Alan L. Lewis
Analytic Investment Management, Inc.

Sheen T. Kassouf
University of California, Irvine

R. Dennis Brehm
Analytic Investment Management, Inc.

Jack Johnston
University of California, Irvine

The Ibbotson-Sinquefield Simulation Made Easy

The Ibbotson-Sinquefield simulation for common stock returns, published in this *Journal* in 1976, can be considerably simplified. The model predicts that a future stock return will be given by a historical risk premium, randomly drawn, added to an expected T-bill rate implied by the yield curve. Actual random drawings are not really necessary; the model can be cast as a geometric random walk. So most results can be understood via the rapid convergence of the probability distribution for the common stock geometric return to a normal distribution.

I. Introduction

The Ibbotson-Sinquefield (I-S) simulation for possible future returns on a variety of financial assets has attracted considerable interest (Ibbotson and Sinquefield 1976). We focus here on common stocks and those variables coupled to common stock returns. Our main observation is that considerable simplification is possible without altering, in any important way, the numerical results. Our emphasis of simplification is at two levels. We first want to understand what are the essential driving equations of the model. On the surface, the model appears to be driven by two autoregression equations, one for the inflation rate and one for the real return on Treasury bills. We demonstrate that, in fact, neither equation plays much of a role in predictions for common stock returns. Once these equations are removed explicitly, we are led to a very simple model for stock returns. The second level of simplification has to do with how the simulation is actually carried out. Ibbotson and Sinquefield do this by making actual random drawings from historical distributions; we show that this is not really necessary

(*Journal of Business,* 1980, vol. 53, no. 2)

0021-9398/80/5302-0002$00.95

because the model for common stock returns is essentially a geometric random walk. Thus, probability distributions can be obtained without making drawings. (A referee has called to our attention two recent papers, by Michaud and by Ibbotson and Sreenivasan, using an approach similar to ours.)

The plan of this paper is as follows. In Part II, we review that portion of the I-S model that pertains to common stocks; we reduce the model to its essentials by the end of the section. Then, in Part III, we present our simplified model and its analysis. Detailed numerical comparisons with I-S results are presented in that section. Some conclusions follow in Part IV.

II. I-S Simulation Model

Ibbotson and Sinquefield consider returns on common stocks, Treasury bills, long-term government bonds, and corporate bonds. They also measure the inflation rate in terms of changes in the Consumer Price Index. We focus on their predictions for common stocks; common stocks are coupled together in their model with T-bills and the inflation rate only. Thus, we will ignore bonds.

We will use a simplified notation, namely,

S_t = annual rate of return on common stocks in year t.
B_t = annual rate of return on Treasury bills in year t.
I_t = annual inflation rate in year t.

With historical data from 1926 to 74, I-S measured S_t with the Standard and Poor's 500 stock index, and B_t with 30-day T-bills. For the simulation, $t = 1, 2, \ldots, 25$ corresponds to 1976, 1977, . . . , 2000.

The real rate of return on each financial asset may then be defined as $s_t = S_t - I_t, b_t = B_t - I_t$, and the risk premium (that is, the excess return on stocks over T-bills) is $R_t = S_t - B_t$. In I-S notation, these six variables are R_{mt}, R_{ft}, R_{It}, $R_{mr,t}$, R_{rt}, R_{pt}. The authors fit 48 years of historical data (1927–74) for the inflation rate and real return on T-bills to first-order autoregression equations, obtaining

