Journal of Financial and Quantitative Analysis
March 1973

THE PREDICTION OF SYSTEMATIC AND SPECIFIC RISK IN COMMON STOCKS

*Barr Rosenberg and Walt McKibben**

Ex ante predictions of the riskiness of returns on common stocks -- or, in more general terms, predictions of the probability distribution of returns -- can be based on fundamental (accounting) data for the firm and also on the previous history of stock prices. In this article, we attempt to combine both sources of information to provide efficient predictions of the probability distribution of returns. We predict two parameters of the distribution of returns for each security in each year: the response to the overall market return (β), and the variance of the part of risk, specific to the security, that is uncorrelated with the market return. A cross section of time series data on returns and accounting variables, taken primarily from the Compustat tape, is used. Several recent developments in statistical methodology are applied.

I. Introduction

Traditionally, accounting variables such as leverage and payout have been used to assess the riskiness of return on common equity in a firm. Recently, the use of historical security prices to predict future risk has been developed in academia and adopted in financial management. In an article in the *Accounting Review*, Beaver, Kettler, and Scholes [3] used accounting variables as instrumental variables in the prediction of future market betas on the basis of past estimated betas and showed that this approach yielded better predictions than the direct use of past betas. Their approach is tantamount to using accounting variables exclusively in forecasting the beta, since the forecast is a linear combination of historical accounting variables, and historical betas are used only to select this linear combination. We believe the correct approach is to use both historical returns and historical accounting variables to predict the distribution of future returns.

Section II presents a familiar parameterization of the joint probability distribution of returns on common stocks. In Section III, a stochastic model of the parameters in the returns distribution is proposed. The statistical methods to be applied to estimate the stochastic model are also explained. In Section IV, the descriptors of accounting variables and historical security prices to be used in the study are

---

*University of California, Berkeley, and University of Western Ontario, respectively. Financial support was provided by National Science Foundation Grant GS-3306 and Associates Workshop for Business Research, University of Western Ontario. We gratefully acknowledge the contributions of Daryl Carlson, Vinay Marathe, and, especially, Michel Houglet to this research. This abridged version of the paper delivered at the WFA Conference reflects several useful suggestions by the discussant, Professor Paul Cootner, to whom we wish to express our thanks.

introduced. Section V gives the results of the statistical analysis. Section VI provides a summary and a discussion of the role of this study in the continuing effort to analyze the probability distribution of investment return.

II. Representation of the Probability Distribution of Security Outcomes

The intent of this study is pragmatic: to quantify the probability distribution of returns to a degree that will be useful in the investment decision. We attempt to model the logarithms of proportional returns computed over calendar years for NYSE and ASE securities, defined as

$$(1) \qquad r_{nt} = \log\left(\frac{P_{nt} + DIV_{nt}}{P_{n,t-1}}\right) = \log\left(P_{nt} + DIV_{nt}\right) - \log P_{n,t-1} \qquad \begin{matrix} n=1,\ldots,N \\ t=1,\ldots,T \end{matrix}$$

where P_t denotes a price as of December 31, DIV denotes total dividends paid in a year, n is the index of the security, and t is the index of the year over which the security is held.

A specification of the probability distribution of the logarithm it returns must be chosen. For each security, a measure of central tendency, which may as well be the mean, is required. It is natural to describe the distribution further in terms of its higher moments, but this approach is valid only if these higher moments exist. There has been some doubt about the existence of finite second moments, but recent studies [13, 14, and 18] suggest strongly that the high kurtosis of empirical frequency distributions of price changes is due, not to infinite variance in the distributions of individual price changes, but rather to fluctuations in the variances of the individual price changes over time. The evidence in [18] suggests that the fluctuations in variance are predictable and that the distribution of the individual logarithms of price relatives, net of these fluctuations, is nearly normal. As is verified later, there are no indications of infinite variance in the sample about to be studied. It is therefore natural to specify the joint probability distribution of the returns on all securities in the sample over a holding period beginning in period t by the vector of *ex ante* means

$$(2) \qquad \underset{\sim}{\xi}_t = \begin{pmatrix} \xi_{1t} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{Nt} \end{pmatrix} \equiv \begin{pmatrix} E[r_{1t}] \\ \cdot \\ \cdot \\ \cdot \\ E[r_{Nt}] \end{pmatrix}$$

and the matrix of variances and covariances,

$$(3) \qquad \underset{\sim}{C}_t = ||c_{mnt}|| \equiv ||E[(r_{mt} - \xi_{mt})(r_{nt} - \xi_{nt})]|| \quad .$$

The problem is simplified by the assumption that the covariances between returns stem from a limited number of observable underlying factors. In this abridged version of the study, only one factor is included: the market factor, M_t, operationalized as $\bar{r}_t$, the average of the logarithmic returns on all securities in our sample.[1] The assumed model for the return on security n in period t is

$$r_{nt} = \alpha + \beta_{nt} M_t + \eta_{nt} \quad . \tag{4}$$

η_{nt} is a random variable with zero mean that is specific to the return on security n in period t, and is uncorrelated with M and with η_{ms} for $m \neq n$ or $s \neq t$. Let σ^2_{nt} be the variance of η_{nt}, or the *specific risk* of security n in period t. In the *ex ante* distribution of future returns on securities, M_t is a random variable with a probability distribution that summarizes anticipated economic and market events. Assume it is distributed with mean f_t and variance F_t. Then,

$$\begin{aligned} \xi_{nt} &\equiv E[r_{nt}] = \alpha + \beta_{nt} f_t \\ C_{nt} &\equiv VAR[r_{nt}] = \beta^2_{nt} F_t + \sigma^2_{nt} \\ C_{mnt} &\equiv COV[r_{mt}, r_{nt}] = \beta_n \beta_m F_t \quad \text{for} \quad m \neq n \quad . \end{aligned} \tag{5}$$

The component of variance stemming from the underlying factor will be called the systematic risk.

