Robert C. Blattberg and
*Nicholas J. Gonedes**

A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices

I. INTRODUCTION

There has been a great deal of discussion about the statistical distribution of rates of return on common stocks. At an early stage the prevalent belief was that distributions of rates of return on common stocks were adequately characterized by the normal distribution. This belief seemed to be consistent with the pioneering work of Bachelier.[1] It was also observed, however, that empirical distributions of such returns had more kurtosis (i.e., "fatter tails") than that predicted by the normal distribution. The evidence provided by Mandelbrot and Fama suggested that one could explicitly account for the observed "fat tails" by using the symmetric-stable distribution.[2]

This article considers another family of symmetric distributions that can also account for the observed "fat tails" of returns distribution. This alternative is the Student (or t) distribution. It will be indicated that this alternative model has implications for empirical and theoretical work that are quite different from those of the symmetric-stable model. The descriptive validity of the Student model, relative to that of the symmetric-stable model, will be assessed using actual daily rates of return.

This article is organized as follows: Section II describes the prop-

* Graduate School of Business, University of Chicago. This research has been financed in part by National Science Foundation grant GS-2347. An earlier version of this paper was presented at the Annual Meetings of the Econometric Society, December 1971. We are indebted to Eugene Fama for supplying the daily rates of return used in this study. The comments of Eugene Fama, Arnold Zellner, Merton Miller, and Fischer Black are gratefully acknowledged. A special note of thanks is due to Harry Roberts for his many detailed and insightful remarks on substance and exposition. Results mentioned but not presented here are available from the authors or in Report No. 7242, issued (under the same title) by the Center for Mathematical Studies in Business and Economics, University of Chicago, 1972.

1. L. J. B. A. Bachelier, *Theorie de la speculation* (Paris: Gauthier-Villars, 1900), reprinted in *The Random Character of Stock Market,* ed. P. H. Cootner (Cambridge, Mass.: M.I.T. Press, 1964), pp. 17–78; *Le jeu, la chance, et le hasard* (Paris: E. Flammarion, 1914).

2. B. Mandelbrot, "The Variation of Certain Speculative Prices," *Journal of Business* 36 (October 1963): 394–419; "The Variation of Some Other Speculative Prices," *Journal of Business* 40 (October 1967): 393–413; "New Methods in Statistical Economics," *Journal of Political Economy* 71 (October 1963): 421–40; and E. F. Fama, "The Behavior of Stock Market Prices," *Journal of Business* 38 (January 1965): 34–105.

erties of the stable and Student distributions and then briefly discusses the importance of (1) explicitly recognizing that distributions of returns are "fat tailed" and (2) some important differences that result when the Student rather than the stable model is used to account for the "fat tails." Section III summarizes the derivations of both the Student and stable models using more basic stochastic processes. Details of the derivations appear in Appendix A. Also included in Section III is a discussion of empirical results for other models similar to the ones considered in this paper. Section IV discusses methods for empirically comparing the Student and stable models. Section V describes the specific estimation tools used for our comparative results. Finally, Section VI presents and discusses the empirical results.

II. PROPERTIES OF THE STUDENT AND SYMMETRIC-STABLE DISTRIBUTIONS

We begin this section by defining and stating some properties of the Student and symmetric-stable distributions. Then, we consider several implications of describing daily rates of return on common stocks with the Student and stable models.

A. Definitions and Properties of the Student and Stable Models

The Student distribution.[3]—The Student density function with location parameter m, scale parameter $H > 0$, and degrees of freedom parameter, $\nu > 0$, is:

$$f(x|m, H, \nu) = \frac{\nu^{(1/2)\nu}}{B\left(\frac{1}{2}, \frac{1}{2}\nu\right)} [\nu + H(x - m)^2]^{-1/2(\nu+1)}\sqrt{H},$$

where $B(\cdot, \cdot)$ is the "beta function," that is, $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$, where $\Gamma(\cdot)$ is the "gamma function." The Student distribution has the following properties: (1) $E(\tilde{x}) = m$, for $\nu > 1$ and $\text{Var}(\tilde{x}) = H^{-1}\, \nu/(\nu - 2)$, for $\nu > 2$; (2) in general, all moments of order $r < \nu$ are finite; and (3) when $\nu = 1$, the Student density function is the Cauchy density function. As $\nu \rightarrow \infty$, the Student distribution converges to the normal distribution.

When a Student random variable with $\nu > 2$, $\tilde{x}$, is standardized by dividing $\tilde{x} - E(\tilde{x})$ by $\sqrt{\text{Var}(\tilde{x})}$, then the density function of the resulting standardized variable has the following properties: (1) it has fatter tails than the density function of a conventional standardized normal random variable (i.e., one with mean equal to zero and variance equal to unity), and (2) it is higher than the standard normal density in the

3. See H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory* (Cambridge, Mass.: Harvard University Press, 1961), chap. 7.

neighborhood of their common mean, zero. Note, however, that the above standardization scheme is not the one used in most tables and discussions of the density function of a "standardized" Student random variable. There, the scaling is by $\sqrt{H^{-1}}$ rather than $\sqrt{\text{Var}(\tilde{x})}$. The density function of $[\tilde{x} - E(\tilde{x})]/\sqrt{H^{-1}}$ exhibits the first property stated above, but it is not higher than the standard normal density in the neighborhood of their common mean, zero; in fact, the density function is lower in this neighborhood.

The differences induced by the different scaling procedures in the maximum ordinate of the scaled random variables' density functions are indicated in table 1 for selected values of ν. The effect on the ordinate

Table 1

Maximum Ordinates of Student Density Functions

ν	$f\left[y=0 \mid m=0, \text{Var}(\tilde{y}) = \frac{\nu}{\nu-2}, \nu\right]$	$f[y=0 \mid m=0, \text{Var}(\tilde{y}) = 1, \nu]$
3	.3676	.6367
4	.3750	.5305
5	.3796	.4900
6	.3827	.4687
7	.3850	.4555
8	.3867	.4465
9	.3880	.4399
10	.3891	.4350
11	.3900	.4311
12	.3907	.4279
13	.3914	.4254
14	.3919	.4233
15	.3924	.4215
16	.3928	.4199
17	.3931	.4184
18	.3934	.4172
19	.3937	.4162
20	.3940	.4153
75	.3976	.4030
∞	.3989	.3989

NOTE.—$\nu = \infty$ denotes normal distribution.

when the argument of the density function equals four is indicated in table 2.

The symmetric-stable distribution.[4]—This distribution is defined by its characteristic function because, in general, its density function is not known. The log characteristic function of a symmetric-stable distribution with location parameter δ, scale parameter $c > 0$, and characteristic exponent $\alpha \in (0, 2)$, is: $\ln\phi_{\tilde{x}}(t) = i\,\delta t - |ct|^{\alpha}$, where t is some real number and $i = \sqrt{-1}$.

4. See B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, trans. K. L. Chung (Reading, Mass.: Addison-Wesley, 1954), chap. 7; and S. J. Press, *Applied Multivariate Analysis* (New York: Holt, Rinehart & Winston, 1972), chap. 6.

Table 2
Ordinates of Student Density Functions When the Argument Equals Four

ν	$f\left[4 \mid m=0, \operatorname{Var}(\tilde{y})=\frac{\nu}{\nu-2}, \nu\right]$	$f[4 \mid m=0, \operatorname{Var}(\tilde{y})=1, \nu]$
3	.0092	.0023
4	.0067	.0022
5	.0051	.0019
6	.0041	.0017
7	.0033	.0015
8	.0028	.0013
9	.0023	.0011
10	.0020	.0011
11	.0018	.0009
12	.0016	.0009
13	.0014	.0008
14	.0013	.0007
15	.0012	.0006
16	.0011	.0006
17	.0010	.0006
18	.0009	.0005
19	.0009	.0005
20	.0008	.0005
75	.0003	.0002
∞	.0001	.0001

NOTE.—$\nu = \infty$ denotes normal distribution.

The symmetric-stable distribution has the following properties. (1) This distribution is the Cauchy distribution if $\alpha = 1$. If $\alpha = 2$, it is the normal distribution. (2) It is true that $E(\tilde{x}) = m$, if $\alpha > 1$. (3) In general, all moments of order $r < \alpha$ are finite except when $\alpha = 2$, in which case moments of all orders are finite. (4) If a sum of independent identically distributed random variables has a limiting distribution, then it must be a stable distribution. Thus, the nonnormal stable distributions generalize the classical Central Limit Theorem (CLT) to cases where the second moments of the summed random variables are infinite. (5) A sum of independent stable random variables will be stable with characteristic exponent α^* if each summand is a stable random variable with characteristic exponent α^*.

When a symmetric-stable random variable is standardized by dividing $\tilde{x} - \delta$ by c, then the density function of the resulting standardized variable has the following properties. (1) If $\alpha < 2$, its tails are fatter than the density function of a normal random variable that is standardized in the same way [i.e., by scaling with $c = \sqrt{1/2 \operatorname{Var}(\tilde{x})}$ rather than $\sqrt{\operatorname{Var}(\tilde{x})}$, which is the conventional procedure]; and (2) if $\alpha < 2$, it is higher than the density function of a similarly scaled normal random variable in the neighborhood of their location parameters' common value, zero. These properties are similar to those of the density function of a Student random variable that is scaled by its standard deviation.

The characteristic function given above only applies to symmetric-stable distribution. The modifications needed to define asymmetric stable distributions are indicated in Appendix A, where an asymmetric stable distribution is used to derive the symmetric-stable model for rates of return.

Almost all of this paper focuses on the symmetric-stable model. Unless otherwise indicated, the label "stable model" will be used for the symmetric-stable model.

B. Some Implications of the Student and Stable Models for Empirical and Theoretical Work

For our purposes, the most important parameters of the Student and symmetric-stable models are ν and α, respectively. Thus, throughout this section, we assume that all distributions are standardized so that $\delta = 0$ and $c = 1$ (for the stable model) and $m = 0$ and $H = 1$ (for the Student model).

The empirical evidence to which we referred in Section I suggests that empirical distributions of daily returns are approximately symmetric with "fatter tails" than the normal distribution. The discussion in the preceding subsection indicates that both the Student and stable models can account for such "tail behavior." Thus, there is a similarity between the Student and stable models when comparisons are made with the normal model. And this similarity suggests the potential descriptive validity of each model for daily rates of return.

While there are similarities between the Student and stable models, these two models have some very different implications for empirical and theoretical work. Several extremely important differences are due to differences in the properties associated with sums of random variables generated by each of the models.

Let $\tilde{x}_{it}$, $i = 1, 2, \ldots$, denote the rate of return, under continuous compounding, on security i for day t, $t = 1, 2, \ldots$ (throughout, tilde denotes a random variable). That is, $x_{it} = \ln[P_{it} + D_{it})/P_{it-1}]$, where P_{it} and D_{it} are the (ex-dividend) price and dividend, respectively, on security i for day t. Consider the cross-temporal sum for the ith security

$$\tilde{S}_i^T = \sum_{t=1}^{T} \tilde{x}_{it}.$$

Under continuous compounding, $\tilde{S}_i^T$ is the rate of return on security i over a period of T consecutive days. For example, if $t = 1$ and $t = T$ are, respectively, the first and last days of some month, then $\tilde{S}_i^T$ is the return on security i for that month. Now, suppose that $(\tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{iT})$ is a sequence of independent random variables. Under the Student model (for daily returns) with $\nu > 2$, the distribution of $\tilde{S}_i^T$ converges to a normal distribution as $T \to \infty$. This convergence result is a consequence of the classical CLT.[5] On the other hand, if $(\tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{jT})$ is a sequence from the stable model with $\alpha = \alpha^* < 2$, then the distribution of $\tilde{S}_i^T$ will not converge to a normal distribution because the classical CLT is not applicable to the stable model with $\alpha < 2$. Instead, the distribution of $\tilde{S}_i^T$ will be stable with $\alpha = \alpha^*$ for all T. Thus, the applicability of the Student model to daily returns implies that there is a time period (greater than 1 day) such that rates of return defined over this time period may be described by the normal distribution.

