Stock Prices, Earnings, and Expected Dividends

John Y. Campbell; Robert J. Shiller

The Journal of Finance, Vol. 43, No. 3, Papers and Proceedings of the Forty-Seventh
Annual Meeting of the American Finance Association, Chicago, Illinois, December
28-30, 1987 (Jul., 1988), 661-676.

Stock Prices, Earnings, and Expected Dividends

JOHN Y. CAMPBELL and ROBERT J. SHILLER*

ABSTRACT

Long historical averages of real earnings help forecast present values of future real dividends. With aggregate U.S. stock market data (1871–1986), a vector-autoregressive forecast of the present value of future dividends is, for each year, roughly a weighted average of moving-average earnings and current real price, with between two thirds and three fourths of the weight on the earnings measure. We develop the implications of this for the present-value model of stock prices and for recent results that long-horizon stock returns are highly forecastable.

IN THIS PAPER WE present estimates indicating that data on accounting earnings, when averaged over many years, help to predict the present value of future dividends. This result holds even when stock prices themselves are taken into account. The data are the real Standard and Poor Composite Index and associated dividend and earnings series 1871–1987. Our estimates indicate to what extent dividend-price ratios and returns on this index behave in accordance with simple present-value models, and allow us to shed new light on earlier claims that stock prices are too volatile to accord with such models (LeRoy and Porter [14], Shiller [20], Mankiw, Romer, and Shapiro [15], Campbell and Shiller [1, 2], and West [23]).

It seems appropriate to consider earnings data for forecasting dividends, since earnings are constructed by accountants with the objective of helping people to evaluate the fundamental worth of a company. However the precise economic meaning of earnings data is not clearly defined; accounting definitions are complicated and change through time in ways that are not readily documented. Because of this, many studies of financial time series have avoided the use of earnings data and have thus omitted relevant information about fundamental value from the analysis.[1]

Our approach is to introduce earnings, measured either annually or as an average over a number of years, as an information variable in a vector-autoregressive (VAR) framework. Any errors in measurement in earnings are accounted for automatically by the estimation procedure, which allows earnings to enter the model only insofar as they are useful in forecasting. The VAR framework,

---

[1] There is a large accounting literature on the response of securities prices to earnings announcements; see Kormendi and Lipe [13] for a list of references. However, with a few exceptions, notably Kormendi and Lipe, this literature does not ask whether the response is consistent with a particular fundamental valuation model for the security price.

developed originally in Campbell and Shiller [1, 2], enables us to answer two questions. First, what component of stock returns can be *predicted* given the information used in the VAR system? Secondly, what component of stock returns can be *accounted for* ex post by news about future dividends? The existing literature addresses the first question, but the second question is also important for evaluating present-value models. As Shiller [21] and Summers [22] have shown, it is possible to construct a model in which only a small fraction of one-period stock returns is predictable, but in which news about fundamental value accounts for only a small part of the variability of ex post returns.

Our approach reveals that stock returns and dividend-price ratios are too volatile to be accounted for by news about future dividends. Further, this excess volatility is closely related to the predictability of multiperiod returns. It has recently been shown that stock returns are more highly predictable when they are measured over intervals of several years, rather than over short intervals of a year or less. Fama and French [5, 6] have made this point most forcefully, although the result can also be found in Flood, Hodrick, and Kaplan [7], and Poterba and Summers [18]. (See also DeBondt and Thaler [4].) These papers found that twenty percent or thirty percent of the variance of four- or five-year stock returns can be explained by variables such as lagged multiyear stock returns or dividend-price ratios. The explained variances are higher when dividend-price ratios are used than when lagged returns are used.

It may be helpful, by way of motivation, to give at the outset a simple story indicating why excess volatility is fundamentally related to this forecastability of multiperiod returns. Let us consider the simplest argument for excess volatility given in the original LeRoy and Porter [14] and Shiller [20] papers. It was argued in those papers that if, as the present-value model asserts, price P_t is the expectation of P_t^*, the present value of actual future dividends, then the data must satisfy the variance inequality: $\text{var}(P_t^*) \geq \text{var}(P_t)$. The proof that the model implies this variance inequality was as follows. Since P_t is known at time t, we may write $P_t^* = P_t + u_t$, where u_t is a forecast error. A forecast error must be uncorrelated with the corresponding forecast, so u_t must be uncorrelated with P_t. Therefore $\text{var}(P_t^*) = \text{var}(P_t) + \text{var}(u_t)$. Since variances cannot be negative, the variance inequality follows. This argument can be reversed to show that if the variance inequality is violated in U.S. data, then it must be that $P_t^* - P_t$ is forecastable. We will show below that $P_t^* - P_t$ may itself be considered a sort of infinite-period return. Hence, excess volatility directly implies forecastability of infinite-period returns.

While the above simple story is illustrative of the nature of our argument, we will restate it below in terms of dividend-price ratios to allow for nonstationary dividends and prices, we will avoid any comparisons of P_t and P_t^* estimated with a terminal condition, we will take account of earnings data, and we will allow for a simple form of time variation in the real discount rate on stock. These advances are made possible by our use of the VAR framework discussed above. In our earlier work using this framework (Campbell and Shiller [2]), we found that our rejection of the hypothesis that one-period returns are unforecastable was much less strong than our rejection of the hypothesis that the dividend-price ratio equals the theoretical dividend-price ratio given the present-value model. We will

see that this is essentially the same result as noted by Fama and French and others that the one-period return is much less forecastable than the multiperiod return. The limit of their excess-return regression, where returns are computed over an infinite period of time, is essentially our test that the stock price equals the expected present value of future dividends. Thus we argue that excess volatility and predictability of multiperiod returns are not two phenomena, but one.

The organization of the paper is as follows. In Section I we discuss our data and show that dividend-price and earnings-price ratios predict stock returns measured over several years. We also present an approximation to the continuously compounded stock return, which we need to use in our VAR analysis. We show that predictability of approximate returns is close to that of exact returns. In Section II we explain our VAR methodology and relate it to research on multiperiod returns. In Section III we present basic VAR results, and in Section IV we use them to compare the historical behavior of stock prices and returns with the behavior implied by the present-value model. Section V checks the robustness of our results to changes in specification. Section VI concludes.