$$I_t = .011 + .599\, I_{t-1} + \epsilon_{I,t} \tag{1}$$

$$b_t = -.0015 + .623\, b_{t-1} + \epsilon_{b,t}, \tag{2}$$

with $\sigma(\epsilon_I) = .039$, $\sigma(\epsilon_b) = .038$, and respective R^2 of .33 and .39. The values of the errors $\epsilon_{I,t}$ and $\epsilon_{b,t}$ can be calculated for each of 48 historical years, so that I-S have a collection of 48 pairs: $\epsilon_{I,1927}, \epsilon_{I,1928}, \ldots, \epsilon_{I,1974}; \epsilon_{b,1927}, \epsilon_{b,1928}, \ldots, \epsilon_{b,1974}$. Now, apart from a very important complication which we explain below, I-S use equations (1) and (2) to simulate future returns in the following way. They begin in 1975 ($t = 0$) with a certain $I_0 = .070$ and $b_0 = -.013$. Then they make 399 random drawings of *pairs* of residuals from the historical collection. Substituting these pairs into equations (1) and (2) with $t = 1$ produces 399 simulated values for both I_{1976} and b_{1976}. Thus, they have a simulated distribution for 1976 and means, variances, etc., can be worked out. Proceeding now to 1977, they inherit 399 values for I_{1976} and b_{1976}. Again 399 drawings are made from the historical residuals; each one of the residuals is paired off with a 1976 return value, using equations (1) and (2). Thus, one then has 399 values for I_{1977} and b_{1977}. Then, 1977 means and variances can be worked out, but also cumulative averages involving both 1976 and 1977 can be computed. This procedure is iterated until the year 2000.

The reason that the residuals are drawn in pairs is that I-S found them to be highly negatively correlated, namely,

$$\text{corr}\ (\epsilon_I,\ \epsilon_b) = -.95. \tag{3}$$

And the reason for this strong correlation is, primarily, that the b_t series contains the I_t series as a negative component. That is, equation (2), plus a constant, is essentially the negative of equation (1).

Two comments about the above procedure are in order. First, it is clear that to make 399 random drawings from a distribution of 48 elements has the effect of essentially reproducing the distribution. Many interesting questions can be answered without bothering to make the random drawing. For example, the expected real return on T-bills in 1976 can be read off equation (2) directly as $\bar{b}_{1976} = -.0015 + (.623)(-.013) = -.0096$. In fact, by averaging their 399 values in 1976, I-S (1976, p. 328) find a simulation average $\bar{b}_{1976} = -.013$. The discrepancy between these two numbers is, of course, due to the fact that while the sum of the 48 residuals is guaranteed to be zero, there is no guarantee that the 399 drawings will add up to zero. Clearly, the first number is to be preferred.

The second comment involves the long-time scale implications of equations (1) and (2). Any such autoregression model, with a slope less than one, forces an exponential decay (or growth) of the expected value to a value close to its historical average. This can be seen by iterating the equation. In other words, regardless of the 1975 value for inflation, equation (1) implies a decay or growth to an asymptotic inflation rate of 2.7%, which is not far from its 49-year historical average of 2.3%. Since, as we shall see, I-S do not really use equation (1), this comment is not worth pursuing.

The next important ingredient of the I-S model is the prescription for selecting simulation values for the risk premium R_t. Ibbotson and Sinquefield select 399 values independently in each year t from the 1926–74 historical distribution. For our purposes in Part II, all we need to know is that this distribution has mean $\bar{R} = .086$ and standard deviation $\sigma(R) = .225$.

Now, the model for common stock returns S_t would seem to follow, at this point, from definitions already given. First,

$$S_t = R_t + B_t. \tag{4}$$

R_t is selected as explained above, and the simulation values for B_t should be obtained by adding together the values for b_t (eq. [2]) and I_t (eq. [1]).

In fact, I-S do not follow this procedure. Instead, they do continue to define B_t as the sum

$$B_t = b_t + I_t. \tag{5}$$

Then, they define

$$I_t = \bar{I}_t + \epsilon_{I,t} \tag{6}$$

and

$$b_t = \bar{b}_t + \epsilon_{b,t}. \tag{7}$$

Comparison of equations (6) and (7) with (1) and (2) would imply that

$$\bar{I}_t = .011 + .599\,\bar{I}_{t-1} \tag{8}$$

and

$$\bar{b}_t = -.0015 + .623\,\bar{b}_{t-1}. \tag{9}$$

But, in fact, I-S do not define $\bar{I}_t$ via equation (8). The reason is that they really do not want to add together $\bar{I}_t$ and $\bar{b}_t$ from equations (8) and (9) to obtain $\bar{B}_t$. Instead, they want to use the U.S. government bond yield curve of December 31, 1975, to deduce a value for $\bar{B}_t$. The argument for this is fairly standard; the yield curve provides directly a typical yield Y_t for a bond maturing in year t. From these yields, forward rates F_t are obtained via