Our own empirical results (see Section VI) for the Student model indicate that estimates of ν for most securities examined are well in excess of 25 for sum sizes of 20, corresponding to monthly returns. Since a Student distribution with $\nu \approx 25$ is almost indistinguishable from a normal distribution, it seems reasonable to approximate distributions of monthly rates of return by normal distributions. Note that if the summed returns, $\tilde{S}_i^T$, on individual securities converge to normality as T increases, then the portfolio return,

$$\sum_j \theta_j \tilde{S}_j^T,$$

also converges to normality as T increases, where θ_j is the proportion of the portfolio invested in security j,

$$\sum_j \theta_j = 1.$$[6]

Whether distributions of summed daily returns converge to normality or to the stable model with $\alpha < 2$ has implications for the appropriateness of estimation tools. If, as a result of convergence, the normal model is applicable to returns defined over a period longer than 1 day (e.g., 1 month), then one may proceed as if these returns were generated from a finite variance process. Consequently, one may use, for example, "least-squares" estimation methods or spectral analysis. If, however, the nonnormal stable model is applicable to these returns, then one should proceed as if these returns were generated from an infinite variance process, for which sample variances (which are always finite) exhibit highly erratic behavior; see, for example, the results presented by Mandelbrot and Fama.[7] Thus, estimation tools relying upon sample second moments may induce misleading (and possibly meaningless) results.[8] In

5. The version of the Central Limit Theorem upon which we are relying assumes independent summands. As indicated in Section VI, rates of return on a given security defined over 1 day (or some longer period) conform reasonably well with this assumption.

6. That is, if the asymptotic distribution of $(\tilde{S}_1^T, \tilde{S}_2^T, \ldots, \tilde{S}_N^T)$ is N-variate normal, then the asymptotic distribution of a linear combination of $\tilde{S}_j^T$, $j = 1, 2, \ldots, N$ is univariate normal. Also note that, if the distribution of $(\tilde{S}_1^T, \tilde{S}_2^T, \ldots, \tilde{S}_N^T)$ is N-variate Student (stable) with $\nu = \nu^*$ $(\alpha = \alpha^*)$, then the distribution of a linear combination of $\tilde{S}_j^T$, $j = 1, 2, \ldots, N$ is univariate Student (stable) with $\nu = \nu^*$ $(\alpha = \alpha^*)$.

7. See n. 2 above.

8. The use of truncated sample second moments may not be a desirable

short, the suitability of measures of dispersion and estimation tools for returns defined over time intervals longer than 1 day depend upon the distributional model applicable to daily rates of returns.

The appropriateness of a distributional model is also important for theories of asset pricing under uncertainty, such as those based upon the mean-variance framework used in Markowitz, Tobin, Lintner, and Sharpe.[9] Suppose that individuals' preferences may be represented by continuous bounded utility functions and that individuals maximize expected utility.[10] Then, use of the mean-variance approach in describing maximum expected-utility strategies is easily justified when utility functions are defined on terminal wealth (or periodic consumption) and when the distributions of investment outcomes (i.e., rates of return) are normal.[11] If, using this approach, one also postulates nondecreasing strictly concave utility functions (i.e., "risk aversion"), then one can demonstrate that the expected utility of an investment varies directly (inversely) with the first (second) moment of that investment's outcome.

Now, for the same types of utility functions, consider the implications of limiting normal and nonnormal stable models for risk-taking behavior. For the stable model, one can show that the expected utility of all outcomes for which $\alpha < 2$ is strictly less than the expected utility of their normal counterparts, that is, outcomes with $\alpha = 2$. More generally, expected utility varies directly with α, $0 < \alpha \leqslant 2$.[12] Thus, if (as a

---

remedy for this problem, relative to the alternative of a more appropriate distributional model; see Mandelbrot (1967) (n. 2 above). This is particularly so if, for example, the appropriate value of α is not exceedingly close to 2.

9. H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments* (New York: John Wiley & Sons, 1959); J. Tobin, "Liquidity Preference as Behavior Towards Risk," *Review of Economic Studies* 26 (February 1958): 65–86; J. Lintner, "Security Prices, Risk, and Maximal Gains from Diversification," *Journal of Finance* 20 (December 1965): 587–615; "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics* 47 (February 1965): 13–37; W. F. Sharpe, "A Simplified Model for Portfolio Analysis," *Management Science* 10 (January 1963): 277–93; "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance* 19 (September 1964): 425–42.

10. The boundedness property is needed to avoid violating utility-theory axioms when dealing with continuous random variables; see the discussion in K. J. Arrow, *Essays in the Theory of Risk Bearing* (Chicago: Markham Publishing Co., 1971), pp. 61–63, which is based upon the analyses of K. Menger, "Das Unsicherheitsmoment in der Wertlehre," *Zeitschrift für Nationalokonomie* 51 (1934): 459–85, reprinted in *Essays in Mathematical Economics in Honor of Oskar Morgenstern,* ed. M. Shubik (Princeton, N.J.: Princeton University Press, 1967). It is particularly important to note that the need for boundedness is not unique to the specific distributional models considered in this report.

11. The mean-variance framework can also be justified when utility functions are assumed to be quadratic in terminal wealth (or consumption). But this assumption has some well-known deficiencies within the context of asset-pricing models; see, for example, Arrow (*Essays,* pp. 96–97), or G. Hanoch and H. Levy, "The Efficiency of Choices Involving Risk," *Review of Economic Studies* 37 (July 1969): 342.

12. Consider two symmetric-stable Paretian distributions with parameters $(\delta_1, c_1, \alpha_1)$ and $(\delta_2, c_2, \alpha_2)$, respectively. In accordance with our assumptions, set

result of convergence) the returns of interest (e.g., monthly rates of return) may be described by the normal model, then using the stable model with $\alpha < 2$ will induce an overstatement (understatement) of risk (expected utility). If this is the case, and if all securities' daily returns cannot be described by the same value of α, one may disprove theoretical propositions on asset pricing simply because those propositions do not provide rankings of assets consistent with optimality conditions. For the problem at hand, this inadequacy is due (once again) to inappropriate distributional assumptions.

Selecting either the Student or stable model (with $\alpha < 2$) for daily returns has implications for empirical work on daily returns as well as returns defined over longer time periods. The Student model allows the use of well-defined density functions. Well-defined density functions for the stable model exist in only two cases: $\alpha = 1$ and $\alpha = 2$. Thus, the likelihood function of the Student model can be expressed in closed form, and maximum-likelihood estimates for all parameters of the model may be obtained. This permits one to use the available statistical theory on

---

$\delta_1 = \delta_2 = 0$ and $c_1 = c_2 = 1$, and suppose $\alpha_1 > \alpha_2$. Let $F(x|\alpha_1)$ and $F(x|\alpha_2)$ denote the distribution functions conditional upon α_1 and α_2, respectively. Finally, let $U(\cdot)$ denote a bounded strictly concave nondecreasing utility function. Consider the change in expected utility given by:

$$E[U(\tilde{x})|\alpha_1] - E[U(\tilde{x})|\alpha_2] = \Delta E[U(\tilde{x})]. \tag{1.1}$$

Upon integrating (1.1) by parts, one gets:

$$\begin{aligned} E[U(\tilde{x})] &= \int_{-\infty}^{\infty} [F(x|\alpha_2) - F(x|\alpha_1)]U'(x)dx, \\ &= \int_{-\infty}^{0} [F(x|\alpha_2) - F(x|\alpha_1)]U'(x)dx \qquad (1.2) \\ &+ \int_{0}^{\infty} [F(x|\alpha_2) - F(x|\alpha_1)]U'(x)dx, \end{aligned}$$

where $U'(x) \equiv \partial U/\partial x$. Since the distributions are symmetric about zero, $F(x|\alpha_1) = 1 - F(-x|\alpha_1)$ and $F(x|\alpha_2) = 1 - F(-x|\alpha_2)$. Thus (1.2) may be rewritten as:

$$\begin{aligned} \Delta E[U(\tilde{x})] = &\int_{-\infty}^{0} [F(x|\alpha_2) - F(x|\alpha_1)]U'(x)dx \qquad (2) \\ &+ \int_{0}^{\infty} \Big\{[1 - F(-x|\alpha_2)] - [1 - F(-x|\alpha_1)]\Big\} U'(x)dx. \end{aligned}$$

Given that $\alpha_1 > \alpha_2$, we have: (1) $F(x|\alpha_2) > F(x|\alpha_1)$ for all $x \in (-\infty, 0)$, (2) $F(0|\alpha_1) = F(0|\alpha_2)$, and (3) $F(x|\alpha_2) < F(x|\alpha_1)$ for $x \in (0, \infty)$. This implies that the first integral in (2) is positive and the second is negative. But, by strict concavity, $U'(-x) > U'(x)$, for $x \in (0, \infty)$. Thus, the first integral in (2) has a value in excess of the absolute value of the second integral. Consequently, $\Delta E[U(\tilde{x})] = E[U(\tilde{x})|\alpha_1] - E[U(\tilde{x})|\alpha_2] > 0$. A more detailed statement of the reasoning used here can be had from Lemma 1 and Theorem 3 of Hanoch and Levy, "Efficiency of Choices Involving Risk." Related material appears in J. Hadar and W. Russell, "Rules for Ordering Uncertain Prospects," *American Economic Review* 59 (March 1969): 25–34.

:imum-likelihood estimators. For the stable model, maximum-likeli-d estimators can be obtained, but the currently available methods for ıg so involve approximations that appear to be quite costly (in terms omputer time).[13] The stable model's parameters may be estimated g other estimators, such as those proposed by Fama and Roll.[14] But e appears to be no statistical theory for the estimator of α proposed ıe latter works.[15] And α is perhaps the most important parameter of ɪtable model. (Note that this entire paragraph says nothing about the :s incurred if one model is used when, in fact, the other is more riptively valid. Thus, our remarks do not fully justify selection of :r distributional model.)

C. *Additional Remarks*

ɪughout, we have emphasized the implications of $\alpha < 2$ and $\nu > 2$ not large) for empirical and theoretical work. We did not emphasize 'reasonableness" or "unreasonableness" of the implied nonexistence eoretical second moments when $\alpha < 2$ for the stable model, or the implied existence of theoretical second moments when $\nu > 2$ for the Student model. It seems to us that emphasizing these factors really "misses the point." The nonnormal stable model's importance does not lie in its not having a finite theoretical second moment, even though this feature often attracts the most attention. The importance of this model does lie in its ability to account for the observed kurtosis of empirical distributions of daily rates of return; it accomplishes this by indicating a distribution whose theoretical second moment is infinite.

One additional comment is in order. In the preceding discussion, we used $\ln[(P_t + D_t)/P_{t-1}]$ to measure a rate of return. An often-used alternative measure is $[(P_t + D_t)/P_{t-1}] - 1$. The former measure was used by Fama, the latter was used by Blume, and both measures were used by Officer.[16] Empirically, these different measurement procedures do not appear to induce important differences in the kinds of estimation results to be considered here. The technical reason for this is that $\ln(1 + r) \approx r$ if r is not very large, such as $|r| \leqslant .15$. For daily returns on common stocks (particularly during the postwar period) this technical result

13. Of particular interest here are the estimation procedures discussed in W. DuMouchel, "Stable Distributions in Statistical Inference" (Ph.D. diss., Department of Statistics, Yale University, 1971).

14. E. F. Fama and R. Roll, "Some Properties of Symmetric Stable Distributions," *Journal of the American Statistical Association* 63 (September 1968): 817–36; "Parameter Estimates for Symmetric Stable Distributions," *Journal of the American Statistical Association* 66 (June 1971): 331–38.

15. The asymptotic distributions of the estimators proposed by Fama and Roll (ibid.) for δ and c are Gaussian because each estimator is a linear combination of estimated fractiles, which have asymptotic normal distributions. Truncated means are used to estimate δ when $\alpha < 2$.

