.723	1.4	a	6

[a]P_F is greater than .05.

semiannual regressions (Table 20). In sixteen of forty-four regressions (or about 36 per cent), R^2 was greater than .80, whereas this was true in only four cross sections when the regressions were run semiannually. In addition, thirty-two R^2's (or about 73 per cent) were greater than .70 when the regressions were run quarterly

TABLE 25

Industrials: Forty-Four Regressions, Yield on Eleven Variables, Number of Times t Was Greater Than +2.00 or Less Than -2.00 and Distribution of Plus and Minus Signs of Coefficients

Coefficient	t Greater Than +2.00	t Less Than -2.00	Distribution of Signs: Plus	Distribution of Signs: Minus	$P_B \leqq$ [a]
b_2	1	8	11	33	.001
b_3	1	6	14	30	.02
b_{4r} [b]	12	0	34	10	.001
b_5	0	11	8	36	.0001
b_6	0	5	14	30	.02
b_7	2	2	28	16	.05
b_8	1	11	14	30	.02
b_{12}	1	4	8	36	.0001
b_{13}	1	1	20	24	.33
b_{15}	0	4	18	26	.15

[a]Probability of obtaining at least larger number of signs (either plus or minus) if true probability equals .50.

[b]Small r indicates redefined version of variable used. See text.

as compared with fifteen (or about 68 per cent) when the regressions were run semiannually. In brief, narrowing the cross section from six to three months has, on the whole, reduced the error, despite the fact that the number of variables has been substantially reduced.

Table 25 summarizes various data bearing on the significance of the variables included in these "reruns." The first two columns indicate the number of times t was greater than +2.00 or less than −2.00. The coefficient on X_2, for example, was greater than +2.00

in one cross section and less than −2.00 in eight cross sections, the coefficient on X_3 was greater than +2.00 in one cross section and less than −2.00 in six cross sections, and so forth for each other variable. The next two columns tabulate the number of plus and minus signs of each coefficient over the forty-four cross-sections, and the last column gives the probability of obtaining the larger number of signs. Thus, of the forty-four coefficients on X_2, thirty-three were negative and eleven positive. The probability of obtaining the thirty-three minus signs, if the true probability is .50, is .001 and so forth for each other variable. Each of the ten variables, except X_{13}, shows significance by either the t's test or the sign test. X_{13} showed trend at $P < .02$.

Importance of Variables

The foregoing analysis has accomplished two things: (1) it has separated the "significant" from the "nonsignificant" variables and (2) it has provided us with a series of forty-four quarterly predictive equations, which will be used below to compute series on yields on direct placements. It has told us nothing, however, about the *importance* of each variable, that is, which variables are capable of having a substantial impact on yield and which are not. Thus, if a coefficient is small and the variability of the associated variable is also small, the impact on yield will not be great, even if the coefficient is highly significant statistically.

To assess the importance of each variable in this sense, three steps were taken. First, an over-all regression was run for the period as a whole on all the significant variables, with X_4 redefined as indicated above. In this regression, time was held approximately constant by using the monthly yield on Aaa corporates (Moody's) as a variable.[16] The results of this over-all regression are given in

[16] We might alternatively have averaged each b_i over its forty-four sample values, weighting each sample value inversely as to its variance, or have used that value of each b_i which, of the forty-four sample values, showed the strongest significance. AAA corporates were used because they seemed to conform, better than other available series, to the movements of direct placements.

TABLE 26

Industrials: "Over-All" Regression, Log Y on Eleven Variables, Regression Coefficients, Standard Errors, and Tests of Significance

Coefficient	$\bar{b}$	$\sigma_{\bar{b}}$	t	$P_t^a \leq$
Intercept	+1.1816	.0361	+32.70	.001
b_1	+ .7995	.0146	+54.85	.001
b_2	- .0475	.0081	- 5.84	.001
b_3	- .0228	.0143	- 1.59	.15
b_{4r}	+ .0683	.0071	+ 9.58	.001
b_5	- .0232	.0028	- 8.32	.001
b_6	- .0178	.0046	- 3.91	.001
b_7	+ .0010	.0003	+ 3.59	.001
b_8	- .0195	.0029	- 6.76	.001
b_{12}	- .0322	.0047	- 6.87	.001
b_{13}	- .0212	.0155	- 1.37	.20
b_{15}	- .0349	.0077	- 4.54	.001

[a]With 1,271 degrees of freedom, two tailed.
For this regression, R^2 = .806, F = 481.1, and P_F = <.001.

Table 26.[17] Second, each of the regression coefficients given by the over-all regression was multiplied by the standard deviation of the appropriate X_i. Finally, the antilog of each product was obtained (without regard to the sign of that product). The results are given in Table 27.