$$(1 + F_1)(1 + F_2) \ldots (1 + F_t) = (1 + Y_t)^t. \tag{10}$$

If a series of T-bills were equivalent to one long-term bond, one would set $\bar{B}_t = F_t$; in fact, the series of bills is more liquid. Thus, one corrects for this by subtracting from F_t an assumed liquidity or maturity premium. Ibbotson and Sinquefield take for this maturity premium the difference between the historical average annual rate on 20-year government bonds (3.4%) and Treasury bills (2.3%). Thus, I-S determine $\bar{B}_t$ via

$$\bar{B}_t = F_t - .011. \tag{11}$$

Now, the trouble is that one cannot have equations (8), (9), and (11) all at the same time. Thus, I-S throw out equation (8) and replace it by

$$\bar{I}_t = F_t - .011 - \bar{b}_t. \tag{8'}$$

Equation (6) is still the simulation equation for I_t, but $\bar{I}_t$ is taken from equation (8′). The difference between equations (8) and (8′) is not small; as we have already pointed out, equation (8) implies inflation rates in the year 2000 around 2.7%. But, the I-S (1976, p. 327) forward rates around year 2000 are about 8.5%, $\bar{b}_t$ always hovers near zero, and so equation (8′) produces year 2000 inflation rates near 7.4% (see I-S 1976, p. 336, fig. 7). Thus, it is important to realize that I-S are really saying that expected inflation rates are determined from the yield curve.

Now, where have all these gyrations gotten us in terms of predictions for common stock returns? The model is equation (4). Let us recapitulate how B_t is determined: B_t is obtained by adding together equations (6) and (7), with $\bar{I}_t$ given by equation (8′) and $\bar{b}_t$ given by equation (9). Thus, the I-S model for stocks can be summarized in three lines:

$$S_t = R_t + B_t, \tag{12}$$

where

$$B_t = F_t - .011 + \epsilon_{B,t} \tag{13}$$

and

$$\epsilon_{B,t} = \epsilon_{I,t} + \epsilon_{b,t}. \tag{14}$$

Notice that the autoregression model for the real return on T-bills, equation (9), has disappeared also. The slope and intercept values in that equation make absolutely no difference to the predictions for stocks (except via the fact that one needs these values to determine the residuals).

Our final comment is that, for practical purposes involving stocks alone, one might as well set $\epsilon_{B,t}$ equal to zero in equation (13). This is because of the high negative correlation between ϵ_I and ϵ_b. We have

$$\text{var}\ (\epsilon_B) = \text{var}\ (\epsilon_I) + \text{var}\ (\epsilon_b) + 2\ \text{cov}\ (\epsilon_I,\epsilon_b). \tag{15}$$

But, since $\sigma(\epsilon_I) = .039$, $\sigma(\epsilon_b) = .038$, and cov $(\epsilon_I,\epsilon_b) = -0.95$, we have

$$\sigma(\epsilon_B) = .015. \tag{16}$$

This small value is dominated, in equation (12), by the much larger standard deviation of R_t: $\sigma(R) = .225$. Our simplified model is exactly equations (12) and (13) with ϵ_B set to zero. We want to pursue this simplified model and show that, indeed, it essentially reproduces the I-S numerical results.

III. Our Simplified Model for Stock Returns

We have argued, in the previous section, that the I-S numerical results for common stock returns should be well approximated by the simple model

$$S_t = R_t + B_t, \tag{17}$$

where

$$B_t = F_t - .011. \tag{18}$$

In words, a reasonable guess at what stocks will return in the future is obtained by adding expected T-bill rates, implied by the yield curve, to historical risk premiums.

A major advantage of this model is that it can be used to predict returns almost completely by hand—we will not need to make random drawings at all. Thus, we will be able to see why some of the graphs in I-S have certain familiar shapes.

Values of B_t, using equation (18) and the forward rates listed in I-S (1976, p. 327), are shown in the first column of table 1.

The cumulative wealth W_t from common stocks is given by

$$W_t = (1 + S_t)W_{t-1}, \tag{19}$$

with a certain dollar $W_0 = 1$ in 1975. Thus, since $\bar{R} = .086$,

$$\bar{W}_t = (1.086 + B_t)\,\bar{W}_{t-1}. \tag{20}$$

TABLE 1 **Various Measures of the Probability Distribution of the Cumulative Wealth from Common Stocks**