I. Predicting Stock Returns Using Prices, Dividends and Earnings

The data set used in this paper consists of annual observations on prices, dividends and earnings for the Standard and Poor Composite Stock Price Index, extended back to 1871 by using the data in Cowles [3]. The series on prices and dividends are also used in Campbell and Shiller [1, 2], and in much of the literature on volatility tests. Campbell and Shiller [2] show that the properties of the post-1926 data are very similar to those of the CRSP series on the value-weighted New York Stock Exchange Index, while Wilson and Jones [25] have carefully analyzed the pre-1926 data. The nominal earnings series for 1926 to 1986 is the Standard and Poor earnings per share adjusted to index, total for the year. For earlier years, our nominal earnings series is earnings-price ratio series R-1 (Cowles [3], pp. 404–5) times the annual average Standard and Poor Composite Index for the year. We deflate nominal series using a January Producer Price Index (annual average before 1900), 1967 = 100.

We write the real price of the stock index, measured in January of year t, as P_t. The real dividend paid on the index during period t is written D_t. The realized log gross return on the portfolio, held from the beginning of year t to the beginning of year $t + 1$, is $h_{1t} \equiv \log((P_{t+1} + D_t)/P_t) = \log(P_{t+1} + D_t) - \log(P_t)$. The realized log gross return over i years, from the beginning of year t to the beginning of year $t + i$, is

$$h_{it} \equiv \sum_{j=0}^{i-1} h_{1,t+j}. \tag{1}$$

We also wish to study excess returns on common stock over short debt. The short-term interest rate we use is the annual return on 4–6 month prime commercial paper, rolled over in January and July. If we write the realized log real return on commercial paper in year t as r_t, and aggregate to a multiperiod return r_{it} in the manner of equation (1), then the excess return on stock over i periods is $h_{it} - r_{it}$. Working with excess returns has the advantage that price

deflators cancel so that results are not contaminated by measurement error in the deflators.

We begin our empirical work by regressing real and excess stock returns on some explanatory variables that are known in advance (at the start of year t). For real returns, we consider the following variables:[2] the log dividend-price ratio, $\delta_t \equiv d_{t-1} - p_t$ (the dividend is lagged one year to ensure that it is known at the start of year t); the lagged dividend-growth rate, Δd_{t-1}; log earnings-price ratio $\epsilon_t \equiv e_{t-1} - p_t$; and two log earnings-price ratios based on moving averages of earnings. The latter two are a ten-year moving average of log real earnings minus current log real price, $\epsilon_t^{10} \equiv ((e_{t-1} + \cdots + e_{t-10})/10) - p_t$, and a thirty-year moving average of log real earnings minus current log real price, $\epsilon_t^{30} \equiv ((e_{t-1} + \cdots + e_{t-30})/30 - p_t$.

The ratio variables are used here with the same motivation that we see in the financial press, as indicators of fundamental value relative to price. The notion is that if stocks are underpriced relative to fundamental value, returns tend to be high subsequently, the converse holds if stocks are overpriced. A moving average of earnings is used because yearly earnings are quite noisy as measures of fundamental value; they could even be negative while fundamental value cannot be negative. The use of an average of earnings in computing the earnings-price ratio has a long history. Graham and Dodd [10] recommended an approach that "shifts the original point of departure, or basis of computation, from the current earnings to the average earnings, which should cover a period of not less than five years, and preferably seven to ten years." (*Security Analysis*, page 452). We push their averaging scheme even further, to thirty years, in recognition of the substantial decadal variability of earnings, under the supposition that fundamental value may be less variable than this decadal variability.

We regress real stock returns on each of these variables individually, and also on the combination (δ_t, Δd_{t-1}, ϵ_t^{30}). For excess stock returns, the procedure is similar except that we use the excess of dividend growth over the commercial-paper rate, $\Delta d_{t-1} - r_{t-1}$, in place of the real dividend-growth rate.

Table I presents regression results for the period 1871–1987 (truncated where necessary at the end of the sample to allow computation of multiperiod returns, and at the beginning of the sample to allow computation of ϵ_t^{10} and ϵ_t^{30}). Returns are measured over one, three, and ten years. The left side of panel A gives results for real returns, and the left side of panel B gives results for excess returns. For each regression the table reports the R^2 statistic, and in parentheses the significance level for a Wald test of the hypothesis that all coefficients (other than a constant) are zero. The Wald test corrects for the moving-average structure of the equation errors when the dependent variable is a multiperiod return, but it does not correct for heteroscedasticity.[3]

The table shows that several of the variables in our list have a striking ability to predict returns on the Standard and Poor Index. This is true whether returns are measured in real terms or as an excess over commercial-paper rates. The

---

[2] In this paper lower-case letters indicate natural logs of the corresponding upper-case letters.

[3] As in our previous paper (Campbell and Shiller [2]), the results are hardly changed by using White's [24] heteroscedasticity correction for standard errors.

Table I

Predicting Stock Returns, 1871–1987[a]

	Exact Returns (expression 1, text)			Discounted Returns (expression 4, text)		
	1-year	3-year	10-year	1-year	3-year	10-year
A. Real Returns						
Explanatory variables						
δ_t	0.039	0.110	0.266	0.048	0.135	0.327
	(0.033)	(0.015)	(0.001)	(0.017)	(0.006)	(0.000)
Δd_{t-1}	0.000	0.004	0.003	0.000	0.004	0.003
	(0.964)	(0.522)	(0.485)	(0.977)	(0.568)	(0.537)
ϵ_t	0.019	0.090	0.296	0.023	0.104	0.303
	(0.143)	(0.027)	(0.000)	(0.100)	(0.017)	(0.000)
ϵ_t^{10}	0.040	0.111	0.401	0.047	0.130	0.423
	(0.036)	(0.031)	(0.000)	(0.022)	(0.019)	(0.000)
ϵ_t^{30}	0.067	0.195	0.566	0.079	0.225	0.615
	(0.013)	(0.008)	(0.000)	(0.007)	(0.004)	(0.000)
$\delta_t,\ \Delta d_{t-1},\ \epsilon_t^{30}$	0.076	0.204	0.637	0.088	0.235	0.667
	(0.073)	(0.046)	(0.000)	(0.041)	(0.022)	(0.000)
B. Excess Returns						
δ_t	0.016	0.080	0.184	0.022	0.101	0.246
	(0.180)	(0.037)	(0.033)	(0.114)	(0.019)	(0.010)
$\Delta d_{t-1} - r_{t-1}$	0.026	0.027	0.000	0.026	0.026	0.001
	(0.082)	(0.127)	(0.811)	(0.082)	(0.134)	(0.758)
ϵ_t	0.011	0.054	0.195	0.015	0.066	0.206
	(0.261)	(0.083)	(0.009)	(0.194)	(0.053)	(0.005)
ϵ_t^{10}	0.052	0.145	0.341	0.060	0.168	0.399
	(0.017)	(0.010)	(0.003)	(0.010)	(0.005)	(0.001)
ϵ_t^{10}	0.051	0.187	0.480	0.074	0.218	0.548
	(0.017)	(0.007)	(0.002)	(0.009)	(0.003)	(0.000)
$\delta_t,\ \Delta d_{t-1} - r_{t-1},\ \epsilon_t^{30}$	0.086	0.195	0.493	0.096	0.229	0.553
	(0.046)	(0.045)	(0.011)	(0.028)	(0.022)	(0.004)