III. A Stochastic Model of the Parameters

A. The probability distribution of return is determined by β_{nt} and σ^2_{nt}, $n=1,\ldots,N$, $t=1,\ldots,T$. Each parameter is given the subscripts n and t to indicate that it can assume a different value for every security in every period. Traditionally, it has been assumed that the parameters are unchanging over time, so that $\beta_{nt} = \beta_{n.}$, $\sigma^2_{nt} = \sigma^2_{n.}$ for all t. But, it is more reasonable to expect that these parameters will vary in response to changes in the characteristics of the firm and in the market's perception of the firm. Let $w_1,,w_J$ be a set of descriptors that represent characteristics of the firm, characteristics of the market's response to the firm, and where appropriate, effects that are a mixture of these two. Suppose these descriptors are defined so that the relationship between them and the parameters is approximately linear. Then,

$$\begin{aligned} \beta_{nt} &= \sum_{j=1}^{J} b_j w_{jnt} + \xi_{nt} = \underset{\sim}{b}' \underset{\sim}{w}_{nt} + \xi_{nt} \\ \sigma^2_{nt} &= \sum_{j=1}^{J} v_j w_{jnt} + \nu_{nt} = \underset{\sim}{v}' \underset{\sim}{w}_{nt} + \nu_{nt} \end{aligned} \tag{6}$$

where $\underset{\sim}{b}$, $\underset{\sim}{v}$, and $\underset{\sim}{w}$ are column vectors.

The random variables ξ_{nt} and ν_{nt} introduce those components of β and of σ^2 that cannot be predicted on the basis of the descriptors alone. The meaning of ξ and ν

---

[1] In the earlier version of this study the realized rate of return in the bond market and the proportional change in the index of industrial production were included as additional systematic components of risk. The differential responsiveness of the securities to these components, as well as variations across securities in the α coefficient, were estimated just as the β coefficient is estimated in this abridged version. We hope to reintroduce these other risk components using monthly price data, when this study is repeated. Notice that total returns, rather than excess returns as suggested by the capital asset pricing models, are being modeled.

can be clarified as follows: if there were no descriptors except a constant term, so that (6) read $\beta_{nt} = \bar{\beta} + \xi_{nt}$ and $\sigma^2_{nt} = \overline{\sigma^2} + \nu_{nt}$, then ξ and ν would introduce all of the dispersion in β and σ^2 across the population of firms. As descriptors are added to the equation, more and more of the dispersion in β and σ^2 is explained by the prior predictions based upon the descriptors. Thus, the variances of ξ and ν, equal to the unexplained variance in β and σ^2, respectively, will fall; but the variance cannot fall to zero unless all variation in the risk of firms can be captured by the descriptors, a near impossibility.

The variable ξ_{nt} is assumed to be independent of $w_1,,w_j$, M, and η. Substituting (6) into (4),

$$r_{nt} = \alpha + \left(\sum_j b_j w_{jnt} + \xi_{nt}\right) M_t + \eta_{nt} = \alpha + \sum_j b_j (w_{jnt} M_t) + u_{nt} \tag{7}$$

where $u_{nt} = \xi_{nt} M_t + \eta_{nt}$. The variables of the form $(W_{jnt} M_t)$ are products of the underlying factor and the descriptors and hence are observed. The random variable u_{nt} includes the noise resulting from the specific risk, and also the noise resulting from the unpredictable variation in β. Under the assumptions stated thus far, $E(u_{nt}) = 0$ for all n and t, and u is independent of the observable variables in (7). Thus, least squares applied to (7) will yield unbiased estimates of $\underset{\sim}{b}$. Let T(n) denote the set of time periods for which data are available on security n, and let τ_n be the number of periods in this set. Then,

$$\hat{\underset{\sim}{b}} = \left(\sum_{n=1}^{N} \sum_{T(n)} \underset{\sim}{\chi}_{nt} \underset{\sim}{\chi}'_{nt}\right)^{-1} \left(\sum_{n=1}^{N} \sum_{T(n)} \underset{\sim}{\chi}_{nt} r_{nt}\right) \quad , \tag{8}$$

where $\underset{\sim}{\chi}_{nt} = (M_t w_{1nt} : \dots : M_t w_{Jnt})'$.

B. Assume that ξ and ν are fixed over the history of the firm. In other words, whatever difference there is between the individual security's β_{nt} and σ^2_{nt} and the values implied by the descriptors will be fixed over time. Then, ξ_{nt} and ν_{nt}, t=1,,T can be replaced by ξ_n and ν_n. It is natural to assume that ξ_n and ν_n are independent across different securities, and that the variance of ξ_n and ν_n is the same for all securities. It is also assumed that ν_n and ξ_n are independent. In this case, the stochastic specification becomes

$$\begin{aligned} &E(\nu_n) = 0, \quad E(\nu_n^2) = \theta, \quad E(\xi_n \nu_n) = 0, \quad E(\nu_n \nu_m) = 0, \quad E(\nu_n \xi_m) = 0, \\ &E(\xi_n) = 0, \quad E(\xi_n^2) = \omega, \quad E(\xi_n \xi_m) = 0 \quad \text{for} \quad m \neq n \quad . \end{aligned} \tag{9}$$

Thus, ω is the variance of the component of β which is not associated with the measured characteristics $w_1,\dots,w_j$. Using (9), it is now possible to specify the moments of the random terms in the regression equation (7)

$$E(u_{nt}) = 0, \quad E(u_{nt} u_{ms}) = \left\{ \begin{matrix} \omega M_t^2 + \sigma^2_{nt} \\ \omega M_s M_t \\ 0 \end{matrix} \right\} \text{ for } \left\{ \begin{matrix} m=n, \; s=t \\ m=n, \; s\neq t \\ m\neq n \end{matrix} \right\} . \tag{10}$$

It is possible to estimate ω by several methods. The maximum likelihood estimator developed in [17] is the most efficient and is directly applicable in this case. However, Rao's approach [15], as generalized by Swamy [23], is applicable under minor simplifying assumptions, and it is easier computationally and much clearer heuristically. Consider the vector of regression residuals for the n^{th} firm

$$(11)\qquad \underset{\sim}{r}_n - \hat{\underset{\sim}{r}}_n = ||r_{nt} - \hat{r}_{nt}|| = ||M_t(\underset{\sim}{b}-\hat{\underset{\sim}{b}})'\underset{\sim}{w}_t + M_t\xi_n + \eta_{nt}|| \quad , \quad t\epsilon T(n) \quad .$$

The error in estimating $\underset{\sim}{b}$ will contribute a small proportion of the variance of the residual, on the order of $N_p/(\sum_{n=1}^{N} \tau_n)$, where N_p is the number of estimated coefficients in the array $\underset{\sim}{b}$. In our regression, this proportion will be less than 1/300, so we have an excellent approximation

$$(12)\qquad r_{nt} - \hat{r}_{nt} \approx M_t\xi_n + \eta_{nt} \quad , \quad t\epsilon T(n)$$

or in matrix form,

$$\underset{\sim}{r}_n - \hat{\underset{\sim}{r}}_n \approx \underset{\sim}{Z}_n\xi_n + \underset{\sim}{\eta}_n \quad ,$$

where $\underset{\sim}{Z}_n$ is the column vector of market returns for time periods $t\epsilon T(n)$.