16. Fama, "Behavior of Stock Market Prices"; M. E. Blume, "Portfolio Theory: A Step Toward Its Practical Application," *Journal of Business* 43 (April 1970): 152–73; R. R. Officer, "A Time Series Examination of the Market Factor of the New York Stock Exchange" (Ph.D. diss., University of Chicago, 1971).

appears to have descriptive validity. All estimation results presented this report are based on returns measured using: $[(P_t + D_t)/P_{t-1}] -$ Given the experiences of others, we are confident that the use of $\ln[(P_t$ $D_t)/P_{t-1}]$ would not have altered our inferences.

III. MODELS FOR RATES OF RETURN

The Student and stable models can be derived as continuous mixtui of a normal distribution. In this derivation, the variance of the norr distribution is a random variable. When the reciprocal of the variance follows a gamma-2 distribution, then the unconditional distribution of returns is Student. When the variance follows a strictly positive stable distribution with $\alpha < 1$, then the unconditional distribution is symmetric stable with $\alpha < 2$. Detailed derivations of the Student and stable models are provided in Appendix A. Some brief remarks on our derivations are provided in Subsection A. Several alternative models for rates of return on common stocks are reviewed in Subsection B.

A. Derivation of the Student and Stable Models: Summary

Our derivation of the Student and stable models uses the notion of a subordinated stochastic process. Here, the following definitions are applicable. Let $[\tilde{x}(s);\ s \geqslant 0]$ and $[\tilde{h}(s);\ s \geqslant 0]$ denote stochastic processes. Define another stochastic process $\{\tilde{Z}(s) = \tilde{X}[\tilde{h}(s)];\ s \geqslant 0\}$. The process $[\tilde{Z}(s)]$ is said to be subordinate to the process $[\tilde{X}(s)]$; the process $[\tilde{h}(s)]$ is the directing process.

Let $\tilde{X}(s)$ denote the rate of return on a common stock over a time interval of length s and suppose that $[\tilde{X}(s)]$ is a stationary Gaussian stochastic process, with $\tilde{X}(s)$ independent of $\tilde{X}(s^*)$ for all $s \neq s^*$. We assume that these returns are expressed as deviations about their mean; hence, $E[\tilde{X}(s)] = 0$. The random variable $\tilde{h}(s)$ may be interpreted as the change in the economic environment (or the change in available information) occurring during the time interval s. It is assumed that $\tilde{h}(s)$ is independent of $\tilde{h}(s^*)$, for all $s \neq s^*$. Before the realization of $\tilde{h}(s)$ is available, the rate of return over the interval s may be written as $\tilde{Z}(s) = \tilde{X}[\tilde{h}(s)]$. This formulation allows the distribution of rates of return, $[\tilde{Z}(s)]$, to incorporate changes in the economic environment over intervals of length s. Specifically, this formulation allows the variance of returns over the interval s to depend on the realization of $\tilde{h}(s)$.

All of the above is used in the derivation of both the Student and stable models. The factor that distinguishes one model from the other is the distribution function of $\tilde{h}(s)$, the directing process. If $\tilde{h}(s)$ follows a strictly positive (asymmetric) stable distribution with $\alpha \in (0, 1)$, then $\tilde{Z}(s)$ will follow a symmetric-stable distribution with $\alpha < 2$. On the other hand, if $[\tilde{h}(s)]^{-1}$ follows a gamma-2 distribution (which is also asymmetric and strictly positive), then $\tilde{Z}(s)$ will follow a Student distribution.

B. Other Stock Price Models

A number of other models that use a mixture of a normal distribution and a distribution on the variance of a normal distribution have been proposed for rates of return on stocks. We will review three of these models in this section.

Press proposed the following model.[17] Let $\tilde{Z}(t)$ denote the log of the price of a given security at time t and assume that $\tilde{Z}(t)$ is a process with stationary and independent increments. Specifically,

$$\tilde{Z}(t) = C + \sum_{k=1}^{N(t)} \tilde{Y}_k + \tilde{X}(t),$$

where $Z(0) = C$ is a known constant; $\tilde{Y}_1, \ldots, \tilde{Y}_k, \ldots,$ is a sequence of mutually independent random variables following a normal distribution with mean θ and variance $\sigma_2{}^2$; $\tilde{N}(t)$ is a Poisson counting process with parameter λt, which represents the number of random events occurring at time t; and $[\tilde{X}(t), t \geqslant 0]$ is a Weiner process independent of $\tilde{N}(t)$ and $(\tilde{Y}_1, \tilde{Y}_2, \ldots)$. The $\tilde{X}(t)$ is normally distributed with mean 0 and variance $\sigma_1{}^2 t$.

To estimate his model, Press used monthly prices of 10 stocks in the Dow-Jones Industrials from 1926 to 1960. He split the data into three periods: (*a*) 1926–50, (*b*) 1926–55, and (*c*) 1926–60, for which his sample sizes were at most 300, 360, and 420 observations, respectively.

Press's method of estimating the model's parameters (θ, λ, $\sigma_1{}^2$, $\sigma_2{}^2$) was cumulant matching since, according to Press, "The method of maximum likelihood estimation does not yield explicit estimators in this problem" (p. 322.). His results showed negative signs for some estimated variances, which he set at zero. He suggests that these anomalous results were caused by insufficient sample sizes. Model inadequacy is another possible explanation.

It is apparent from Press's study that a major problem with his model is estimating the parameters. However, since Press's article appeared, the use of a mixture of a normal process and a distribution for the variance of the normal process has been found more frequently in the literature on distributions of returns on stocks.

Clark considered a mixture of a normal distribution and a lognormal distribution for the variance of the normal distribution.[18] Clark labels the unconditional distribution the "lognormal-normal distribution."

17. S. J. Press, "A Compound Events Model for Security Prices," *Journal of Business* 41 (July 1968): 317–35.

18. P. K. Clark, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices" (Discussion Paper no. 1, Center for Economic Research, University of Minnesota, April 1971).

This distribution has no closed-form expression and, thus, must be expressed in integral form. This presents some obvious difficulties for empirical work.

After defining the lognormal-normal distribution, Clark compares this distribution with the stable distribution (and others, though no results for the others are presented in his paper) using changes in the prices of cotton futures for two periods: 1947–50 and 1951–55.

The data series Clark uses makes it difficult to compare models. First, cotton contracts do not have 4-year lives. Thus, Clark had to splice series across contract lives. Consequently, additional noise may have been added to his series. Second, the time period from 1951 to 1955 was preceded by a suspension of trading due to existing price controls. This could have affected the amount of variation in the series at the beginning of the period, just as trading began. Finally, the open interest (similar to "shares outstanding") is not fixed. Volume and price fluctuations may be influenced by changes in open interest.

Another problem with Clark's results is that he uses the Kolmogorov-Smirnov test (K-S test) to test goodness-of-fit of the stable and lognormal-normal models. The critical values of the K-S test used by Clark assume the parameter values are known, but he estimated the parameters. Thus, the critical values he uses are inappropriate. Kendall and Stuart point out:[19] "Nothing is known in general about the behavior of the D_n statistic [D_n denotes the K-S statistic for a sample size of n] when parameters are to be estimated in testing a composite hypothesis of fit. . . . It will clearly not remain distribution free."[20] In light of the criticisms presented above, it seems that Clark's conclusions that the lognormal-normal model fits much better than the stable model must be viewed with caution.

Praetz studied a mixture of a normal distribution and a gamma-2 distribution for the variance of the normal distribution, which results in a Student distribution.[21] This is the same model that we are considering in this paper. However, the models Praetz compares, the data, his methods of estimation, and his method of comparison differ from ours.

The models Praetz compares are the: (1) Student distribution, (2)

19. M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics,* 2d ed. (New York: Hafner Publishing Co., 1968), 3: 458.

20. A distributional hypothesis is a simple hypothesis if the type of distribution and all values of the distribution's parameters are specified in the test of fit. If estimates of the hypothesized type of distribution need to be used in a test of fit, then the distributional hypothesis is a composite hypothesis. Some simulation results on the K-S test of a composite hypothesis of normality are available; see H. W. Lilliefors, "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown," *Journal of the American Statistical Association* 62 (June 1967): 399–402. However, these results do not directly apply to the composite hypotheses in Clark's paper, the lognormal-normal and nonnormal stable hypotheses.

21. P. D. Praetz, "The Distribution of Share Price Changes," *Journal of Business* 45 (January 1972): 49–55.

normal distribution, (3) compound events model (Press's model, described earlier), and (4) stable model. The compound events model is not included in our study.

The data Praetz uses consist of weekly observations on 17 share-price index series from the Sydney stock exchange from 1958 to 1966. We use both daily and weekly observations for each of the 30 stocks in the Dow-Jones Industrials. (It is well to note that such indices often exhibit artificial serial correlation.)

To estimate the parameters in the models, Praetz first standardized each data series by subtracting the sample mean and dividing by the sample standard deviation. He then grouped the standardized data into 26 intervals and estimated the unknown parameters by selecting the parameters values that minimized the χ^2 statistic, Σ [observed-expected]-/ expected. In contrast, we used maximum-likelihood estimation for the Student model and the estimators proposed by Fama and Roll for the stable model.

Praetz's method of comparing the models involves comparing the minimum χ^2 statistics of the models. Our method is to test whether the data converge to normality or are stable and to compute the likelihood ratio for the two models.

In evaluating Praetz's study, we first note that his method of "standardizing" the data is inappropriate. Praetz standardizes by subtracting the sample mean and dividing by the sample standard deviation. The simulation results in Fama and Roll suggest the sample standard deviation is an extremely bad estimator of the scale parameter for stable data.[22] This may explain why the stable distribution does not fit his data well.

Another point is based on the statement in Praetz's paper that "the stable distribution always provides a better fit than the normal" (p. 54). The stable model will never provide a worse "fit" than the normal model because the normal distribution is a special case of the stable Paretian distribution. To estimate α in the stable distribution, Praetz minimizes the χ^2 statistic with respect to α, with α being constrained to lie within the interval [1, 2]. The χ^2 statistic for the normal distribution assumes $\alpha = 2$. Thus, the χ^2 statistic will never be less for the normal distribution than for the stable distribution.

The above discussion casts serious doubts on Praetz's results. In this paper, we hope to show, using more appropriate statistical procedures, that the Student model appears to describe rates of return data better than the stable model.

IV. METHODS FOR MODEL COMPARISON

This section discusses two methods for discriminating between the stable and Student models. They are: (1) calculate the likelihood ratio, and

22. Fama and Roll, "Parameter Estimates," p. 332.

(2) determine whether the distributions of rates of returns are stable under addition.

A. The Likelihood Ratio

The likelihood function of n observations is their joint density function evaluated at the n observations using a given set of parameter values. Suppose the data are generated by one of two distributions but we are uncertain about which one. Each distribution has a different likelihood function. By evaluating each likelihood function at the n observations and the appropriate estimates, we can compute the ratio of the values of these two likelihood functions. The likelihood ratio can then be used to determine the distribution for which the odds are greater. The following proposition assures us that as the sample becomes large, the ratio of the values of the two likelihood functions will indicate which distribution has generated the data.