[a] The numbers reported are the R^2 in the regression of return on the explanatory variables, and in parentheses the significance level of a Wald test of the hypothesis that all coefficients in the regression are zero. The Wald test adjusts for overlapping data in regressions with multiperiod returns, but does not adjust for heteroscedasticity. The sample period is 1871–1987, truncated at the end where necessary to compute multiperiod returns.

variables with predictive power are those that include the stock price itself: the log dividend-price ratio δ_t, and the three earnings-price ratios ϵ_t, ϵ_t^{10} and ϵ_t^{30}. The forecasting power of these variables is statistically significant at conventional levels for one-period returns, but the fraction of variance explained is modest at this horizon: 3.9% of the variance of one-year real returns is explained by the log dividend-price ratio, for example. As the number of years used to compute the return increases, however, the fraction of variance explained also increases, and the constant-expected-return model is rejected more strongly. The log dividend-price ratio explains 26.6% of the variance of ten-year real returns, for example, and the thirty-year moving-average earnings-price ratio explains 56.6% of this

variance. These results confirm and extend the findings of Fama and French [6] for a longer data set, and establish that a very high proportion of multiperiod returns is forecastable using a long moving average of earnings.[4]

The lagged rate of dividend growth, by contrast, does not predict stock returns at any horizon. This is true whether we deflate it with a price index or use the commercial-paper rate. Also the system of three variables does not achieve an R^2 statistic that is much greater than that for ϵ_t^{30} alone.

In what follows, we will be concerned with the relationship between the realized log one-period return h_{1t}, the dividend-growth rate Δd_t, and the log dividend-price ratio δ_t. The exact relationship between these variables is nonlinear. It takes the form:

$$h_{1t} = \log(\exp(\delta_t - \delta_{t+1}) + \exp(\delta_t)) + \Delta d_t. \qquad (2)$$

However this equation can be linearized by a first-order Taylor expansion around the point $\delta_t = \delta_{t+1} = \delta$. We argued in Campbell and Shiller [2] that the log dividend-price ratio follows a stationary stochastic process, so that it has a fixed mean that can be used as the expansion point δ. We will also define the interest rate implicit in the chosen δ as $r = g + \ln(1 + \exp(\delta))$, where g is the mean Δd. We obtain

$$h_{1t} \approx \xi_{1t}$$

$$\xi_{1t} \equiv \delta_t - \rho\delta_{t+1} + \Delta d_t + k = (1 - \rho)d_t + \rho p_{t+1} - p_t + k, \qquad (3)$$

where $\rho = 1/(1 + \exp(\delta)) = \exp(-(r - g))$, and $k = \log(1 + \exp(\delta)) - \delta \exp(\delta)/(1 + \exp(\delta))$.

Equation (3) says that the log one-period return on the stock portfolio, h_{1t}, can be approximated by a variable ξ_{1t} that is linear in the log dividend-price ratios δ_t and δ_{t+1} and the dividend-growth rate Δd_t. The approximation in (3) replaces $\log(P_{t+1} + D_t)$ with $\rho\log(P_{t+1}) + (1 - \rho)\log(D_t)$, where ρ is a parameter related to the mean ratio of prices to dividends.

We now define a multiperiod extension of (3). For the purpose of showing the relation between the excess-volatility literature and the multiperiod-return forecasting literature, it is helpful to define this slightly differently than would be natural given (1). We define the discounted i-period return ξ_{it} as:

$$\xi_{it} \equiv \sum_{j=0}^{i-1} \rho^j \xi_{1,t+j}. \qquad (4)$$

The variable ξ_{it} is the *discounted* sum of approximate returns from t to $t + i - 1$. It has the convenient property that it depends only on δ_t, δ_{t+i}, and dividend-growth rates from t to $t + i$; log dividend-price ratios for times between t and $t + i$ do not appear. While the summation in (1) approaches infinity as i

---

[4] When we use the Fama-French sample periods, 1927–86, we find that the dividend-price ratio explains 21.9% of the variance of exact four-year real returns. (Four years was the longest horizon they reported.) This roughly confirms their estimated R^2 of 29%. The 30-year average of earnings does only slightly better than the dividend-price ratio over this sample period and return horizon, explaining 22.6% of the variance of returns. When we extend the horizon to ten years, however, the 30-year earnings average explains 47.5% and the dividend-price ratio only 24.8% of the variance of returns.

increases, the summation in (4) instead approaches (under the assumption that δ_t and Δd_{t-1} are jointly stationary) a well-defined limit, a stationary stochastic process. We can thus speak of an infinite-period log return, which we will see below is related to the log dividend-price ratio; this is why use of the definition (4) ties the multiperiod-return literature to our own earlier study of the behavior of the dividend-price ratio.

One interpretation of the discounted i-period return ξ_{it} is that it is (up to a constant term that depends on i) a linearization of an exact i-period log return H_{it} where dividends paid are reinvested not in the stock itself but in an instrument that pays a fixed real return.[5] H_{it} can be written in terms of the log dividend-price ratio and log dividend-growth rates:

$$H_{it} = \ln\{\exp(\delta_t - \delta_{t+i} + \sum_{j=0}^{i-1} \Delta d_{t+j}) + \sum_{j=0}^{i-1} \exp(\delta_t + \sum_{k=0}^{j} \Delta d_{t+k} + r(i-j-1))\}.$$

The first term inside the curly brackets is the price relative P_{t+i}/P_t. The subsequent terms give the terminal value of total dividends received between t and $t + i - 1$ divided by P_t. Note that since reinvestments are not made in the stock, dividend-price ratios between t and $t + i$ do not enter the expression, as also with (4). Let us linearize the above expression around $\delta_t = \delta$ and $\Delta d_{t+j} = g$, for all j. This gives us the discounted i-period return ξ_{it} defined in equation (4), plus a constant that increases with i.