Now, consider the estimator

$$(13)\qquad \hat{\xi}_n \equiv (\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}\underset{\sim}{Z}_n'(\underset{\sim}{r}_n-\hat{\underset{\sim}{r}}_n) \approx \xi_n + (\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}\underset{\sim}{Z}'\underset{\sim}{\eta}_n \quad .$$

The estimator $\hat{\xi}_n$ will be a virtually unbiased estimator of the unpredictable component of beta. With regard to its second moment, from (13)

$$E[\hat{\xi}_n^2] \approx E[\xi_n^2] + (\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}\underset{\sim}{Z}_n'E[\underset{\sim}{\eta}_n\underset{\sim}{\eta}_n']\underset{\sim}{Z}_n(\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}$$

$$(14)\qquad = \omega + (\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}\underset{\sim}{Z}_n'\begin{pmatrix} \sigma_{n1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{n\tau_n}^2 \end{pmatrix}\underset{\sim}{Z}_n(\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1} \quad .$$

The results will be little affected if the variances, σ_{nt}^2, $t\epsilon T(n)$, are replaced by their average value $\sigma_{n.}^2$, especially since the variance for the individual security will be nearly constant over time (the component ν_n is constant and the characteristics w_{jnt} that determine the specific variance will change relatively slowly for most firms). With this simplification,

$$(15)\qquad E[\xi_n^2] \approx \omega + (\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}\underset{\sim}{Z}_n'\sigma_{n.}^2\underset{\sim}{I}\underset{\sim}{Z}(\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1} = \omega + \sigma_{n.}^2(\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1} \equiv \omega + U_n \quad .$$

An approximately unbiased estimator $\hat{\sigma}_{n.}^2$ of $\sigma_{n.}^2$ is provided by summing the squares of the elements of the vector $(\underset{\sim}{I}-\underset{\sim}{Z}_n(\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}\underset{\sim}{Z}_n')(\underset{\sim}{r}_n-\hat{\underset{\sim}{r}}_n)$ (the residuals from regression (13)) and by dividing by the number of degrees of freedom (τ_n-1). Then, $\hat{\xi}_n^2 - \hat{U}_n$, where $\hat{U}_n = \hat{\sigma}_{n.}^2(\underset{\sim}{Z}_n'\underset{\sim}{Z}_n)^{-1}$ is an approximately unbiased estimator of ω. Averaging over all N firms, we have an estimator of ω given by

$$(16)\qquad \hat{\omega} = \sum_{n=1}^{N} (\hat{\xi}_n^2 - \hat{U}_n)/N \quad .$$

This coincides with Swamy's estimate for ω, except for the imperfect correction for estimation errors in $\hat{\underset{\sim}{b}}$. In this application, the method has the added weakness of ignoring the admitted changes in σ_{nt}^2 over time. Nevertheless, with the very large sample size available, the estimator of ω should have satisfactory properties.

There are two benefits to estimating ω. The first is that the estimates of ω and σ_{nt}^2 yield, through (10), a description of the heteroscedasticity in (7); the resulting Aitken's Generalized Least Squares Estimators for $\underset{\sim}{b}$ will be more efficient.

The second advantage is more important. If ω turns out to be nonzero, it follows that the components ξ_n for all firms are fixed over time and are typically nonzero. That is to say, the prediction $\hat{\beta} = \hat{\underset{\sim}{b}}'\underset{\sim}{w}_{nt}$ will be in error, not only because of the estimation error (if any) in $\hat{\underset{\sim}{b}}$, but also because of the error ξ_n which is fixed over time. Since ω tells us the variance of this component, it indicates how much, on the average, an estimate of ξ_n will allow the prediction to be improved. Moreover, an estimate of ω is useful in constructing a minimal-mean-square-error estimator of ξ_n. A good estimator of ξ_n, an approximation to the empirical Bayes estimator, is provided by

$$\tilde{\xi}_n = (\omega^{-1} + \hat{U}_n^{-1})\hat{U}_n^{-1}\hat{\xi}_n \quad . \tag{17}$$

(This estimator is equivalent to the empirical Bayes estimator derived in [17], assuming the variance of estimation errors for $\hat{\underset{\sim}{b}}$ is so small that it can be ignored.) The estimator $\tilde{\xi}_n$ is reduced toward zero relative to $\hat{\xi}_n$ to reflect the prior information that ξ_n is drawn from a population with zero mean and variance ω. Therefore, $\tilde{\xi}_n$ will be biased toward zero, but its (estimated) mean square error $(\hat{\omega}^{-1} + \hat{U}_n^{-1})^{-1}$ will always be smaller than the (estimated) mean square error, $\hat{U}_n$, associated with the use of $\hat{\xi}_n$. It will also be smaller than the (estimated) mean square error, $\hat{\omega}$, associated with the predictor $\hat{\beta} = \hat{\underset{\sim}{b}}'\underset{\sim}{w}_{nt}$ that assumes ξ_n is zero.

To conclude this subsection, the suggested predictor for security n for a period s subsequent to the sample is

$$\hat{\beta}_{ns} = \hat{\underset{\sim}{b}}'\underset{\sim}{w}_{ns} + \tilde{\xi}_n \quad . \tag{18}$$

In the absence of a history of the firm, information to predict ξ_n does not exist, and $\tilde{\xi}_n$ is set to zero.

