Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence
Author(s): Marjorie A. Flavin
Source: *The Journal of Political Economy*, Vol. 91, No. 6 (Dec., 1983), pp. 929-956
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/1831198
Accessed: 21/08/2009 15:20

Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence

Marjorie A. Flavin

University of Virginia

Numerous authors, including Shiller, LeRoy and Porter, and Singleton, have reported empirical evidence that stock prices and long interest rates are more volatile than can be justified by standard asset-pricing models. This paper shows that in small samples the "volatility" or "variance-bounds" tests tend to be biased, often severely, toward rejection of the null hypothesis of market efficiency. Thus the apparent violation of market efficiency may be reflecting the sampling properties of the volatility measures, rather than a failure of the market efficiency hypothesis itself. The paper also reports some unbiased estimates of the bounds on holding period yields and long interest rates. Much of the evidence of excess volatility disappears when the tests are corrected for small sample bias.

In recent papers, Shiller (1979) and LeRoy and Porter (1981) have reported empirical evidence that stock prices and long interest rates are more volatile than can be justified within the standard asset-pricing models. Further empirical evidence on excess volatility in the financial markets has been reported in numerous studies, including Pesando (1979), Amsler (1980), Singleton (1980), Grossman and Shil-

I am grateful for comments and suggestions received during workshops at the Board of Governors, Brown, Chicago, CUNY, Duke, Michigan, University of Rochester, and Virginia Polytechnic Institute. For helpful discussions and comments on earlier drafts I also thank Richard Ashley, Robert Barro, Edwin Burmeister, Matthew Cushing, Wake Epps, Robert Hall, James Hamilton, Lars Hansen, Stephen LeRoy, Ronald Michener, Richard Porter, Robert Shiller, Kenneth Singleton, Jonathan Skinner, James Stock, William Poole, and Charles Whiteman. Earlier versions of the paper were circulated under the title "Small Sample Bias in Tests of Excess Volatility in the Financial Markets."

[*Journal of Political Economy*, 1983, vol. 91, no. 6]

ler (1981), Shiller (1981*a*, 1981*b*, 1981*c*), and Blanchard and Watson (1982). According to the empirical evidence reported in these papers, the variance of stock prices, holding yields on long-term bonds, and long interest rates exceed the upper bounds implied by the variance of dividends and short interest rates. Further, the variances of stock prices and long interest rates exceed their estimated upper bounds by very large margins in many cases.

This paper argues that in small samples the "volatility" or "variance-bounds" tests tend to be strongly biased toward rejection of the null hypothesis of no excess volatility. Thus the apparent violation of the market efficiency hypothesis may be reflecting the sampling properties of the volatility measures in small samples rather than a failure of the market efficiency hypothesis.

The innovative tests developed by Shiller and LeRoy and Porter are formulated according to the following line of reasoning. If stock prices are modeled as the present discounted value of rationally forecasted future dividends, the volatility, or variance, of the stock price is limited by the volatility, or variance, of the dividend series. Similarly, under the expectations theory of the term structure of interest rates, which asserts that the long-term interest rate is equal to an average of rationally expected future short-term interest rates, the variance of the long rate is limited by the variance of the short rate. The upper bound on the volatility of long rates, or stock prices, has been tested either (1) by comparing a point estimate of the upper bound with a point estimate of the variance being bounded or (2) by calculating both point estimates and the asymptotic covariance matrix of the estimates and testing whether the estimated variance of long rates or stock prices exceeds the estimated upper bound by an amount that is statistically significantly greater than zero. In either procedure, the test statistics may be misleading if, for samples of the size typically used in the variance-bounds tests, the point estimates are biased or, more generally, if the asymptotic distributions are not close approximations to the finite sample distributions. This paper argues that the estimate of the upper bound in these tests is biased downward in small samples and that the magnitude of the bias is large enough to provide a potential explanation of the apparent violation of the bounds.

To see intuitively why the variance-bounds tests tend to be biased against the null hypothesis, consider the basic bound on the volatility of long interest rates: $\text{var}\,(R_t) < \text{var}\,(R_t^*)$, where R_t is the actual long rate and R_t^* is the perfect-foresight long rate, defined as the value the long rate would take if agents had perfect foresight concerning the path of the short rate. Under the market efficiency hypothesis, R_t is equal to the expectation of R_t^* conditional on currently available information and therefore must have a variance smaller than the variance of R_t^*. If the population means of R_t and R_t^* were known a priori, unbiased estimates of var (R_t) and var (R_t^*) could be obtained by taking squared deviations of R_t and R_t^* from their population means. The empirical applications of the variance-bounds tests have relied on sample variances of R_t and R_t^* that were computed by taking deviations from sample means. Taking deviations from the sample mean induces downward bias in the sample variance, however, since the sample mean has the following property: the sample variance of a data series, expressed in deviations from some constant, is minimized when that constant is set equal to the sample mean. The greater the variance of the sample mean, the greater is the extent to which the sample mean will "fit" some of the stochastic components of the data series and the greater is the bias in the sample variance. Because R_t^* is a long moving average of a variable (the short rate), which is itself highly serially correlated, the variance of $\bar{R}_t^*$ tends to exceed the variance of $\bar{R}_t$, and as a result the sample variance of R_t^* tends to be more strongly downward biased than the sample variance of R_t. Since var (R_t^*) is the upper bound on var (R_t), the net effect is that the difference $\hat{\text{var}}$ (R_t^*) − $\hat{\text{var}}$ (R_t) is biased toward rejection of the null hypothesis of no excess volatility. This bias toward rejection of the null hypothesis also arises in tests of the upper bound on the variance of stock prices and the variance of holding period yields on long-term bonds.

Section I considers an economy in which the short rate is generated by an AR1 process with the autoregressive parameter equal to 0.95 for quarterly observations. Investors are risk neutral and form expectations of future short rates rationally, with the result that yields on 20-year discount bonds are generated exactly as hypothesized by the pure (i.e., no liquidity premium) expectations hypothesis. The exact finite sample distributions of the sample statistics, $\hat{\text{var}}$ (R_t), $\hat{\text{var}}$ (R_t^*), and $\hat{\text{var}}$ (R_t^*) − $\hat{\text{var}}$ (R_t), are calculated for a sample of size 100 in such an economy.

In Section II, some of the test procedures implemented in Shiller (1979, 1981*b*) and Singleton (1980) are reviewed in light of the findings concerning the small sample distributions of the variance-bounds statistics. Depending on the bound being tested and the estimation method used, the bias toward rejection of no excess volatility ranges from modest to strong to severe. Section II also reports some unbiased estimates of the bounds on the variances of holding period yields and long interest rates. Much of the evidence of excess volatility in the bond market disappears when the tests are corrected for small sample bias.

I. Comparison of Small Sample and Asymptotic Distributions in an Efficient-Market Economy

In order to keep the problem as simple as possible, the model economy studied in this section is one in which the short rate follows a first-order autoregressive (AR1) process:

$$r_t = \rho r_{t-1} + \epsilon_t, \tag{1}$$

where r_t is the short-term interest rate, expressed in deviations from the mean, and ϵ_t is an independently and identically distributed disturbance; $\epsilon \sim N(0, \sigma_\epsilon^2)$.

According to the expectations hypothesis of the term structure, the linearized long rate on a pure discount bond is simply the average of current and future expected short rates:

$$R_t = \frac{1}{n}\sum_{j=0}^{n-1} {}_t r_{t+j}, \tag{2}$$

where R_t is the n-period long rate and ${}_t r_{t+j}$ is the expectation, in period t, of r_{t+j}. Note that the long rate does not include a liquidity premium.

Using the assumption that the short rate follows an AR1 process, we see that all expected future short rates are proportional to the current short rate:

$${}_t r_{t+j} = \rho^j r_t = \rho^j \sum_{i=0}^{\infty} \rho^i \epsilon_{t-i}. \tag{3}$$

Thus the long rate is also proportional to the short rate:

$$R_t = \frac{1-\rho^n}{n(1-\rho)} r_t = \frac{1-\rho^n}{n(1-\rho)} \sum_{i=0}^{\infty} \rho^i \epsilon_{t-i}. \tag{4}$$

Define the "perfect-foresight" long rate, R_t^*, as the value the long rate would take if agents had perfect foresight concerning the path of the short rate:[1]

$$R_t^* = \frac{1}{n}\sum_{j=0}^{n-1} r_{t+j}. \tag{5}$$

By straightforward but tedious manipulation, the perfect-foresight long rate can be expressed as a linear combination of past, current, and future disturbances:

[1] Shiller uses the terminology "ex post rational" long rate in referring to this variable.

$$R_t^* = \frac{1}{n(1-\rho)} \sum_{j=1}^{n-1} (1 - \rho^{n-j})\epsilon_{t+j} + \frac{1-\rho^n}{n(1-\rho)} \sum_{i=0}^{\infty} \rho^i \epsilon_{t-i}. \quad (6)$$

Equation (6) reflects the basic hypothesis of efficiency in the bond market: the actual long rate, R_t, is the expectation of the perfect-foresight long rate, R_t^*, conditional on all available information.