Naturally equations (3) and (4) do not give actual log returns exactly; since they were derived from a linearization, there is some approximation error. In Campbell and Shiller [2], we presented considerable evidence that in practice the error is quite small for one-period returns. Here we supplement that analysis by repeating the regressions discussed above using discounted multiperiod returns ξ_{it} rather than exact returns h_{it}. We treat the parameter ρ as fixed, and set it equal to 0.936 following Campbell and Shiller [2].[6]

The results are given in the right side of Table I. They are generally similar to those discussed before; while there is a slightly greater tendency to reject the constant-expected-return model with discounted returns (indicating that the approximation error is correlated with the explanatory variables), the difference is relatively minor. This confirms that we can speak of our definition of multiperiod returns (4) as roughly interchangeable, for present purposes, with the definition (1) used by Fama and French [5, 6] and others.

II. A Vector-Autoregressive Approach

In the previous section we derived an approximation to the log return on stock that is linear in log dividend-price ratios and dividend-growth rates. We now exploit this linearity in analyzing stock price movements.

First, we write the discounted i-period log return as an explicit linear function

---

[5] We assume this reinvestment rate of return is equal to the rate of return r implicit in the ρ used in the linearization, that is, $r = g - \ln(\rho)$.

[6] In that paper we showed that varying ρ in a plausible range did not greatly affect our conclusions. Here, too, when we set $\rho = 1$ in equation (4) (but retain $\rho = 0.936$ in equation (3)), so that ξ_{it} becomes the simple sum which approximates h_{it}, we obtain very similar results to those reported.

of δ_t, δ_{t+i} and Δd_{t+j}, $j = 0, \ldots, i - 1$. From equations (3) and (4) we have:

$$\xi_{it} = \delta_t - \rho^i \delta_{t+i} + \sum_{j=0}^{i-1} \rho^j \Delta d_{t+j} + k(1 - \rho^i)/(1 - \rho). \qquad (5)$$

Equation (5) shows that the discounted i-period return is higher, the higher the dividend-price ratio is when the investment is initiated, the lower the dividend-price ratio is when the investment is terminated, and the higher dividend growth is between those two dates.[7]

We can also use this equation to see the relationship between multiperiod returns and the literature on price volatility. If we take the limit of (5) as i increases, assuming that $\lim_{i \to \infty} \rho^i E_t \delta_{t+i} = 0$ (which follows from the stationarity of δ_t), we find that we have

$$\lim_{i \to \infty} \xi_{it} = (1 - \rho) \sum_{j=0}^{\infty} \rho^j d_{t+j} - p_t + k/(1 - \rho).$$

The first term on the right-hand side of this expression is the present discounted value of log dividends, which is a log-linearization of P_t^*, while the second term is the log of P_t. Thus, as noted in the introduction, the infinite-period discounted log return is a log-linearization of the variable $P_t^* - P_t$, which is the subject of the volatility literature. Moreover, for finite i, ξ_{it} is a log-linear representation of $P_t^* - P_t$ where P_t^* is computed under the assumption that the present value in period $t + i$ of dividends from $t + i$ onwards equals P_{t+i}. This assumption was used in the volatility literature to obtain an estimate of P_t^* with a finite record of dividends.

Equation (5) makes it easy to compute the implication of a returns model for the dividend-price ratio. For example, suppose our model is that expected real one-period stock returns are constant: $E_t \xi_{1t} = r$. Then $E_t \xi_{it} = r(1 - \rho^i)/(1 - \rho)$. Taking conditional expectations of the left- and right-hand sides of (5) and rearranging, we have

$$\delta_t = - \sum_{j=0}^{i-1} \rho^j E_t \Delta d_{t+j} + \rho^i E_t \delta_{t+i} + (r - k)(1 - \rho^i)/(1 - \rho). \qquad (6)$$

This equation says that the log dividend-price ratio at time t is determined by expectations of future real dividend growth over i periods, by the i-period-ahead expected dividend-price ratio, and by the constant required return on stock. If we take the limit as i increases, assuming as before that $\lim_{i \to \infty} \rho^i E_t \delta_{t+i} = 0$, we obtain

$$\delta_t = - \sum_{j=0}^{\infty} \rho^j E_t \Delta d_{t+j} + (r - k)/(1 - \rho). \qquad (7)$$

Equation (7) expresses the log dividend-price ratio as a linear function of expected real dividend growth into the infinite future.

A similar approach can be used when our returns model is that expected *excess* returns on stock, over some alternative asset with return r_t, are constant: $E_t \xi_{1t} = E_t r_t + c$. In our empirical work, we take r_t to be the real return on commercial

[7] Note that as i grows larger, less weight is given in (5) to the terminal dividend-price ratio δ_{t+i}, and hence to the terminal price. One might wonder why the terminal price is downweighted in an approximate expression for log total return over t to $t + i$. The reason is that as i is increased the component of total return due to reinvestment of intervening dividends at the fixed rate grows larger, causing the slope of the log function at the point of linearization to approach zero as i is increased.

paper. For this model we have

$$\delta_t = \sum_{j=0}^{i-1} \rho^j \mathrm{E}_t[r_{t+j} - \Delta d_{t+j}] + \rho^i \mathrm{E}_t \delta_{t+i} + (c - k)(1 - \rho^i)/(1 - \rho), \quad (6')$$

and taking the limit as i increases,

$$\delta_t = \sum_{j=0}^{\infty} \rho^j \mathrm{E}_t[r_{t+j} - \Delta d_{t+j}] + (c - k)/(1 - \rho). \quad (7')$$

This relation is what Campbell and Shiller [2] call the "dividend-ratio model". It may also be described as a dynamic Gordon model, after the simple growth model proposed by Myron Gordon [9], which makes the dividend-price ratio equal the interest rate minus the growth rate of dividends. The original Gordon model did not specify how the dividend-price ratio should change through time if interest rates or growth rates change through time: equation (7′) says that the dividend-price ratio is related to a present value of expected one-period interest rates and dividend-growth rates.