C. The next step is to estimate the coefficients of the linear function $\sigma_{nt}^2 = \underset{\sim}{v}'\underset{\sim}{w}_{nt} + \nu_n$. Since σ_{nt}^2 is not observed, it is impossible to regress it directly on the descriptors. However, the expected value of the squared residual $(r_{nt}-\hat{r}_{nt})^2$ in the earlier regression is closely related to σ_{nt}^2, and this relationship provides the basis for a linear regression to estimate $\underset{\sim}{v}$. Ignoring the negligible effect of the estimation error $(\hat{\underset{\sim}{b}}-\underset{\sim}{b})$, the residual simplifies as in (12) and, as a result, the expected value of the squared residual is approximately

$$E[(r_{nt}-\hat{r}_{nt})^2] \approx M_t^2\omega + \sigma_{nt}^2 = c_t + \underset{\sim}{v}'\underset{\sim}{w}_{nt} + \nu_n \quad , \quad t\epsilon T(n), \tag{19}$$

where $c_t = M_t^2\omega$. The earlier estimate $\hat{\omega}$ yields the estimates $\hat{c}_t = M_t^2\hat{\omega}$, t=1,...,T. Because of the large number of observations on which $\hat{\omega}$ is based, $\hat{c}_t \approx c_t$, and

$$E[y_{nt}] = \underset{\sim}{v}'\underset{\sim}{w}_{nt} + \nu_n, \text{ where } y_{nt} = (r_{nt}-\hat{r}_{nt})^2 - \hat{c}_t \quad ,$$

or

$$y_{nt} = \underset{\sim}{v}'\underset{\sim}{w}_{nt} + \nu_n + \S_{nt} \quad , \tag{20}$$

where $\S_{nt}$ is the difference between the squared residual and its expected value. Both $\S_{nt}$ and ν_n have mean values of zero. Therefore, although $\S_{nt}$ is jointly dependent with $\underset{\sim}{v}'\underset{\sim}{w}_{nt}$, a regression of y on $w_1, \ldots, w_J$ will yield consistent estimators of the parameters $\underset{\sim}{v}$, as has been explained in [18].

A consistent and asymptotically unbiased estimator of Θ, the variance of the population from which the ν_n are drawn, is then provided by a method analogous to that used to estimate ω. Specifically,

$$\hat{\Theta} = \frac{\sum_{n=1}^{N}\left(\hat{\nu}_n^2 - \hat{\psi}_n\right)}{N} \quad , \text{ where } \hat{\nu}_n = \frac{\sum_{T(n)} (y_{nt}-\hat{y}_{nt})}{\tau_n} \tag{21}$$

$$\hat{\psi}_n = \frac{\sum_{T(n)} (y_{nt}-\hat{y}_{nt}-\hat{\nu}_n)^2}{(\tau_n-1)\tau_n} \quad , \text{ and } \hat{y}_{nt} = \hat{\underset{\sim}{v}}'\underset{\sim}{w}_{nt} \quad .$$

Here $\hat{\nu}_n$ is the estimator of the random variable ν, and $\hat{\psi}_n$ is an estimate of the estimation error variance in this estimator. $\hat{\Theta}$ provides an estimate of the heteroscedasticity in the regression $y_{nt} = \underset{\sim}{v}'\underset{\sim}{w}_{nt}$ and, hence, a more efficient estimator for $\underset{\sim}{v}$ through Aitken's Generalized Least Squares. Also, an empirical Bayes predictor for ν_n is provided by

$$\tilde{\nu}_n = (\hat{\Theta}^{-1}+\hat{\psi}_n^{-1})^{-1}\hat{\psi}_n^{-1}\hat{\nu}_n \quad . \tag{22}$$

This estimator is drawn closer to zero than $\hat{\nu}_n$ and accordingly is biased toward zero, but the (estimated) mean square error, $(\hat{\Theta}^{-1}+\hat{\psi}_n^{-1})^{-1}$, associated with $\hat{\sigma}_{nt}^2 = \underset{\sim}{v}'\underset{\sim}{w}_{nt} + \tilde{\nu}_n$, is again smaller than the (estimated) mean square error, $\hat{\psi}_n$, associated with $\hat{\sigma}_{nt}^2 = \hat{\underset{\sim}{v}}'\underset{\sim}{w}_{nt} + \hat{\nu}_n$, and is also smaller than the (estimated) mean square error, $\hat{\Theta}$, associated with $\hat{\sigma}_{nt}^2 = \hat{\underset{\sim}{v}}'\underset{\sim}{w}_{nt}$.

Thus, the suggested predictor for σ_{nt}^2 for periods subsequent to the sample is

$$\hat{\sigma}_{nt}^2 = \hat{\underset{\sim}{v}}'\underset{\sim}{w}_{nt} + \tilde{\nu}_n \quad . \tag{23}$$

IV. Descriptors of the Individual Securities

The descriptors of the individual securities that are used in the study are explained in detail in [19]. Exact computational formulas for all descriptors are given in the appendix to that paper. The 32 descriptors, and mnemonics for them, are listed below. The notation $(\overline{5})$ indicates that the descriptor is an average over the past five years. The right-hand columns give the sign of the effect that we predicted each descriptor would have on β and σ^2. (A blank indicates that no strong effect was anticipated.)

	Accounting-Based Descriptors	β	σ^2
1.	*Standard deviation of a per-share earnings growth measure* (σE)	+	+
2.	*Accounting beta*, or covariability of earnings with overall corporate earnings (Aβ)	+	
3.	*Latest annual proportional change in per-share earnings* (ΔE)	+	
4.	*Dividend payout ratio* $(\bar{5})$ (PAY)	-	-
5.	*Logarithm of mean total assets* $(\bar{5})$ (SIZE)	-	-
6.	*Standard & Poor's quality rating* (QUAL)	-	-
7.	*Estimated probability of default on fixed payments* (DEF)		
8.	*Liquidity* (the quick ratio) (LIQ)		+
9.	*Absolute magnitude of per-share dividend cuts* $(\bar{5})$ (CUT)		+
10.	*Mean leverage* (senior securities/total assets) $(\bar{5})$, (LEV)		+
11.	*Smoothed operating leverage* (fixed charges/operating income) (LEV*)		
12.	*Standard deviation of per-share operating income growth* $(\bar{5})$ (σO)		
13.	*Growth measure for per-share operating income* $(\bar{5})$ (ΓO)		
14.	*Operating profit margin* (MG)		
15.	*Retained earnings per dollar of total assets* $(\bar{5})$ (REA)		
16.	*Growth measure for total assets* $(\bar{5})$ (ΓA)		
17.	*Growth measure for total net sales* $(\bar{5})$ (ΓTS)		
18.	*Growth measure of per-share earnings available for common* $(\bar{5})$ (ΓE)		+
19.	*Nonsustainable growth estimate* (NSG)		
20.	*Gross plant per dollar of total assets* (PLANT)		-