Substituting equation (4) into equation (6) yields

$$R_t^* = R_t + \theta_t, \quad (7)$$

where

$$\theta_t = \frac{1}{n(1-\rho)} \sum_{j=1}^{n-1} (1 - \rho^{n-j})\epsilon_{t+j}.$$

R_t, which depends only on current and past disturbances, and θ_t, which depends only on future disturbances, are distributed independently, with the implication that

$$\text{var}(R_t^*) = \text{var}(R_t) + \text{var}(\theta_t). \quad (8)$$

Since the variance of the forecast error must be nonnegative, the variance of R_t^* constitutes an upper bound on the variance of R_t:

$$\text{var}(R_t^*) \geq \text{var}(R_t). \quad (9)$$

The upper bound on the variance of the long rate, equation (9), is, of course, a restriction on the population moments of R_t^* and R_t. Assuming that r_t, and therefore R_t^* and R_t, are stationary and ergodic time-series processes, the population variances of R_t^* and R_t can be consistently estimated from a single realization of the process over time.[2]

The upper bound on the volatility of long rates, or stock prices, has been tested either by comparing point estimates of var (R_t^*) and var (R_t) or by calculating both point estimates of var (R_t^*) and var (R_t) and the asymptotic covariance matrix of the estimates and testing whether the difference $\hat{\text{var}}(R_t^*) - \hat{\text{var}}(R_t)$ is significantly less than zero. In either procedure, the test statistics may be misleading if, for samples of the size typically used in the variance-bounds tests, the point estimates are biased or, more generally, if the asymptotic distributions are not close approximations to the finite sample distributions.

[2] If the short rate is nonstationary (i.e., if $\rho \geq 1$), the variances of r_t, R_t, and R_t^* are undefined and the theoretical variance bounds must be reformulated. Almost all of the empirical volatility literature, including LeRoy and Porter (1981), Shiller (1979, 1981*b*), and Singleton (1980), has been based on the assumption that the short rate (or dividends) is a stationary process.

This section of the paper studies the properties of the variance-bounds statistics in samples of 100 quarterly observations on the yields of 20-year discount bonds and 3-month bills. These observations are assumed to be drawn from an efficient-market economy in which the short rate is generated by an AR1 process with $\rho = 0.95$. The exact small sample distributions that the sample statistics, $\hat{\text{var}}\,(R_t^*)$, $\hat{\text{var}}\,(R_t)$, and $\hat{\text{var}}\,(R_t^*) - \hat{\text{var}}\,(R_t)$, would have in such an economy are then calculated and compared to the asymptotic distributions.

Calculation of the Small Sample Distributions[3]

In order to avoid having to refer to "the variance of the variance of R_t," the following notation will be used: $V = \hat{\text{var}}\,(R_t)$, $V^* = \hat{\text{var}}\,(R_t^*)$, and $D = \hat{\text{var}}\,(R_t^*) - \hat{\text{var}}\,(R_t)$.

In equations (1)–(6), the first observation on the short rate, r_1, was expressed as a function of disturbances from the infinite past. For the purpose of calculating the small sample distributions, it is more convenient to model the first observation on the short rate as a random draw from the stationary distribution of the short rate:

$$r_1 = \frac{\epsilon_1}{\sqrt{1 - \rho^2}}, \tag{10}$$

where $\epsilon_1 \sim N(0, \sigma_\epsilon^2)$. By modeling r_1 as a drawing from the stationary distribution of the short rate, a sample of T observations on r can be expressed as a function of T disturbances rather than an infinite number. For the purpose of characterizing the distributions of V, V^*, and D, the stochastic specification of r_1 given by equation (10) is completely equivalent to the specification of r_1 as a function of disturbances from the infinite past.

With this modification, each of the random variables V, V^*, and D can be expressed as a quadratic form in the disturbance of the short-rate process, $\boldsymbol{\epsilon}$. To construct the quadratic form representing the variance of the long rate, V, recall that in this example the long rate is proportional to the short rate,

$$R_t = \alpha r_t, \tag{11}$$

where $\alpha = (1 - \rho^n)/n(1 - \rho)$. The vector of T observations on the short rate can be expressed as a linear transformation of the disturbances:

[3] I am grateful to Robert Hall for suggesting this approach for computing the exact finite sample distributions of the variances.

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \cdot \\ \cdot \\ \cdot \\ r_T \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{1-\rho^2}} & 0 & 0 & 0 & \dots & 0 \\ \frac{\rho}{\sqrt{1-\rho^2}} & 1 & 0 & 0 & \dots & 0 \\ \frac{\rho^2}{\sqrt{1-\rho^2}} & \rho & 1 & 0 & \dots & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ \frac{\rho^{T-1}}{\sqrt{1-\rho^2}} & \rho^{T-2} & \rho^{T-3} & \rho^{T-4} & \dots & 1 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_T \end{bmatrix}. \quad (12)$$

Using the notation S (for short rate) for the $T \times T$ matrix in equation (12) and $\boldsymbol{\epsilon}$ for the T-element column vector of disturbances, the sample variance of the long rate, V, can be expressed as a quadratic form in $\boldsymbol{\epsilon}$,

$$V = \boldsymbol{\epsilon}' A \boldsymbol{\epsilon}, \quad (13)$$

where A is the $T \times T$ symmetric matrix, $A = \alpha^2 T^{-1} S'S$. At this point, the mean of the short rate, which is also the mean of the long rate, is assumed to be known a priori; the quadratic form A represents the variance of R_t around the population mean. The variance of R_t around its sample mean will be studied later.

The quadratic form representing the sample variance of the perfect-foresight long rate, V^*, will be of order $T + n - 1$, where n is the number of periods in the long rate, because the last observation on R_t^* depends on r_T and $n - 1$ subsequent observations on the short rate. Let L denote the $T \times T + n - 1$ matrix that transforms the $T + n - 1$ observations on the short rate into T observations on the perfect-foresight long rate:

$$\begin{bmatrix} R_1^* \\ R_2^* \\ R_3^* \\ \cdot \\ \cdot \\ \cdot \\ R_T^* \end{bmatrix} = \frac{1}{n} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 1 & 1 & \dots & 1 & 0 & \dots & 0 \\ \cdot & & & & & & & & & \\ \cdot & & & & & & & & & \\ \cdot & & & & & & & & & \\ 0 & 0 & 0 & \dots & 0 & 1 & 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \cdot \\ \cdot \\ \cdot \\ r_{T+n-1} \end{bmatrix}, \quad (14)$$

where the width of the band of ones is n.

The sample variance of R_t^* around its population mean can be expressed as the quadratic form of order $T + n - 1$,

$$V^* = \boldsymbol{\epsilon}' B \boldsymbol{\epsilon}, \tag{15}$$

where $B = T^{-1}S'L'LS$; S is the lower-triangular matrix defined as in equation (12), except with order $T + n - 1$ instead of T, and $\boldsymbol{\epsilon}$ is the $T + n - 1$ element column vector of disturbances.

The difference between the variances, $D = V^* - V$, is given by the difference of the quadratic forms for V^* and V:

$$D = \boldsymbol{\epsilon}'[B - A]\boldsymbol{\epsilon}. \tag{16}$$

(Of course the $T \times T$ matrix A as defined above must be augmented by adding $n - 1$ rows and $n - 1$ columns of zeros so that it conforms with the matrix B.)

The problem now becomes one of calculating the distribution of a quadratic form in normal deviates. Let Λ denote the diagonal matrix with the eigenvalues of the quadratic form A on the main diagonal, and P the matrix of eigenvectors; $P'AP = \Lambda$. Since $PP' = I$, $\boldsymbol{\epsilon}'A\boldsymbol{\epsilon} = \boldsymbol{\epsilon}'PP'APP'\boldsymbol{\epsilon} = \boldsymbol{\epsilon}'P\Lambda P'\boldsymbol{\epsilon}$. Define a new disturbance term $\boldsymbol{\eta}$ such that $\boldsymbol{\eta} = P'\boldsymbol{\epsilon}$. Since P is an orthonormal matrix, the new disturbances are independently distributed, $\eta \sim N(0, \sigma_\epsilon^2)$, with the same variance as the original disturbance, ϵ. Thus the sample variance of the long rate, V, is a weighted sum of squared normal deviates:

$$V = \boldsymbol{\epsilon}'A\boldsymbol{\epsilon} = \boldsymbol{\eta}'\Lambda\boldsymbol{\eta} = \sum_{j=1}^{T} \lambda_j \eta_j^2, \tag{17}$$

where λ_j are the eigenvalues of the quadratic form A. The characteristic function of $\Sigma_{j=1}^{T} \lambda_j \eta_j^2$ is

$$\phi(t) = \prod_{1}^{T}(1 - 2i\lambda_j\sigma_\epsilon^2 t)^{-1/2}. \tag{18}$$

By inverting the characteristic function, one can obtain the cumulative distribution function of the random variable, V. The value of the distribution function, evaluated at x, is given by

$$F(x) = 1/2 - \frac{1}{\pi}\int_0^\infty t^{-1}I[e^{-itx}\phi(t)]dt, \tag{19}$$

where $I[\cdot]$ denotes the imaginary part of the expression in the brackets.

All of the small sample distributions reported in the next section were computed by the following procedure: (1) The symmetric matrix defining the quadratic form was generated, after assigning numerical values to the parameters T, n, and ρ; (2) the eigenvalues of the matrix were obtained using numerical methods; and (3) the inversion formula (eq. [19]), which is a function of the eigenvalues, was integrated numerically.[4]

Asymptotic Distributions

Unlike the small sample distributions, the asymptotic distributions can be derived analytically. Let V, V^*, and D denote the sample statistics calculated by taking deviations from the population mean, and V_s, V_s^*, and D_s denote the corresponding statistics computed by taking deviations from the sample mean. The bias induced in V_s, V_s^*, and D_s by taking deviations from the sample mean is of order $1/T$ (Anderson 1971, p. 463). Similarly, $T \cdot \text{var}(V_s)$, $T \cdot \text{var}(V_s^*)$, and $T \cdot \text{var}(D_s)$ differ from $T \cdot \text{var}(V)$, $T \cdot \text{var}(V^*)$, and $T \cdot \text{var}(D)$, respectively, by terms of order $1/T$ (Anderson 1971, p. 471). Thus, in deriving the means and variances of the asymptotic distributions of V_s, V_s^*, and D_s, we can analyze the simpler case in which the statistics are calculated by taking deviations from the population mean.