The linearity of these relationships makes it possible to test them as restrictions on a vector autoregression. This procedure has several advantages over the straightforward multiperiod-regression approach discussed in the previous section. First, one need only estimate the VAR once: then one can conduct Wald tests of (6) or (6′) for any i, without reestimating the system. Secondly, as i increases, the regression approach forces one to shorten the sample period. This becomes quite serious when returns are calculated over five to ten years. The VAR, by contrast, can be estimated over the whole sample. Thirdly, the VAR can be used to test the restrictions of (7) or (7′), which are the limits of (6) and (6′) as i increases. This is important because (7) and (7′) directly state the implications of the returns model for the dividend-price ratio. Finally, the VAR approach enables us to characterize the historical behavior of the dividend-price ratio in relation to an unrestricted econometric forecast of future dividends and discount rates. It is important to note that if the present-value model is correct, then this unrestricted forecast, which we call δ_t', should equal the log dividend-price ratio δ_t no matter how much information market participants have. The reason for this is that δ_t, which is included in the VAR system, is a sufficient statistic for market participants' information about the present value of future dividends.

A detailed account of the VAR framework is given in Campbell and Shiller [1, 2]. Here we briefly summarize it for the constant-expected-returns case. Consider estimating a VAR for the variables δ_{t+1}, Δd_t and ϵ_{t+1}^{30}. The last variable, a moving-average earnings-price ratio, is included only as a potential predictor of stock returns. If the VAR has only one lag, then the system estimated is

$$\begin{bmatrix} \delta_{t+1} \\ \Delta d_t \\ \epsilon_{t+1}^{30} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \delta_t \\ \Delta d_{t-1} \\ \epsilon_t^{30} \end{bmatrix} + \begin{bmatrix} u_{1t+1} \\ u_{2t+1} \\ u_{3t+1} \end{bmatrix} \quad (8)$$

where the variables in the vector are demeaned. This can be written more compactly, in matrix form, as $z_{t+1} = A z_t + v_{t+1}$.

Now a first-order vector autoregression has the desirable property that to forecast the variables ahead k periods, given the history $H_t = \{z_t, z_{t-1}, \dots\}$, one

just multiplies z_t by the kth power of the matrix A:

$$\mathrm{E}[z_{t+k} \mid H_t] = A^k z_t. \tag{9}$$

This makes it easy to translate equations (6) and (7) into restrictions on the VAR. First, define vectors $e1 = [1\ 0\ 0]'$, so that $e1'z_t = \delta_t$, and $e2 = [0\ 1\ 0]'$, so that $e2'z_t = \Delta d_{t-1}$. Next, take the expectation of equation (6), conditional on H_t:

$$\delta_t = -\sum_{j=0}^{i-1} \rho^j \mathrm{E}[\Delta d_{t+j} \mid H_t] + \rho^i \mathrm{E}[\delta_{t+i} \mid H_t] + (r - k)(1 - \rho^i)/(1 - \rho). \tag{6''}$$

The left-hand side is unaffected, because δ_t is in the information set H_t, and the right-hand side becomes an expectation conditional on H_t.

Finally, apply the multiperiod forecasting formula (9):

$$e1'z_t = -\sum_{j=0}^{i-1} \rho^j A^{j+1} e2'z_t + \rho^i A^i e1'z_t + (r - k)(1 - \rho^i)/(1 - \rho). \tag{10}$$

If (10) is to hold for arbitrary z_t, we must have

$$e1'(I - \rho^i A^i) = -e2'A(I - \rho A)^{-1}(I - \rho^i A^i). \tag{11}$$

These are complicated nonlinear restrictions on the coefficient-matrix A, but they do simplify in two special cases, which are emphasized in Campbell and Shiller [2]. First, if $i = 1$ then we have a set of *linear* restrictions that one-period returns are unpredictable: $e1'(I - \rho A) = -e2'A$. In terms of the individual coefficients, the restrictions are $a_{21} = \rho a_{11} - 1$, $a_{22} = \rho a_{12}$ and $a_{23} = \rho a_{13}$. The coefficients in the equation for the earnings-price ratio, a_{31}, a_{32}, and a_{33}, are unrestricted. Secondly, if $i = \infty$ then we have a set of nonlinear but simple restrictions that the log dividend-price ratio δ_t equals the unrestricted VAR forecast of real dividend growth into the infinite future, which we will call δ_t'. The restrictions are $\delta_t \equiv e1'z_t = -e2'A(I - \rho A)^{-1}z_t \equiv \delta_t'$, which requires that $e1' = -e2'A(I - \rho A)^{-1}$. We will compare the historical behavior of δ_t', the VAR forecast of future real dividend growth, with that of the log dividend-price ratio δ_t.

Of course, the restrictions for all i are algebraically equivalent. If $e1'(I - \rho A) = -e2'A$, then one can postmultiply by $(I - \rho^i A^i)$ for any i to get the i-period restriction. The reverse is also possible since stationarity of the VAR guarantees nonsingularity of $(I - \rho^i A^i)$. This algebraic equivalence reflects the fact that if one-period returns are unpredictable, then i-period returns must also be, and vice versa. Nevertheless, Wald tests on the VAR may yield different results depending on which value of i is chosen, just as regression tests did in Table I.

The VAR approach can easily be modified to handle different specifications. To test the model in which expected excess returns are constant, one simply replaces Δd_t with $\Delta d_t - r_t$ and proceeds as before. To handle higher order VAR behavior, one estimates the higher order system and then stacks it into first-order "companion" form as discussed by Sargent [19] and Campbell and Shiller [1, 2]. When z_t, A, $e1$ and $e2$ are suitably redefined, the restriction (11) remains correct.

III. Results of the VAR Procedure

In Table II we apply the VAR method to our data on stock prices, dividends and earnings over the period 1871–1987. The sample period is truncated at the beginning to allow for construction of a thirty-year moving average of earnings, but it need not be truncated at the end even though we will test for unpredictability of multiperiod returns. We estimate first-order VARs, using real dividend growth in panel A (to test the constant-expected-*real*-return model), and the excess of dividend growth over the commercial-paper rate in panel B (to test the constant-expected-*excess*-return model). We devote most of our attention to the results in panel A, discussing the panel-B results briefly in Section V.