Column 1 of Table 27 lists the various $\bar{b}_i$'s; column 2, the σ of each X_i; and column 3 the product of each $\bar{b}$ multiplied by the

[17] Two variables which were deemed significant in the cross sections, showed only slight significance over-all—X_3 and X_{18}. The coefficients on these variables fluctuated around zero (see Chart D-1). The cross-section analysis does not, of course, hold expectations as to the future course of interest rates constant and perhaps the signs of these two coefficients were sensitive to such expectations. See Appendix A.

TABLE 27

Industrials: Percentage Impact of Each Variable on Yield When that Variable Increased by One Standard Deviation

Variable	b_i (1)	σ_{x_i} (2)	$b_i\sigma_{x_i}$ (3)	Antilog of Col. 3[a] (4)
X_2	-.0475	1.7523	-.08323	1.09
X_3	-.0228	.3855	-.00879	1.01
X_{4r}	+.0683	1.7018	+.11623	1.12
X_5	-.0232	.8237	-.01911	1.02
X_6	-.0178	.4857	-.00865	1.01
X_7	+.0010	8.5396	+.00854	1.01
X_8	-.0195	1.4262	-.02781	1.03
X_{12}	-.0322	1.7823	-.05739	1.06
X_{13}	-.0212	.3369	-.00714	1.01
X_{15}	-.0349	.5410	-.01888	1.02

[a]Signs ignored.

corresponding σ of each X. Column 4 gives the antilog of this product and represents the approximate percentage change in yield which would be produced by an increase in X_i of an amount equal to one standard deviation.

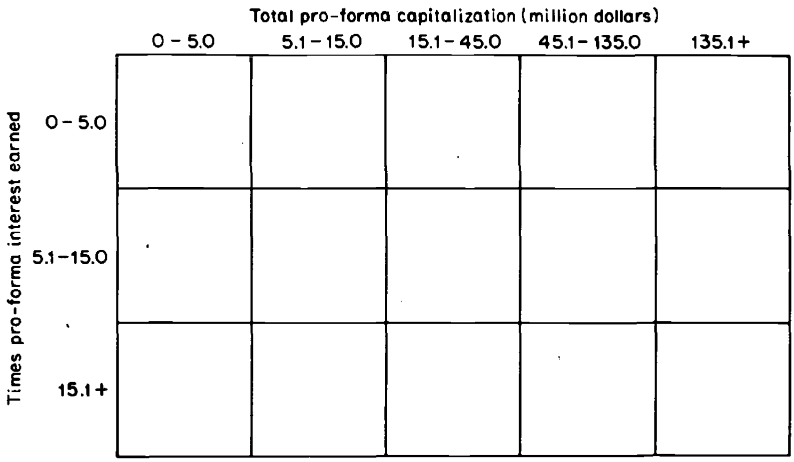
The ten variables, ranked in the order of their respective percentage impact on yield, are X_{4r}, X_2, X_{12}, X_8, X_5, X_{15}, X_3, X_6, X_7, and X_{13}.

The Yield Series

With the foregoing results in hand, two sets of series were constructed: a set of *cross-classified* series and a set of *computed* series. The cross-classified series hold (approximately) constant just two significant variables—albeit two of the most important—

CHART 6

Industrials: System of Cross Classification Used to Construct Seven Cross-Classified Yield Series, 1951–61



whereas the computed series hold rigidly constant all ten significant variables. The cross-classified series are relatively easy to understand and construct and provide "reasonable homogeneity" through time.[18]

THE CROSS-CLASSIFIED SERIES

The procedure used to construct the cross-classified series involved four main steps.

First, the X_4 variable was put back into its original form—as a measure of coverage. Doing so made it possible to take account of three rather than just two of the most important variables.

Quarterly regressions were then run on X_4, as so defined, and on X_2 alone. Weighted averages struck over the set of forty-four coefficients for each variable were, respectively, −.0288 and −.0231.[19]

[18] On the average over the forty-four quarters, X_2 and X_4 together explain about 45 per cent of variation in yield. The ten significant variables explain about 75 per cent.

[19] Both coefficients showed trend over the period but the relationship between

Second, class intervals were established over X_2 and X_4, based on the foregoing weighted average regression coefficients. These class intervals are given in Chart 6. The fact that the regression coefficients on X_2 and X_4 were close to being equal meant that the effect on yield of a given percentage *increase* in (e.g.) X_2 would be almost precisely offset by a corresponding percentage *decrease* in X_4 and this, in turn, meant that the cells lying along any given left-to-right diagonal (see Chart 6) could be regarded, other things being equal, as representing roughly equivalent "quality." Thus, all the issues falling in each of the cells lying along diagonal number 4 were presumed to represent equivalent "quality" in this sense—and correspondingly with each of the other four left-to-right diagonals. This is equivalent to saying that the sum of the logarithms (i.e., $\log X_2 + \log X_4$), taken at the means of the two classes, is approximately equal from one cell to another along any left-to-right diagonal. The *size* of each class interval was chosen so as to allow a "spread," measured at the mean of each interval, of about 5 per cent in yield between diagonals.[20]

Third, the observations were deposited each into their appropriate cells. Simple arithmetic averages were then struck along each diagonal for each quarter. This procedure produced seven quarterly yield series.