Using equations (4) and (6) for R_t and R_t^*, respectively, we see that straightforward calculation of the means of the asymptotic distributions of V, V^*, and D yields

$$\mu_V = E(R_t^2) = \frac{\alpha^2\sigma_\epsilon^2}{1-\rho^2},$$

$$\mu_{V^*} = E(R_t^{*2}) = \left(\frac{\alpha^2}{1-\rho^2} + \sum_{i=1}^{n-1} \alpha_i^2\right)\sigma_\epsilon^2, \tag{20}$$

$$\mu_D = E[(R_t^* - R_t)^2] = \sigma_\epsilon^2 \sum_{i=1}^{n-1} \alpha_i^2,$$

where $\alpha = (1 - \rho^n)/n(1 - \rho)$ and $\alpha_i = [1 - \rho^{(n-i)}]/n(1 - \rho)$.

Singleton (1980) showed that the sample statistics V, V^*, and D are consistent estimators of μ_V, μ_{V^*}, and μ_D. Further, V, V^*, and D are asymptotically normally distributed, with variances given by

$$\lim_{T\to\infty} T \text{ var}(V) = 2\sum_{s=-\infty}^{\infty} [E(R_t R_{t+s})]^2,$$

$$\lim_{T\to\infty} T \text{ var}(V^*) = 2\sum_{s=-\infty}^{\infty} [E(R_t^* R_{t+s}^*)]^2, \tag{21}$$

$$\lim_{T\to\infty} T \text{ var}(D) = 2\sum_{s=-\infty}^{\infty} [E(\theta_t \theta_{t+s})]^2.$$

(Recall that $\theta_t = R_t^* - R_t$.)

[4] The eigenvalues were computed using the International Math and Science Library (IMSL) routine EIGRS; the numerical integration was performed using IMSL routine DCADRE.

Evaluating these variances for the model posited in this paper gives

$$\operatorname{var}(V) = \frac{2\alpha^4(1+\rho^2)}{T(1-\rho^2)^3}\,\sigma_\epsilon^4,$$

$$\begin{aligned}\operatorname{var}(V^*) = \frac{2}{T}\Bigg\{&\left(\frac{\alpha^2}{1-\rho^2} + \sum_{i=1}^{n-1}\alpha_i^2\right)^2 \\ &+ 2\sum_{j=1}^{n-2}\left(\sum_{i=1}^{n-1-j}\alpha_i\alpha_{i+j} + \frac{\rho^j\alpha^2}{1-\rho^2}\right. \\ &\left.+ \sum_{i=1}^{j}\alpha\alpha_i\rho^{j-1}\right)^2 \\ &+ \frac{2}{1-\rho^2}\left[\frac{\rho^{(n-1)}\alpha^2}{1-\rho^2} + \sum_{i=1}^{n-1}\alpha\alpha_i\rho^{(n-1-i)}\right]^2\Bigg\}\,\sigma_\epsilon^4,\end{aligned} \tag{22}$$

$$\operatorname{var}(D) = \frac{2}{T}\left[\left(\sum_{i=1}^{n-1}\alpha_i^2\right)^2 + 2\sum_{j=1}^{n-2}\left(\sum_{i=1}^{n-1-j}\alpha_i\alpha_{i+j}\right)^2\right]\sigma_\epsilon^4.$$

Comparison of Small Sample and Asymptotic Distributions

Recall that the numerical example was constructed to mimic quarterly data in which the short rate was a 3-month rate and the term of the long rate was 20 years. The autoregressive parameter of the short-rate process, ρ, was set at 0.95.[5] For the small sample distributions, the sample size, T, was set at 100 (quarterly) observations. Table 1 reports the means of the asymptotic distributions of V, V^*, and D, for $\rho = 0.95$, $T = 100$, and $n = 80$. The variance of the short-rate innovation is normalized at one ($\sigma_\epsilon^2 = 1$).

The asymptotic standard deviations reported in table 1 were obtained by evaluating the expressions (22) for the asymptotic variances of V, V^*, and D for a sample size of 100. Before turning to the calculations of the actual small sample distributions, it should be noted that table 1 itself contains some evidence that the asymptotic distributions are not close approximations to the small sample distributions for samples of 25 years of quarterly data. Because V and V^* are both sample variances, neither random variable can take on negative values. Using asymptotic distribution theory to approximate the distribution of V^*, however, one would conclude that V^* is normally distributed with mean 3.802 and standard deviation 4.619, implying that V^* is less than zero for over 20 percent of its distribution.

[5] In a first-order autoregression of quarterly observations on 3-month Treasury bill yields (sample period 1950:I–1982:I), the estimated autoregressive parameter was 0.953, with a standard error of .03. (These data were obtained from Salomon Brothers, Inc. 1982.)

TABLE 1

MEANS AND STANDARD DEVIATIONS OF ASYMPTOTIC DISTRIBUTIONS

Variable	Mean	Asymptotic Standard Deviation
V^*	3.802	4.619
V	.620	.150
D	3.182	3.139

The actual small sample distributions of V, V^*, and D are plotted in figure 1. In each of the three panels of figure 1, the distributions labeled *a* represent the small sample distributions of V^*, V, and D, assuming that the mean of the short-rate process is known a priori. The distributions labeled *b* represent the small sample distributions of V_s^*, V_s, and D_s when R_t and R_t^* are each expressed in deviations from their respective sample means instead of the population mean. The distributions labeled *c* in panels 1 and 3 represent the small sample distributions of V_s^* and D_s, respectively, when R_t^* is calculated using a terminal condition R_T^* and both R_t^* and R_t are expressed in deviations from the sample mean. The distributions labeled *d* represent the small sample distributions of spectral estimates. The distributions *b*, *c*, and *d* are explained more fully below.

Consider the distributions labeled *a*. If the mean of the underlying process is known, the sample variance is an unbiased estimate of the population variance (Anderson 1971, p. 448).[6] However, even when the mean is known, the sample variances are not closely approximated by the normal distribution. All three random variables have strongly skewed small sample distributions; the probability that V will take on a value less than its mean is 60 percent, and V^* and D each have a 65 percent probability of taking on values less than their respective means.

It is important to keep in mind that two unrealistically strong assumptions concerning the information available to the econometrician have been maintained in computing the small sample distributions represented by the *a* curves. First, the mean of the short-rate process has been assumed to be known a priori. Second, the perfect-foresight variance V^* has been calculated assuming that all of the $n - 1$ postsample observations on the short rate, r_{T+1} to r_{T+n-1}, are available, enabling the econometrician to construct the perfect-foresight

[6] The means of the small sample distributions of V, V^*, and D were calculated making use of the fact that the mean of a quadratic form in normal deviates is equal to the sum of the eigenvalues of the quadratic form. For each of the three quadratic forms, the sum of the eigenvalues matched the analytically derived population mean (reported in table 1) to four decimal places.