Let $w_1, \ldots, w_n$ be observations on a sequence of n independent random variables with common density function $f_1(\cdot)$. (In the following discussion we will omit the parameters when denoting the density function.) Then their joint density function will be

$$f_n(w_1, \ldots, w_n) = \prod_{i=1}^{n} f_1(w_i).$$

Consider any other (well-defined) density function $g_i(\cdot)$ and the corresponding joint density function

$$g_n(w_1, \ldots, w_n) = \prod_{i=1}^{n} g_1(w_i).$$

Finally, define the sequence of likelihood ratios

$$\Lambda_n = \frac{g_n(w_1, w_2, \ldots, w_n)}{f_n(w_1, w_2, \ldots, w_n)},$$

$$\Lambda_{n+1} = \frac{g_{n+1}(w_1, w_2, \ldots, w_n, w_{n+1})}{f_{n+1}(w_1, w_2, \ldots, w_n, w_{n+1})},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\Lambda_{n+s} = \frac{g_{n+s}(w_1, w_2, \ldots, w_n, \ldots, w_{n+s})}{f_{n+s}(w_1, w_2, \ldots, w_n, \ldots, w_{n+s})}.$$

Doob proves that the

$$\lim_{k\to\infty} \Lambda_k = 0$$

with probability one, unless $f(\cdot)$ and $g(\cdot)$ are identical, in which case

$$\lim_{k\to\infty} \Lambda_k = 1$$

with probability one.[23]

23. J. L. Doob, *Stochastic Processes* (New York: John Wiley & Sons, 1953), p. 349.

Given suitable regularity conditions for $g_1(\cdot)$ and $f_1(\cdot)$ and large samples, the ratio of likelihood functions evaluated at maximum-likelihood estimates may be given a useful Bayesian interpretation Specifically, assuming equal prior probabilities for models $g_1(\cdot)$ and $f_1(\cdot)$, this ratio represents the asymptotic posterior odds of model $g_1(\cdot)$ relative to $f_1(\cdot)$. Detailed discussions of this issue may be found in Jeffreys, Lindley, and Zellner.[24]

Parameter values are required to calculate the value of the likelihood function for each model. The true parameter values are unknown. A reasonable alternative is to use maximum-likelihood estimates because of their large-sample properties. The samples used in most of this paper are approximately 1,300 observations; consequently, we will assume we are in a "large-sample" situation. For the Student distribution, the maximum-likelihood estimates will be used. For the stable distribution maximum-likelihood estimates are very expensive to compute, particularly for a large number of securities. Fortunately, Fama and Roll have developed a low-cost method for finding parameter estimates for the stable distribution.[25] Their estimates are "fairly good" in large samples (see Fama and Roll), and, thus, for the stable distribution we will use the Fama-Roll estimators.[26]

Using parameter estimates which maximize the likelihood function for the Student distribution but not for the stable distribution causes problems when we are using the likelihood ratio to discriminate between models. Obviously, the likelihood ratio will favor the Student model more frequently than if both likelihood functions were evaluated using maximum-likelihood estimates. In Section VI we discuss simulation results that were used to assess the severity of this problem. For these results, the distribution of the data is known to be either stable or Student. The likelihood functions were evaluated using the same estimating techniques used for actual rates of return. The results for stable data with α equal to 1.65 or 1.80 (see table 11) indicate that for 1,300 observations, not using maximum-likelihood estimates for the parameters of the stable model does not cause incorrect classification (except in one case). For $\alpha = 1.50$ there are a number of incorrect classifications (seven out of 20). However, stock prices are generally found to have estimates of α between 1.65 and 1.80 and, therefore, the results for $\alpha = 1.65$ and $\alpha = 1.80$ are the most relevant for actual rates of return. For smaller sample sizes (260 or less), the results for stable data seem to indicate the likeli-

24. H. Jeffreys, *Theory of Probability,* 3d ed. (Oxford: Clarendon Press, 1961), pp. 193–94; D. V. Lindley, "The Use of Prior Probability Distributions in Statistical Inference and Decisions," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probabilities,* ed. J. Newman (Berkeley: University of California Press, 1961), 1:453–68; A. Zellner, *An Introduction to Bayesian Inference in Econometrics* (New York: John Wiley & Sons, 1971), pp. 31–33.

25. Fama and Roll, "Properties."

26. Ibid., and Fama and Roll, "Parameter Values."

hood ratio frequently incorrectly classifies the data as coming from the Student model. (Incorrect classification does not happen nearly so often for small sample sizes when the data are Student.) For this reason, we will use the likelihood ratio to compare the two distributions for the original observations ("unsummed") so that the sample sizes will be sufficiently large to overcome the problem of not using maximum-likelihood estimates for the parameters of the stable model. A more detailed discussion of these results is provided in Section VI.

B. *Stability*

A theoretical property which differentiates the stable and Student distributions is stability. Stability simply means that if we have n identically and independently distributed random variables, the distribution of their sum will differ from the distribution of each random variable only by location and scale parameters. The Student distribution is not stable, but the symmetric-stable distribution is. For degrees of freedom greater than 2, the distribution of sums of independent identically distributed Student random variables will tend to a normal distribution. Consequently, we can use the property of stability to discriminate between the two models.

For rates of return, we can take sums of daily rates of return. If a daily series follows a stable distribution with a specific characteristic exponent, then so should the summed series. On the other hand, if a daily series follows a Student distribution with fixed degrees of freedom, the summed series will not follow a Student distribution (see Ruben).[27] The distribution for the summed series will have a smaller tail area and be less peaked than the distribution for the daily series. Therefore, if we estimate both the degrees-of-freedom parameter and the characteristic exponent for series with sum sizes of, say, 1, 5, 10, and 20 and observe that these parameter estimates increase as the sum size increases, then we can infer that the daily series is nonstable.

V. ESTIMATION OF THE MODEL'S PARAMETERS

The likelihood function for both statistical models contains three unknown parameters. In order to calculate the value of the likelihood function, we must find estimates of these parameters. Obvious candidates are maximum-likelihood estimates (M.L.E.) since we are interested in the ratio of the likelihoods. We will use M.L.E. for the Student distribution. However, as has been discussed in the previous sections, getting M.L.E. for the stable distributions is both difficult and costly (in terms of computer time). The best method (to date) for finding M.L.E. for the stable distribution is given in Du Mouchel.[28] However, his method

27. H. Ruben, "On the Distribution of the Weighted Difference of Two Independent Student Variables," *Journal of the Royal Statistical Society,* ser. B., 22 (1960): 188–94.

28. DuMouchel, "Stable Distributions."

still takes considerable computer time. As an alternative, we will use estimators devised by Fama and Roll (see n. 14). This section outlines these two estimation procedures. The properties of these estimators are evaluated (in Section VI) using simulation results.

A. M.L.E. for the Student Distribution

The likelihood function for the Student distribution with location m, scale H, degrees of freedom ν, and sample size n is

$$L(m, H, \nu; x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{\nu^{(1/2)\nu}}{B\left(\frac{1}{2}, \frac{1}{2}\nu\right)} \cdot [\nu + H(x_i - m)^2]^{-(1/2)(\nu+1)}\sqrt{H}.$$

The usual method for finding the M.L.E. of m, H, and ν is to differentiate the likelihood function with respect to each parameter, set the resulting three equations equal to zero, and solve for m, H, and ν, the M.L.E. Unfortunately, finding analytical solutions to these three equations for m, H, and ν is extremely difficult. Therefore, we numerically searched for the parameter values which maximize the likelihood function. (Our numerical search routine uses a Fibonacci search on m and ν and a Newton search on H.)

B. Fama-Roll Estimates for the Stable Distribution

The finite-sample properties of several estimators for the characteristic exponent (α), the dispersion parameter (c), and the location parameter (δ) of a symmetric-stable distribution are described in Fama and Roll (see n. 14); these procedures are briefly described in this section.

Estimator for δ.—The estimator for δ is the .75 truncated mean computed by: (1) ordering the data from largest to smallest, (2) truncating the data so that 12.5 percent is discarded from each extreme, (3) calculating the mean for the remaining 75 percent of the ordered sample observations. Fama and Roll (see n. 14) indicate that for $\alpha \approx 1.70$, which is a common estimate of α for rates of return, the .75 truncated mean is a more efficient estimator than the sample mean, the median, the .25, or the .50 truncated means.

Estimator of c.—Let $\hat{x}_f$ and x_f denote the f sample and theoretical fractile, respectively, from a sample of N observations on $\tilde{x}$. For the Cauchy distribution ($\alpha = 1$) the dispersion parameter, c, equals the semi interquartile range, $1/2(x_{.75} - x_{.25})$. For $1 < \alpha \leqslant 2$, c is approximately equal to the semi interquartile range, suggesting that c might be estimated by $1/2(\hat{x}_{.75} - \hat{x}_{.25})$. Fama and Roll recommend an estimator that is similar to the estimator of the semi interquartile range but one that has a smaller asymptotic bias. This estimator, which we will use, is

$$\hat{c} = \frac{1}{2} \frac{\hat{x}_{.72} - \hat{x}_{.28}}{.827}.$$

Fama and Roll (see n. 14) indicate for $1 \leqslant \alpha \leqslant 2$, $\hat{c}$ has an asymptotic bias of less than .4 percent. Additional finite sample properties of $\hat{c}$ are given in Fama and Roll (see n. 14).

Estimator of α.—Consider a symmetric-stable random variable, $\tilde{x}$, with location δ, scale c, and characteristic exponent α. Let $\tilde{z} = (\tilde{x} - \delta)/c$. Then $\tilde{z}$ will be a standardized symmetric-stable random variable with location 0, scale 1, and characteristic exponent α.

The tail areas for the distribution of $\tilde{z}$ are sensitive to changes in α. We will exploit this to find an estimate for α by studying the f fractile of the distribution of $\tilde{z}$, z_f, with f chosen so that it is in the tail of the distribution. Tables of z_f for different values of α are available. By using estimates of z_f for different values of α and comparing them to actual values of z_f, we can find an estimate of α.

The estimate we will use for z_f is:

$$\hat{z}_f = \frac{\hat{x}_f - \hat{x}_{1-f}}{2\hat{c}},$$

where $\hat{c}$ is our estimate of c and $\hat{x}_f$ and $\hat{x}_{1-f}$ are sample fractiles computed from our observed data.

Our estimating procedure for α is then: (1) compute $\hat{z}_f$ from our rates-of-return series, (2) search for the value of α in the tables of the standardized symmetric-stable distributions which makes $\hat{z}_f$ closest to z_f. This value of α will be our estimate. Again, Fama and Roll (see n. 14) provide sampling results for the properties of α. Their results suggest that a suitable value of f for $\alpha \approx 1.7$, which is applicable to rates of return, is $f = .97$.

VI. ESTIMATION RESULTS

The results discussed in this section consist of: (1) results from a Monte Carlo simulation study based upon simulated observations from the stable and Student distributions, and (2) results based upon actual daily rates of return. The simulation study was done to gain some insight into the finite-sample properties of the estimation and model-comparison methods that we used for the actual data.

A. The Actual Data

The actual data used are consecutive daily rates of return for each of the 30 securities in the Dow-Jones Industrial Average over the period 1957–62. The daily rates of return were computed by deducting unity from daily price relatives adjusted for dividends and capitalization changes (e.g., stock splits). The time periods of the observations are not identical for all 30 securities. Typically, the time period is from about the end of 1957 to September 26, 1962. The actual date of the first observation for a security varies from January 1956 to April 1958; the date of the last observation is the same for all 30 securities. These data are the same as those used by Fama (see n. 2).

Recent work on security returns has emphasized forms of cross-sectional dependence that results in cross-sectional correlation among securities' returns. We did not adjust the data in our sample for such correlation because it appears that, given our objectives, such an adjustment is not necessary.[29]

We will not present evidence regarding the serial independence of daily returns. Results on this topic, for the 30 securities in our sample, are provided by Fama (see n. 2). In general, it appears that daily returns, and returns defined over longer time intervals, do not strongly violate the assumption of serial independence. The most frequently observed inconsistency is the tendency for extreme values of daily returns (of unpredictable sign) to succeed extreme values of daily returns.