These series, however, showed a fairly large number of inconsistencies, roughly half of which were due to the small number of observations in some series in some quarters. Finally, therefore, the seven classes were consolidated into three, so as to provide a strong "middle" series containing about 50 per cent of the observations and "lower" and "upper" series each containing roughly 25 per cent of the observations. The three resulting series are set forth

them was, for all practical purposes, stable. The ratio of b_2 to b_4 was 1.25 in the first quarter of 1951 and 1.20 in the last quarter of 1961.

[20] In addition, some 200 observations, which were not included in the original regressions, were added at this point. These observations were not included originally because in each case some data were missing. After the regressions were run, however, these 200 observations were found to be complete in the necessary respects.

TABLE 28

Industrials: Yields on Direct Placements, Cross Classified, by Class, Quarterly, 1951–61

Year and Quarter	Class I (1)	Class II (2)	Class III (3)	Composite[a] (4)
1951				
1	3.21	3.78	4.11	3.70
2	3.48	3.93	4.63	4.01
3	3.55	3.99	4.84	4.13
4	3.70	4.16	4.75	4.20
1952				
1	3.88	4.33	4.73	4.31
2	3.78	4.27	4.75	4.27
3	3.72	4.31	4.86	4.30
4	3.88	4.20	5.06	4.38
1953				
1	3.94	4.59	4.77	4.43
2	4.29	4.53	4.86	4.56
3	4.48	4.67	5.37	4.84
4	4.45	4.57	5.10	4.71
1954				
1	4.03	4.57	5.02	4.54
2	4.01	4.26	4.57	4.28
3	3.72	4.07	4.68	4.16
4	3.73	4.25	4.34	4.11
1955				
1	3.73	4.17	4.79	4.23
2	3.83	4.23	5.16	4.41
3	3.87	4.24	5.17	4.43
4	4.27	4.55	4.83	4.55
1956				
1	4.46	4.49	4.78	4.58
2	4.48	4.64	4.85	4.66
3	4.30	4.91	5.32	4.84
4	4.81	5.20	5.63	5.21

(continued)

TABLE 28 (concluded)

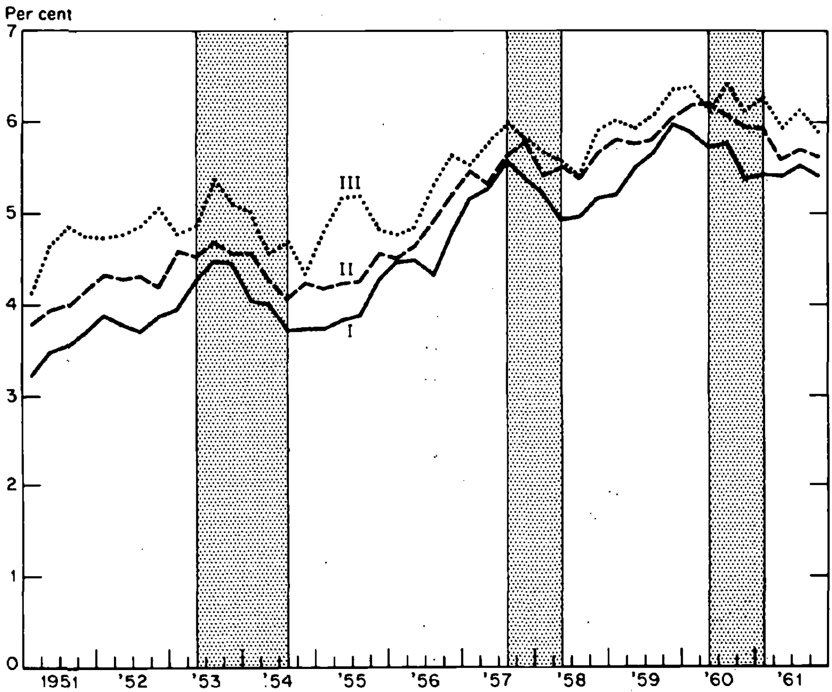
Year and Quarter	Class I (1)	Class II (2)	Class III (3)	Composite[a] (4)
1957				
1	5.16	5.45	5.54	5.38
2	5.27	5.32	5.75	5.45
3	5.58	5.61	5.98	5.72
4	5.38	5.78	5.83	5.66
1958				
1	5.24	5.42	5.67	5.44
2	4.92	5.50	5.57	5.33
3	4.95	5.37	5.40	5.24
4	5.16	5.67	5.90	5.58
1959				
1	5.20	5.80	6.00	5.67
2	5.50	5.75	5.93	5.73
3	5.65	5.80	6.07	5.84
4	5.98	6.02	6.34	6.11
1960				
1	5.89	6.17	6.38	6.15
2	5.72	6.20	6.11[b]	6.01
3	5.76	6.07	6.39	6.07
4	5.36	5.94	6.13	5.81
1961				
1	5.43	5.92	6.27	5.87
2	5.40	5.59	5.94	5.64
3	5.53	5.70	6.13	5.79
4	5.41	5.62	5.92	5.65

[a]This series is equal to the sum of Classes I, II, and III divided by 3 (the arithmetic mean).