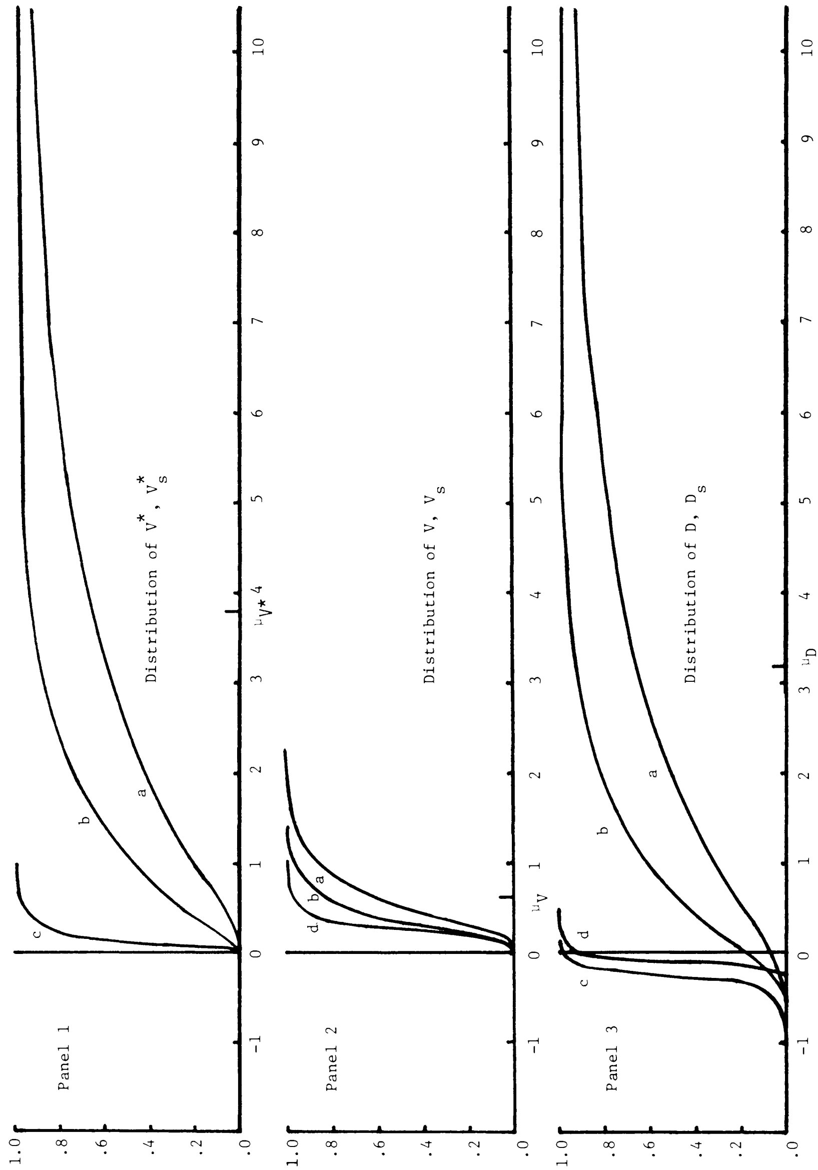

FIG. 1.—Small sample distributions

long-rate series, R_1^* to R_T^*, without having to resort to any form of extrapolation of the short-rate data. Even under these unrealistically favorable assumptions concerning the availability of prior information and data, there is a 6.6 percent chance that the sample variance, V, will exceed its upper bound, V^*, if the null hypothesis of market efficiency is true.

In practice, the variance-bounds tests have been implemented using data on r_t and R_t in deviations from their sample means rather than deviations from population means. Let A_s denote the quadratic form representing the variance of R_t expressed in deviations from its sample mean.[7] The distribution of $V_s = \epsilon' A_s \epsilon$ is given by curve b in panel 2 of figure 1. When the population mean of R_t is not known a priori and the sample variance is expressed in deviations from the sample mean, the sample variance is a downward biased estimator of the population mean of var (R_t); the mean of V_s is 0.4251, as compared to the population mean of var (R_t) of 0.6200.

Similarly, the curve labeled b in panel 1 is the small sample distribution of V_s^* in which the vector of observations on the short rate is expressed in deviations from the sample mean before constructing the series on R_t^*. Again, expressing the data in deviations from the sample mean creates a downward bias: the mean of the distribution b is 1.537, less than half the value of the population moment, var (R_t^*), of 3.802. Because V_s^* is more strongly downward biased than V_s, expressing the data in deviations from sample means results in a net downward bias to $D_s = V_s^* - V_s$. When the data are expressed in deviations from sample means, D_s has a mean of 1.112, as compared to a population value of var (R_t^*) − var (R_t) of 3.182. Further, there is a 16.8 percent chance that the sample variance, V_s, will exceed its upper bound, V_s^*, if the null hypothesis of market efficiency is true.

II. Review of Previous Tests of Excess Volatility

In this section some of the test procedures implemented by Shiller and Singleton are reviewed in light of the findings concerning the small sample distributions of the variance-bounds test statistics.

[7] To construct the quadratic form A_s, take the matrix S and calculate the sum of the elements in each column. Denote the sum of the elements in the jth column as $m(j)$. Subtract $m(j)/100$ from each element of the jth column of the original matrix S, for $j = 1, 2, \ldots, 100$. Premultiplying this matrix by its transpose and multiplying by the scalar $\beta^2 T^{-1}$ gives the quadratic form A_s. In modifying A to form A_s, the degrees of freedom correction is automatically incorporated into A_s, since the rank of the quadratic form is reduced by one by the modification. Thus A has T nonzero eigenvalues, while A_s has $T - 1$ nonzero eigenvalues.

Shiller's Approach

In his empirical work on the volatility of stock prices and long interest rates (Shiller 1979, 1981*a*, 1981*b*, 1981*c*; Grossman and Shiller 1981), Shiller not only examines the basic upper bound on the variance of the long interest rate or stock prices, var $(R_t) <$ var (R_t^*), but also formulates and presents empirical evidence on an upper bound on the variance of holding period yields.

In illustrating the apparent excess volatility of long interest rates in his 1979 paper, Shiller graphs actual AAA utility bond yields against a perfect-foresight long rate constructed from data on the 4–6-month prime commercial paper rate. In these graphs, the perfect-foresight long rate moves smoothly and remains within the band between 6.25 percent and 6.75 percent, while the actual long rate moves sharply and varies between 4.5 percent and 11.5 percent over the same period (1966:I–1977:III). Since the variance of the perfect-foresight long rate places an upper bound on the variance of the actual long rate, these graphs do appear to "stand in glaring contradiction" (p. 1213) to the efficient markets model.

In the absence of actual data on the postsample values of the short rate, Shiller computed the perfect-foresight long-rate series, R_t^*, recursively from an assumed terminal value, R_T^*,

$$R_t^* = \gamma R_{t+1}^* + (1 - \gamma) r_t, \tag{23}$$

where γ is a constant close to but less than one.[8] In the case of a pure discount bond, $\gamma = (n - 1)/n$. The terminal value, R_T^*, was assumed to be equal to the average short rate over the sample period. It is a simple matter to grind out the distribution of Shiller's approximation to the perfect-foresight long rate using the methods of the previous section. Let $\tilde{R}_t^*$ denote Shiller's approximation to R_t^* and $\tilde{V}^*$ the sample variance of $\tilde{R}_t^*$ expressed in deviations from the sample mean; $\tilde{V}^* = \boldsymbol{\epsilon}'\tilde{B}_s\boldsymbol{\epsilon}$, where $\tilde{B}_s$ is a symmetric matrix of order 100.[9]

[8] The parameter γ arises in Shiller's linearization of the basic term structure equation relating the long rate to future expected short rates. If a coupon bond is selling near par, the mean of the long rate ($\bar{R}$) will be approximately equal to the coupon rate (C). Shiller takes Taylor series expansions around $R_t = \bar{R} = C$ and $r_{t+j} = \bar{R} = C$, $j = 0, 1, \ldots, n - 1$, to obtain $R_t = [(1 - \gamma)/(1 - \gamma^n)] \Sigma_{j=0}^{n-1} \gamma^j {}_t r_{t+j}$, where $\gamma = 1/(1 + R_0)$ and R_0 is the point around which the equation is linearized. In practice, Shiller sets $R_0 = \bar{R}$ and thus linearizes around the mean of the long rate. For a pure discount bond, $C = 0$ and $\gamma = 1$ so that the long rate is a simple (unweighted) average of future short rates in the linearized model: $R_t = (1/n) \Sigma_{j=0}^{n-1} {}_t r_{t+j}$ (see Shiller 1979, pp. 1194–99).

[9] The quadratic form representing V^* was generated in the following way. The 100×100 matrix, $\tilde{L}$, which transforms the 100 observations on the short rate into 100 observations on R_t^*, was constructed by setting its first 21 rows equal to the first 21 rows of the previously defined matrix L. (This reflects the fact that the first 21 observations on R_t^* do not depend on the assumed terminal value $\tilde{R}_T^*$.) The assumption that R_T^* is equal to the average short rate over the sample period is imposed by setting each

The distribution of the quadratic form $\epsilon'\tilde{B}_s\epsilon$, which represents the sample variance of the perfect-foresight long rate constructed using Shiller's assumption concerning the terminal value of R_t^*, is given by the curve c in panel 1 of figure 1. The small sample variance $\epsilon'\tilde{B}_s\epsilon$ has a mean of 0.1521 and a zero probability of taking on values greater than 1.1. As illustrated in panel 1, Shiller's method of obtaining an approximate series for R_t^*, when applied to this numerical example, results in an estimated variance of R_t^* that is severely biased downward. Not only is the expectation of $\epsilon'\tilde{B}_s\epsilon$ far below the population mean of var (R_t^*) of 3.802, there is a zero probability that $\epsilon'\tilde{B}_s\epsilon$ will take on a value even one-third the value of the population mean of var (R_t^*).

Curve c in panel 3 shows the distribution of $\epsilon'[\tilde{B}_s - A_s]\epsilon$, which represents the difference between Shiller's approximation to the variance of R_t^* and the variance of R_t, both expressed in deviations from sample means. The measure of the perfect-foresight variance is more strongly biased than the sample variance of R_t, and the difference between the sample variances is negative throughout 99.9 percent of its distribution. The mean of $\epsilon'[\tilde{B}_s - A_s]\epsilon$ is -0.2729, as compared to the population mean of the difference var (R_t^*) − var (R_t) of 3.182. Thus, even though markets are efficient in this example and the *population* variance of R_t^* is several times the population variance of R_t, *estimating* the variance of R_t^* by imposing the terminal condition that R_T^* equal the sample mean of the short rate induces so much downward bias that the sample variance of R_t exceeds its estimated upper bound with probability .999.