B. *The Design of the Monte Carlo Study*

The design of our Monte Carlo Study is as follows: (1) 26,000 random numbers uniformly distributed over the interval (0, 1) were generated from a uniform random-number generator (see Appendix B for the properties of our uniform random numbers). (2) The random numbers from (1) were used to generate 26,000 random numbers from each of several stable distributions and Student distributions with prespecified parameter values. Each set of 26,000 values (in their original order of appearance) from the stable and Student distributions was partitioned into 20 non-overlapping samples of 1,300 observations. (3) The estimation proce-

29. The usual method of removing cross-sectional correlation is to base all estimation results on the residuals of the "market-model," $\tilde{R}_{it} = \alpha_i + \beta_i \tilde{I}_t + \tilde{\epsilon}_{it}$, where $\tilde{R}_{it}$ is the ith security's rate of return for period t, $\tilde{I}_t$ is the rate of return on the market index for period t, and $\tilde{\epsilon}_{it}$ is a serially independent disturbance term, with $E(\tilde{\epsilon}_{it}) = 0$. For the ith and jth securities $\text{cov}(\tilde{R}_{it} \cdot R_{jt}) \neq 0$ because both securities contain the same factor, $\tilde{I}_t$. Once this is removed, it is usually assumed that the residuals are uncorrelated, namely, $E(\tilde{\epsilon}_{it} \cdot \tilde{\epsilon}_{jt}) = 0$, for all i and j. The empirical results of Officer, "Time Series Examination," for the stable model showed that estimates of α based directly upon daily rates of return and estimates of α based upon the market model's residuals for these daily returns were essentially the same. The primary reason for this result appears to be that the cross-sectional correlation among daily return is close to zero. It is known, however, that the cross-sectional correlation for returns increases as the time interval over which returns are defined increases. When considering sums of 20 daily returns, Officer's estimates of α based directly upon sums of daily returns and those based upon sums of the market model's residuals for these returns were slightly different. However, this result showed that the direction in which $\hat{\alpha}$ moves as a function of the sum sizes was the same for sums of residuals and sums of daily returns. This suggests that our inferences about convergence to normality based upon estimates of α for increasing sum sizes will be unaffected by cross-sectional correlation. B. F. King, in "Market and Industry Factors in Stock Price Behavior," *Journal of Business* 39 (January 1966): 139–90, shows that there is also an industry factor along with the market factor. Thus, if the ith and jth security are in the same industry, then $\text{cov}(\tilde{\epsilon}_{it}, \tilde{\epsilon}_{jt}) \neq 0$. However, King found that the industry factor was small and, therefore, the $\text{cov}(\tilde{\epsilon}_{it}, \tilde{\epsilon}_{jt})$ will not be large.

dures used for the actual price data were used on the random numbers from the stable and Student distributions for parameter estimates. It should be noted that the estimating procedure for the Student model was applied to the simulated stable data and vice versa. Thus, we have sampling distributions for estimators of the stable (Student) model's parameters based upon data actually generated from Student (stable) distributions as well as data actually generated from stable (Student) distributions. The purpose of applying the stable model to Student data, and vice versa, is to offer another means of comparing the models.

For each distributional model, part (2) requires prespecified parameter values. We decided to use values of the location and dispersion parameters that are consistent with values of estimated parameters observed for daily rates of return from a pilot study. Since estimates of ν (for the Student model) and α (for the stable model) are particularly important for our objectives, we decided to prespecify sets of parameter values for the Student and stable models that differed only with respect to their values of ν and α, respectively; that is, neither the values of location parameters (m and δ) nor those of dispersion parameters (H and c) were varied. The prespecified values of the location and dispersion parameters are: $\delta = m = .0003$, $c = .00725$, and $H = 6{,}000$. For the stable model, three values of α were specified: 1.5, 1.65, and 1.8. For the Student model, three values of ν were specified: 3, 5, and 8.

According to the procedure described in (2), the ith observation from each of the six distributions is associated with the same value of the uniform number, $\tilde{u}(0, 1)$. Thus, the simulated observations for different distributions are not independent. This lack of independence is consistent with our objectives, namely, evaluating results for (1) different values of ν and α for the Student and stable models, respectively, and (2) different distributional models, holding other things constant [such as the underlying values of $\tilde{u}(0, 1)$].

C. *Discussion of the Simulation Results*

In this section, the Monte Carlo simulation results are discussed and then used to develop guidelines for interpreting the results for actual rates of return.

Results for estimating ν and α.—Our results for estimating ν are given in tables 3, 4, and 6. For a sample size of 1,300 and 20 replications, we see the following (table 3): (1) when $\nu = 3$ and $\nu = 8$, the estimates of ν deviate slightly from the true value, though the estimates are within 2 S.E. of the true value; (2) the sample standard deviation, and the mean squared relative deviation increase as ν increases.[30] A

30. We used the sample standard deviation (among other things) to measure the precision of our estimates of ν. We could just as easily have used the standard error of ν by dividing all the sample standard deviations by $\sqrt{20}$. The standard deviation and standard error are directly proportional and so either can be used without changing the results.

Table 3
Estimates of ν for Simulated Student Data
(Sample Size = 1,300)

Replication Number	True Value of ν		
	$\nu = 3$	$\nu = 5$	$\nu = 8$
1	2.80	4.84	7.66
2	3.06	5.61	10.21
3	3.05	5.61	10.71
4	3.31	6.62	13.01
5	3.05	5.34	9.43
6	2.54	4.33	6.37
7	3.30	6.37	12.75
8	2.80	4.85	7.90
9	2.80	4.84	7.65
10	3.05	5.09	7.90
11	3.05	5.34	8.93
12	3.31	6.12	10.97
13	3.06	5.35	8.66
14	2.53	3.83	5.60
15	2.53	3.83	5.60
16	2.80	4.84	7.66
17	2.79	4.58	7.14
18	2.29	3.82	5.35
19	2.79	4.33	6.38
20	2.54	4.33	6.38
Mean	2.87	4.99	8.31
Standard deviation	.28	.80	2.21
$\sqrt{\text{Mean squared relative deviation}}$*	.10	.16	.27

* $\sqrt{(1/N)\Sigma_i[(\hat{\nu}_i - \bar{\hat{\nu}}_i)/\bar{\hat{\nu}}]^2}$, where $\bar{\hat{\nu}}$ is the average estimate.

possible reason for this decrease in precision as ν increases is that the rate of change in the density function decreases as the degrees-of-freedom parameter increases. For example, a change from 2 to 3 df changes the shape of the density function quite perceptibly, whereas a change from 49 to 50 df results in an almost insignificant change in the shape of the density function. Thus, the likelihood function is much flatter for large values of ν than for small values, resulting in much less precise estimates of ν when the actual degrees-of-freedom parameter is large.

Table 5 shows how the properties of estimates of ν change as the sample size decreases. For "small" samples, $n = 65$, the estimates of ν for all three values of the degrees-of-freedom parameter appear to be extremely inaccurate. Our estimates for $\nu = 5$ and $\nu = 8$ also have large standard deviations and mean-squared relative deviations for sample sizes of 260. When using maximum-likelihood estimation for ν, very large sample sizes are needed. Even sample sizes as large as 260 appear to be too small to obtain accurate estimates. A reason for the need for large sample sizes is that the tail area offers much of the information for detecting differences in ν. For sample sizes of 260, there are still a small number of observations in the tails. Also, as ν increases it becomes more

Table 4
Estimates of ν for Simulated Student Data
(Sample Size = 260)

Replication Number	True Value of ν		
	$\nu = 3$	$\nu = 5$	$\nu = 8$
1	4.08	7.65	14.03
2	2.54	4.34	6.89
3	2.79	4.84	8.66
4	4.58	11.23	55.14
5	5.34	14.03	60.00*
6	4.08	9.95	35.48
7	3.06	5.85	10.98
8	4.34	20.42	60.00*
9	3.31	5.86	9.95
10	2.54	4.08	5.86
11	3.56	6.62	13.27
12	3.31	7.65	23.99
13	2.28	3.57	5.10
14	2.28	3.56	4.85
15	2.79	4.85	7.90
16	2.03	3.05	3.83
17	2.28	3.31	4.59
18	3.05	5.34	10.20
19	2.53	4.07	5.87
20	3.06	5.34	9.43
Mean	3.19	6.78	18.00
Standard deviation	.87	4.18	18.52
$\sqrt{\text{Mean squared relative deviation}}$	.27	.62	1.04

* Upper bound established in estimation procedure used for simulated data.

difficult to notice changes in the tails of the density. Thus, even larger samples are needed.

Our results for estimating α are given in table 6. For sample sizes of 1,300 and 20 replications we observe a slight downward bias in our estimates of α. Similar results were found in Fama and Roll (see n. 14). For sample sizes of 260, we can study table 7 because sums of independent stable random variables each with characteristic α are also stable with characteristic exponent α. The results in table 7 indicate that for smaller samples, the downward bias is more pronounced. For a more extensive study of estimating α, see Fama and Roll (see n. 14).

Results for sums of stable and student random variables.—As indicated earlier, a method of discriminating between the stable and Student models is to test for convergence to normality. To decide whether the distribution of rates of return is converging, we must know how quickly sums of Student random variables converge to normality. To determine the rate of convergence, we generated Student random variables with known degrees of freedom and took sums of sizes 5, 10, and 20. We then computed an estimate of the degrees-of-freedom parameter for each sum size. Table 8 gives the results for 20 replications and a sum size of 5. We did not use the results for smaller sum sizes because the estimates of

Table 5
Comparisons of Estimates of ν for Simulated Student Data for Different Sample Sizes*

Sample Size	True Value of ν: $\nu = 3$	$\nu = 5$	$\nu = 8$
1,300:			
Mean	2.87	5.00	8.31
Standard deviation	.28	.80	2.21
$\sqrt{\text{Mean squared relative deviation}}$	.10	.16	.27
260:			
Mean	3.19	6.78	18.00
Standard deviation	.87	4.19	18.52
$\sqrt{\text{Mean squared relative deviation}}$	.27	.62	1.04
65:			
Mean	9.74	17.28	26.04
Standard deviation	17.2	22.08	24.45
$\sqrt{\text{Mean squared relative deviation}}$	1.77	1.28	.94

* The maximum-likelihood estimation program contained an upper bound of 60 on estimates of ν for the simulated data. The number of truncated estimates underlying the results summarized in this table is as follows:

Sample Size	True Value of ν	Number of Truncated Estimates of ν
260	8	3
65	5	2
65	8	6

ν appear to be extremely erratic; these results are available from the authors.

Before giving the simulation results, we should note that sums of independent identically distributed Student random variables do not follow a Student distribution. We shall use the estimates of ν only as a descriptive measure of convergence since ν can no longer be interpreted as a Student parameter.

For Student data with $\nu = 3$, and sums of size five, the average estimate of ν increases from 2.87 to 4.70. Unfortunately, the standard deviation and mean-squared relative deviation also increase quite noticeably (.281 to 1.52 and .098 to .322, respectively).[31] Faster convergence seems to result for $\nu = 5$ and $\nu = 8$ with the average value of ν changing from 5.00 to 16.87 and 8.31 to 34.36, respectively. Again, our estimates of ν are extremely inaccurate. From the results just given, we see that because of sampling error, it is difficult to determine the exact rate of convergence to normality.

Our alternative to the Student distribution—the stable distribution—

31. Our estimates for ν contain a number of values at 60.00. This is due to a truncating in our estimating procedure. The effect of this truncation is to bias our estimates of the standard deviation downward. For the actual daily returns, the upper bound on estimates of ν was usually set at 89.984 (see the nn. to the tables).

Table 6
Estimates of α for Simulated Stable Data
(Sample Size = 1,300)

Replication Number	True Values of α		
	$\alpha = 1.50$	$\alpha = 1.65$	$\alpha = 1.80$
1	1.50	1.65	1.80
2	1.51	1.67	1.85
3	1.46	1.62	1.76
4	1.52	1.67	1.83
5	1.52	1.60	1.88
6	1.40	1.63	1.76
7	1.50	1.66	1.83
8	1.52	1.67	1.84
9	1.54	1.70	1.88
10	1.55	1.72	1.80
11	1.48	1.63	1.76
12	1.61	1.78	1.80
13	1.51	1.66	1.80
14	1.43	1.57	1.70
15	1.45	1.60	1.75
16	1.52	1.67	1.83
17	1.40	1.65	1.70
18	1.36	1.51	1.64
19	1.46	1.60	1.74
20	1.46	1.61	1.76
Mean	1.48	1.64	1.78
Standard deviation	.057	.055	.060
$\sqrt{\text{Mean squared relative deviation}}$	.04	.03	.03

implies lack of convergence to normality. To develop guidelines for applying the stable model to actual data, we generated stable numbers with known characteristic exponent. We then summed these random numbers and estimated the characteristic exponent for the sums. Table 8 gives the results for 20 replications on sums of size five. (Results for sums of sizes 10 and 20 are available from the authors.)