[b]Inconsistency.

CHART 7

Industrials: Yields on Direct Placements, Cross Classified, by Class, Quarterly, 1951–61



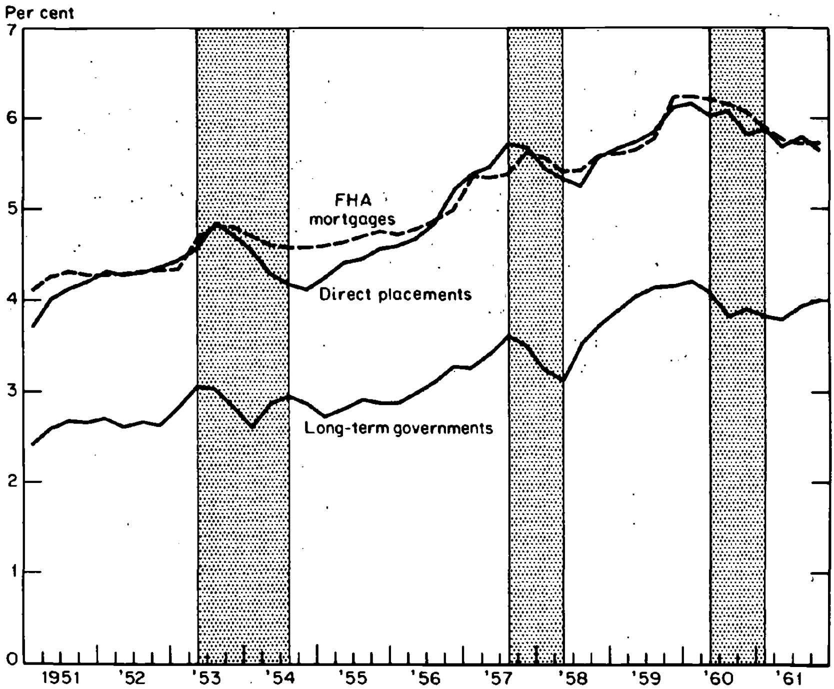
Shaded areas represent business contractions; white areas, expansions.
SOURCE: Table 28.

in Table 28 and Chart 7. Grade I is made up of classes 1, 2, and 3; Grade II, of classes 4 and 5, and Grade III of classes 6 and 7.[21]

[21] A variety of alternative procedures was available for the purpose of reducing the number of series from seven to three. The means of adjacent *diagonals* could have been averaged. This procedure would, of course, have given equal weight to each diagonal. The mean in each *cell,* along each of the three diagonals, could have been calculated and these means in turn averaged. This procedure seemed better than the first because it would have given less weight to an extreme fluctuation along any diagonal. The mean for any given quarter could simply have been struck, over all the *observations* lying along the diagonals to be consolidated. This procedure would have given equal weight to *each observation* and would have further reduced the likelihood that the result in

CHART 8

Industrials: Yields on Direct Placements, Composite Cross Classified, Compared with Yields on FHA Mortgages and Yields on Long-Term Governments, Quarterly, 1951–61



Shaded areas represent business contractions; white areas, expansions.
SOURCE: Table 28; *Federal Reserve Bulletin; Treasury Bulletin.*

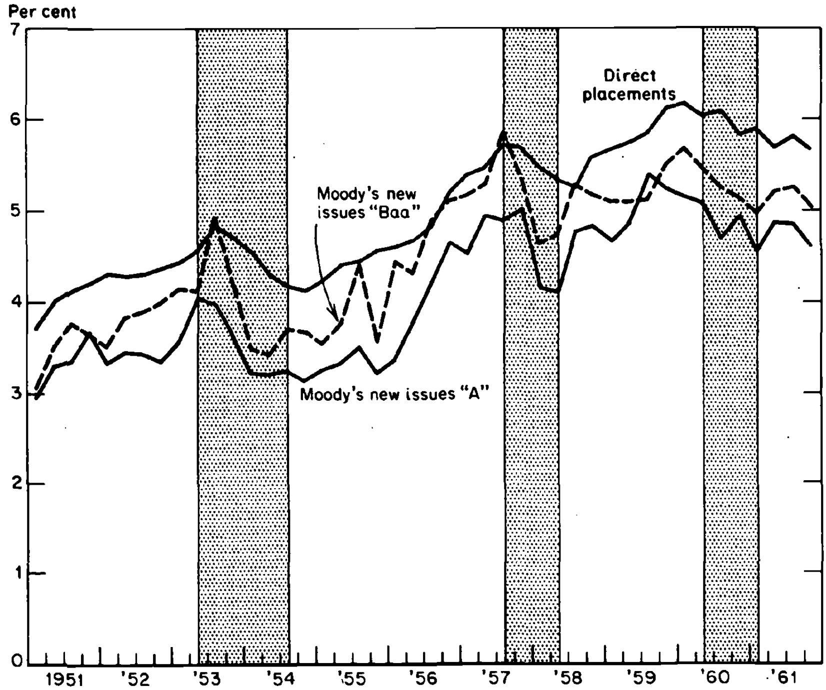
A "composite" series, which is simply an arithmetic average, quarterly, of the three series set forth in Table 28, is given in the last column of that table. Charts 8 and 9 compare the movements of this "composite" series with the movements of yields on FHA