It is important to point out that Shiller's 1979 paper uses the constructed variable $\tilde{R}_t^*$ only for the purpose of illustrating the notion of excess volatility of long interest rates; none of his formal statistical tests of market efficiency in the bond market use the constructed variable. In his subsequent paper addressing the volatility of stock prices (1981*b*), however, Shiller does use a perfect-foresight stock price variable that parallels $\tilde{R}_t^*$ in its construction. That is, the perfect-foresight stock price variable is constructed by assuming a terminal value equal to the sample mean of the (detrended) actual stock price

element of the last (100th) row equal to 0.01. Using $\tilde{L}(i, j)$ to denote the element in the ith row and the jth column of $\tilde{L}$, rows 99–22 were generated recursively by setting $\tilde{L}(i, j) = (79/80)\tilde{L}(i + 1, j)$ for $i \neq j$ and $\tilde{L}(i, j) = (79/80)\tilde{L}(i + 1, j) + (1/80)$ for $i = j$. The 100 observations on $\tilde{R}_t^*$ are given by $\tilde{R}_t^* = \tilde{L}S\epsilon$, where S is the previously defined matrix that transforms the vector of disturbances into the vector of realizations on the short rate. In order to reflect the fact that $\tilde{R}_t{}^*$ is expressed in deviations from the sample mean, take the matrix product $\tilde{L}S$ and calculate, for each column, the sum of the elements in the column. Denote the sum of the elements in the jth column as $m(j)$. Subtract $m(j)/100$ from each element of the jth column of the original matrix product $\tilde{L}S$, for $j = 1, 2, \ldots, 100$. Premultiplying this matrix by its transpose and dividing by T gives the matrix $\tilde{B}_s$.

series and solving backward recursively. While the numerical values in this paper were chosen with the bond market rather than the stock market in mind, the two problems are similar enough that the evidence of downward bias in Shiller's estimate of the perfect-foresight long rate may indicate that his estimate of the standard deviation of the perfect-foresight stock price index could be seriously downward biased as well.

The formal statistical evidence of excess volatility of long interest rates presented in Shiller's 1979 paper was based primarily on a comparison of the variance of the holding period return of a long-term bond with the variance of the short interest rate. As derived by Shiller, the linearized holding yield ($\tilde{H}_t$) is given by

$$\tilde{H}_t = \frac{R_t - \gamma_n R_{t+1}}{1 - \gamma_n}, \tag{24}$$

where $\gamma_n = \gamma(1 - \gamma^{n-1})/(1 - \gamma^n)$, γ is as previously defined, and n is the number of periods in the long-term bond. Directly from equation (24), var ($\tilde{H}_t$) can be expressed as a function of var (R_t) and cov (R_t, R_{t+1}). The cov (R_t, R_{t+1}) term is then substituted out to obtain an expression for var ($\tilde{H}_t$) in terms of var (R_t), var (r_t), γ_n, and ρ_{rR} (the correlation coefficient between R_t and r_t). The upper bound on var ($\tilde{H}_t$) is obtained by maximizing this expression with respect to var (R_t); thus the bound itself is not a function of var (R_t):

$$\max_{\text{var}(R_t)} \text{ var}(\tilde{H}_t) = \frac{\text{var}(r_t)\rho_{rR}^2}{1 - \gamma_n^2}. \tag{25}$$

Since ρ_{rR}^2 must be less than one, the variance of the short rate places an upper bound on the variance of $\tilde{H}_t$. Shiller's basic inequality restriction is

$$\sigma(\tilde{H}_t) \leq a\sigma(r_t), \tag{26}$$

where $a = 1/\sqrt{1 - \gamma_n^2}$.

Using the fact that $\tilde{H}_t$ is approximately serially uncorrelated (and assuming that $\tilde{H}_t$ is normally distributed), Shiller uses the χ^2 distribution to compute a one-sided 95 percent confidence interval for the sample statistic $\hat{\sigma}(\tilde{H}_t)$. A lower bound on $\hat{\sigma}(\tilde{H}_t)$, denoted $\sigma_m(\tilde{H}_t)$, is then calculated from the confidence interval. Since the small sample distribution of the estimated standard deviation of the short rate, $\hat{\sigma}(r_t)$, is not known, Shiller does not conduct a formal statistical test of the hypothesis that the standard deviation of the holding period return satisfies the upper bound in equation (26). However, comparison of the point estimates of the standard deviation of the short rate with the lower bound on the holding period yield seems discouraging from the point of view of proponents of market efficiency. In four of the six data sets studied by Shiller, the lower bound on the variability of holding period yields, $\sigma_m(\tilde{H}_t)$, was twice as large as the point estimate of its upper bound. For the other two data sets, $\sigma_m(\tilde{H}_t)$ was narrowly within the estimated upper bound.[10]

To see that inequality (26) will tend to be biased toward rejection of the null hypothesis of market efficiency, note that under Shiller's assumptions an unbiased estimate of the variance of the holding period yield can be obtained simply by taking the sum of squares of the deviations of $\tilde{H}_t$ from its sample mean and dividing by degrees of freedom $(T - 1)$ rather than sample size (T). However, the short rate, r_t, is highly serially correlated, so that the same correction for degrees of freedom will not eliminate the downward bias in the sample variance of r_t. Recall that in the numerical example studied in this paper the actual long rate is proportional to the short rate. Thus the small sample distribution of the variable $V = \epsilon' A_s \epsilon$ characterizes the sample variances of either R_t or r_t, under different normalizations of the error variance, σ_ϵ^2. For the numerical values examined above, the downward bias was substantial even for a sample of 25 years worth of data; the sample standard deviation of r_t is $\sqrt{.4251/.6200} = 82.8$ percent of the population standard deviation of r_t.

Using the notation var (r_t) for the population variance of r_t and vâr (r_t) for the sample variance of r_t (computed by taking deviations from the sample mean and dividing by T), we denote the relative bias of vâr (r_t) by[11]

$$\frac{E[\hat{\text{var}}\,(r_t)]}{\text{var}\,(r_t)} = 1 - \frac{\text{var}\,(\bar{r}_t)}{\text{var}\,(r_t)}, \tag{27}$$

where var $(\bar{r}_t)$ is the variance of the sample mean of r_t. If the short rate is generated by an AR1 process, equation (27) can be evaluated by straightforward algebra:

$$\frac{E[\hat{\text{var}}\,(r_t)]}{\text{var}\,(r_t)} = 1 - \frac{1 + \rho}{(1 - \rho)T} + \frac{2\rho(1 - \rho^T)}{(1 - \rho)^2 T^2}. \tag{28}$$

[10] In the paper discussed above, Shiller's long-term interest rate data consisted of data on bonds with very long terms to maturity; in some data sets, the bonds were 25-year bonds, in other data sets, the bonds were consols. In a subsequent paper (1981*c*), Shiller reports the sample standard deviations of the 6-month Treasury bill rate and the holding period yield on medium term bonds (1-year–4.5-year Treasury notes). For the sample period 1955:II–1972:II, the sample standard deviation of the holding period yield did not exceed the point estimate of its upper bound for Treasury notes with 1 year or 1.5 years to maturity. For Treasury notes with 2–4.5 years to maturity, the sample standard deviation of the holding period yield did exceed the point estimate of the upper bound, but the violation was smaller in magnitude than the violations reported in Shiller (1979), based on the very long-term bonds.

[11] I am grateful to James H. Stock for suggesting the closed-form expressions (eqq. [27] and [28]) for the bias of the sample variance of the short rate.

TABLE 2

Data Set	T (1)	n (2)	γ_n (3)	ρ (4)	$\hat{\sigma}(r)$ (5)	$a\hat{\sigma}(r)$ (6)	k (7)	$\frac{a\hat{\sigma}(r)}{k}$ (8)	$\hat{\sigma}(\tilde{H})$ (9)	$\sigma_m(\tilde{H})$ (10)
1. U.S., quarterly 1966:I–1977:II	46	100	.978	.95	1.78	8.55	.70	12.21	19.5	16.5
2. U.S., monthly 1969:1–1974:1	61	288	.992	.983	1.77	14.03	.53	26.47	27.4	23.6
3. U.S., annual 1960–76	17	25	.925	.815	1.39	3.66	.79	4.63	9.82	7.65
4. U.S., annual 1919–58	40	25	.940	.815	1.86	5.44	.90	6.04	5.48	4.58
5. U.K., quarterly 1956:I–1977:II	86	∞	.980	.95	2.84	14.3	.81	17.65	34.4	30.4
6. U.K., annual 1824–1929	106	∞	.968	.815	1.17	4.66	.96	4.85	4.95	4.43

NOTE.—Explanation of symbols: T = sample size; n = number of periods in the long-term bond; γ_n = constant involved in the linearization of the model; ρ = autoregressive parameter in the short-rate process; $\hat{\sigma}(r)$ = estimated standard deviation of the short rate, calculated by taking deviations from the sample mean and dividing by $T - 1$; $a\hat{\sigma}(r)$ = estimated upper bound, where $a = 1/(1 - \gamma_n^2)^{1/2}$; k = relative bias of $\hat{\sigma}(r)$; $\hat{\sigma}(\tilde{H})$ = estimated standard deviation of the linearized holding period yield; $\sigma_m(\tilde{H})$ = critical value for the lower 5 percent tail of $\hat{\sigma}(\tilde{H})$, assuming that $\hat{\sigma}^2(\tilde{H})$ is distributed χ^2.

Using equation (28) we can calculate and correct for the bias to the upper bound of the holding period yield. The accuracy of the bias calculation depends, of course, on the validity of the assumption concerning the time-series process generating the short rate. Table 2 calculates the bias to the upper bound on the holding period yield for each of the six data sets studied in Shiller's 1979 paper. In calculating the bias, I retain the assumption that quarterly data on short-term interest rates (3–6-month maturities) are well approximated by an AR1 process with an autoregressive parameter of 0.95. I also assume that the AR1 parameter for monthly, as opposed to quarterly, observations on the short rate is $0.95^{1/3} = 0.983$ and that the AR1 parameter for annual observations is $(0.95)^4 = 0.815$.