The VAR coefficients, a_{ij} for $i, j = 1, 2, 3$, are reported at the top of the table. Below each coefficient is an asymptotic standard error in parentheses. The coefficients in the second row (the dividend-growth equation) are perhaps of special interest; they show that the dividend-price ratio has strong forecasting power for dividend growth, and the earnings-price ratio ϵ_t^{30} is also highly significant. These results suggest that some improvement is possible in the dividend-growth equation proposed by Marsh and Merton [16, 17], which does not use the long average of earnings variable.

The hypothesis that expected real returns on stock are constant restricts the coefficients in the first two rows, the equations for the dividend-price ratio and real dividend growth respectively. We should have $a_{21} = \rho a_{11} - 1$, $a_{22} = \rho a_{12}$ and $a_{23} = \rho a_{13}$. As before, we fix the parameter ρ at 0.936.

These restrictions do not hold exactly, and the differences $a_{21} - \rho a_{11} + 1$, $a_{22} - \rho a_{12}$ and $a_{23} - \rho a_{13}$ are the coefficients obtained in a regression of ξ_{1t} on the VAR explanatory variables. Coefficients from such a regression are reported in Table II below the VAR results. (This regression was also used in Table I, panel A).

Wald tests of the model restriction (11), for $i = 1, 2, 3, 5, 7, 10$ and ∞, are reported next in Table II. The test statistic for $i = 1$ is numerically identical to the statistic obtained from the regression of ξ_{1t} on the VAR explanatory variables; its significance level of 0.041 is therefore identical to the one reported in Table I, panel A. When $i > 1$, the exact equivalence of the regression test and the VAR is broken, but the general nature of the results is the same. The VAR tests, like the multiperiod regression tests, reject more and more strongly as the return horizon increases. In the limit, at $i = \infty$, the null hypothesis is that the log dividend-price ratio δ_t equals the unrestricted VAR forecast of the present value of future real dividend growth δ_t'. This hypothesis can be rejected at better than the 0.1% level.

IV. Comparison of Historical and Theoretical Stock Prices and Returns

In this section we use the VAR estimates in Table II to compare actual stock prices and returns with their theoretical counterparts. We find that with the

Table II

One-Lag VAR Results, 1871–1987[a]

A. Real Returns

Dependent variable	Explanatory variable			
	δ_t	Δd_{t-1}	ϵ_t^{30}	R^2
δ_{t+1}	0.610	0.210	0.086	0.503
	(0.134)	(0.175)	(0.093)	
Δd_t	−0.418	0.332	0.209	0.361
	(0.067)	(0.087)	(0.046)	
ϵ_{t+1}^{30}	0.008	−0.104	0.874	0.791
	(0.125)	(0.163)	(0.087)	
ξ_{1t}	0.011	0.135	0.129	0.088
	(0.118)	(0.154)	(0.082)	

Significance levels for VAR tests of unpredictability of returns:
Number of years over which returns are computed

1	2	3	5	7	10	∞
0.041	0.023	0.012	0.002	0.000	0.000	0.000

Some implications of the VAR estimates:

$$\delta_t' = 1.032\,\delta_t - 0.078\,\Delta d_{t-1} - 0.776\,\epsilon_t^{30}$$
$$(0.076) \qquad (0.046) \qquad (0.101)$$

$$\sigma(\delta_t')/\sigma(\delta_t) = 0.672 \qquad \mathrm{corr}(\delta_t',\delta_t) = 0.175$$
$$(0.074) \qquad\qquad (0.146)$$

$$\sigma(\xi_{1t}')/\sigma(\xi_{1t}) = 0.269 \qquad \mathrm{corr}(\xi_{1t}',\xi_{1t}) = 0.915$$
$$(0.067) \qquad\qquad (0.064)$$

B. Excess Returns

Dependent variable	Explanatory variable			
	δ_t	$\Delta d_{t-1} - r_{t-1}$	ϵ_t^{30}	R^2
δ_{t+1}	0.619	0.482	0.087	0.541
	(0.126)	(0.164)	(0.087)	
$\Delta d_t - r_t$	−0.393	0.235	0.179	0.339
	(0.066)	(0.086)	(0.045)	
ϵ_{t+1}^{30}	−0.024	0.256	0.908	0.796
	(0.121)	(0.158)	(0.083)	
ξ_{1t}	0.028	−0.216	0.097	0.096
	(0.114)	(0.149)	(0.078)	

Significance levels for VAR tests of unpredictability of returns:
Number of years over which returns are computed

1	2	3	5	7	10	∞
0.028	0.013	0.005	0.000	0.000	0.000	0.000

Some implications of the VAR estimates:

$$\delta_t' = 0.927\,\delta_t + 0.046(\Delta d_{t-1} - r_{t-1}) - 0.634\,\epsilon_t^{30}$$
$$(0.144) \qquad (0.086) \qquad\qquad (0.217)$$

$$\sigma(\delta_t')/\sigma(\delta_t) = 0.580 \qquad \mathrm{corr}(\delta_t',\delta_t) = 0.309$$
$$(0.136) \qquad\qquad (0.341)$$

$$\sigma(\xi_{1t}')/\sigma(\xi_{1t}) = 0.485 \qquad \mathrm{corr}(\xi_{1t}',\xi_{1t}) = 0.733$$
$$(0.044) \qquad\qquad (0.188)$$

[a] Results are for vector autoregressions with three-element vector including ϵ_t^{30}. The first group of numbers reported are regression coefficients, with standard errors in parentheses. (In the δ_t' column the numbers are implied coefficients from the VAR, with asymptotic standard errors calculated numerically). Also reported are R^2 statistics from the regressions. Below this are significance levels for Wald tests of restrictions (11), with $i = 1, 2, 3, 5, 7, 10$ and ∞. The Wald test at $i = \infty$ is a test of the hypothesis that $\delta_t = \delta_t'$. Below this are some implied statistics computed from the VAR, with asymptotic standard errors calculated numerically in parentheses.

constant-expected-real-return model, the log dividend-price ratio δ_t has only a weak relation to its theoretical counterpart δ_t', a result that strongly contradicts the model. The variables δ_t' and δ_t have a correlation of only 0.175 (this estimate has a standard error of 0.146), and δ_t' is less variable than δ_t (see the bottom of Table II, panel A). Its standard deviation is 0.672 times that of δ_t, with a small standard error of 0.074. This would suggest that the dividend-price ratio is unrelated to the theoretical value implied by the constant-expected-real-return model. However, a plot of δ_t and δ_t' (Figure 1) shows a suggestion of *short-run* coherence, even though the overall correlation between the two is virtually zero. Our VAR results also indicate that the dividend-price ratio helps to forecast *short-run* dividend changes.