any given quarter in any given class would be affected unduly by an extreme observation.

All three procedures were tried. The first produced three inconsistencies and the others, two each. The series did not differ greatly one from another and the one represented by the third procedure, above, was used.

CHART 9

Industrials: Yields on Direct Placements, Composite Cross Classified, Compared with Moody's Yields on New Issues Grades A and Baa, Quarterly, 1951–61



Shaded areas represent business contractions; white areas, expansions.
SOURCE: Table 28 and Moody's Bond Survey.

mortgages, long-term governments, and Moody's new issues "A" and "Baa."

THE COMPUTED SERIES

Next, values were established, closely approximating the mean values of the three cross-classified series (see Table 29).[22] These

[22] Of course, any set of arbitrary values might have been used for this purpose—and any reader who wishes to experiment may do so, subject to the caveat that the coefficients should not be used outside the range of observations from which they were derived. The mean values of the cross-classified series

TABLE 29

Industrials: Mean Values Used to Obtain Computed Series, by Class

Variable	Unit	Series Ic	IIc	IIIc
X_2	Million dollars	150.0	20.0	4.0
X_3	Years	10.0	8.0	7.0
X_{4r}	Million dollars	1.2	0.3	0.1
X_5	a	1.5	1.5	1.3
X_6	b	0.6	0.6	0.7
X_7	Years	4.0	5.0	3.0
X_8	Million dollars	16.0	4.0	1.3
X_{12}	Million dollars	20.0	4.0	0.9
X_{13}	Years	18.0	15.0	13.0
X_{15}	Dollars of long-term debt per dollar of total capital	0.31	0.31	0.28

[a]For this variable, first mortgage bonds = 0; second mortgage bonds or debts secured by securities or lease = 1; senior debentures = 2; and subordinated debentures = 3. The figures here are an average of these code numbers.

[b]For industrial classification, durables = 0 and nondurables = 1. The figures here are an average of these code numbers.

values were held rigidly constant for each of the three series through the period. Quarterly yields for each series were then computed, using the forty-four regression equations given by the rerun on the ten significant variables.[23] These computed yields are given in the first three columns of Table 30, and in Chart 10.

were used because (a) comparisons with the cross-classified series might be enlightening and (b) the cross-classified series were equally spaced in terms of "quality" and using their mean values suggested that the computed series would be approximately equally spaced also.

1951–61 values were used for the *individual* series instead of 1956 values (as, below, for the *composite* series) because the 1956 values for the individual series did not appear to be representative.

[23] The series were centered in the middle of the second month of each quarter.

TABLE 30

Industrials: Yields on Direct Placements, Computed, by Class, Quarterly, 1951–61

Year and Quarter	Class Ic (1)	Class IIc (2)	Class IIIc (3)	Composite[a] (4)
1951				
1	3.27	3.81	4.33	3.80
2	3.26	3.73	4.21	3.73
3	4.17	4.36	5.05	4.53
4	3.29	4.24	5.20	4.24
1952				
1	3.49	4.41	5.34	4.41
2	3.82	4.54	5.26	4.54
3	3.66	4.94	6.27	4.96
4	3.47	4.04	4.57	4.03
1953				
1	3.96	4.50	4.93	4.46
2	4.02	4.84	5.65	4.84
3	4.30	4.55	4.76	4.54
4	4.12	4.65	5.14	4.64
1954				
1	4.26	4.57	4.77	4.53
2	3.56	4.19	4.76	4.17
3	3.70	4.23	4.77	4.23
4	3.55	3.98	4.42	3.98
1955				
1	3.37	4.09	5.10	4.19
2	3.52	3.96	4.37	3.95
3	3.58	4.14	4.72	4.15
4	4.08	4.68	5.29	4.68
1956				
1	4.33	4.46	4.60	4.46
2	4.53	4.57	4.49[b]	4.53
3	4.61	4.77	5.05	4.81
4	4.68	5.33	5.83	5.28

(continued)

TABLE 30 (concluded)