Columns 2, 3, 5, 6, 9, and 10 of table 2 reproduce certain columns of table 1 in Shiller (1979).[12] Column 7 reports the value of k, the relative bias of $\hat{\sigma}(r)$; $E[\hat{\sigma}(r)] = k\sigma(r)$, which was calculated as

$$k = \left\{\left(\frac{T}{T-1}\right)\left[1 - \frac{1+\rho}{(1-\rho)T} + \frac{2\rho(1-\rho^T)}{(1-\rho)^2T^2}\right]\right\}^{1/2}. \tag{29}$$

As indicated by column 7, the small sample bias in $\hat{\sigma}(r)$ ranges from trivial for the data set with a sample period of 100 years ($k = 0.96$) to substantial for the monthly data set with a sample period of 5 years

[12] Shiller also reports a "tighter" upper bound on $\sigma(\tilde{H})$ as well as bounds that are applicable when r_t is nonstationary in his table 1.

(k = 0.53). Column 8 reports the estimated upper bound, corrected for small sample bias: $a\hat{\sigma}(r)/k$.

Without the bias correction, the upper bound on the volatility of holding period yields was violated on the basis of the sample statistics, that is, $\sigma_m(\tilde{H}) > a\hat{\sigma}(r)$, for four of the six data sets. Further, in the four data sets that violate the bound, $\sigma_m(\tilde{H})$ is roughly twice its estimated upper bound.

The bias correction changes the result of only one data set (data set 2) from violation to nonviolation of the bound. However, the three data sets that still violate the bound are not independent observations against the null hypothesis since they cover substantially the same historical period: each of the three contains the period 1966–76.

After the upper bound has been corrected for bias, the evidence of excessive volatility of holding period yields is considerably less dramatic: for three virtually nonoverlapping sample periods (U.S., 1969–74; U.S., 1919–58; and U.K., 1824–1929), the standard deviation of holding period yields is narrowly within the upper bound. The standard deviation of holding period yields exceeds the upper bound by a margin of 35–75 percent for the three data sets that contain the period 1966–76.[13]

Singleton's Approach

While Singleton (1980) reports some results based on holding period returns, his paper focuses primarily on the upper bound on the long interest rate:

$$\text{var}\ (R_t) \leq \text{var}\ (R_t^*). \tag{30}$$

Using spectral analysis, Singleton computes consistent estimates not only of var (R_t) and var (R_t^*) but also of the covariance matrix of the estimates of var (R_t) and var (R_t^*). Like Shiller, Singleton finds that his point estimates of var (R_t) exceed the point estimates of the upper bound, var (R_t^*). Further, Singleton conducts asymptotic tests of whether the variance of the long rate satisfies the upper bound. For

[13] When the bound on the volatility of holding period yields is applied to the stock market, Shiller finds that the estimated standard deviation of the holding period yield is more than five times the upper bound, even with a sample period of over one hundred years (Shiller 1981*b*). Considering the sample size, correcting for the bias induced by eliminating the sample mean will not substantially change the magnitude of the violation. However, the exponential trend in the stock price series had been removed from the stock price data as well as the dividend data in order to achieve stationarity. One would have to assess the biases potentially induced by the detrending in order to reliably interpret the strength of the evidence against the market efficiency hypothesis in the context of the stock market. Shiller emphasized that the results could be sensitive to detrending, stating that "assumptions about public knowledge or lack of knowledge of the long run growth path are important" (1981*b*, p. 421, n. 2).

each of the three data sets analyzed, the violation of the upper bound is statistically significant at the 5 percent level.

Singleton does not use postsample data on the short rate to literally construct a data series on the perfect-foresight long rate R_t^*. Instead, his estimates of var (R_t^*) are computed on the basis of observations on the short rate and on the theoretical relationship between the short rate and the perfect-foresight long rate. The perfect-foresight long rate in Shiller's linearized model is given by

$$R_t^* = \delta \sum_{s=0}^{n-1} \gamma^s r_{t+s}, \tag{31}$$

where γ is a constant (see n. 8), and $\delta = (1 - \gamma)/(1 - \gamma^n)$. Based on the known linear relationship between R_t^* and r_t,[14] the spectral density of R_t^* can be expressed as a function of the spectral density of r_t:

$$S_{R^*}(\lambda) = g^2(\lambda)S_r(\lambda), \quad |\lambda| \leq \pi, \tag{32}$$

where $g^2(\lambda) = \delta^2[1 - 2\gamma^n \cos(n\lambda) + \gamma^{2n}]/[1 + \gamma^2 - 2\gamma \cos(\lambda)]$ and $S_{R^*}(\lambda)$ and $S_r(\lambda)$ are the spectral density functions of R_t^* and r_t, respectively. The variance of R_t^* is equal to the integral of the spectral density of R_t^*:

$$\text{var } R_t^*) = \int_{-\pi}^{\pi} S_{R^*}(\lambda)d\lambda. \tag{33}$$

Singleton estimated the variance of the perfect-foresight long rate by estimating $S_r(\lambda)$ from the short-rate data, calculating the function $S_{R^*}(\lambda)$ implied by equation (32), and integrating the estimate of $S_{R^*}(\lambda)$ over the interval $-\pi$ to π. Before computing the spectral densities of r_t and R_t, Singleton removed the sample mean and sample (linear) trend from the data, which consisted of monthly observations over the sample period 1959:1–1971:6.

In his appendix, "Model Restrictions on the Spectral Densities of Interest Rates," Shiller discusses the theoretical relationship (eq. [32]) between the spectral densities of R_t^* and r_t in the limiting case in which the long bond is a consol. He notes that $g^2(\lambda)$ is equal to one at $\lambda = 0$ and declines monotonically as λ increases. Further, $g^2(\lambda)$ drops rapidly as λ increases for γ close to one.

The fact that $g^2(\lambda)$ declines monotonically implies that the low fre-

[14] For purposes of illustration, I assume that the parameters as well as the form of the relationship between R_t^* and r_t are known. In the context of the term structure of interest rates, treating γ as known is, in my view, a sensible interpretation of the model, since γ is defined as $\gamma = 1/(1 + R_0)$, where R_0 is the point around which the linearization is taken. In practice, both Shiller and Singleton set $R_0 = \bar{R}_t$ and linearize around the sample mean of the long rate. Empirical results reported below indicate that the results are perceptibly but not dramatically affected by varying the value of γ.

quency components of $S_r(\lambda)$ account for a greater proportion of the variance of R_t^* than of the variance of r_t. A sample mean and sample trend will tend to "fit" much of the low frequency movement in a small sample of time-series data. Thus, taking deviations from the sample mean and trend will bias a sample variance downward by underestimating the low frequency movements of the series. This downward bias in the sample variance of the short rate will be "amplified" by the filter function $g^2(\lambda)$ to create a (proportionally) greater downward bias in the estimate of var (R_t^*).[15] Elimination of the sample mean and sample trend will also create a downward bias in the sample variance of the actual long rate, R_t. However, because the spectrum of the actual long rate is much less concentrated at the low frequencies than the perfect-foresight long rate, the bias in the sample variance of R_t will tend to be smaller than the bias to the estimated variance of R_t^*.

The Appendix describes a procedure for constructing quadratic forms that represent the sample statistics calculated by applying the spectral estimators to 100 quarterly observations generated by the model economy specified in Section I.[16] The finite sample distributions of these quadratic forms are plotted in figure 1. The distribution of the spectral estimate of var (R_t^*) was so close to the distribution of Shiller's estimator that it could not be plotted distinctly from curve c in panel 1. The mean of the spectral estimate of var (R_t^*) was 0.183, as compared to the population value of var (R_t^*) of 3.802. The curve labeled d in panel 2 gives the small sample distribution of the spectral estimate of var (R_t), which has mean 0.288, as compared to the population value of var (R_t) of 0.62. The small sample distribution of the spectral estimate of the difference is given by curve d in panel 3. In the numerical example, μ_D = var (R_t^*) − var (R_t) = 3.2. However, because of the severe bias to the spectral estimate of var (R_t^*), the spectral estimate of D has mean −0.104 and has a 92.12 percent probability of taking on values less than zero. That is, the estimated variance of R_t exceeds the estimated upper bound 92 percent of the time, even though the null hypothesis of market efficiency holds, by construction, in the numerical example.

[15] The point that Singleton's estimate of the upper bound was biased downward by the detrending is stated in Shiller (1981c): "Perhaps [Singleton's] more dramatic results stem from his decision to subtract linear trends from the data, and in effect assume the trends were known by the market in advance. Any such assumption has the effect of reducing the uncertainty about future interest rates and thus reducing the permissible volatility of long rates according to the expectations model. Ultimately the inequality tests must hinge on our priors as to the reasonableness of such assumptions" (p. 76).

[16] Because the long rate in the numerical example is assumed to be the yield on a pure discount bond, the parameters in the linearized term structure relation (eq. [31]) can be specified a priori; γ = 1 and δ = $1/n$.

TABLE 3

ESTIMATED NONCENTRAL SECOND MOMENT OF R_t^* AND R_t

TERM OF LONG BOND	SAMPLE PERIOD	ESTIMATED NONCENTRAL SECOND MOMENT OF R_t^*	ESTIMATED NONCENTRAL SECOND MOMENT OF R_t
10 years	1950:I–1973:I	22.03	19.00
20 years	1950:I–1963:I	17.79	10.90

NOTE.—Noncentral second moments were calculated as $(1/T)\ \Sigma_{t=1}^{T} x_t^2$, where T denotes the number of observations.