	Market-Based Descriptors	β	σ^2
21.	*Historical beta*, a regression of stock return on market return over preceding calendar years in the sample, assuming alpha equals zero (Hβ)	+	+
22.	*Standard error of residual risk* (deviations from regression (21)) $(\bar{5})$ (RESRISK)		+
23.	*Marketability*, measured as ratio of annual dollar volume of trading to mean annual dollar volume for all securities (MKT)		
24.	*Negative semi-deviation of returns* $(\bar{5})$ (NEG)		
25.	*Share turnover as a percentage of shares outstanding* (STO)	+	+
26.	*Logarithm of unadjusted share price* (LP)	-	-
27.	*Dummy variable equal to one if stock is listed on the NYSE in latest period; equal to zero, otherwise* (NYSE)		

	Market-Valuation Descriptors	β	σ^2
28.	*Smoothed dividend yield* (YIELD)		
29.	*Earnings/price ratio* (E/P)		
30.	*Book value of common equity per share/price*	-	
31.–32.	*Estimates of misvaluation based on naive growth forecasts* (G) and (G2)		

The Standard & Poor's quality ratings for each year were taken from the previous year's December issue of *Security Owners' Stock Guide* [22]. All other data were taken from two issues of the Compustat Annual Industrial Tape [6], dated September 4, 1969, and November 24, 1971. The 578 securities that were included in the data file from 1950 to 1971 were selected as the sample.

All descriptors used as explanatory variables in a given calendar year use data available before January 1 of that year. Data from annual reports are used only if the fiscal year of the report terminated four or more months prior to January 1, since a delay of up to three months in publication of the report is legally permissible and quite common. Data relating to the market are taken from previous calendar years.

Four years of lagged data are required to compute some of the descriptors. Consequently, the first calendar year return that can be used as a dependent variable for a firm is the sixth year of Compustat data, except that the fifth year can be used with firms having July or August fiscal years. Hence, the regression period is 1954-1970, although data are available for only a few of the firms in 1954.

The first 13 years of returns (1954-1966) were chosen to estimate the relationship between the descriptors and the response parameters. All told, this period provided 5,875 observations on 558 firms. The four years of later data were reserved for testing the predictive ability of the model and were then added into a final regression using all 17 years of data (8,055 observations on 578 firms) to improve the parameter estimates.

V. The Empirical Results

The 32 descriptors were selected, without any prior fitting to the data, on the basis of studies reported in the literature and the authors' intuition. The experimental design described above was fully conceptualized, as were the statistics to be used to evaluate the results, before the preparation of the data was completed. Thus, the results reported here are the outcome of a relatively pure experiment. The first regressions, using several other components of variance in addition to the market component, are reported in [19]. Because of the multicollinearity among these components over the limited number (13 or 17) of time periods observed, the present report was prepared using only the market component. The regressions reported below were conducted in a predesigned sequence. First, all 32 descriptors were included in regressions of form (8) for β and of form (20) for σ^2 over the 13-year period. (The variable *SIZE* had to be omitted from the regressions for β because of a computational error.) Those descriptors insignificant at the 90 percent level were deleted. (This cutoff was selected on the basis of the results.) The regressions after these deletions are reported in the left-hand columns of Table 1. Then the descriptors that capture historical price behavior (*historical beta* and *residual risk*) were deleted and the regressions were repeated. These results are reported in the center columns of Table 1. Finally, the regressions were repeated for the full 17-year period. These are reported in the right-hand columns.

A. With regard to the regressions results for β, note first that the overall constant, α, was nowhere significant. The $\bar{R}^2$ measures the ability of the regression to explain the variance of returns (not the variance of estimated betas). A benchmark is provided by the $\bar{R}^2$ yielded by the assumption that $\beta = 1$ for all securities, which is .360 for the 17-year period. Thus, the additional explanatory power provided by the ability of the descriptors to explain variation in β is 2.2 percent of the variance of logarithms of price relatives. The F-statistic for this added explanatory power is $F(13,8040) = 22.0$, which is highly significant since the 99 percent confidence point is 2.15. This F-statistic is upward biased, since variables were rejected from the earlier regression on the basis of the data, but the F-statistic for the earlier regression, including 31 descriptors, was 9.8 against a 99 percent confidence point of 1.69, also highly significant. Thus, the data base provides strong evidence for the existence of a relationship between β and the descriptors.

TABLE 1

ESTIMATED RELATIONSHIPS BETWEEN
β AND σ^2, RESPECTIVELY, AND THE DESCRIPTORS[a]

	Regressions for β			Regressions for σ^2		
	1954-1966		1954-1970	1954-1966		1954-1970
$\bar{R}^2$	0.364	0.363	0.382	0.061	0.057	0.070
α	-0.0010	-0.0009	-0.0012			
	(-0.280)	(-0.240)	(-0.369)			
Descriptor						
0 Constant	0.3919	0.4503	0.9996	0.0863	0.1015	0.0918
	(2.063)	(2.390)	(6.345)	(4.892)	(5.409)	(6.466)
1 σE	0.0996	0.1043	0.0975	0.0147	0.0159	0.0113
	(5.07)	(5.334)	(5.822)	(7.732)	(8.158)	(6.967)
2 Aβ				-0.0046	-0.0048	-0.0036
				(-2.753)	(-2.824)	(-2.538)
3 ΔE	-0.0543	-0.0585	-0.0615			
	(-2.946)	(-3.193)	(-3.9)			
4 PAY				-0.0085	-0.0100	-0.0076
				(-4.897)	(-5.827)	(-5.089)
5 SIZE				-0.0143	-0.0155	-0.0116
				(-8.326)	(-8.930)	(-7.922)
6 QUAL	0.0542	0.0519	0.0174	-0.0059	-0.0062	-0.0082
	(2.995)	(2.876)	(1.128)	(-3.285)	(-3.213)	(-5.357)
8 LIQ	0.0465	0.0458	0.0231	0.0016	0.0017	0.0005
	(2.969)	(2.925)	(1.675)	(0.928)	(0.995)	(0.317)
9 CUT	-0.0322	-0.0338	-0.0233			
	(-1.844)	(-1.939)	(-1.561)			
10 LEV	0.0383	0.0391	0.0423	0.0074	0.0079	0.0073
	(2.161)	(2.205)	(2.755)	(4.192)	(4.452)	(4.77)
17 ΓTS	0.0270	0.0358	0.0057			
	(1.385)	(1.872)	(0.342)			
18 ΓE	0.0663	0.0719	0.0404			
	(3.174)	(3.463)	(2.341)			
20 PLANT	-0.0543	-0.0542	-0.0517			
	(-3.391)	(-3.385)	(-3.778)			
21 Hβ	0.0377		0.0505			
	(2.315)		(3.562)			
22 RESRISK				0.0079		0.0134
				(4.358)		(8.469)
24 NEG				0.0014	0.0022	-0.0009
				(0.854)	(1.285)	(-0.618)
25 STO	0.0653	0.0707	0.0347	0.0039	0.0052	0.0080
	(3.332)	(3.633)	(2.239)	(2.204)	(2.988)	(5.133)
26 LP	-0.0524	-0.0502	-0.0994			
	(-2.722)	(-2.611)	(-6.04)			
30 B/P	0.0803	0.0708	0.066			
	(4.019)	(3.617)	(3.784)			

[a]All descriptors are standardized so that the regression coefficient gives the change in β or σ^2 from a one-standard-deviation change in the descriptor; t-statistics are given in parentheses.