One-period returns ξ_{1t} are about four times as variable as they should be given the model. To see this, we computed a variable $\xi_{1t}' \equiv \delta_t' - \rho\delta_{t+1}' + \Delta d_t$. This is our estimate of what the one-year return on stock *would* be, if the constant-expected-real-return model held so that δ_t equalled δ_t'. Note that ξ_{1t} should equal ξ_{1t}' even if the market has superior information not available to econometricians. We find that ξ_{1t}' has a standard deviation only 0.269 as large as that of ξ_{1t}. This appears to be a strong result, as the standard error on this ratio of standard deviations is only 0.067. This result is good evidence that returns on stocks are far too volatile to accord with the constant-expected-real-return present-value model, confirming the earlier claims of the volatility literature.

Although returns seem to be too volatile, we do estimate a remarkably high correlation coefficient between actual returns ξ_{1t} and their theoretical counterparts ξ_{1t}', equal to 0.915. Returns may be too volatile, but they appear to be on track in the sense that they correlate very well with their theoretical values.

This result is due to the same feature of the data that gives the short-run coherence between δ_t and δ_t' observed in Figure 1. It is easy to see where the result comes from if we use the derived equation defining δ_t', as shown in Table
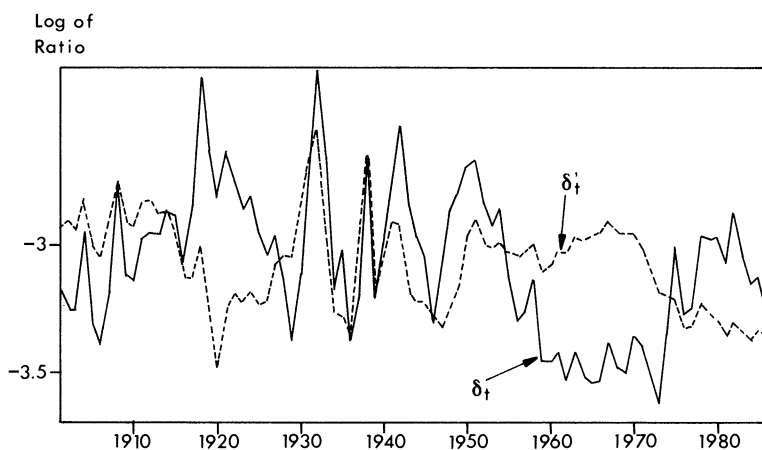

Figure 1. Log dividend-price ratio δ_t (solid line), and theoretical counterpart δ_t' (dashed line), 1901–86. The variable δ_t' is the optimal forecast of the present value of future real dividend-growth rates (constant discount rate), based on the vector-autoregressive model as given in Table II panel A. That is, $\delta_t' = -e2'A(I - \rho A)^{-1}z_t = 1.032\delta_t - 0.078\Delta d_{t-1} - 0.776\epsilon_t^{30}$.

II, panel A. This equation defines δ_t' as $\delta_t' = 1.032\delta_t - 0.078\Delta d_{t-1} - 0.776\epsilon_t^{30}$. Let us define p_t' as the theoretical log real price implied by the model, $p_t' = d_{t-1} - \delta_t'$. The present-value model implies that p_t' should equal p_t, even if economic agents have superior information not observed by econometricians. By contrast, our estimates imply that $p_t' = 0.776e_t^{30} + 0.256p_t + 0.046d_{t-1} - 0.078d_{t-2}$, where e_t^{30} is the thirty-year moving average of log real earnings. This shows that p_t' is essentially three fourths of the long moving average of real log earnings plus one fourth of the current price. It is a weighted average of the moving average of log real earnings and of log real price with most of the weight on the moving average.

A plot of p_t and p_t' over the period 1901–1986 is shown in Figure 2. The variable p_t' is strikingly smoother than p_t and at the same time shows short-run movements that are highly correlated with it. This is as we would expect: the long moving average of real earnings is very smooth, since long moving averages smooth out the series averaged. Hence, most of the short-run fluctuations in p_t are seen, in an attenuated form, in p_t'. Since returns ξ_{1t} and ξ_{1t}' are essentially changes in p_t, their behavior is dominated by the short-run movements in the series so that they are highly correlated with each other. Dividend-price ratios δ_t and δ_t', on the other hand, are determined by the levels of p_t and p_t' and are not very correlated.

V. How Robust Are the Results to Changes in Specification?

In panel B of Table II, we repeat all these exercises using dividend growth deflated by the commercial-paper rate rather than the inflation rate of the producer price index. The null hypothesis here is that expected excess returns on stock over commercial paper are constant. We obtain results that are similar to, though for the most part somewhat less dramatic than, those in panel A. The
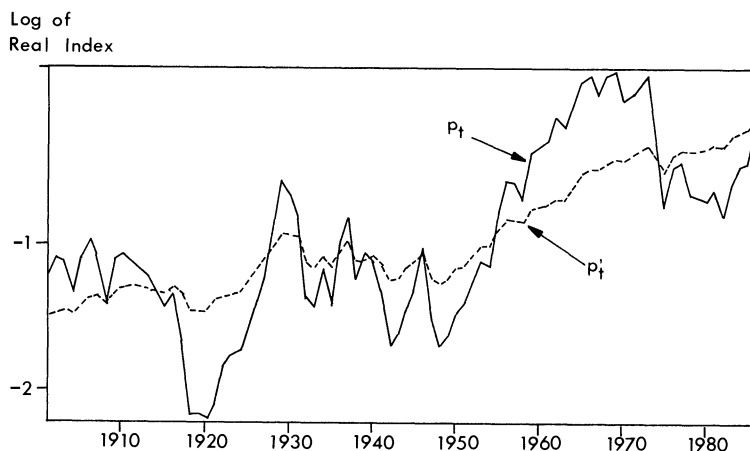
Figure 2. Log real stock price index p_t (solid line) and theoretical log real price index p_t' (dashed line), 1901–86. The theoretical log real price index p_t' is the optimal forecast of the log-linearized present value (constant rate of discount) of real dividends based on the vector-autoregressive forecasting model presented in Table II panel A. The variable p_t' is computed as $d_{t-1} - \delta_t'$ where δ_t' is the series plotted in Figure 1.

correlation between δ_t and δ_t' is small, at 0.309. The standard deviation of ξ_{1t}' is just under half that of ξ_{1t}, and the two have a substantial correlation, of 0.733. The implied variable p_t' now places a weight of 0.634 on e_t^{30} and 0.293 on p_t. Again, the long moving average of earnings dominates the stock price in forecasting dividend growth adjusted for commercial-paper rates.