Some Unbiased Estimates

The paper closes by reporting some unbiased estimates of the variance-bounds statistic var (R_t^*) − var (R_t). Following a suggestion of Richard Porter, the noncentral second moments of R_t^* and R_t were calculated. Assuming that the long rate does not contain a liquidity premium, R_t^* and R_t have the same mean, μ. Since

$$E(R_t^{*2}) = \text{var}\ (R_t^*) + \mu^2$$

and (34)

$$E(R_t^2) = \text{var}\ (R_t) + \mu^2,$$

the difference between the noncentral moments is an unbiased estimate of var (R_t^*) − var (R_t).

Data series on R_t^* for 10- and 20-year Treasury bonds were constructed from data on the 3-month Treasury bill rate using Shiller's linearized term structure relation:

$$R_t^* = \frac{1-\gamma}{1-\gamma^n} \sum_{s=0}^{n-1} \gamma^s r_{t+s}, \quad \gamma = \frac{1}{1+R_0}, \tag{35}$$

where $n = 40$ for the 10-year bond and $n = 80$ for the 20-year bond. The term structure equation was linearized around the point $R_0 = 0.01$. Data were available on the short rate for 1950:I–1982:IV.[17] In order to avoid using a terminal condition, the R_t^* series for the 10-year bond was calculated for t = 1950:I–1973:I and the R_t^* series for the 20-year bond was calculated for t = 1950:I–1963:I. The estimated noncentral second moments of R_t^* and R_t are reported in table 3.[18]

[17] The data were from Salomon Brothers, Inc. (1982).

[18] If the term structure relation is linearized around $R_0 = 0.02$ instead of 0.01, the sample noncentral second moment of R_t^* is 21.19 for the 10-year bond and 15.22 for the 20-year bond. With quarterly data, R_0 is a quarterly (nonannualized) interest rate.

According to table 3, the difference between the estimated non-central second moments of R_t^* and R_t, which is an unbiased estimate (assuming no liquidity premium) of var (R_t^*) − var (R_t), is 3.03 for the 10-year bonds and 6.89 for the 20-year bonds.[19] The empirical finding that var (R_t) is within the upper bound imposed by var (R_t^*) will not be reversed if the model is generalized to include a (constant) positive liquidity premium in the long rate. If R_t does contain a liquidity premium, the population mean of R_t would exceed the population mean of R_t^*. Thus the assumed absence of a liquidity premium biases the estimate of var (R_t^*) − var (R_t) downward.

The average of the squared observations on the short rate for the sample, 1950 : I–1982 : IV, was 34.53. Since r_t, R_t, and R_t^* all have the same population mean in the absence of a liquidity premium, these estimates of the noncentral second moments imply that the variances are in the order predicted by the efficient markets model:

$$\text{var}\ (r_t) > \text{var}\ (R_t^*) > \text{var}\ (R_t). \tag{36}$$

Since var (R_t^*) − var (R_t) = var (θ_t), table 3 provides estimates of the market's standard error in predicting R_t^*. For the 10-year bonds, the market's standard error in forecasting R_t^* was 174 basis points; for 20-year bonds, the standard error was 262 basis points.

The postwar quarterly data on the 3-month Treasury bill rate, 10-year Treasury bond yield, and perfect-foresight Treasury bond yield are plotted in figure 2. The perfect-foresight 10-year bond yield, denoted by the dotted line, starts to rise steeply in the early 1970s, reflecting the unusually high short rates in 1979–82. The rise in the long rate may have appeared, in 1969, to indicate overreaction to the contemporaneous rise in short rates, or excess volatility. However, history has clearly exonerated the sharp rise in the long rate in the late 1960s and early 1970s. In studying figure 2, one is struck, not by the volatility of the long rate, but by the accuracy of the long rate in predicting the explosion of short rates in the early 1980s.

[19] In calculating the noncentral second moment of R_t, the sample period was limited to the exact sample available for the corresponding perfect-foresight long rate; i.e., the 1973 : II–1982 : IV data on the 10-year long rate and the 1963 : II–1982 : IV data for the 20-year long rate were not used. If the noncentral second moment of R_t itself were of primary interest, using all of the available data would be efficient. However, in the variance-bounds problem, one is primarily interested in obtaining a precise estimate of the *difference* of the two moments. The sampling variability of the difference of the two sample moments is an increasing function of the variance of the sample second moment of R_t and a decreasing function of the covariance of the sample second moments of R_t and R_t^*. Including the additional data on R_t reduces the variance of the sample moment of R_t but also reduces the covariance of the two sample moments. Based on the conjecture that the effect of the covariance term dominates, the additional observations available for R_t were excluded.

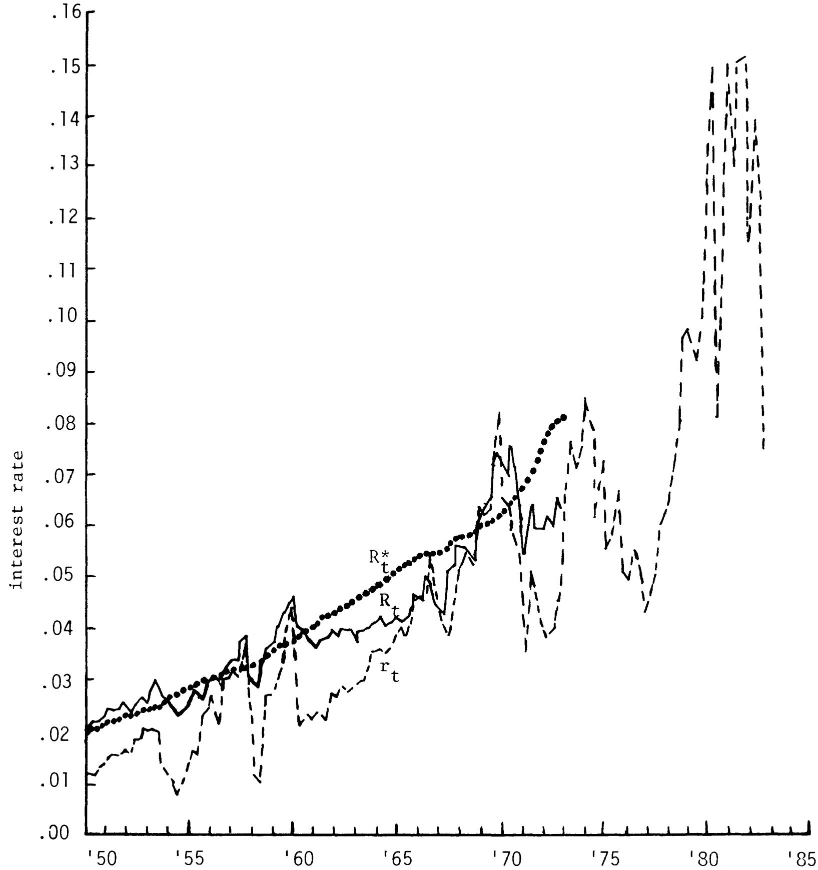

FIG. 2.—Plot of r_t, R_t, and R_t^*, U.S. Treasury securities: r_t = 3-month Treasury bill rate, denoted by dashed line; R_t = 10-year Treasury bond yield, denoted by solid line; R_t^* = perfect-foresight 10-year rate, denoted by dotted line. The sample period for the short-rate data is 1950:I–1982:IV. The perfect-foresight long rate was computed using Shiller's linearized term structure relation, $R_t^* = [(1 - \gamma)/(1 - \gamma^n)] \Sigma_{s=0}^{n-1} \gamma^s r_{t+s}$, for $\gamma = 1/(1 + R_0)$, where R_0 is the point around which the term structure relation is linearized. For the R_t^* series plotted in figure 2, R_0 = 0.01 and n = 40. The series on R_t^* was computed only up to 1973:I, the last observation for which the necessary postsample data were available. The data, from Salomon Brothers (1982), consist of observations taken during January, April, July, and October of each year.

III. Conclusions

The basic problem addressed by this paper—that the upper bound on the volatility of long interest rates or stock prices is difficult to measure in small samples—was certainly recognized by the authors who formulated the variance-bounds tests. In fact, Shiller refrains from conducting formal statistical tests of the hypothesis that the holding period yield on long interest rates is within its upper bound on the grounds that the small sample distribution of the upper bound is unknown. In the conclusion to his 1979 paper, Shiller acknowledges that he cannot rule out the possibility that the population variance of the short rate exceeds the sample variance by a sufficiently large margin to exonerate the market efficiency hypothesis, "since we have no real information in small samples about possible trends or long cycles in interest rates. Indeed, some would claim that short-term interest rates may be unstationary and hence have infinite variance. The fact that the lower bound on the left-hand side exceeds the sample value of the right-hand side may be interpreted as safely telling us, then, that we must rely on such unobserved variance or expected explosive behavior of short rates if we wish to retain expectations models" (pp. 1213–14).