Year	B_t	$\bar{W}_t$	$\tilde{W}_t$	W_t^{min}	W_t^{max}	$g_t^{min}(\%)$	$g_t^{max}(\%)$
1975	...	1.00	1.0	1.0	1.0	...	...
1976	.0491	1.13	1.1	.80	1.60	−22.8	47.0
1977	.0627	1.30	1.3	.78	2.08	−10.7	36.8
1978	.0661	1.50	1.4	.79	2.63	−8.1	32.3
1979	.0709	1.74	1.6	.81	3.26	−5.4	29.6
1980	.0662	2.00	1.8	.84	4.00	−3.5	27.7
		(2.05)	(1.83)	(.73)	(4.16)	(−6.2)	(28.5)
1981	.0678	2.31	2.1	.88	4.86	−2.1	26.3
1982	.0638	2.65	2.3	.92	5.88	−1.1	25.3
1983	.0638	3.05	2.6	.98	7.07	−.2	24.4
1984	.0629	3.50	3.0	1.04	8.47	.5	23.7
1985	.0637	4.02	3.4	1.11	10.1	1.1	23.1
1986	.0637	4.63	3.8	1.19	12.1	1.6	22.6
1987	.0709	5.35	4.3	1.28	14.3	2.0	22.2
1988	.0721	6.19	4.8	1.37	17.0	2.4	21.8
1989	.0733	7.18	5.4	1.48	20.1	2.8	21.4
1990	.0730	8.32	6.1	1.59	23.7	3.1	21.1
1991	.0756	9.66	6.9	1.72	28.0	3.4	20.8
1992	.0751	11.21	7.8	1.86	32.9	3.6	20.6
1993	.0725	12.99	8.8	2.01	38.8	3.9	20.3
1994	.0712	15.02	10.0	2.17	45.7	4.1	20.1
1995	.0716	17.39	11.3	2.36	53.5	4.3	19.9
1996	.0720	20.13	12.7	2.57	62.7	4.5	19.7
1997	.0724	23.31	14.3	2.79	73.6	4.7	19.5
1998	.0728	27.01	16.2	3.03	86.3	4.8	19.4
1999	.0732	31.30	18.3	3.30	100.9	5.0	19.2
2000	.0736	36.28	20.6	3.60	117.9	5.1	19.1
		(38.3)	(21.7)	(3.55)	(130.1)	(5.1)	(19.5)

NOTE.—Ibbotson-Sinquefield numbers for the years 1980 and 2000 are in parentheses.

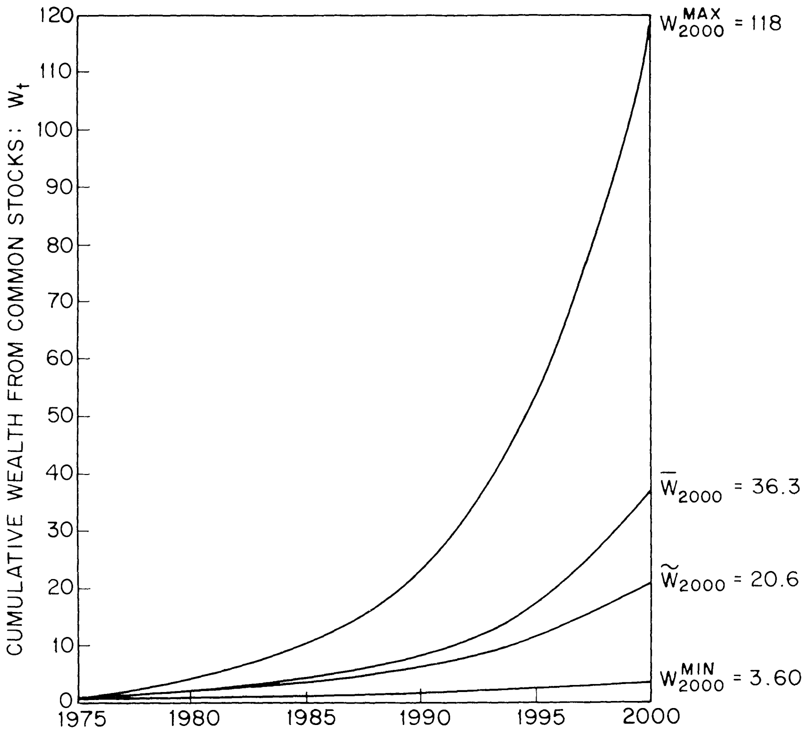

FIG. 1.—The mean, median, and 95% confidence levels for the cumulative wealth from common stocks.