The estimate of the dispersion of β_n about unity, based upon the 17-year sample, was .0104, implying a standard deviation of .102. However, much of this dispersion is explained by the accounting descriptors. In fact, the estimates of residual ω were *negative* in all the regressions reported in Table 2. A negative estimate for a variance is, of course, nonsensical. However, estimators of variance components such as the one used in this study do admit the possibility of negative estimates, when the true variance is small relative to the standard error of the estimator. When this problem is encountered, the universally recommended procedure is to set the estimate equal to zero. Thus, the estimate of $\hat{\omega} = 0$ was used in subsequent stages of the analysis.

When one examines the regression coefficients giving the estimated response of β to the descriptors, there are numerous significant t-statistics. However, the pattern of signs is not as anticipated: of the 13 coefficients, 4 have the anticipated sign, 3 have the opposite sign, and 6 correspond to effects that were not expected to be strong. In retrospect, some of the arguments leading to the anticipated signs appear questionable. The results suggest that β increases as the specific risk in the anticipated earnings stream increases.

B. Turning to the regressions for σ^2, the dependent variable in each of the three regressions was constructed from the square of the residual in the corresponding regression to explain returns, so that the results are dependent on the set of descriptors that are permitted to affect β. However, the pattern of signs and significance in the regression for σ^2 never changed, despite the use of several alternative models for return, including the naive model that $\beta_n = 1$ for all n. In every case, *of the 8 significant coefficients, 7 have the anticipated sign,* and the other sign is plausible but corresponds to an effect that was not expected to be significant. All of the significant descriptors are traditionally associated with riskiness, with the exception of *accounting beta*, and all of the signs of the estimated effects are as would be expected. The t-statistics are consistently high. The $\bar{R}^2$ is the proportion of the variance of the squared residuals that is explained by the estimated relationship. Since these squared residuals exhibit a great deal of noise about their mean values, which are the variances to be modeled, a small $\bar{R}^2$ is not surprising.

The estimate of Θ, the variance of the systematic random component in σ^2, yielded by the second regression, which excludes the descriptor *RESRISK*, is .000678. (The estimates of Θ for the other two regressions are not meaningful, because the presence of *RESRISK* as a descriptor implies an implausible and badly specified dynamic pattern in the random component of the individual firm's specific risk.) This implies a standard deviation for v_n, persistent random component in specific risk, equal to .0260, or 46 percent of the mean value of σ^2 in this 13-year period, $\overline{\sigma^2} = .0561$. Thus, information on the variance of historical price behavior, by providing an estimate of this component, should give us additional predictive power over and above that provided by the descriptors.

C. The next step is to evaluate the forecasting performance of the estimators over the four-year period, 1967-1970, reserved for this purpose. Only those firms in the initial sample that had at least ten years of acceptable observations over the historical period prior to 1966 were used, so that naive predictors for β and σ^2 based on the history of price behavior would be meaningful. There were 410 of these firms, with 4945 observations in the historical period and 1633 observations in the period reserved for forecast evaluation.

In evaluating a predictor for β, the evaluation criterion is the mean square error in the set of forecasts $\hat{r}_{ns} = \hat{\alpha}_{ns} + \hat{\beta}_{ns}\bar{r}_s$, where s is the time subscript varying over the years 1967 to 1970, and n is the individual firm subscript. The following prediction rules were evaluated:

(i) *VOID* $\hat{\alpha}_{ns} = 0,\ \hat{\beta}_{ns} = 0$;

(ii) *NAIVE* $\hat{\alpha}_n = \alpha^*_n,\ \hat{\beta}_n = \beta^*_n$, where α^*_n and β^*_n are taken from the regression $r_{nt} = \alpha + \beta\bar{r}_t$, fitted over the years prior to 1966;

(iii) *NAIVE*, $\alpha = 0$ $\hat{\alpha} = 0,\ \hat{\beta}_n = \beta^*_n$;

(iv) *NAIVE BAYES* $\hat{\alpha} = 0,\ \hat{\beta}_n = \overline{\beta^*} + (\mathrm{VAR}(\beta^*_n - \beta_n)^{-1} + \mathrm{VAR}(\beta)^{-1})^{-1}\ \mathrm{VAR}(\beta^*_n - \beta_n)^{-1}\ (\beta^*_n - \overline{\beta^*})$

(v) *UNIT BETA* $\hat{\alpha} = 0,\ \hat{\beta}_{ns} = 1$; and

(vi) *PREDICTED BETA* $\hat{\alpha}_{ns} = 0,\ \hat{\beta}_{ns} = \sum_{j=1}^{J} \hat{b}_j w_{nsj}$.

The mean square error about the *VOID* predictor, which is just the mean square logarithmic return, provides a standard of comparison. The mean square error about the *UNIT BETA* predictor, which is the sum of squared deviations of individual logarithmic returns about the market logarithmic return, provides a stricter standard. To improve upon this, it is necessary to predict successfully the variability in beta. The *NAIVE* and *NAIVE*, $\alpha = 0$ predictors are straightforward. The *NAIVE BAYES* procedure was carried out as follows: the variance of the naive estimators of beta across the sample of firms was computed, $\mathrm{VAR}(\beta^*) = .1779$. Then the average estimation error variance for these naive estimators was computed, $\overline{(\mathrm{VAR}(\beta^*_n - \beta_n))} = .1491$. The difference, or .0288, is an unbiased estimate of the variance of the underlying values of β_n, $\mathrm{VAR}(\beta)$. (This procedure is exactly analogous to the methods used previously to estimate ω and θ.) These estimators yield the empirical Bayes adjustment in the *NAIVE BAYES* predictor. This approach was suggested by Vasicek [24]. Finally, the *PREDICTED BETA* forecast is based on the estimated regression coefficient in the second column of Table 2.