For stable data, our average estimate of α decreases when we take sum sizes of five (1.485 to 1.46, 1.64 to 1.62, and 1.79 to 1.73). The standard deviations do not increase as much as they did for the Student data. The decrease in the estimate of α is not due to sampling error, but it is probably a downward bias due to the smaller sample sizes used to estimate α for sum sizes of five. The downward bias in estimates of α was first found in the extensive simulations of Fama and Roll (see n. 14). For actual data and approximately 1,300 observations, if the estimate of α does not decrease when going from sum sizes of one to sum sizes of five, this suggests convergence.

Table 9 gives a summary comparison of the results from tables 7 and 8 as well as a cross comparison of estimates of ν for stable data and α for Student data. The results indicate that if we estimate α for the Student data using the Fama-Roll estimating technique, we observe an increase in the estimate of α as the sum sizes go from one to five. For

Table 7
Estimates of α for Sums of Size Five from Simulated Stable Data
(Sample Size = 260 per Replication)

Replication Number	True Value of α		
	$\alpha = 1.50$	$\alpha = 1.65$	$\alpha = 1.80$
1	1.45	1.58	1.68
2	1.46	1.65	1.80
3	1.50	1.65	1.72
4	1.60	1.68	1.77
5	1.53	1.60	1.75
6	1.40	1.57	1.65
7	1.60	1.77	1.77
8	1.50	1.67	1.80
9	1.45	1.57	1.78
10	1.50	1.66	1.76
11	1.53	1.65	1.75
12	1.52	1.66	1.97
13	1.57	1.74	1.80
14	1.52	1.65	1.65
15	1.28	1.43	1.50
16	1.30	1.54	1.65
17	1.40	1.63	1.74
18	1.36	1.40	1.64
19	1.33	1.51	1.61
20	1.46	1.70	1.80
Mean	1.46	1.62	1.73
Standard deviation	.09	.09	.09
$\sqrt{\text{Mean squared relative deviation}}$	.06	.06	.06

stable data the estimates of α decrease, on average, as we go from sum sizes of one to five. On the other hand, the estimates of ν increase for both types of data as we go from sum sizes of one to five. This suggests that a more conservative approach in examining convergence to normality is to use estimates of α to indicate convergence. Thus, when studying actual rates of return, we will rely more heavily upon estimates of α to indicate convergence than estimates of ν.

The likelihood ratio.—In this section, we present results for the (natural) log-likelihood ratios. The log-likelihood ratios allow one to determine which model has the higher likelihood, given the data. For large samples these log-likelihood ratios provide measures of relative degrees of belief in each distributional model (see Section IV-A).

The results for the log-likelihood ratios are summarized in tables 10 and 11. The number computed is the log of: the value of the likelihood function for the Student distribution divided by the value of the likelihood function for the stable distribution.[32] The results are only given for sum sizes of one. For larger sum sizes, the fact that the number of times the value was in the wrong direction (i.e., was greater than zero when the

32. The approximation for the stable density function is given in H. Bergstrom, "On Some Expansions of Stable Distribution Functions," *Arkiv for Matematik* 2, no. 18 (1952): 375–78.

Table 8

Estimates of ν for Sums of Size Five from Simulated Student Data (Sample Size = 260 per Replication)

Replication Number	True Value of ν		
	$\nu = 3$	$\nu = 5$	$\nu = 8$
1	3.82	7.14	10.21
2	4.85	14.80	60.00*
3	5.61	15.56	47.23
4	5.09	7.90	9.94
5	5.35	9.43	12.24
6	3.56	7.90	26.81
7	5.35	14.55	41.36
8	6.12	34.72	60.00*
9	4.57	15.56	60.00*
10	3.06	5.35	8.42
11	6.38	60.00*	60.00*
12	9.44	60.00*	60.00*
13	5.86	20.67	60.00*
14	4.08	10.71	25.02
15	2.79	5.10	8.16
16	4.08	9.19	17.36
17	4.08	15.83	60.00*
18	3.06	6.11	12.51
19	3.31	8.92	27.31
20	3.56	7.90	20.67
Mean	4.70	16.87	34.36
Standard deviation	1.52	15.82	21.20
$\sqrt{\text{Mean squared relative deviation}}$	.322	.938	.617

* Upper bound established in estimation program used for simulated data.

data came from a stable distribution) makes it difficult to use these results (which are available from the authors).

For sums of size one, when the actual data are Student, the log-likelihood ratios are greater than zero for every replication except one (when $\nu = 8$). For stable data, the ratio is less than zero for all replications when $\alpha = 1.65$ and $\alpha = 1.80$. For $\alpha = 1.50$, the ratio is greater than zero seven out of 20 times. Since estimates of α for daily returns are usually around 1.65, the inaccuracies that exist for $\alpha = 1.50$ should not interfere with our using this ratio to discriminate between the two distributions.

From our simulation results we see that the log-likelihood ratio is an exceptionally good discriminator between the Student and stable distributions when the samples contain 1,300 observations (approximately the number of observations available in each series of actual daily returns). For $\nu =$ 3, 5, 8 and $\alpha =$ 1.65, 1.80 we had only one incorrect classification from 100 trials.[33]

33. The simulation results for all the separate values of α (and ν) use the same uniform numbers to generate the Student and stable-random numbers. The results, therefore, reflect the influence of changes in ν and α, holding everything else constant.

Table 9

Estimates for ν and α for Sum Sizes of One and Five for Simulated Student and Stable Data

Data Type	Estimation Procedure	Sum Size	True Parameter Value		
			ν		
			$\nu = 3$	$\nu = 5$	$\nu = 8$
Student	Student	1:			
		Mean	2.87	5.00	8.31
		Std. dev.	.281	.795	2.209
		5:			
		Mean	4.70	16.87*	34.36†
		Std. dev.	1.52	15.82	21.20
			$\hat{\alpha}$		
			$\nu = 3$	$\nu = 5$	$\nu = 8$
Student	Stable	1:			
		Mean	1.54	1.70	1.80
		Std. dev.	.046	.063	.083
		5:			
		Mean	1.68	1.80‡	1.86§
		Std. dev.	.088	.170	.104
			$\hat{\alpha}$		
			$\alpha = 1.50$	$\alpha = 1.65$	$\alpha = 1.80$
Stable	Stable	1:			
		Mean	1.49	1.64	1.79
		Std. dev.	.057	.055	.060
		5:			
		Mean	1.46	1.62	1.73
		Std. dev.	.090	.091	.095
			$\hat{\nu}$		
			$\alpha = 1.50$	$\alpha = 1.65$	$\alpha = 1.80$
Stable	Student	1:			
		Mean	2.41	2.93	4.40
		Std. dev.	.803	.309	.740
		5:			
		Mean	2.40	3.10	4.97
		Std. dev.	.60	.63	1.94

* Two estimates of ν truncated at $\nu = 60$.
† Seven estimates of ν truncated at $\nu = 60$.
‡ Three estimates of α truncated at $\alpha = 2$.
§ Three estimates of α truncated at $\alpha = 2$.

Summary of the simulation results.—(*a*) The M.L.E. of ν for the Student model is not very good for sample sizes of 260 and 65 but is fairly accurate for sample sizes of 1,300. (*b*) As the degrees of freedom increase, the standard deviation and mean-squared relative deviation of the M.L.E. of ν increase. (*c*) To test for convergence, an increase in α appears to be a better measure of convergence than an increase in ν. (*d*)

Table 10
Log-Likelihood Ratios for Simulated Student Data

Replication Number	True Value of ν		
	$\nu = 3$	$\nu = 5$	$\nu = 8$
1	18	16	10
2	19	13	7
3	32	18	21
4	21	21	10
5	8	11	8
6	8	9	13
7	21	18	12
8	7	15	6
9	19	11	5
10	9	3	−1
11	13	21	21
12	4	3	6
13	16	16	22
14	18	14	26
15	6	9	9
16	9	11	6
17	15	18	15
18	23	19	20
19	7	14	11
20	12	18	5

NOTE.—Log-likelihood ratio = Log_e [likelihood of Student model ÷ likelihood of stable model]. Digits to the right of the decimal point are omitted. The likelihood ratio for any entry, x, is $y = e^x$.

Table 11
Log-Likelihood Ratios for Simulated Stable Data

Replication Number	True Value of α		
	$\alpha = 1.5$	$\alpha = 1.65$	$\alpha = 1.80$
1	3	−5	−7
2	−19	−4	36
3	28	−1	−1
4	0.9	−1	−0.3
5	−9	−7	−2
6	−77	−11	−12
7	7	−1	−2
8	−4	−12	−8
9	−2	−10	−9
10	−41	−15	−18
11	2	−5	−2
12	−4	−9	−4
13	−11	−7	−8
14	8	−7	−5
15	−140	−10	−9
16	−11	−9	−9
17	−2	−8	−7
18	−957	−3	−2
19	−10	−14	−12
20	14	−14	−17

NOTE.—Log-likelihood ratio = Log_e [likelihood of Student model ÷ likelihood of stable model]. Digits to the right of the decimal point are omitted for log-likelihood ratios greater than one in absolute value. The likelihood ratio for any entry, x, is $y = e^x$.

The log-likelihood ratios will almost always indicate the true distribution when comparing the Student and stable distributions for sum sizes of one, 1,300 observations, and $\nu = 3, 5, 8$ or $\alpha = 1.65, 1.80$. In these cases, the (approximate) posterior odds in favor of the true distribution are usually quite high.

D. Results for Rates of Return

In the previous section, we saw that the best methods of discriminating between the Student and stable models were: (1) tests of convergence to normality using sum sizes of five, and (2) the value of the log-likelihood ratios for the daily rates of return. In this section, we present results from applying these two discriminators to actual rates of return for the 30 securities described in Section VI*A*. Our sample sizes range from 1,111 observations for Union Carbide to 1,693 observations for both Standard Oil of California and General Electric. Sample sizes for every security are listed in table 12.

Table 12
Sample Size for the Thirty Securities Used in the Study

	Sum Size	
	1	5
Union Carbide	1,111	223
Dupont	1,243	248
Proctor and Gamble	1,447	289
Sears	1,236	247
Standard Oil of California	1,693	338
Standard Oil of New Jersey	1,156	231
Swift	1,446	289
Texaco	1,159	231
Bethlehem Steel	1,200	240
Chrysler	1,692	338
Eastman Kodak	1,238	247
United Aircraft	1,200	240
U.S. Steel	1,200	240
Westinghouse	1,446	289
General Electric	1,693	338
General Foods	1,408	281
General Motors	1,446	289
Goodyear	1,162	232
International Harvester	1,200	240
International Nickel	1,243	248
International Paper	1,447	289
Johns Manville	1,205	241
Allied Chemical	1,223	244
Alcoa	1,190	238
American Can	1,219	243
American Telephone and Telegraph	1,219	243
American Tobacco	1,283	256
Anaconda	1,193	238
Woolworth	1,445	289
Owens Illinois	1,237	247

The results from estimating the degrees-of-freedom parameter for daily observations are given in table 13. The estimates range from 2.53