The mean cumulative wealths, calculated from equation (20), are shown in column 2 of the table; I-S simulation averages for 1980 and the year 2000 are shown in parenthesis under ours in the table. $\bar{W}_t$ is also graphed in figure 1; this figure should be compared with I-S figure 2 (1976, p. 333).

For purposes of error analysis and confidence levels, note that the stream of expected T-bill rates (col. 1 of the table) is fairly flat (because the December 31, 1975, yield curve was flat); thus, approximate all B_t by their mean $\bar{B}_t = .069$. Then, with $X_t \equiv \ln W_t$ and $G_t \equiv \ln (1.069 + R_t)$, equation (19) becomes

$$X_t = X_{t-1} + G_t, \tag{21}$$

with $X_0 = 0$. Equation (21) says that X_t is undergoing an arithmetic random walk with successive independent jumps G_t. Thus, with rapid convergence, the probability distribution for X_t is normal with mean $t\bar{G}$, and variance $t\sigma^2(G) = t[\overline{G^2} - (\bar{G})^2]$. We could obtain these numbers from the R_t distribution, but it is a lot easier to approximate G_t by three terms of its Taylor expansion about $1 + \bar{B}_t$, that is,

$$G_t \approx \ln(1.069) + R_t/(1.069) - \frac{1}{2} R_t^2/(1.069)^2. \tag{22}$$

Then, $\bar{G}$ is determined entirely in terms of $\bar{R}$ and $\sigma^2(R)$. Similarly, $\sigma^2(G)$ is given, in this same approximation, by $\sigma^2(R)/(1.069)^2$. In this way, we obtain

$$\bar{G} \approx .121, \quad \text{and} \quad \sigma(G) \approx .213. \tag{23}$$

The median of the W_t distribution, which we denote by $\tilde{W}_t$, is then given by $\tilde{W}_t = \exp(\bar{G}t)$; these values are shown in column 3 of the table and figure 1.

Confidence levels are obtained from

$$\text{Prob}\,(W_t \leqslant W_t^{\max}) = \Phi[(\ln W_t^{\max} - t\bar{G})/\sqrt{t}\,\sigma(G)], \tag{24}$$

where $\Phi(x)$ is the cumulative normal distribution from $-\infty$ to x. In this way for 95% confidence [Prob $(W_t \leqslant W_t^{\max}) = .95$], one has

$$W_t^{\max} = \exp[t\bar{G} + 1.64\sqrt{t}\,\sigma(G)], \tag{25}$$

and, similarly,

$$W_t^{\min} = \exp[t\bar{G} - 1.64\sqrt{t}\,\sigma(G)], \tag{26}$$

where $W_t^{\min}$ is defined by Prob $(W_t \geqslant W_t^{\min}) = .95$. Again, these numbers are shown in the table and figure 1.

Define a cumulative annual growth rate $g(t)$ via $W_t = \exp[g(t)t]$. Thus, $g(t)$ is given by $g(t) = X_t/t$. Ibbotson and Sinquefield define a number they call the geometric annualized return R_t^G. Their definition of R_t^G is that $W_t = (1 + R_t^G)^t$. Thus, the relationship between our $g(t)$ and their R_t^G is that $(1 + R_t^G) = \exp[g(t)]$. In the year 2000, I-S find $R_t^G = .130$; since we have $g(t) = \bar{G} = .121$, our value for R_t^G is $\exp(.121) - 1 = .129$. Confidence levels for $g(t)$ can also be obtained immediately from the normal distribution. Again, for 95% confidence, one has

$$g^{\max}(t) = \bar{G} + 1.64\,\sigma(G)/\sqrt{t}, \tag{27}$$

$$g^{\min}(t) = \bar{G} - 1.64\,\sigma(G)/\sqrt{t}. \tag{28}$$

These numbers are also shown in the table and are graphed in figure 2. Figure 2 is our analogue of I-S figure 1 (1976, p. 332), keeping in mind that one translates via $(1 + R_t^G) = \exp[g(t)]$. As one sees from equations (27) and (28), the narrowing of $g(t)$ distribution with time is no more than the conventional law of large numbers result. Ibbotson and Sinquefield explain this narrowing as an "averaging effect"—we now see that it is indeed the averaging effect explicitly described by the central limit theorem.