The results appear in the first row of Table 2. Notice that only the predictor based on the accounting descriptors does better than the assumption that $\beta = 1$. The information contained in the naive estimate of β, which is based on 13 calendar-year price movements, is worse than useless. The predictor based on the accounting descriptors does improve over the assumption that $\beta = 1$, although the improvement is slight (a 2 percent reduction in the unexplained variance of logarithmic returns).

Next, four predictors of specific risk were evaluated:

(i) *VOID* $\hat{\sigma}^2_{ns} = \text{constant}$;

(ii) *NAIVE* $\hat{\sigma}^2_n$ = residual mean square from regression for α^*_n and β^*_n ;

TABLE 2

MEAN SQUARE FORECAST ERRORS FOR LOGARITHMS OF CALENDAR-YEAR RETURN FOR DIFFERENT FORECAST METHODS AND DIFFERENT FORECAST VARIANCE ADJUSTMENTS
(Years 1966-1970, 410 Firms, 1633 Observations)

Variance Adjustment	Forecast Method					
	(i) Void	(ii) Naive	(iii) Naive $\alpha = 0$	(iv) Naive Bayes	(v) Unit Beta	(vi) Predicted Beta
(i) Void	.1169	.0786	.0809	.0742	.0739	.0725
(ii) Naive	.1176	.0781	.0789	.0737	.0734	.0703
(iii) Predicted	.1066	.0722	.0732	.0679	.0683	.0650
(iv) Empirical Bayes	.1101	.0736	.0749	.0697	.0699	.0666

TABLE 3

KURTOSIS OF FORECAST ERRORS FOR LOGARITHMS OF CALENDAR-YEAR RETURN FOR DIFFERENT FORECAST METHODS AND DIFFERENT FORECAST VARIANCE ADJUSTMENTS
(Years 1966-1970, 410 Firms, 1633 Observations)

Variance Adjustment	Forecast Method					
	(i) Void	(ii) Naive	(iii) Naive $\alpha = 0$	(iv) Naive Bayes	(v) Unit Beta	(vi) Predicted Beta
(i) Void	3.92	4.18	4.06	4.09	4.04	4.08
(ii) Naive	3.81	3.96	4.00	4.02	3.90	3.96
(iii) Predicted	3.25	3.51	3.41	3.44	3.43	3.48
(iv) Empirical Bayes	3.72	3.66	3.59	3.63	3.59	3.65

(iii) *PREDICTED* $\hat{\sigma}^2_{ns} = \sum_{j=1}^{J} \hat{v}_j w_{nsj}$; and

(iv) *EMPIRICAL BAYES* $\hat{\sigma}^2_{ns} = \sum_{j=1}^{J} \hat{v}_j w_{nsj} + \tilde{v}_n$.

The predictions based on the accounting descriptors were computed using the second estimated regression in Table 1. These formulas can yield zero or negative variance predictions, since they are linear functions of the descriptors. To avoid this problem, for methods (ii), (iii), and (iv), the predicted variance was set at a floor of one-fourth of the average variance for all firms in the historical period, or .0140, whenever the prediction fell below this level. This adjustment biases the prediction rule, and it can only be regarded as a temporary substitute for a systematic approach to estimating a nonlinear prediction rule for specific risk. Nevertheless, the adjustment rule was chosen prior to examination of the data, so evaluation of the resulting prediction rule is a true test of the information content of the predictors.

The criteria to be used in evaluating the predictors are rather subtle. Let each set of variance predictions be standardized so that the geometric average, taken over all n and all s, is equal to unity. Let the standardized predictions be denoted by s^2_{ns}, and let the errors in predicting logarithmic return be denoted by $r_{ns} - \hat{r}_{ns} = e_{ns}$. If the predictions of specific risk do contain any information, then the variables e_{ns}/s_{ns}, which are corrected for the predicted standard deviations, should be more nearly uniformly distributed than the variables e_{ns}. This increase in uniformity can be measured in two important ways. First, the weighted sum of squared errors

$\sum_n \sum_s \frac{e^2_{ns}}{s^2_{ns}}$ will be reduced relative to $\sum_n \sum_s e^2_{ns}$. (In fact, if the random variables e_{ns} are normally distributed, this reduction is the basis for the likelihood ratio test of alternative variance predictions.) Second, the kurtosis of the transformed variables e_{ns}/s_{ns} will be smaller than the kurtosis of the untransformed variables e_{ns}.

The weighted sums of squares for the alternative variance predictions appear in the last three rows of Table 2, and the kurtoses of the untransformed and transformed forecast errors are given in Table 3. The results strongly confirm the usefulness of the specific risk predictions based on the accounting descriptors. These predictions perform better than all three alternatives regardless of the rule that is used to forecast return. The naive prediction does somewhat better than the void prediction; the empirical Bayes predictor, which in theory should be optimal, performs substantially better than the naive predictor; and the predictor based upon accounting descriptors alone performs best of all.

VI. Discussion

This study is in an intermediate stage. There are numerous improvements and extensions that we intend to make in the future, the most important of which are the following:

(1) To repeat the study using excess returns rather than total returns, and to add additional systematic risk factors.

(2) To include all available observations on securities listed on the NYSE and ASE. This step will allow the number of securities to be tripled and the number of observations to be doubled.

(3) To use monthly rather than calendar-year price data. This will allow a more sensitive analysis of the probability distribution of returns and will also provide a more informative history of price behavior, thereby increasing the chances that the *historical beta* and *residual risk* will be useful in prediction.

(4) To study the usefulness of the methods in predicting the return distributions of portfolios, as opposed to individual securities.

Despite the shortcomings of the present analysis, the results that have been achieved do suggest some interesting conclusions with regard to the distribution of security prices and the methodology that should be used in studying it.

A. The Security "Beta"

In this study we have suggested that both accounting information and historical price behavior should be used to predict beta, and we have indicated a statistically appropriate procedure for combining these sources of information. We have found that the limited information on historical price behavior provided by the history of calendar-year returns is useless: the prediction of beta based on the history of calendar-year returns alone is worse than the naive prediction that $\beta = 1$, if the criterion of evaluation is mean square error in forecasting returns. Moreover, the additional information provided by the historical beta, over and above that provided by the accounting descriptors, is small. This is shown in two ways: first, by the relatively small t-statistics for historical beta in the first and third regressions in Table 1; second by the zero estimates for ω associated with all regressions in the study.