Shiller's subsequent papers on long interest rate volatility (1981*a*, 1981*c*) reach the same general conclusion: the observed volatility of long interest rates can be justified as the rational response to new information about future short-term interest rates only if the population variance of short interest rates is much larger than the sample variance. In addition to random sampling error, Shiller cites several situations in which the population variance of the short rate would tend to exceed the sample variance: the short-rate process is nonstationary; the short-rate process is stationary but inappropriately detrended;[20] or the short-rate data suffer from what Krasker (1980) has termed the "peso problem"—market participants rationally perceived the possible occurrence of a major disturbance that was not realized within the sample period.[21]

[20] LeRoy and Porter (1981) also recognized the importance of the treatment of trends. They write, "The question remains whether the resulting series [earnings and price data for Standard and Poor's Composite Index, ATT, GE, and GM, corrected for inflation and earnings retention] can be assumed to obey the stationarity requirement. There appears to be some evidence of downward trends, although they are not clearly significant. We have decided to neglect such evidence and simply assume that the series are stationary since otherwise it is necessary to address such difficult questions as ascertaining to what degree stockholders can be assumed to have foreseen the assumed trend in earnings. It seems preferable to assume instead that there exist long cycles in the earnings series, implying that a sample of only a few decades may well appear nonstationary. . . . We do not argue that this treatment is entirely adequate, nor do we in any way minimize the problem of nonstationarity; the dependence of our results on the assumption of stationarity is probably their single most severe limitation" (pp. 568–69).

[21] Krasker (1980) examines an "apparent" failure of market efficiency in which the forward rate for Mexican pesos persistently underpredicted the future spot rate. Krasker argued that market participants rationally perceived a significant probability that the peso would be devalued. Since the devaluation did not occur within the sample period, the rational discounting of the peso in forward contracts gave rise to strong serial correlation in the spot-rate forecast errors.

This paper is focused primarily on the small sample properties of the variance-bounds test when the variances are expressed in deviations from the sample mean rather than the population mean. Even in the absence of other problems, such as nonstationarity, inappropriate detrending, and the peso problem, the results of the paper indicate that the variance-bounds test statistics tend to be seriously biased toward rejection of the null hypothesis of market efficiency when the variances are computed in deviations from sample means. For samples of the size typically used in the variance-bounds tests, the magnitude of the bias is substantial. The strategy of focusing on the consequences of taking deviations from the sample mean was not motivated by a judgment that other potential problems with the data, such as the peso problem, are empirically unimportant. To the contrary, my guess is that, for some data sets, the peso problem is probably very important. However, the effects of the peso problem are extremely difficult to assess empirically since by definition it involves the effects of unrealized possible outcomes.

By taking into account the small sample properties of the variance-bounds statistics, the evidence of excess volatility of holding period returns and long rates is attenuated along several dimensions. First, the upper bound on the variance of 10- and 20-year long rates is not violated in the postwar U.S. quarterly data. Second, the violation of the upper bound on holding period yields is not robust with respect to sample period. Third, in data sets for which the variance of holding period yields still exceeds the upper bound, the magnitude of the violation is smaller, and no evidence has been presented that the violation of the upper bound is statistically significant.

Appendix

Procedure for Obtaining the Small Sample Distributions of the Spectral Estimates

A vector of 101 observations on the short rate is given by $S\epsilon$, where S is a square matrix of order 101 as given in equation (12) and ϵ is a vector of 101 observations on the disturbance. Before computing the spectra, Singleton transformed the data by removing the sample mean and sample trend and by prewhitening by the filter $1 - .85L$. Defining X as the 101×2 matrix consisting of a column of ones and the column vector $[1, 2, 3, \ldots, 101]'$, and I the 101×101 identity matrix, construct $M = [I - X(X'X)^{-1}X']$. Construction of a 100×101 matrix, denoted H, which quasi differences the data with the filter $1 - .85L$ is straightforward. A vector of 100 observations on the transformed short-rate data is represented as $HMS\epsilon$. Denote the matrix product HMS as the 100×101 matrix D: $D = HMS$.

Construct the complex-valued matrix P,

$$P = \frac{1}{\sqrt{2\pi T}} \begin{bmatrix} e^{i1\lambda_0} & e^{i2\lambda_0} & \dots & e^{iT\lambda_0} \\ e^{i1\lambda_1} & e^{i2\lambda_1} & \dots & e^{iT\lambda_1} \\ \vdots & \vdots & & \vdots \\ e^{i1\lambda_{T-1}} & e^{i2\lambda_{T-1}} & \dots & e^{iT\lambda_{T-1}} \end{bmatrix}, \tag{A1}$$

where T is the sample size (in this case 100) and $\lambda_j = 2\pi j/T, j = 0, \ldots, T - 1$. The Fourier transform of the data is given by $PD\epsilon$.

The unweighted sum of the periodogram ordinates could be obtained by computing $\epsilon' D' \overline{P}' PD\epsilon$, where $\overline{P}$ is the conjugate of P. Singleton's estimates were weighted averages of the periodogram ordinates, however; he first smoothed the ordinates with an inverted V window (of width nine ordinates) and multiplied by a filter function, denoted $f(\lambda)$. The filter function is used to "recolor" the data and, in the case of the perfect-foresight long rate, also incorporates the theoretical filter $g^2(\lambda)$ given in equation (32). Denote the column vector of T periodogram ordinates as $\mathbf{z}$. Applying the window can be represented by premultiplying $\mathbf{z}$ by a square matrix V of order T. Define a diagonal matrix F, also of order T, in which the kth diagonal element is $f(\lambda_j)$, $\lambda_j = 2\pi(k - 1)/T$. The sum of the weighted periodogram ordinates is given by $\mathbf{u}FV\mathbf{z}$, where $\mathbf{u}$ is a T-element row vector of ones. Denote the T-element row vector $\mathbf{u}FV$ as $\mathbf{w}$ and construct the $T \times T$ diagonal matrix W, in which the ith diagonal element of W is the ith element of the vector $\mathbf{w}$. Thus the weighted sum of periodogram ordinates is given by

$$\epsilon' D' \overline{P}' WPD\epsilon. \tag{A2}$$

Multiply the matrix product $D'\overline{P}'WPD$ by the scalar $2\pi/T$ and denote the resulting matrix as C:

$$C = \frac{2\pi}{T} D'\overline{P}'WPD. \tag{A3}$$

The matrix C is a real, symmetric matrix of order 101.

Thus the sum of the weighted spectral density function can be expressed as a quadratic form in normal deviates. The small sample distributions of the spectral estimates of the variances can be obtained by applying the procedure described in Section I to the quadratic form $\epsilon' C\epsilon$.

All that remains is to specify the filters used to weight the smoothed periodogram ordinates. In the case of the spectral estimate of var (R_t^*), the filter was

$$f_{r^*}(0) = \frac{1}{(1 - .85)^2},$$
$$f_{r^*}(\lambda_j) = \frac{1 - \cos(n\lambda_j)}{n^2(1 - \cos\lambda_j)[1 - 2(.85)\cos\lambda_j + (.85)^2]}, \tag{A4}$$

where $\lambda_j = 2\pi j/T, j = 1, 2, \ldots, T - 1$, and $n = 80$.

To obtain the spectral estimate of var (R_t), the filter was

$$f_r(\lambda_j) = \frac{\alpha^2}{1 - 2(.85)\cos\lambda_j + (.85)^2}, \tag{A5}$$

where $\lambda_j = 2\pi j/T, j = 0, 1, \ldots, T - 1$, and α is the factor of proportionality between the short rate and the long rate in the numerical example; $\alpha = (1 - \rho^n)/n(1 - \rho)$.

References

Amsler, Christine. "The Term-Structure of Interest Rates in an Expanded Market Model." Ph.D. dissertation, Univ. Pennsylvania, 1980.

Anderson, Theodore W. *The Statistical Analysis of Time Series.* New York: Wiley, 1971.

Blanchard, Olivier J., and Watson, Mark W. "Bubbles, Rational Expectations, and Financial Markets." Working Paper no. 945. Cambridge, Mass.: Nat. Bur. Econ. Res., July 1982.

Flavin, Marjorie A. "Time Series Evidence on the Expectations Hypothesis of the Term Structure." *Carnegie-Rochester Conference Series on Public Policy,* vol. 20. Amsterdam: North-Holland, in press.

Grossman, Sanford J., and Shiller, Robert J. "The Determinants of the Variability of Stock Market Prices." *A.E.R. Papers and Proc.* 71 (May 1981): 222–27.

Krasker, William. "The 'Peso Problem' in Testing the Efficiency of Forward Exchange Markets." *J. Monetary Econ.* 6 (April 1980): 269–76.

LeRoy, Stephen F., and Porter, Richard D. "The Present-Value Relation: Tests Based on Implied Variance Bounds." *Econometrica* 49 (May 1981): 555–74.

Michener, Ronald W. "Variance Bounds in a Simple Model of Asset Pricing." *J.P.E.* 90 (February 1982): 166–75.

Nerlove, Marc; Grether, David M.; and Carvalho, José L. *Analysis of Economic Time Series: A Synthesis.* New York: Academic Press, 1979.

Pesando, James. "Time Varying Term Premiums and the Volatility of Long-Term Interest Rates." Unpublished manuscript. Toronto: Univ. Toronto, Dept. Econ., July 1979.

Salomon Brothers, Inc. *Analytical Record of Yields and Yield Spreads.* New York: Salomon Brothers, April 1982.

Shiller, Robert J. "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure." *J.P.E.* 87 (December 1979): 1190–1219.

———. "The Use of Volatility Measures in Assessing Market Efficiency." *J. Finance* 36 (May 1981): 291–304. (*a*)

———. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *A.E.R.* 71 (June 1981): 421–36. (*b*)

———. "Alternative Tests of Rational Expectations Models: The Case of the Term Structure." *J. Econometrics* 16 (May 1981): 71–87. (*c*)

Singleton, Kenneth J. "Expectations Models of the Term Structure and Implied Variance Bounds." *J.P.E.* 88 (December 1980): 1159–76.