We also checked to see whether our VAR results are robust to increases in the lag length of the VAR. We estimated VARs of order 1 through 5. Except for the fact that the significance levels in the one-period-return regression decline with lag length, the conclusions for the most part do not seem to be very sensitive to the order of the estimated VAR. We also checked to see whether a shorter ten-year moving average of earnings gives similar results to those reported in Table II. We estimated that $p_t' = 0.423e_t^{10} + 0.445p_t + 0.246d_{t-1} - 0.114d_{t-2}$. The correlation between δ_t' and δ_t is higher than in Table II, but the correlation drops dramatically as lag length is increased towards 5. Finally, we estimated the VAR system in Table II for the shorter sample 1927–86. We obtained results that were very similar to those for the full sample period.

VI. Conclusion

Our results indicate that a long moving average of real earnings helps to forecast future real dividends. The ratio of this earnings variable to the current stock price is a powerful predictor of the return on stock, particularly when the return is measured over several years. We have shown that these facts make stock prices and returns much too volatile to accord with a simple present-value model. Yet annual returns do seem to carry some information and are correlated with what they should be given the model.

Whenever a new variable is introduced into an analysis, in this case the long moving average of earnings, and the new variable plays an important role in the results, it is natural for critics to wonder if the new variable really belongs in the analysis. There is always the possibility that many different variables were attempted, until the results changed, and only the one that changed the results was reported. However, we think that it can be argued that a long moving average of earnings is a very natural variable to use to represent fundamental value, and that there are not many competitors for this role. We note also that we found evidence of excess volatility in earlier research (Campbell and Shiller [2]) that did not use the information in earnings.

In evaluating our results, it should also be borne in mind that (disregarding small-sample considerations) if we find one variable that destroys the model, then introducing new variables can never save the model. Since the log dividend-price ratio δ_t is in the information set assumed, it should get a unit coefficient and all other variables should get zero coefficients in the equation for the theoretical log dividend-price ratio δ_t'. Adding more variables can never bring us back to this situation, so long as the earnings variable is included. Another way to put this, recalling our argument that excess volatility is the same as forecastability of multiperiod returns, is that once a forecasting variable is found that predicts multiperiod returns, adding new forecasting variables can never make them unforecastable.

REFERENCES

1. John Y. Campbell and Robert J. Shiller. "Cointegration and Tests of Present Value Models." *Journal of Political Economy* 95 (October 1987), 1062–88.
2. ———. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." Unpublished paper, Princeton University and Yale University, revised October 1987.
3. Alfred Cowles 3rd and Associates. *Common-Stock Indexes*, 2nd ed. Bloomington, Indiana: Principia Press, 1939.
4. Werner DeBondt and Richard Thaler. "Does the Stock Market Overreact?" *Journal of Finance* 60 (July 1985), 793–805.
5. Eugene F. Fama and Kenneth R. French. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy* 96 (April 1988), 246–73.
6. ———. "Dividend Yields and Expected Stock Returns." Unpublished paper, University of Chicago, July 1987.
7. Robert P. Flood, Robert J. Hodrick, and Paul Kaplan. "An Evaluation of Recent Evidence on Stock Market Bubbles." NBER Working Paper No. 1971, July 1986.
8. Christian Gilles and Stephen F. LeRoy. "The Variance-Bounds Tests: A Critical Survey." Unpublished paper, University of California Santa Barbara, August 1987.
9. Myron J. Gordon. *The Investment, Financing and Valuation of the Corporation.* Homewood, Illinois: Irwin, 1962.
10. Benjamin Graham and David L. Dodd. *Security Analysis*, 1st ed. New York: McGraw-Hill, 1934.
11. Allan W. Gregory and Michael R. Veall. "Formulating Wald Tests of Nonlinear Restrictions." *Econometrica* 53 (November 1985), 1465–1468.
12. Allan W. Kleidon. "Variance Bounds Tests and Stock Price Valuation Models." *Journal of Political Economy* 94 (October 1986), 953–1001.
13. Roger C. Kormendi and Robert Lipe. "Earnings Innovations, Earnings Persistence and Stock Returns." *Journal of Business* 60 (1987), 323–345.
14. Stephen F. LeRoy and Richard D. Porter. "Stock Price Volatility: Tests Based on Implied Variance Bounds." *Econometrica* 49 (1981), 97–113.
15. N. Gregory Mankiw, David Romer, and Matthew D. Shapiro. "An Unbiased Reexamination of Stock Market Volatility." *Journal of Finance* 40 (May 1985), 677–87.
16. Terry A. Marsh and Robert C. Merton. "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices." *American Economic Review* 76 (June 1986), 483–498.
17. ———. "Dividend Behavior for the Aggregate Stock Market." *Journal of Business* 60 (January 1987), 1–40.
18. James M. Poterba and Lawrence H. Summers. "Mean Reversion in Stock Prices: Evidence and Implications." Unpublished paper, Harvard University, revised March 1987.
19. Thomas J. Sargent. "A Note on the Estimation of the Rational Expectations Model of the Term Structure." *Journal of Monetary Economics* 5 (1979), 133–43.
20. Robert J. Shiller. "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (June 1981), 421–436.
21. ———. "Stock Prices and Social Dynamics." *Brookings Papers on Economic Activity* 2 (1984), 457–498.
22. Lawrence H. Summers. "Does the Stock Market Rationally Reflect Fundamental Values?" *Journal of Finance* 41 (July 1986), 591–601.
23. Kenneth D. West. "Dividend Innovations and Stock Price Volatility." *Econometrica* 56 (January 1988), 37–61.
24. Halbert White. *Asymptotic Theory for Econometricians.* Orlando, Florida: Academic Press, 1984.
25. Jack W. Wilson and Charles P. Jones. "A Comparison of Annual Common Stock Returns: 1871–1925 with 1926–85." *Journal of Business* 60 (April 1987).