Another innovation in this study is quite important. By using the regression (8), in which artificial variables equal to the product of the market return and the descriptor appear as the regressors, it is possible to estimate the relationship between beta and the accounting descriptors in a regression with security return as the dependent variable. This approach allows each security beta to vary over time as it should when the accounting descriptors are changing, and it is therefore more efficient than using estimated betas as dependent variables, which can be efficient only if beta is constant over the regression period in which it is estimated. This approach also has the virtue of emphasizing the frequency distribution of returns rather than the frequency distribution of estimated betas. In this same spirit, we have used as our criterion for beta prediction the mean square error in predicting security returns, rather than the mean square error in predicting estimated betas.

B. Specific Risk

The method proposed for the analysis of specific risk seems to have been highly successful. The estimated relationships are significant, and the signs of all significant coefficients correspond with *a priori* intuition. Moreover, the estimated predictors based on accounting descriptors provide an important improvement in forecast performance over all alternative methods that were tried.

It appears that historical price behavior should provide additional information, over and above that provided by the accounting descriptors. The computed *RESIDUAL RISK*, which is a relatively poor measure of historical riskiness, is highly significant in the regressions in Table 2, and the estimates of Θ are positive and significant. Nevertheless, the empirical Bayes predictor, using historical price behavior as well as accounting information, did not forecast as well as the predictor using accounting information alone.

It will be interesting to observe whether monthly price data provide enough additional information about historical price behavior to improve upon the accounting-based prediction for beta and specific risk.

C. The Moments of the Distribution of Security Prices

At the beginning of the paper, the controversy over the existence of higher moments for the distribution of security prices was mentioned. The results of the study at this stage throw some further weight in the direction of finite moments. The kurtosis of the sample of 1633 logarithmic returns from the years 1967-1970 used to evaluate the forecasting performance of the method was only 3.92, and this was reduced to 3.25 when the predictions of specific risk were employed. These values are much closer to 3, the kurtosis of the normal distribution, than those that have usually been associated with security price distributions. From the viewpoint that regards high kurtosis as the consequence of fluctuations in variance rather than of infinite variance in the underlying distribution itself (see [18]), these figures imply (a) that there were relatively small differences in variance across securities and across time periods in the sample and (b) that whatever differences there were are largely explained by the predictions of specific risk.

REFERENCES

[1] Babcock, Guilford C. "The Trend and Stability of Earnings Per Share." Seminar on the Analysis of Security Prices, University of Chicago, November 1970.

[2] Beaver, William. "Financial Ratios as Predictors of Failure." Supplement to *Journal of Accounting Research*, Autumn 1966, pp. 71-111.

[3] Beaver, William; Paul Kettler; and Myron Scholes. "The Association Between Market Determined and Accounting Determined Risk Measures." *The Accounting Review*, vol. 45 (October 1970), pp. 654-682.

[4] Bower, Dorothy H., and Richard S. Bower. "Test of a Stock Valuation Model." *Journal of Finance*, vol. 25 (May 1970), pp. 483-492.

[5] Compustat Annual Industrial Tapes with cutoff dates September 4, 1969, and November 24, 1971.

[6] Compustat *Manual*. Denver, Colo.: Investors Management Sciences, April 2, 1972.

[7] Dun and Bradstreet. *Dun's Review*. Published monthly.

[8] Hickman, W. B. *Corporate Bond Quality and Investor Experience*. Princeton, N. J.: Princeton University Press for the National Bureau of Economic Research, 1958.

[9] Horrigan, James. "The Determination of Long Term Credit Standing with Financial Ratios." Supplement to *Journal of Accounting Research*, Autumn 1966, pp. 44-62.

[10] Malkiel, Burton G. "Equity Yields, Growth, and the Structure of Share Prices." *American Economic Review*, vol. 53 (December 1963), pp. 1004-1030.

[11] McKibben, Walt. "Econometric Forecasting of Common Stock Investment Returns: A New Methodology Using Fundamental Operating Data." *Journal of Finance*, vol. 27 (May 1972), pp. 371-380.

[12] Nerlove, Marc. "Factors Affecting Differences Among Rates of Return on Investments in Individual Common Stocks." *Review of Economics and Statistics*, vol. 50 (August 1968), pp. 312-331.

[13] Praetz, Peter D. "The Distribution of Share Price Changes." *Journal of Business*, vol. 45 (January 1972), pp. 49-55.

[14] Press, S. James. "A Compound Events Model for Security Prices." *Journal of Business*, vol. 40 (July 1967), pp. 317-335.

[15] Rao, C. R. "The Theory of Least Squares When the Parameters Are Stochastic." *Biometrika*, vol. 52 (1965), pp. 447-458.

[16] Rosenberg, Barr. "Varying-Parameter Regression in the Analysis of a Cross Section of Time Series." Working Paper No. IP-165. Berkeley, Calif.: University of California, Institute of Business and Economic Research, 1971.

[17] ______________. "Linear Regression with Randomly Dispersed Parameters." *Biometrika*, April 1973.

[18] ______________. "The Behavior of Random Variables with Nonstationary Variance and the Distribution of Security Prices." Submitted to the *Journal of the American Statistical Association*.

[19] Rosenberg, Barr, and McKibben, Walt. "The Prediction of Systematic and Specific Risk in Common Stocks." Working Paper No. 12, Research Program in Finance. Berkeley, Calif.: Institute of Business and Economic Research, University of California, 1972.

[20] Sharpe, William F. "A Simplified Model for Portfolio Analysis." *Management Science*, vol. 9 (January 1963), pp. 277-293.

[21] Sorter, George H.; Selwyn W. Becker; T. Ross Archibald; and William H. Beaver. "Accounting and Financial Measures as Indicators of Corporate Personality--Some Empirical Findings." *Research in Accounting Management*. American Accounting Association, 1966.

[22] Standard & Poor's Corporation. *Security Owners' Stock Guide*. Published monthly.

[23] Swamy, P. A. V. B. "Efficient Inference in a Random Coefficient Regression Model." *Econometrica*, vol. 38 (March 1970), pp. 311-323.

[24] Vasicek, Oldrich A. "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas." *Journal of Finance* (forthcoming).