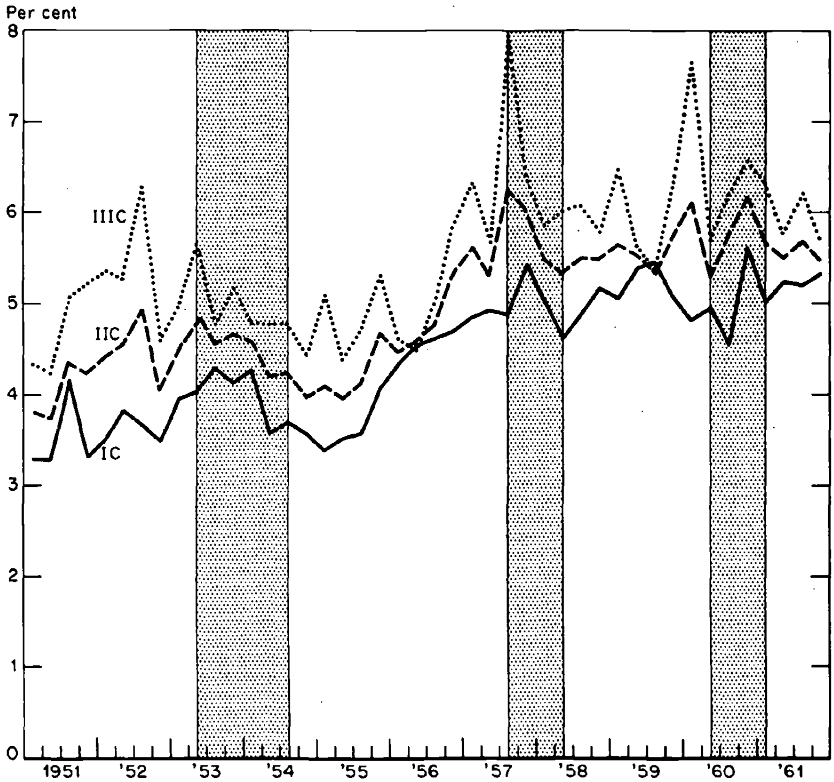
Year and Quarter	Class Ic (1)	Class IIc (2)	Class IIIc (3)	Composite[a] (4)
1957				
1	4.85	5.59	6.30	5.58
2	4.90	5.28	5.64	5.27
3	4.87	6.26	7.94	6.36
4	5.43	6.03	6.40	5.95
1958				
1	5.03	5.48	5.84	5.45
2	4.59	5.31	6.00	5.30
3	4.89	5.50	6.07	5.49
4	5.17	5.47	5.77	5.47
1959				
1	5.03	5.65	6.48	5.72
2	5.38	5.51	5.62	5.50
3	5.44*	5.34	5.36	5.38
4	5.06	5.77	6.32	5.72
1960				
1	4.80	6.11	7.64	6.18
2	4.95	5.28	5.67	5.30
3	4.53	5.73	5.99	5.42
4	5.62	6.17	6.56	6.12
1961				
1	4.99	5.66	6.32	5.66
2	5.23	5.48	5.77	5.49
3	5.19	5.69	6.21	5.70
4	5.33	5.47	5.71	5.50

[a]This series is equal, for each quarter, the sum of Classes Ic, IIc, and IIIc, divided by 3 (arithmetic mean).

[b]Inconsistency.