Table 13
Estimation Results for Daily Rates of Return

	Estimates of Degrees of Freedom for Student Model		Estimates of the Characteristic Exponent for the Stable Model		
	Sum Sizes		Sum Sizes		Log-Likelihood Ratios for Sums of Size One*
	1	5	1	5	
Union Carbide	7.6562	23.315	1.71	1.80	14.46
Dupont	6.1215	8.267	1.67	1.76	12.81
Proctor and Gamble	3.3021	5.937	1.52	1.80	14.64
Sears	2.8021	4.089	1.55	1.62	9.61
Standard Oil of California	4.8368	5.206	1.62	1.84	17.72
Standard Oil of New Jersey	3.5694	27.040	1.61	1.76	7.85
Swift	4.3281	14.417	1.60	1.78	17.03
Texaco	5.3455	9.657	1.65	1.83	15.48
Bethlehem Steel	4.7830	6.175	1.64	1.77	10.75
Chrysler	6.3715	9.666	1.73	1.77	10.04
Eastman Kodak	5.3542	3.598	1.72	1.47	7.88
United Aircraft	4.8455	10.394	1.66	1.89	8.71
U.S. Steel	13.2600	13.452	1.87	1.80	2.99
Westinghouse	6.1128	8.929	1.73	1.75	10.50
General Electric	4.8368	7.287	1.66	1.70	14.38
General Foods	5.0955	5.281	1.67	1.77	10.19
General Motors	5.0955	6.493	1.68	1.78	9.29
Goodyear	4.8368	11.162	1.65	1.77	8.24
International Harvester	5.1042	8.570	1.72	1.73	4.72
International Nickel	3.8194	6.044	1.58	1.65	11.30
International Paper	5.1042	8.096	1.68	1.60	10.78
Johns Manville	5.8542	8.783	1.77	1.72	6.54
Allied Chemical	5.0417	89.984†	1.73	1.94	7.43
Alcoa	4.8368	5.725	1.67	1.86	8.16
American Can	3.3198	4.735	1.65	1.61	7.89
American Telephone and Telegraph	2.5347	2.349	1.45	1.45	10.29
American Tobacco	2.8021	3.351	1.49	1.59	14.29
Anaconda	8.9323	6.731	1.76	1.61	11.27
Woolworth	3.3194	2.561	1.60	1.45	12.61
Owens Illinois	4.5781	9.182	1.60	1.66	14.10

* $\log_e$ [likelihood of Student model ÷ likelihood of Stable model]. The odds ratio for any entry, x, is $y = e^x$.
† Upper bound established in the estimation procedure used for actual data.

for American Telephone and Telegraph to 13.26 for U.S. Steel with the average estimate 4.79. The estimates of α for the same daily observations are given in table 13. The estimates range from 1.45 for American Telephone and Telegraph to 1.87 for U.S. Steel. The mean estimate of α is 1.65, which is approximately the value of α found in other studies of daily returns using the stable model.

The results for estimates of ν and α for sum sizes of five are also given in table 13. The average estimated value of ν is 11.22, and the average estimated value of α is 1.72. Both show an increase. If we have simulated stable data with $\alpha = 1.65$ and we estimate ν for sum sizes of one and then for sum sizes of five, the results in table 9 show that these estimates of ν increase, on average, from 2.93 to 3.10, a very small increase. If we have Student data with $\nu = 5$, and we estimate ν for sum sizes of one and five, our estimate of ν increases, on average, from 5.00 to 16.87, a substantial increase. Our average estimate of ν for the actual data changes from 4.79 to 11.22; the magnitude of this increase is more consistent with the case of the Student model applied to Student data than the Student model applied to stable data.

Similar statements can be made about the changes in estimates of α for sums of sizes one and five. For simulated stable data with $\alpha = 1.65$, the average estimate of α for sum sizes of one is 1.64; for sum sizes of five, it is 1.62, a slight decrease which is consistent with the downward bias in our estimator of α as the sample size decreases. For simulated Student data with ν equal to 5.00, the average estimate of α for sum sizes of one is 1.70; for sum sizes of five it is 1.80. For actual rates of return, the average estimate of α changes from 1.65 to 1.72 as we go from sum sizes of one to five. This increase in the average estimate of α suggests a process that is not stable, but one which is much closer to a process converging to normality.

Our basic inference from the results about convergence is that the data appear to converge to normality. This conclusion is based upon how the estimates of α and ν change when we go from sum sizes of one to sum sizes of five.[34] (Results for sum sizes of 10 and 20 are available from the authors.)

Our second empirical comparison of the two models is based upon the log-likelihood ratios. The results for the daily returns on the 30 securities are given in table 13. In every case, the value of the log-likelihood ratio is greater than zero, indicating the Student model provides a better description of the data than the stable model. Our simulation results given in tables 10 and 11 show that this ratio almost always indicates the correct model. Only when $\alpha = 1.50$ were there any major misclassifications. Thus, observing that all values are greater than zero strongly indicates that the Student model provides a better empirical description of the data than the stable model.

For large sample sizes, log-likelihood ratios can be interpreted as log-odds (see Section VI); we shall use this interpretation for the results

34. Our results for estimates of α are consistent with those presented by Officer who applied the stable model to daily rates of return from the period July 2, 1962 through July 11, 1969 for a sample of 50 common stocks listed on the New York Stock Exchange. (Officer's sample does not include the 30 securities whose daily returns were used for our own results.) Officer secured estimation results for sum sizes of 1–5, 7, 10, 15, and 20. The upshot of his results is identical to that of our own; as the sum size increases, the estimates of α increase, on average, or remain almost unchanged. As indicated earlier, this result is inconsistent with the stable model, conditional on the procedures used to estimate α. (The procedures used by Officer to estimate the stable model's parameters are identical to those that we used.)

for sums of size 1. When we use table 13, it can be seen that the log-odds in favor of the Student model are quite high. The lowest log-odds is 2.99 (for U.S. Steel), which corresponds to odds of about 20 to one in favor of the Student model. Since we did not use maximum-likelihood estimates for the stable model, the reported log-odds are inflated. The adjustments needed to remove this inflation are unknown. Since, however, the odds are so strongly in favor of the Student model, any seemingly reasonable adjustment would not, we believe, alter our basic inference: the Student model appears to describe the data better than does the stable model. This inference is given additional support by the evidence indicating convergence to normality.

The log-odds reported in table 13 pertain to all observations available for each security. These results do not indicate which fractiles of the empirical distributions contribute most to the apparent superiority of the Student model, as reflected in the log-odds. In order to consider this issue, we applied the following procedure to the unsummed observations for three securities: (1) the observations were arranged in ascending order, (2) the value of the log-odds was then computed for each observation, and (3) the value of the log-odds was also computed for the first 2 percent of the observations (after the rearrangement in [1]), the second 2 percent, the third 2 percent, etc. The results for each security indicated that the superiority of the Student model is primarily due to the observations in the tails of the empirical distributions. A representative plot from step (3) is provided in figure 1. This plot is based upon the observed returns for Bethlehem Steel. The sample size for this security is 1,200. Thus, each plotted value of the log-odds is based upon $(.02)(1{,}200) = 24$ ordered observations. (The estimated median would fall between the twenty-fifth and twenty-sixth set of grouped observations.) The plot has a U-shape, indicating that the Student model is heavily favored in the tails of the empirical distribution, whereas the stable model is more heavily favored around the median of the distribution.

We inferred from our results that the Student model provides a better empirical description than the stable model. This does not mean that the rates of return do, in fact, follow a Student model. It only indicates that the latter provides a better empirical fit than the stable model. The Student model has fat tails as does the stable model, but converges to normality for large sum sizes. The stable model does not converge to normality. Some implications of this important difference for theoretical and empirical work were discussed in Section II.

Even though the Student model provides a good fit to actual rates of return, there are some empirical results that it does not describe adequately. First, it has been observed (in, for example, Fama [see n. 2]) that large rates of return tend to be succeeded by large returns of unpredictable sign. This suggests a dependency in rates-of-return series. The Student (and stable) model studied here assumes returns are independent, which is contrary to available empirical evidence. Second, when we

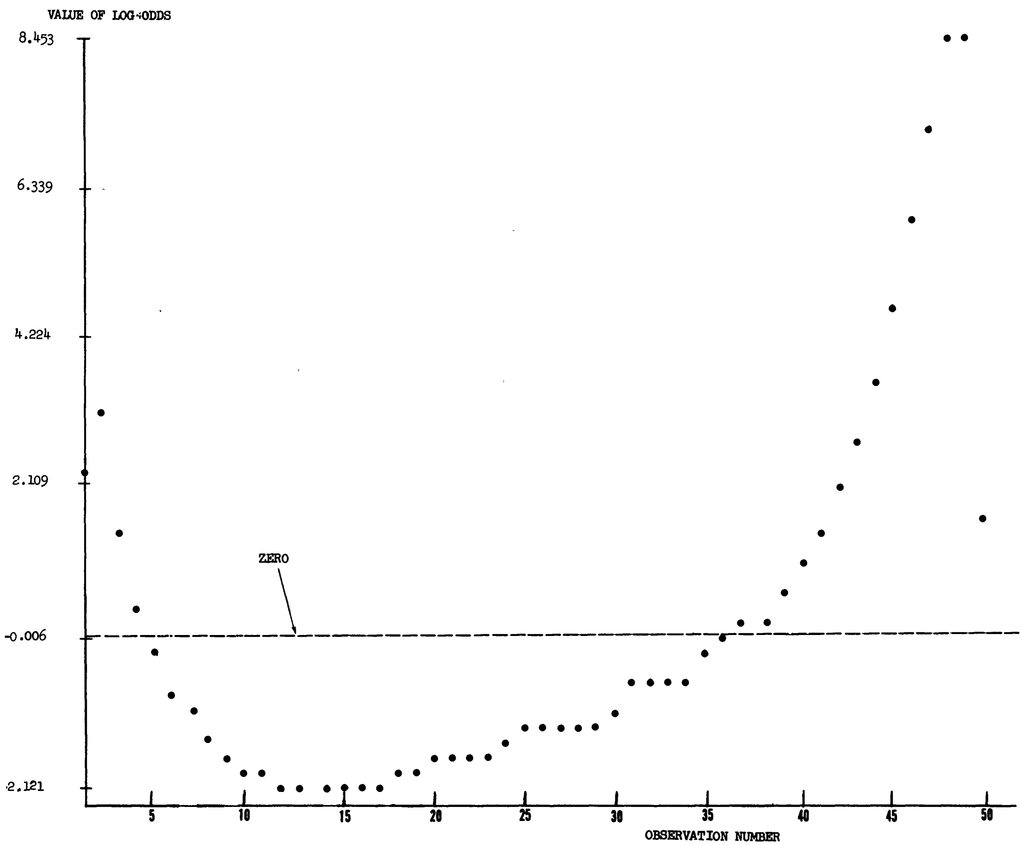

FIG. 1

went from sum sizes of one to sum sizes of five in our simulation study, the estimates of ν always increased. Using table 13, we observe a decrease in four of the 30 estimated values of ν when going to sum sizes of five (Eastman Kodak, American Telephone and Telegraph, Anaconda, Woolworth). This decrease may have occurred because large changes are, on average, followed by large changes. The dependency among large returns may cause the tail areas for sum sizes of five to be larger than if the returns were independent. However, in spite of these empirical deviations from theoretical expectations, we feel that the Student model offers a superior alternative to the stable model.

VII. SUMMARY

The Student and symmetric-stable distributions, as models for daily rates of return on common stocks, were discussed and empirically evaluated. Both models were derived using the framework of subordinated stochastic processes. Some important theoretical and empirical implications of these models were also discussed. The descriptive validity of each model, relative to the other, was assessed by applying each model to actual daily rates of return. Our interpretations of our empirical results were guided by results from a Monte Carlo investigation of the properties of our estimators and model-comparison methods. The major inference of this report is that, for daily rates of return, the Student model has greater descriptive validity than the symmetric-stable model.

APPENDIX A
DERIVATIONS OF THE STUDENT AND STABLE MODELS

Before beginning with the derivations, we state the definition of a subordinated stochastic process. Let $[\tilde{X}(s); s \geqslant 0]$ and $[\tilde{h}(s); s \geqslant 0]$ denote stochastic processes. Define another stochastic process $\{\tilde{Z}(s) = \tilde{X}[\tilde{h}(s)]; s \geqslant 0\}$. The process $[\tilde{Z}(s)]$ is said to be subordinate to the process $[\tilde{X}(s)]$; the process $[\tilde{h}(s)]$ is the directing process.[35] We shall present models below in which the Student and symmetric-stable distributions are subordinated stochastic processes with $[\tilde{X}(s)]$ being a stationary Gaussian stochastic process.

Let $\tilde{X}(\Delta t)$ denote the rate of return over an interval of time $[t, t + \delta]$ for $t \geqslant 0$ and fixed $\delta > 0$. We assume that the rates of return are expressed as deviations about their mean.