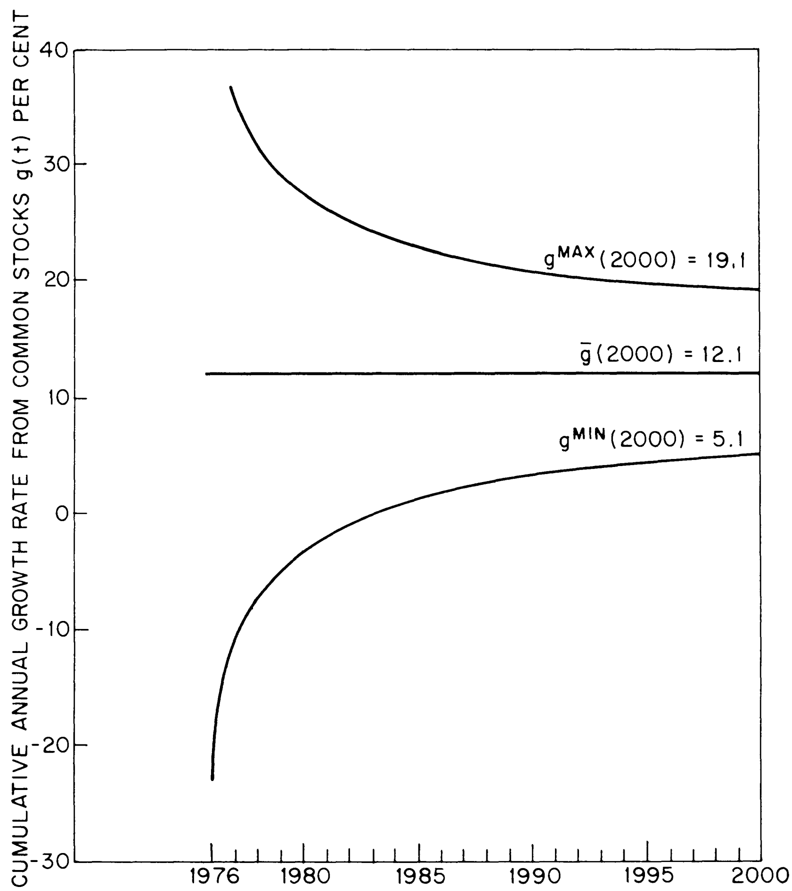

FIG. 2.—The mean and 95% confidence levels for the cumulative annual growth rate from common stocks.

IV. Conclusions

We have reduced the I-S simulation model to its bare bones and shown that their numerical results could be reproduced without making drawings. However, we have not addressed ourselves to the question of how reasonable this basic model is, and whether or not substantial improvements might be made. Let us conclude with some brief remarks in this regard.

First, if we ignore the use made of the yield curve, it seems apparent that the inflation model of equation (1) can be substantially improved upon. Clearly, one would want to make use of available economic data for predictions instead of just last year's inflation rate. To predict that current high inflation will simply decay to its long-time historical average would seem unrealistic.

Next, there is the question of whether or not the yield curve really is a sensible predictor of interest rates. Clearly, the error in prediction, $\epsilon_{B,t}$, must grow with time. The nature of this growth could be estimated by historical studies of the predictive power of the yield curve. It is this inherent error in the yield curve which should really appear in equation (14).

Finally, the basic model for stock returns, equation (12), would have been of dubious value over the past decade, when higher bill rates managed to induce lower stock returns. Thus, over the short term, it may make more sense to have stocks respond not only to current interest rates, but also to deviations of the current rates from expectations implicit in B_{t-1}. That is, one may want to explore a model of the type $S_t = S_t(R_t, B_t, B_{t-1})$. Since bill rates and inflation have been tracking each other closely, such a model would have the possibility of explaining not only the long-run success of common stock returns as an inflation hedge, but their short-term failure in the last decade.

References

Ibbotson, Roger G, and Sinquefield, Rex A. 1976. Stocks, bonds, bills, and inflation: simulations of the future. *Journal of Business* 49, no. 3 (July): 313–37.

Ibbotson, Roger G, and Sreenivasan, Aruna. 1977. The portfolio mix: a multiperiod investment planning model. *Proceedings of the Center for Research in Security Prices* (May).

Michaud, Richard. 1976. Risk and compound return. *Proceedings of the Center for Research in Security Prices* (May).