CHART 10

Industrials: Yields on Direct Placements, Computed, by Class, Quarterly, 1951–61

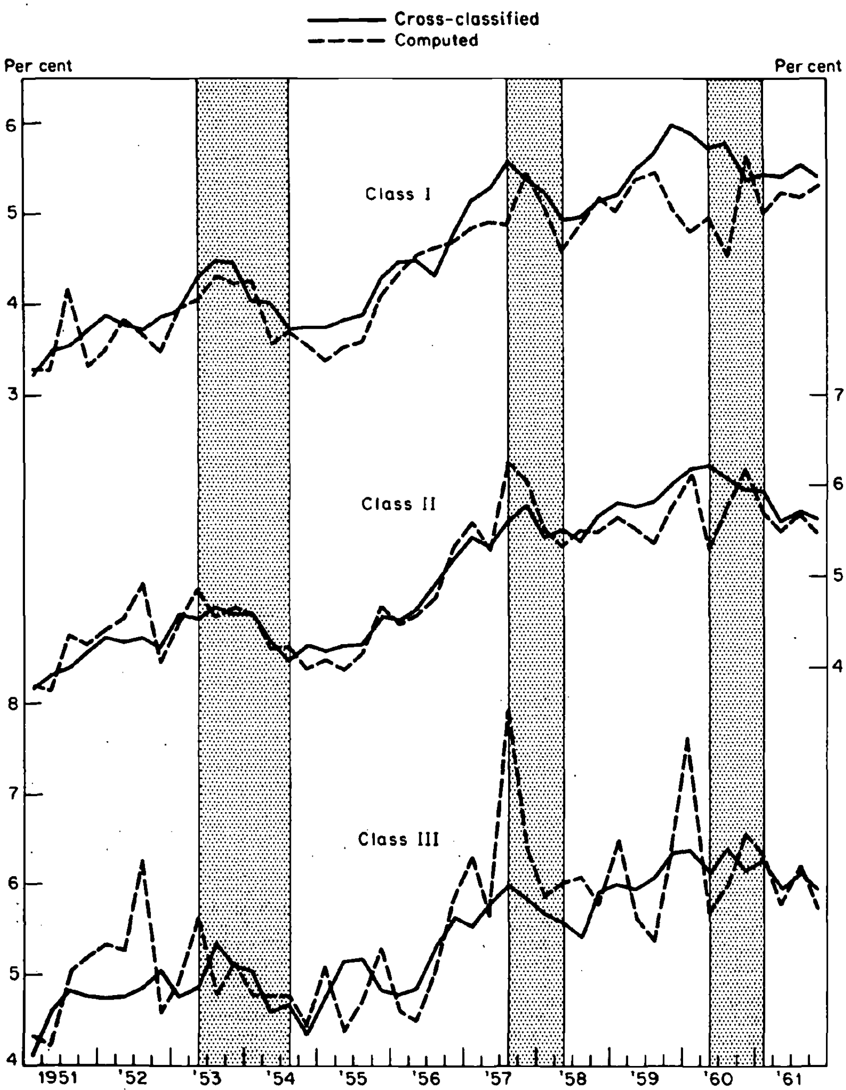


Shaded areas represent business contractions; white areas, expansions.
SOURCE: Table 30.

Chart 11 compares each of the three computed series with its corresponding cross-classified series. On the whole, the general movements of both class I series, except for 1960, and the general movements of both class II series, except for the second and third quarters of 1960, are much the same—although, of course, they need not have been, simply because the cross-classified series do not hold everything constant. There are sharp differences, however, between the two class III series—especially in the first three quarters

CHART 11

Industrials: Cross-Classified Yield Series Compared with Computed Yield Series, by Class, Quarterly, 1951–61



Shaded areas represent business contractions; white areas, expansions.
SOURCE: Tables 28 and 30.

TABLE 31

Industrials: Mean Values Used to Obtain Computed Composite Series

Variable	Unit	Value
X_2	Million dollars	23.2
X_3	Years	9.8
X_{4r}	Million dollars	0.3
X_5	a	1.5
X_6	b	0.6
X_7	Years	6.0
X_8	Million dollars	2.9
X_{12}	Million dollars	2.8
X_{13}	Years	15.3
X_{15}	Dollars of long-term debt per dollar of total capital	.29

[a]See note a, Table 29.

[b]See note b, Table 29.

of 1952, the third quarter of 1957, and the first quarter of 1960. In these quarters, the cross-classified series is lower than it should be simply because observations were missing in the bottom class.[24]

[24] One reader has suggested that the computed series are more erratic than the cross-classified series. This certainly seems to be true at least of the class III series, and this fact may mean that even though residuals were approximately normally distributed, the regression coefficients may not be very reliable at the low end of the distribution in some quarters—primarily because of missing actuals at the low end in those quarters. That is, in some quarters the regression coefficients are being used, in computing the class III series, beyond the range of the observations from which they were derived.

Further, the computed series of course reflect the effect of sampling error in the quarterly regression coefficients. This problem might have been dealt with by computing coefficients (as was done for financials in Chapter 5), for periods of three or four years, holding time constant, by using the yield on Aaa corporates (or a similar series) as a variable, Unhappily, additional runs along these lines were beyond the resources of the present study. They might well be worthwhile, however, as part of any attempt to bring and keep the present series up to date.

TABLE 32

Industrials: Three Composite Yield Series Compared with Each Other and with Average Actual Yields in Sample, Quarterly, 1951–61

Year and Quarter	C_1 (1)	C_2 (2)	C_3 (3)	C_4 (4)
1951				
1	3.70	3.80	3.73	3.69
2	4.01	3.73	3.25	3.83
3	4.13	4.53	4.32	4.28
4	4.20	4.24	3.93	4.05
1952				
1	4.31	4.41	4.17	4.21
2	4.27	4.54	4.36	4.10
3	4.30	4.96	4.27	4.34
4	4.38	4.03	4.00	4.20
1953				
1	4.43	4.46	4.35	4.46
2	4.56	4.84	4.53	4.54
3	4.84	4.54	4.40	4.76
4	4.71	4.64	4.58	4.58
1954				
1	4.54	4.53	4.38	4.47
2	4.28	4.17	4.21	4.28
3	4.16	4.23	4.16	4.21
4	4.11	3.98	4.20	4.04
1955				
1	4.23	4.19	3.86	4.12
2	4.41	3.95	4.20	4.24
3	4.43	4.15	4.29	4.35
4	4.55	4.68	4.43	4.61
1956				
1	4.58	4.46	4.43	4.57
2	4.66	4.53	4.65	4.64
3	4.84	4.81	4.88	4.76
4	5.21	5.28	5.20	5.13