Assume that $[\tilde{X}(\Delta t); t \geqslant 0]$ is a stationary Gaussian process with the following properties:[36]

$$E[\tilde{X}(\Delta t)] = 0, \tag{A1}$$

$$\text{Var}\,[\tilde{X}(\Delta t)] = \sigma^2 \Delta t, \tag{A2}$$

$$\tilde{X}(\Delta t) \text{ is independent of } \tilde{X}(\Delta t') \text{ for all } t' > t + \delta, \text{ and any choice of } t \text{ and } \delta. \tag{A3}$$

Next introduce another stochastic process $[\tilde{h}(\Delta t); t \geqslant 0]$ with the properties:

$$\tilde{h}(\Delta t) > 0, \tag{B1}$$

$$\tilde{h}(\Delta t) \text{ is independent of } \tilde{h}(\Delta t') \text{ for all } t' > t + \delta \text{ and any choice of } t \text{ and } \delta. \tag{B2}$$

Finally, consider the subordinated process $\{\tilde{Z}(\Delta t) = \tilde{X}[\tilde{h}(\Delta t)]; t \geqslant 0\}$. Given the assumed properties of $[\tilde{X}(\Delta t)]$ and $[\tilde{h}(\Delta t)]$, the process $[\tilde{Z}(\Delta t)]$ has the following properties:

$$E[\tilde{Z}(\Delta t)] = 0, \tag{C1}$$

$$\text{Var}\,[\tilde{Z}(\Delta t) | h(\Delta t)] = \sigma^2 h(\Delta t), \tag{C2}$$

35. Subordinated stochastic processes are discussed in W. Feller, *An Introduction to Probability Theory and Its Applications,* 2d ed. (New York: John Wiley & Sons, 1971), 2: chap. 17, Sec. 7.

36. A stochastic process $Y(s)$ is (strictly) stationary if for every finite sequence of points, $s_1, s_2, \ldots, s_n$, the distribution function $F[Y(s_1), Y(s_2), \ldots, Y(s_n)]$ is identical to $F[Y(s_1 + k), Y(s_2 + k), \ldots, Y(s_n + k)]$ for all values of the translation parameter k. Additional details are given in Doob.

$$\tilde{Z}(\Delta t) \text{ is independent of } \tilde{Z}(\Delta t') \text{ for all } t' > t + \delta, \text{ and any choice of } t \text{ and } \delta, \quad \text{(C3)}$$

$$[\tilde{Z}(\Delta t)] \text{ is a stationary stochastic process.} \quad \text{(C4)}$$

For a given realization, $h(\Delta t)$, of $\tilde{h}(\Delta t)$, the probability density function of $Z(\Delta t)$ is a conditional normal density function with mean zero and variance $\sigma^2 h(\Delta t)$, denoted by $f_N[Z(\Delta t) \mid 0, \sigma^2 h(\Delta t)]$. To find the unconditional distribution of $Z(\Delta t)$ we need a distribution for $h(\Delta t)$.

Define $g(\Delta t) = [h(\Delta t)]^{-1}$. Suppose that $g(\Delta t)$ follows a gamma-2 distribution with parameters $s^2\Delta t$ and ν denoted by $f_{\gamma 2}[g(\Delta t) \mid s^2\Delta t, \nu]$.[37] Then the unconditional distribution for $Z(\Delta t)$ is:

$$D[Z(\Delta t)|0, H(\Delta t), \nu] = \int_0^\infty f_N[Z|0, g(\Delta t)\sigma^2] f_{\gamma 2}[g(\Delta t)|s^2\Delta t, \nu]\, dg(\Delta t), \quad (5)$$

where $H(\Delta t) = 1/s^2(\Delta t)\ 1/\sigma^2$. The distribution in (1) is a Student distribution with mean 0, scale $H(\Delta t)$, and degrees of freedom ν.

This derivation of the Student model allows the distribution of rates of returns, $[\tilde{Z}(\Delta t)]$, to incorporate changes in the "states of nature" over the interval $[t, t + \delta]$. The state of nature at t is a complete description of the economic environment up to time t; $\tilde{h}(\Delta t)$ may be interpreted as the change in state over $[t, t + \delta]$. The change in the "state of nature" may be viewed as a change in the stock of information available to the capital market transactors.

Observe that the kind of "state" dependency provided by the Student model is not the same as that considered by, for example, Radner, Arrow, and Hirshleifer.[38] In the latter works, all uncertainty attaches to the "state," or change thereof, and given the state, the value of the state-dependent random variable is known with certainty. In our scheme, only the distribution function of $\tilde{Z}(\Delta t)$ is known with certainty when the change of state is known.

Prior to introducing the subordinated process $[\tilde{Z}(\Delta t)]$, the variance of the process, $\sigma^2\Delta t$, only depended upon the length of the interval of time, Δt, over which the rate of return was defined. It seems reasonable, however, to let the number and importance of events occurring within the fixed time interval cause changes in the level of variability in rates of returns. We accomplish this by introducing the directing process $[\tilde{h}(\Delta t)]$.

37. Raiffa and Schlaifer (p. 227) give the mean and variance of the gamma-2 density, $f_{\gamma 2}[g(\Delta t)|s^2\Delta t, \nu)$, as $E(\tilde{g}(\Delta t)] = 1/s^2\Delta t$ and

$$\text{Var}\,[\tilde{g}(\Delta t)] = \frac{1}{1/2\ \nu(s^2\Delta t)^2}.$$

38. R. Radner, "Competitive Equilibrium Under Uncertainty," *Econometrica* 36 (January 1968): 31–58; K. J. Arrow, "The Role of Securities in the Optimal Allocation of Risk Bearing," *Review of Economic Studies* 32 (April 1964): 91–96; J. Hirshleifer, "Investment Decision under Uncertainty: Choice Theoretic Approaches," *Quarterly Journal of Economics* 79 (November 1965): 509–36; "Investment Decision under Uncertainty: Applications of the State Preference Approach," *Quarterly Journal of Economics* 80 (May 1966): 252–77.

We derived the Student distribution by assuming that $\tilde{g}(\Delta t) = [\tilde{h}(\Delta t)]^{-1}$ followed a gamma-2 distribution. The stable model can be derived by assigning a particular asymmetric-stable distribution to $\tilde{h}(\Delta t)$. We have not yet defined the characteristic function of an asymmetric-stable distribution, so we tend to this task first.

The log characteristic function of a stable distribution (symmetric or asymmetric) is:

$$\ln \phi_{\tilde{x}}(t) \doteq i\delta t - |ct|^{\alpha}[1 + i\beta \frac{t}{|t|} w(t, \alpha)],$$

where t is some real number; $i = \sqrt{-1}$; β is the skewness parameter, with $\beta \in [-1, 1]$; δ is the location parameter; $c > 0$ is the scale parameter; α is the characteristic exponent, with $\alpha \in [0, 2]$; and

$$w(t, \alpha) = \begin{cases} \tan(\pi \alpha/2), & \text{if} \quad \alpha \neq 1 \\ \dfrac{2\, ln(|t|)}{\pi}, & \text{if} \quad \alpha \neq 1. \end{cases}$$

If $\beta > 0$ and $\alpha < 2$, the distribution is skewed to the right; it is skewed to the left if $\beta < 0$ and $\alpha < 2$. A symmetric-stable distribution is defined by $\beta = 0$. A strictly positive stable random variable is defined by $\alpha < 1$ and $\beta = 1$.

Now, suppose $\tilde{h}(\Delta t)$ follows a strictly positive stable distribution with location parameter equal to zero, that is, $\beta = 1$ and $\alpha \in (0, 1)$. The characteristic function of this process is:

$$\phi_{\tilde{h}(\Delta t)}(u) = \exp\{-\gamma \Delta t|u|^{\alpha}[1 + i(u/|u|) \tan(\pi\alpha/2)]\}. \qquad (6)$$

Consider the subordinated process $[\tilde{Z}(t)]$ as defined above. If the characteristic function for $\tilde{h}(\Delta t)$ is defined by (6), then, as Mandelbrot and Taylor [1967] demonstrated, the unconditional distribution of $[\tilde{Z}(t)]$ is a symmetric-stable distribution with characteristic exponent $\alpha^* = 2\alpha < 2$, where α is the characteristic exponent of the distribution of $\tilde{h}(\Delta t)$.[39] This can be shown as follows.

$$\begin{aligned} \phi_{\tilde{Z}(\Delta t)}(u) &= E_{\tilde{Z}}\left\{\exp[iu\tilde{Z}(\Delta t)]\right\}, \\ &= E_{\tilde{h}}\left[E_{\tilde{X}}\left(\exp\{iu\tilde{X}[h(\Delta t)]\}|\tilde{h}(\Delta t) = h(\Delta t)\right)\right], \qquad (7) \\ &= E_{\tilde{h}}\left\{\text{Exp}[-1/2\, u^2\sigma^2\tilde{h}(\Delta t)]\right\}, \end{aligned}$$

since the characteristic function of $\tilde{X}[h(\Delta t)]$ is that of a normal distribution with mean 0 and variance $\sigma^2 h(\Delta t)$.

Now, let $w = i\,u\,\sigma^2/2$. Then (7) becomes

$$\phi_{\tilde{Z}(\Delta t)}(u) = E\left\{\exp[iw\tilde{h}(\Delta t)]\right\} = \phi_{\tilde{h}(\Delta t)}(w).$$

39. An alternative statement of this result appears in W. Feller, p. 348. The latter work, as well as B. Mandelbrot and H. M. Taylor, "On the Distribution of Stock Price Differences," *Operations Research* 15 (1967): 1057–62, considers processes with stationary independent increments. Such a framework can be used for rates of return under continuous compounding by considering increments of the log-price relative process.

This becomes

$$\phi_{\tilde{Z}(\Delta t)}(u) = \exp\{-\gamma\Delta t(\sigma^2/2)^\alpha|w|^{2\alpha}[1 - \tan(\pi\alpha/2)]\},$$
$$= \exp(-\hat{\gamma}\Delta t|w|^{2\alpha}), \qquad (8)$$

where $\hat{\gamma} = \gamma(\sigma^2/2)^\alpha[1 - \tan(\pi\alpha/2)]$. Equation (8) is simply the characteristic function for a symmetric-stable distribution with characteristic exponent $(2\alpha) < 2$.

Note that the methods used to derive the Student and stable models are identical. The difference between the two models is in the distributional assumptions for the directing process $\tilde{h}(\Delta t)$. In the Student model $[\tilde{h}(\Delta t)]^{-1}$ follows a gamma-2 distribution; in the stable model, $\tilde{h}(\Delta t)$ follows a strictly positive stable distribution with $\alpha \in (0,1)$.

APPENDIX B

PROPERTIES OF THE UNIFORM RANDOM NUMBERS USED IN THE SIMULATIONS

The quality of our simulation results depends, of course, upon the properties of our sample drawn from the uniform distribution $\tilde{u}(0, 1)$, in particular: the consistency of our sample's properties with those of the theoretical distribution of $\tilde{u}(0, 1)$ and the randomness of our sample. In this regard, we note the following:

i. The first and second moments of $\tilde{u}(0, 1)$ are .5 and $(1/12) \approx .0833$, respectively. The first and second moments of our sample are .498 and .0835, respectively.
ii. For a serially independent series with $N = 26{,}000$ observations, the asymptotic distribution of the estimated first-order serial correlation coefficient, $\hat{\rho}$, has mean and standard deviation equal to: $-(1/N-1)) = -.38 \times 10^{-4}$ and $\sqrt{1/(N-1)} = .0062$, respectively. For our sample, $\hat{\rho} = .00496$.
iii. For a χ^2 test of the sample cumulative probabilities against $\tilde{u}(0, 1)$, we had a χ^2 statistic of 13.514. This statistic is based upon 20 subintervals of (0, 1), each of length .05. For a χ^2 random variable with 19 df, $\tilde{\chi}^2(19)$, $\Pr[\tilde{\chi}^2(19) \geqslant 13.7)] = .80$.

From these (and other) results, we infer that our samples properties are consistent with the distribution of $\tilde{u}(0, 1)$ and randomness.