(continued)

TABLE 32 (concluded)

Year and Quarter	C_1 (1)	C_2 (2)	C_3 (3)	C_4 (4)
1957				
1	5.38	5.58	5.21	5.36
2	5.45	5.27	5.22	5.34
3	5.72	6.36	5.65	5.63
4	5.66	5.95	5.79	5.63
1958				
1	5.44	5.45	5.51	5.36
2	5.33	5.30	5.16	5.16
3	5.24	5.49	5.08	5.22
4	5.58	5.47	5.32	5.41
1959				
1	5.67	6.05	5.76	5.59
2	5.73	5.50	5.52	5.62
3	5.84	5.38	5.82	5.74
4	6.11	5.72	5.75	6.05
1960				
1	6.15	6.18	5.33	6.02
2	6.01	5.30	5.88	5.93
3	6.07	5.42	6.03	5.92
4	5.81	6.12	5.82	5.86
1961				
1	5.87	5.66	5.53	5.84
2	5.64	5.49	5.50	5.63
3	5.79	5.70	5.77	5.89
4	5.65	5.50	5.71	5.61

Source: Col. 1, see Table 28, column 4; col. 2, see Table 30, column 4; col. 3, computed at 1956 mean values for ten significant variables given in Table 31; col. 4, arithmetic average over all actual yields on industrials in sample.

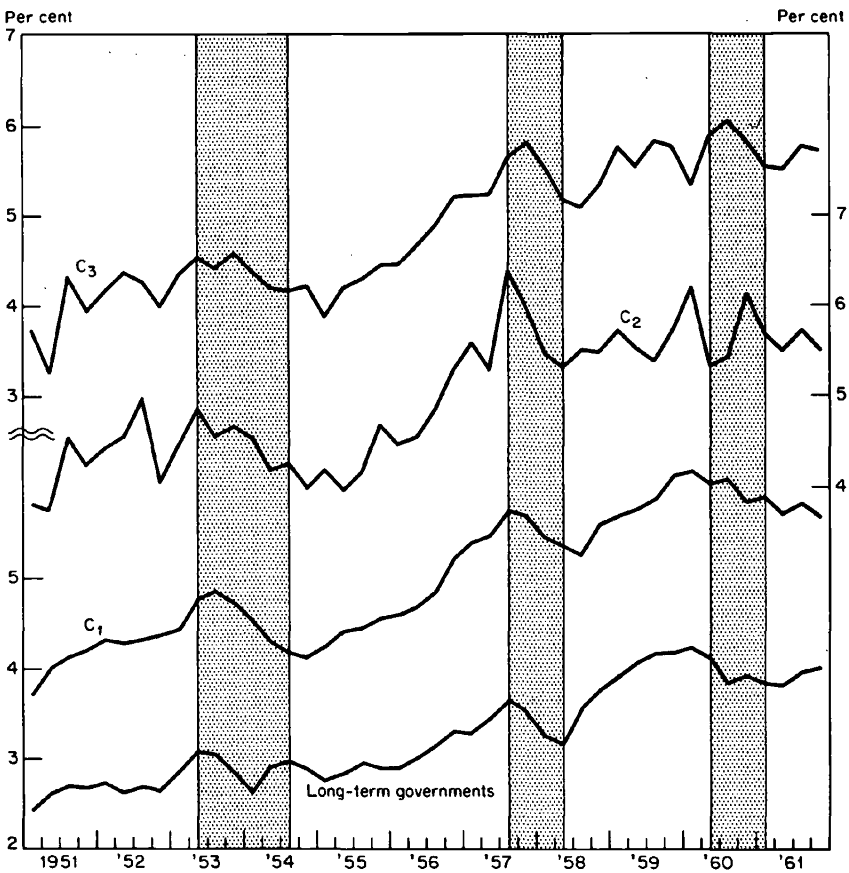
Two additional composite series for industrials were also obtained.

For the first, the three computed series were simply averaged arithmetically (Column 4 of Table 30 and Column 2 of Table 32).

Second, a composite computed series was obtained by using

CHART 12

Industrials: Three Composite Yield Series Compared with Each Other and with Movements of Yields on Long-Term Governments, Quarterly, 1951–61



Shaded areas represent business contractions; white areas, expansions.
SOURCE: Table 32 and *Treasury Bulletin.*

mean values for the significant variables for 1956—the midyear of the period (Table 31). The forty-four regression equations (given by the rerun on the significant variables) were evaluated at these mean values. This series is given in column 3 of Table 32.

All three series are compared in Chart 12. The composite cross-classified series and the computed series based on 1956 mean values are closest together and both conform fairly closely to NBER turning points in business cycles. The arithmetic average of the three computed series (C_2) shows five or six erratic movements.[25]

[25] Whether the "fixed characteristics" series really represent constant quality remains to be seen. This problem will be dealt with in the study of the quality of direct placements, at present under way.